The Journal of Open Source Software

CNATool - Complex Network Analysis Tool

Roberto Luiz Souza Monteiro [1,2¶], **Renata Souza Freitas Dantas Barreto** [2], **Andréia Rita da Silva** [2], **Alexandre do Nascimento Silva** [2,3], **José Roberto de Araújo Fontoura** [2], **Marcos Batista Figueredo** [2], and **Hernane Borges de Barros Pereira** [1,2]

1 SENAI CIMATEC University Center, Brasil
2 Universidade do Estado da Bahia, Brasil
3 Universidade Estadual de Santa Cruz, Brasil
¶ Corresponding author

DOI: 10.21105/joss.05373

Software

- Review
- Repository
- Archive

Editor: Charlotte Soneson

Reviewers:

- @tomalrussell
- @ati-ozgur

Submitted: 14 January 2023
Published: 04 August 2023

Summary

CNATool is an innovative online program designed for analyzing complex and social networks. It provides researchers with a convenient and streamlined approach to network graph analysis, accessible from any Internet-connected device. With support for Pajek and JSON formats, CNATool offers a range of features, including the creation of artificial networks, customizable layout algorithms, and detailed visualization of network properties. Researchers can explore essential network properties such as average degree, density, clustering coefficient, shortest path, diameter, and efficiency. CNATool also provides in-depth analysis of vertex properties, enables graph saving in multiple formats (Pajek, JSON, SVG), and generates comprehensive HTML reports. The tool caters to researchers involved in social and biological network analysis, offering a user-friendly platform for exploring network properties and uncovering patterns. Potential applications include studying social networks, analyzing time-varying networks, investigating functional architecture, examining memory deterioration, exploring the impact of centrality, analyzing accident risks, modeling opinion diffusion, and studying collaborative networks among criminal members. CNATool empowers researchers to delve deeper into the intricacies of complex networks, facilitating meaningful insights and discoveries.

Introduction

The analysis of complex networks has gained significant importance across various domains, including social, biological, and technological fields. Many phenomena in these domains can be represented as network diagrams, enabling visual analysis and leveraging tools developed for graph theory applications (Newman, 2003). Researchers have utilized network theory and computational tools to explore cooperation and collaboration processes in local productive arrangements (Barros Pereira et al., 2007), study collaboration between researchers (Vieira et al., 2016), analyze the evolution of species in affinity networks (Monteiro et al., 2014), identify linkages between stock exchanges (Area Leão Pereira et al., 2019), monitor coupled risks (Zhou et al., 2019), and more.

Health studies have also benefited from network theory, with research on the structure and functions of the brain (Bullmore & Sporns, 2009) and analysis of the visual cortex in monkeys and cats (Hilgetag et al., 2000).

These examples demonstrate the wide range of applications and tools developed for social and complex network analysis. Various software programs, including CNATool, offer researchers a convenient and streamlined approach to network graph analysis, enabling efficient exploration of network properties and uncovering patterns and insights.

CNATool: Innovative Network Analysis Program

CNATool is an innovative online program designed specifically for analyzing complex and social networks. With its user-friendly interface and accessibility from any internet-connected device, CNATool provides researchers with a convenient and streamlined approach to network graph analysis. Key Features:

Graph Visualization: CNATool allows the visualization of graphs, from the graphical interface, thanks to the sigma.js (Sciences-Po médialab and OuestWare, 2023) library, including resources such as dynamic adjustment of the layout of the graph, either through algorithms such as ForceAtlas2, or manually dragging and positioning the vertices.

Graph File Formats: CNATool supports Pajek and JSON formats, allowing researchers to import network graphs effortlessly.

Artificial Networks: Researchers can create artificial networks within CNATool, enabling customizable experimental analysis.

Customizable Layout Algorithms: CNATool offers customizable layout algorithms, facilitating the visualization of network structures.

Network Property Visualization: Detailed visualization of network properties, such as average degree, density, clustering coefficient, shortest path, diameter, and efficiency, helps researchers understand network characteristics.

Vertex Analysis: CNATool provides in-depth analysis of vertex properties, enabling researchers to gain insights into individual elements within the network.

Export and Reports: Researchers can save graphs in multiple formats, including Pajek, JSON, and SVG, and generate comprehensive HTML reports for documentation purposes.

Statement of Need

Newman (2003) presents the main concepts involved in complex and social network analysis. The author discusses the types of networks, topologies, local and global properties. Regarding the types of networks, Newman highlights social, informational, technological and biological networks. These networks, despite having different natures, present common properties such as number of vertices, number of edges, density (Chatterjee & Sinha, 2007; Pereira et al., 2016), average degree, average clustering coefficient (Schank & Wagner, 2005), average shortest path (Johnson, 1977), diameter (Razzaque et al., 2008) and efficiency (Latora & Marchiori, 2001). And even at micro scale, similar parameters are observed, highlighting the clustering coefficient and the closeness (Bhardwaj et al., 2011; Freeman, 1978; Freeman et al., 1979) and betweenness (Barthelemy, 2004; Brandes, 2001; Curado et al., 2022; Freeman, 1977) centralities. With regard to topologies, networks of apparently different natures, such as social and biological, often present phenomena common to small-world (Bakshy et al., 2011; Emmert-Streib, 2006; Marchiori & Latora, 2000; Watts & Strogatz, 1998) and scale-free networks (Barabasi et al., 2002; Crucitti et al., 2003).

Based on this view of the literature, we designed the mind map presented in Figure 1. Moreover, CNATool implements some properties not found in other software, for example incidence-fidelity index (Teixeira et al., 2010). Table 1 presents a summary of the main features presented by each of the analyzed programs. An important observation is that the Cytoscape program, although it does not allow its execution directly on the Web, as an application, it allows, through the Cytoscape.js (National Institute of General Medical Sciences (NIGMS), 2023) library, the creation of Web programs for visualization and analysis of network graphs. Other popular and powerful software such as igraph (The igraph core team, 2023) and NetworkX (NetworkX developers, 2023) were not compared as they are not user applications but programming libraries.

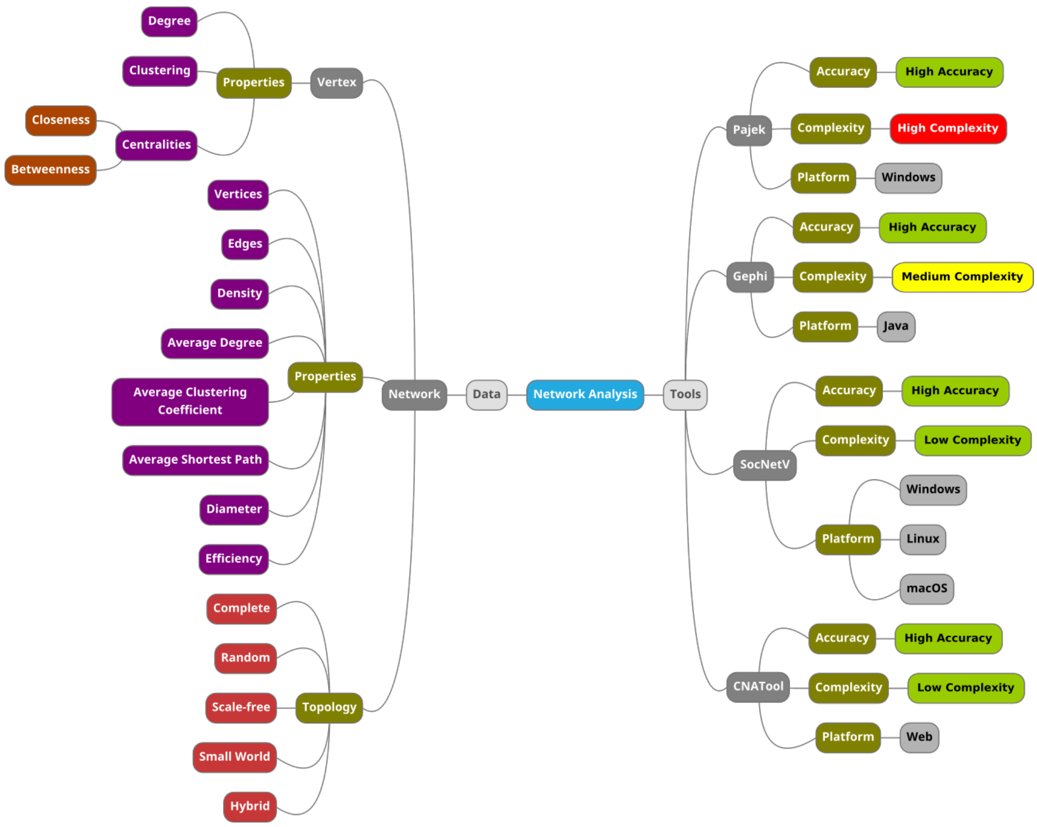

Figure 1: Mind map of concepts involving analysis of complex networks. For calculation accuracy and interface complexity, green means desirable feature, yellow means acceptable feature, and red means undesirable feature.

This program is intended for researchers working in the analysis of social and biological networks. Knowledge of network analysis and familiarity with scripting languages is essential.

Table 1: Summary of the main features presented by each of the analyzed programs.

Feature	CNATool	Gephi	Pajek	SocNetV	Cytoscape
Has high accuracy	Yes	Yes	Yes	Yes	Yes
Is user-friendly	Yes	Yes	No	No	Yes
Runs on Windows	Yes	Yes	Yes	Yes	Yes
Runs on Linux	Yes	Yes	No	Yes	Yes
Runs on macOS	Yes	Yes	No	Yes	Yes
Runs on Android	Yes	No	No	No	No
Runs on iOS	Yes	No	No	No	No
Runs direct on web	Yes	No	No	No	Yes*
Calculates Efficiency	Yes	No	No	No	No
Calculates Incidence-fidelity	Yes	No	No	No	No

Feature	CNATool	Gephi	Pajek	SocNetV	Cytoscape
Uses the GPU to speed up calculations	Yes	No	No	No	Yes

Limitations

As a minor limitation, CNATool does not support the GEXF and GraphML formats, however, the tools that support these formats are generally able to export graphs in Pajek (.NET) format, allowing the interoperability of CNATool with these softwares.

Conclusions

The CNATool tool was developed having in mind the analysis of complex and social networks from any device connected to the Internet. This application offers a friendly and intuitive user interface, while providing accurate results and detailed reports of global and local network properties being analyzed.

The program allows exporting of results in some common file formats, like Pajek and JSON and also provides a command line tool that allows batch processing, contributing to speed up analysis processes when an experiment requires calculation of properties of a large number of networks.

References

Area Leão Pereira, E. J. de, Ferreira, P. J. S., Silva, M. F. da, Miranda, J. G. V., & Pereira, H. B. B. (2019). Multiscale network for 20 stock markets using DCCA. *Physica A: Statistical Mechanics and Its Applications*, *529*, 121542. https://doi.org/10.1016/j.physa.2019.121542

Bakshy, E., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011). Everyone's an influencer: Quantifying influence on twitter. *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, 65–74. https://doi.org/10.1145/1935826.1935845

Barabasi, A.-L., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., & Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica A: Statistical Mechanics and Its Applications*, *311*(3-4), 590–614. https://doi.org/10.1016/S0378-4371(02)00736-7

Barros Pereira, H. B. de, Freitas, M. C., & Sampaio, R. R. (2007). Fluxos de informação e conhecimentos para inovações no arranjo produtivo local de confecções em salvador, bahia. *DataGramaZero – Revista de Ciência Da Informação*, *8*(4).

Barthelemy, M. (2004). Betweenness centrality in large complex networks. *The European Physical Journal B*, *38*(2), 163–168. https://doi.org/10.1140/epjb/e2004-00111-4

Bhardwaj, S., Niyogi, R., & Milani, A. (2011). Performance analysis of an algorithm for computation of betweenness centrality. In B. Murgante, O. Gervasi, A. Iglesias, D. Taniar, & B. O. Apduhan (Eds.), *Computational science and its applications - ICCSA 2011* (pp. 537–546). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-21934-4_44

Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, *25*(2), 163–177. https://doi.org/10.1080/0022250X.2001.9990249

Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, *10*, 186–198. https://doi.org/10.1038/nrn2575

Chatterjee, N., & Sinha, S. (2007). Understanding the mind of a worm: Hierarchical network structure underlying nervous system function in C. elegans. In R. Banerjee & B. K. Chakrabarti (Eds.), *Models of brain and mind* (Vol. 168, pp. 145–153). Elsevier. https://doi.org/10.1016/S0079-6123(07)68012-1

Crucitti, P., Latora, V., Marchiori, M., & Rapisarda, A. (2003). Efficiency of scale-free networks: Error and attack tolerance. *Physica A: Statistical Mechanics and Its Applications*, *320*, 622–642. https://doi.org/10.1016/S0378-4371(02)01545-5

Curado, M., Rodriguez, R., Tortosa, L., & Vicent, J. F. (2022). Anew centrality measure in dense networks based on two-way random walk betweenness. *Applied Mathematics and Computation*, *412*, 126560. https://doi.org/10.1016/j.amc.2021.126560

Emmert-Streib, F. (2006). Influence of the neural network topology on the learning dynamics. *Neurocomputing*, *69*(10-12), 1179–1182. https://doi.org/10.1016/j.neucom.2005.12.070

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 35–41. https://doi.org/10.2307/3033543

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215–239. https://doi.org/10.1016/0378-8733(78)90021-7

Freeman, L. C., Roeder, D., & Mulholland, R. R. (1979). Centrality in social networks: II. Experimental results. *Social Networks*, *2*(2), 119–141. https://doi.org/10.1016/0378-8733(79)90002-9

Hilgetag, C., Burns, G. A. P. C., O'Neill, M. A., Scannell, J. W., & Young, M. P. (2000). Anatomical connectivity defines the organization of clusters of cortical areas in the macaque and the cat. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *355*(1393), 91–110. https://doi.org/10.1098/rstb.2000.0551

Johnson, D. B. (1977). Efficient algorithms for shortest paths in sparse networks. *Journal of the ACM*, *24*(1), 1–13. https://doi.org/10.1145/321992.321993

Latora, V., & Marchiori, M. (2001). Efficient behavior of small-world networks. *Physical Review Letters*, *87*, 198701. https://doi.org/10.1103/PhysRevLett.87.198701

Marchiori, M., & Latora, V. (2000). Harmony in the small-world. *Physica A: Statistical Mechanics and Its Applications*, *285*(3-4), 539–546. https://doi.org/10.1016/S0378-4371(00)00311-3

Monteiro, R. L. S., Fontoura, J. R. A., Carneiro, T. K. G., Moret, M. A., & Pereira, H. B. B. (2014). Evolution based on chromosome affinity from a network perspective. *Physica A: Statistical Mechanics and Its Applications*, *403*, 276–283. https://doi.org/10.1016/j.physa.2014.02.019

National Institute of General Medical Sciences (NIGMS). (2023). *Cytoscape* (Version 3.10.0). https://cytoscape.org

NetworkX developers. (2023). *NetworkX* (Version 3.1). https://networkx.org

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*, *45*(2), 167–256. https://doi.org/10.1137/S003614450342480

Pereira, H. B. B., Fadigas, I. S., Monteiro, R. L. S., Cordeiro, A. J. A., & Moret, M. A. (2016). Density: A measure of the diversity of concepts addressed in semantic networks. *Physica A: Statistical Mechanics and Its Applications*, *441*, 81–84. https://doi.org/10.1016/j.physa.2015.08.024

Razzaque, A., Hong, C. S., Abdullah-Al-Wadud, M., & Chae, O. (2008). A fast algorithm to calculate powers of a boolean matrix for diameter computation of random graphs. *International Workshop on Algorithms and Computation*, *4921*, 58–69. https://doi.org/10.1007/978-3-540-77891-2_6

Schank, T., & Wagner, D. (2005). Approximating clustering coefficient and transitivity. *Journal of Graph Algorithms and Applications*, *9*(2), 265–275. https://doi.org/10.7155/jgaa.00108

Sciences-Po médialab and OuestWare. (2023). *Sigma.js* (Version 1.2.1). https://www.sigmajs.org

Teixeira, G. M., Aguiar, M. dos S. F. de, Carvalho, C. F. de, Dantas, D. R., Cunha, M. do V., Morais, J. H. M. de, Pereira, H. B. de B., & Miranda, J. G. V. (2010). Complex semantic networks. *International Journal of Modern Physics C*, *21*(03), 333–347. https://doi.org/10.1142/S0129183110015142

The igraph core team. (2023). *Igraph* (Version 0.10.6). https://igraph.org

Vieira, R. P., Monteiro, R. L. S., Pereira, H. B. B., Andrade, J. B. de, & Guarieiro, L. L. N. (2016). Redes de colaboração científica do INCT de energia e ambiente. *Revista Virtual Química*, *8*(4), 1234–1248. https://doi.org/10.21577/1984-6835.20160088

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature*, *393*(6684), 440–442. https://doi.org/10.1038/30918

Zhou, Y., Li, C., Ding, L., Sekula, P., Love, P. E. D., & Zhou, C. (2019). Combining association rules mining with complex networks to monitor coupled risks. *Reliability Engineering & System Safety*, *186*, 194–208. https://doi.org/10.1016/j.ress.2019.02.013