Computational Methods for Differential Equations
http://cmde.tabrizu.ac.ir
Vol. 7, No. 4 (Special Issue), 2019, pp. 635-645

Using AHP to complex network analysis tools selection

Rojiar Pir Mohammadiani
Department of Computer Engineering, Faculty of Engineering,
University of Kurdistan, Sanandaj, Iran
E-mail: r.pirmohammadiani@uok.ac.ir

Abstract The analysis of complex networks become more popular through the easily access of huge network data resources in the last years. Researchers have developed techniques and models to help understanding and predicting the behaviour of complex network systems. This advanced analysis is not possible without proper softwares and tools. A large number of tools are available with specific features for analysing and visualizing network systems and we can use a software or a set of suitable tools based on these features and capabilities for the project. Understanding the features of tools and softwares help to achieve better results from network analysis. In this paper, first we review the structure of different types of networks. Based on Wenjun paper, the complex networks are divided into four categories: information networks; social networks; Biological networks and Technological networks [17]. Then we define some functional indicators including: Basic Functionalities, Graph type Support, File Formats Support, Indicator Supports, Visualization Layouts Support, and Community Detection Support. In the next step, by using analytic hierarchical processing (AHP) and truly definable criteria try to evaluate main complex network analysis (CNA) softwares. Eventually, an opportunity is provided using AHP to identify, understand, and evaluate completely four main CNA softwares objectively before identifying and selecting the most efficient CNA software.

Keywords. Complex network analysis, Analysis hierarchical processing.
2010 Mathematics Subject Classification. 65L05, 34K06, 34K28.

1. Introduction

The past decade, ideas from network science have been applied to the analysis of network types. In the context of network theory, Most real graphs have been known as complex networks, display substantial non-trivial topological features, with patterns of connection between their elements that are neither purely regular nor purely random [12]. Analysis of complex networks help us to design scalable and robust communication networks both wired and wireless, and a broad range of other practical issues, develop vaccination strategies for the control of disease, so analyzing and interpreting complex networks is the major subject in complex systems [5]. In this study, we begin by surveying some of the network properties. We will be studying structure of different types in complex network. We can summarize network types with feature such as, Directed, Undirected, Bipartite, Multigraph, Temporal and Labeled. The analyzing software of complex network facilitates quantitative or qualitative analysis by describing features of a network through numerical or visual representation [15]. In this paper, we review tools that are currently available for visualization and analyzing of complex networks. Including tools to visualize large networks, to analyse their topology, to find patterns in data, to study cascade behaviour in networks, and to study clustering, classification, and community structures. By referring to [3, 6, 9, 11, 13, 15], all of the features and attributes of each of these softwares are explained and the most main challenges for researchers along with appropriate software have been indicated. Thus, in this paper, authors aim at utilizing AHP, collecting data as well as systemizing all the opinions and preferred criteria of experts. As a result, they are provided with proper data to make the best selection among different softwares. Additionally, in order to create the comprehensive attitude, six parameters are considered as the required criteria in the analysis of different types of complex networks. These six parameters are Basic Functionalities, Graph type Support, File Formats Support, Indicator Support, Visualization Layouts Support and Community Detection Support. Then different softwares are evaluated by using AHP and some Known criteria. Moreover, we define some functionality indicators then different main tools have been compared base on these indicators. We provide a broad view to choose proper tools for analysis complex network. The article is organize as follows: Section 2 outlines types of networks, section 3 surveys different software tools, section 4 explains indicators to compare tools and finally a conclusion is given in section 5.

2. Related Concepts

2.1. **Type of Networks.** The simplest type of network is only a set of vertices connected by edges; so there are many ways in which networks be more complex than this. For example, there might be more than one distinctive kind of vertex in a network, or more than one distinctive kind of edge. Moreover, vertices or edges may be defined by one or more specifications [12]. Graphs with the directed edges are named directed graphs or sometimes digraphs. Directed graphs can be either cyclic or acyclic. Graphs possibly naturally divided in different ways, bipartite graph for instance. However, networks across a variety of domains show common structure at a qualitative level. Accordingly, we can summarize network types with distinctive features. For example, Directed: network has directed edges, Undirected: network that has undirected edges; Bipartite: bipartite network, Multigraph: network has multiple edges between a pair of nodes, Temporal: for each node/edge we know the time when it showed up in the network, Labeled: network contains labels (weights, attributes) on nodes as well as edges [16] Figure 1 shows various types of neworks.

2.2. **Network in the Real World.** Complex networks are divided into four categories: social networks, technological networks, biological networks and information networks. A. Social Networks may be a set of individuals or group of people with some pattern of common activities. The patterns of friendships between people, business connection among organization, and intermarriages between families, collaboration networks and email communication networks are altogether instances of social

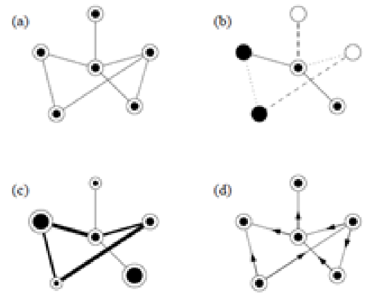

FIGURE 1. Examples of diverse network types: (a) an undirected network with only one type of vertex and a single type of edge; (b) a network with a number of discrete vertex and edge types; (c) a network with varying vertex and edge weights; (d) a directed network in which all edges have a direction.

networks [2, 4, 5]. B. Technological Networks; our second class of networks is technological networks, man-made networks designed typically for distribution of some commodity or resource, such as electricity or information. The electric power grid the network of airline routes and networks of roads, railways and pedestrian traffic are good examples. C. Biological Networks; a number of biological systems can conveniently be considered as networks. Classic example of a biological network is the network of metabolic pathways. D. Information Networks (additionally now and then called "knowledge networks"). The classic example of an information network is an internet network.

2.3. **Complex Network Analyzing Tools.** Networks consist of everything including families, project teams, citation networks, metabolic Networks, membership on networking websites like Twitter or Face book. There are many softwares to analyze networks enabling us to exploit attributes. Different social network analysis softwares and their features are shown in table 1. Since a vast variety of input and output formats are used then different formats are classified as follows.

(1:Graphviz(.dot), 2:Graphlet(.gml), 3:Guess(.gdf), 4:leda(.gml), 5:NetWorkX(.graphml,.net), 6:NodeXL(.graphml,.net), 7:Pajek(.net,.gml), 8:ucinet(.dl), 9:yEd(.gml), 10:gephi(.gexf), 11:edgelist(.csv), 12:databases, 13:Adjacency, 14:Email, 15:csv(text), 16:xsl,xslt(excel), 17:pdf, 18:sonivis(.graphml), 19:tulip(.tlp, .dot), 20:Xml, 21:jar, 22:graph, 23:svg, 24:.png, 25:Matrix, 26:canon,cmap, .eps, .fig, .gd, .gd2, .gif, .gtk, .ico, .imap, cmapx,.ismap,.ps2,.svgz,.tif,.vml,.vmz,.vrml,.wbmp,.xlib,27:bmp,.ps, .jpeg, 28:gnome,dia,graph6/sparse6,29:gelist,ncol,lgl,graphml,dimacs,gml,dot,leda

2.4. **Analytical Hierarchy Process (AHP).** The Analytical Hierarchy Process (AHP) is a decision making tool enabling us to structure the multiple choice criteria

Table 1. different social network analysis softwares and their features

Software Capabilitis	Gephi	SocNetV	R	NodeXL	pajek	graphviz	tulip	Cytoscape	networkx	Igraph
Language implementation	java	C++	R	C #.net	java	C, C++, java	C++	java	python	GNU R Python
Open source	yes	no	yes	yes	no	yes	no	yes	yes	yes
platform	Windows, Linux, MacOs x	Windows, Linux, MacOs	Windows, Linux, Mac	Windows, xp, vista, 7	Windows, Linux Linux	Windows, Mac, Linux	Windows, Mac,	Windows, Linux	Windows, Linux	MS, Windows, Linux, MacOs x, Sun OS,
Input formats	1,2,3, 4,5, 6,7,8,9,10, 11,12,18, 19,	1,5,7,8,11, 13,	Almost all formats	5,7,8,14, 15, 16	8,16,17, some molecular formats,	1 9 ,15	1,7,19,	15,20,21	1.4, 5,7,11,22,	11,
Output formats	3,10,23, 24	5,7,17, 24,25	Most popular formats	5,7,8,15, 16	7,8	17,23, 25,26,27	9,19	17,24,23, 27	1,2, 5,	7,29
The volume of incoming data	Node too big, edge $10 \wedge 6$	Big Data	Big data	Big data	Very big data	Size medium volume 1000 node, big data	Big data	$10 \wedge 7$ node and edge	Big data	Big data
Parallel execution	no	no	no	no	no	no	no	no	no	no
Download	free	free	free	free	free	free	free	free	free	free

into a hierarchy, assessing the relative importance of these criteria, comparing alternatives for each criterion, and determining an overall ranking of the alternatives [14]. The AHP usually involves three stages: decomposition; comparative judgments; and synthesis of priorities. The decomposition principle calls for the construction of a hierarchical network to represent a decision problem according to its most important objective being the first and etcetera. In comparative judgments, users are asked to set up a comparison matrix at each hierarchy by comparing pairs of criteria or sub criteria by ranking their importance. In order to indicate users preference, a scale of values ranging from 1 (indifference) to 9 (extreme most preference) is used. Then, a composite weight is calculated for each alternative after synthesizing the priorities, which in turn are based on derived preferences from the comparison matrix. The knowledge and information of the issue and the views of experts and researchers can provide better and more valuable data for choosing some suitable softwares. A general framework of the proposed approach of ranking softwares in the network analysis are shown in Table 2.

Table 2. A brief summary of all process involved in AHP application

Formulate the decision hierarchy by specifying a hierarchy of interrelated decision elements.
Collect input data by performing a pair wise comparison of the decision elements.
Estimate the relative weights of decision elements by using an eigenvalue method.
Aggregate the relative weights up the hierarchy to obtain a composite weight that represents the Decision maker's opinion of the relative importance of each decision alternative.

3. Softwares Comparison Criteria

As it can be seen in table 1, there exist many softwares for analyzing complex networks each of which is with a specific capability. Six criteria of analyzing different

types of social networks including: Basic Functionalities; Graph type Support; File Formats support; Indicators support; Visualization layouts support; Community detection support , are explained here .
Firstly, Basic Functions point to general attributes of software such as Platform, License, Expectable Computing Time, Tractable Number of Nodes, Time to load 10^5 Nodes and Time to Load 10^6 Edges. File Formats Support is another criterion that indicates which software creates required highlights from raw network data. The raw network data are organized in an edge list, nearness list or adjacency matrix (likewise called socio matrix) and joined with (individual/node level) property data regularly. However, most of network analysis softwares utilize a plain text ASCII data format; some software packages have additional feature of using relational databases to import and store network features [9]. Moving towards data integration introduced a number of common file formats and standard languages for storing information. Datasets that are stored in a standardized format can easily be incorporated into a tool that supports the same format without any need for reprogramming or comprehension of the file format. Common file formats are RDF, csv, .net, garphml. Other commonly used file formats and standard languages are open-source XML-based languages most notably BioPAX, SBML and PSI-MI, rely on levelled approaches, meaning that they contain various levels of complexity and specificity [15, 3].Therefore, one of the other criteria to select proper software by researchers is format of backup files, which is created by different software.
Third, Graph type Support, as we expressed in part 1-2, there are distinctive types of network such as Temporality, Two-mode graphs (bipartite graphs), Multi-relational graphs; which is so important that which one of softwares is much more compatible with these graphs.
Forth, Indicator based network description that many quantitative indicators have defined on networks [16]. The descriptors at the network level have used for comparing the proportion of vertexes to edges, or evaluating properties of the graph like the randomness or small world distributions. On the other hand, the descriptors at the node level are valuable for recognizing the nodes strategically set in the network or featuring those that take vital part in communication for example bridges or hubs. The Centrality measures of node-based analytics are much more essential and increasingly current centrality measures are Degree centrality; Closeness centrality; Betweenness centrality; PageRank and HITS. The number of indicators, which is calculated by a software, is also one of the criteria affecting the selection process of appropriate software by researchers.
With respect to the fifth criterion, Visualization, Tools for visualization and analysis of complex networks are becoming pivotal in the researches on complex networks, shifting from data generating experimental stage to the data analysis and visualization stage. Representing visually can help researchers to explore and find interesting features in networks.
Widely used and well-documented GUI packages include Pajek (freeware), GUESS, InfoVis Cyber infrastructure, Gephi, UCINet, GUESS, ORA, and Cytoscape. Private GUI packages directed at business customers include Orgnet. In addition, Other CNA platforms such as Medusa, Cytoscape, BioLayout Express3D, Osprey, ProViz,

Ondex, PATIKA, and PIVOT have been specifically developed for biotechnological networks [1, 7, 9].
Visual representations of social networks are important to understand network data and convey the analysis results. Visualization often facilitates qualitative interpretation of network data as well. Regarding visualization, network analysis tools are used to change the layout, colours, size and other properties of the network representation [13]. Among all Visualization algorithms, Fruchterman Reingold, is a widely-used force-based algorithm for graph visualization [8]. Another alternative is Kamada-Kawai algorithm [10], which has a faster convergence compared to Fruchterman Reingold, but its results are not as good as Fruchterman Reingold. It can be envisaged to use Kamada-Kawai to calculate the first placement of the vertices. These two methods are categorized in a group of algorithms called spring algorithms.
Clustering or Community Detection -the sixth parameter- refers to clustering, as a method of detecting groups of nodes with dense connections within the groups and sparser connections between the groups. Generally, two main approaches can be distinguished from different methods of detecting communities; Hierarchical approach and Partitioned clustering. With respect to Hierarchical approach, the nodes are aggregated in a hierarchy of clusters from the discrete partition to the whole network, while the partitioned clustering consists in directly dividing the network into a pre-defined number of groups. The six defined criteria, which explained above, have shown in the figure 2.
Therefore, to select appropriate software to analyse networks, the main key is identifying and prioritizing considerable criteria.
Therefore, to select appropriate software to analyse networks, some functionality indicators are defined in this paper, for comparing different main tools. Then, we try to use analytic hierarchical processing techniques to collect and systematize experts opinions and criteria to make much more suitable data to decide on softwares. To create a comprehensive view, all the features and criteria should be considered for analyzing various types of social networks. A definable criteria and AHP are applied to evaluate a software properly leading a better view for selecting appropriate tools to analyze complex networks.

Network analysis softwares generally consists of either packages which are dependent on graphical user interfaces (GUIs), or packages which are used for scripting/ programming languages. It is simpler to learn GUI packages but scripting tools are more dominant and extensible. Based on the reports all around, commonly scripting tools utilized for network analysis include NetMiner with Python scripting engine, the statnet suite of packages for the R statistical programming language, igraph, which has packages for R and Python, the NetworkX library for Python. In spite of being hard to learn, these open source packages are developing as fast as functionality and features than secretly maintained softwares, and extensive documentation and tutorials are accessible [1, 9, 11].

We consider commonly network analysis tools based on scripting/programming languages including igraph, networkX and software-based graphical user interfaces (GUIs) including Pajek and Gephi as decision alternative in AHP model.

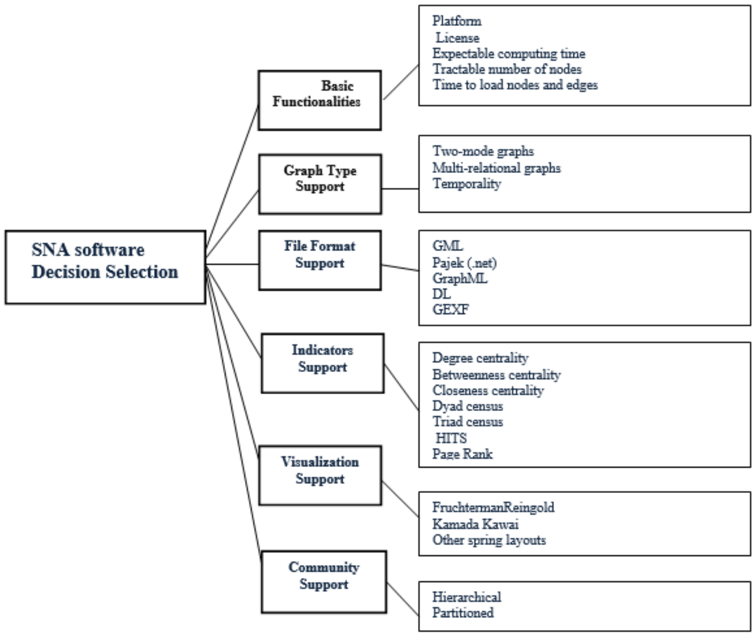

Figure 2. Criteria for comparison social network analysis softwares

Table 3. Basic functionalities of selected softwares

Software Version	Pajek 1.26	Gephi 0.7 alpha	NetworkX 0.6	Igraph 0.53
Type	Stand-alone Software	Stand-alone Software	Library	Library
Platform	Windows	Java	Paython	R/Paython/C libraries
License	Free for non-commercial use	GNU GPL	BSD License	GNU GPL
Expectable computing time	Fast(C)	Medium (Java)	Fast (C, Python)	Fast (C)
Tractable number of nodes	500,000 nodes	150,000 nodes	1,000,000 nodes	>1.9 milion relations (without attributes)
Time to load 10^5 nodes and 10^6 edges	24 seconds	40 seconds	137 seconds	11seconds

4. Applying the AHP Approach

Based on the process involved in AHP application in the first step, we defined six main criteria including: Basic Functionalities, Graph type Support, File Formats

Support, Indicator Supports, Visualization Layouts Support, and Community Detection Support for selecting four main network analysis tools. Then after the exciting criteria is evaluated by using pairwise comparisons. In order to compare the criteria, numerical judgment questionnaires have been used to gather the opinions of 6 experts. The preference of the criteria is indicated by a number which can be chosen subjectively between 1 to 9 where 1 denotes the least importance and 9 denotes the highest degree of favouritism. The possible judgments and their respective degree of importance are shown in Table 4. Table 5 shows pairwise comparison matrix for selection decision corresponding to one of these experts.

Table 4. Pairwise comparison judgments between element X and element Y

Judgment	Value
X is equally preferred to Y	1
X is equally to moderately preferred over Y	2
X is moderately preferred over Y	3
X is moderately to strongly preferred over Y	4
X is strongly preferred over Y	5
X is strongly to very strongly preferred over Y	6
X is very strongly preferred over Y	7
X is very strongly to extremely preferred over Y	8
X is extremely preferred over Y	9

Table 5. Pairwise comparison matrix for selection decision

Selection decision	Community detection support	Graph type support	Indicator support	Visualization layouts support	Basic functionalities	File formats support
Community detection support	1	1.3	1.4	1.3	5	7
Graph type support	3	1	1.3	1.4	5	6
Indicator support	4	3	1	2	7	8
Visualization layouts support	3	4	1.2	1	6	8
Basic functionalities	1.5	1.5	1.7	1.6	1	5
File formats support	1.7	1.6	1.8	1.8	1.5	1

The importance of each criterion is determined according to the matrix of comparison. The steps of AHP algorithm is shown in Figure 3.

Read decision variable $a_1, a_2, \cdots a_n$

produce pairwise comparison (pwc) matrix

Suppose that we have four complex network analysis tools, and six criteria for se-

loop k 1 to n-1
if imp(n) > imp(n-1) then
pwc (n⟶n-1) =r (r=2,$\cdots$,9)
else
if imp(n) =imp(n-1) then
pwc (n⟶n-1) =1
else
c=1/r
end if
end if
pwc (n⟶n) =1
endloop
endloop
, compute eigenvalues and eigenvector
Loop i = 1 to n
Loop j= 1 to n
Sum(j)=sum $(a_{1j}, a_{2j}, a_{3j}, , a_{nj})$
E(ij)=a_{ij}/ sum(j) (E=Eigenvalues)
End loop
$\lambda(i)$ = avg $(a_{i1}, a_{i2}, a_{i3}, \cdots, a_{in})$ (λ = eigenvector)

FIGURE 3. Steps of AHP algorithm

lecting the best analysis tool.

First of all we need to determine the importance of the selected criteria followed calling the comparison matrix to Expert Choice software and then calculating the corresponding relative weights for the alternative elements with respect to the criterion element by using this AHP software. Once all the relative weights have been calculated, a composite weight is determined, for each decision choice. Then, we collected and computed the related external metrics of the selected characteristics (Basic Functionalities Graph type Support, File Formats support, Indicators support, Visualization layouts support, Community detection support) for each of the four complex network analysis tools, and finally we can ranking of the 4 alternative complex network analysis tools. The results are shown in Table 7.

TABLE 6. Importance (relative weights) of variables in influencing final CAN selection

Selection decision	Member 1	Member 2	Member 3	Member 4	Member 5	Member 6
Community detection support	0.121	0.201	0.168	0.151	0.181	0.182
Graph type support	0.161	0.152	0.191	0.130	0.140	0.211
Indicator support	0.354	0.293	0.270	0.283	0.304	0.281
Visualization layouts support	0.273	0.221	0.250	0.204	0.278	0.223
Basic functionalities	0.056	0.0857	0.107	0.121	0.0219	0.064
File formats support	0.0247	0.289	0.123	0.116	0.62	0.040

TABLE 7. The Ranking of the four complex network analysis tools.

complex network analysis tools	Ranking
Pajek	2
gephi	3
NetworkX	4
igraph	1

5. Conclusion

Social network analysis has shownto be a powerful method for understanding the importance of relationships in networks. Complex network analysis software facilitates quantitative or qualitative analysis of complex networks by describing features of a network either numerical or visual representation.

The tools represented in this review are applicable to a wide range of problems and their distinct features make them suitable for a wide range of applications. Most of the tools discussed in this review can cope with datasets of up to around 5000 nodes without compromising too extremely on speed and ease of use. In fact, we should consider the features of networks and capabilities of various network tools before selecting a software or set of tools could be more suitable to our project. Therefore, indicators defined here help to select proper tools for our project. Thus, corresponding to different types of network, we should use knowledgeable experts in distinctive fields to determine appropriate types of soft- ware to analyze different kinds of social networks.

References

[1] H. Agrawal, A. Thakur, R. slathia and K. Sumangali, *A Comparative Analysis of Social Networking Analysis Tools*, School of information technology and Engineering, VIT University Vellore, Tamilnadu, 632014, India, *5* (2015). DOI 10.4172/2165-7866.1000157.

[2] N. Akhtar, *Social Network Analysis Tools, Fourth International Conference on Communication Systems and Network Technologies*, (2014), 388–392. DOI 10.1109/CSNT.2014.83.

[3] I. A. Apostolato, *An Overview of Software Applications for Social Network Analysis*, Department of Sociology University of Bucharest, *3* (2013), 71–77.

[4] V. Batagelj, and A. Mrvar *Some analyses of Erdos collaboration graph*, Social Networks *22* (2000), 173–186.

[5] A. L. Barabasi, H. Jeong, E. Ravasz,Z. Neda, A. Schuberts, and T. Vicsek, *Evolution of the social network of scientic collaborations*, Physic A, *3* (2014), 590–614.DOI 10.1016/S0378-4371(02)00736-7.

[6] D. Combe, Ch. Largeron, E. Egyed-Zsigmond and M. Géry. *A comparative study of social network analysis tools*, Web Intelligence & Virtual Enterprises, *1* (2010) 1–13

[7] S. Fortunato, *Community detection in graphs*, Physics Reports, *486* (2010) 75–174.

[8] T. M.J. Fruchterman and E. M. Reingold, Graph Drawing by Force-directed Placement, Software: Practice and Experience, *21* (1991), 1129–1164.

[9] N. Jokar, *The review of social networks analysis tools*, *85* (2016), 329–339.

[10] T. Kamada and S. Kawai, *An algorithm for drawing general undirected graphs*, Information processing letters, *31* (1989), 7–15.

[11] A. Nadeem, *Social Network Analysis Tools*, Department of Computer Engineering, Zakir Hussain College of Engineering & Technology Aligarh Muslim University, India, IEEE, 978-1-4799-3070 (2014), 8–14.

[12] M. E. J. Newman, *The structure and function of complex networks*, SIAM Rev.

[13] P. Shah and R. Mehta, *Optimizing Graph Mining Attributes based Inference for Social Network Analysis, International Journal of Innovative Research in Computer and Communication Engineering*, *4* (2016).

[14] M. Vedanayaki, *Graph Mining Techniques,Tools and Issues-A study*, Department of Mathematics, Amet University ,Kanthur, Chennai -603 112, *7* (2014), 188–190.

[15] J. Sussin, and N. Drakos, *Hype Cycle for Social Software*, Gartner, (2008).

[16] S. Wasserman, and K. Faust, *Social Network Analysis*, Cambridge University Press, Cambridge (1994).

[17] Z. Wenjun, Fundamentals of network theory, World scientific, (2003).