Online at: https://researchtrendsjournal.com ISSN No: 2584-282X

INTERNATIONAL JOURNAL OF TRENDS IN EMERGING RESEARCH AND DEVELOPMENT

Volume 1; Issue 1; 2023; Page No. 160-166

Received: 01-10-2023
Accepted: 06-11-2023

Social network analysis from graph theory

[1]Gunamma Prasadarao and [2]Dr. Manish KR Singh

[1]Research Scholar, Sunrise University, Alwar, Rajasthan, India
[2]Associate Professor, Sunrise University, Alwar, Rajasthan, India

DOI: https://doi.org/10.5281/zenodo.12663089

Corresponding Author: Gunamma Prasadarao

Abstract

Research in social network analysis, which draws on mathematical concepts from graph theory and theoretical frameworks from sociology, focuses on the role of structure in health outcomes. Relationship patterns among persons, communities, and organisations are what we mean when we talk about structure. Social network analysis is based on the central idea that the features and structure of the network substantially impact the outcome of interest.

Keywords: Networks, graph theory, social, mathematical, sociology

Introduction

The use of social networking sites like LinkedIn, Facebook, Twitter, and Instagram has skyrocketed in the past few years, and these sites have quickly become ingrained in our daily lives. In addition, social media platforms' importance in the information cycle has expanded in the last several years. Big Data and Business Intelligence settings are so required to process the massive amounts of data. Prescriptive data for businesses (like data gathered from graph theory) can be obtained in this way. One definition of social networks is the web of ties that people create and maintain with one another.

Every point, or node, in this diagram stands for a different individual, and every edge, or line of connection, shows how the friendship between any two of those people came to be. Thus, we can investigate patterns and increase business intelligence with the help of graph theory.

Applying mathematical concepts from graph theory and theoretical frameworks from sociology, social network analysis investigates how structure affects health. Relationship patterns among persons, communities, and organisations are what we mean when we talk about structure. It is believed that the structure of the network and its attributes greatly affect the outcome of interest when social network analysis is conducted.

Materilas and Methods

Amina Adadi *et al*. (2018) [1] The rapid and pervasive integration of AI into our everyday lives is ushering in the fourth industrial revolution and hastening the transition to a society that is increasingly dependent on algorithms. A major problem with AI-based systems is their lack of transparency, which persists despite these remarkable improvements. True, these systems' black-box nature permits strong forecasts, but it defies direct explanation. A fresh discussion on explainable AI (XAI) has been sparked by this matter. Improving the trustworthiness and openness of AI-powered systems is an area of study that shows great potential. In order for AI to keep making uninterrupted progress, this is considered a necessary condition. If you are a researcher or practitioner interested in the emerging field of XAI, this survey is a great place to start learning about the basics. We survey previous works on the subject, talk about current trends in the field, and outline the main avenues for future study by analysing the literature.

Svitlana Popereshnyak *et al*. (2021) [2] This study takes into account the most common kinds of social networks as well as techniques for analysing them. In this social network analysis, we look at the different kinds of links and the issues with them. It was determined that there is a connection between social network analysis and graph theory. It was created as a standard database. A model for the suggested approach to cyber-tie analysis was constructed using the ideas of "centrality" and "intermediate centrality" from graph theory.

G.H.J. Lanel *et al*. (2020) [3] This study intends to provide a descriptive analysis of LinkedIn, Facebook, and Twitter's online social media networks using concepts from graph theory. Part one of the research involves cataloguing the aforementioned networks' activities and building a proper graph model to depict such activities on each social media platform. Part two of this research will focus on assessing the built models of each social media platform independently to find out how well they can analyse user behaviour and traits. This study concludes by suggesting a way to share user-generated content from online social networks with third parties in a way that respects their privacy and allows them to make informed decisions.

Alvida Mustika Rukmi *et al*. (2022) [4] A COVID-19 pandemic is threatening the globe in 2020. To combat its spread, the disciplines of epidemiology and networks are essential. When it comes to preventing the transmission of contagious diseases, tracing individuals through contact is crucial. The network of contacts illustrates the potential pathways for the spread of sickness. Social network analysis refers to the study of structural structures employing graph theory with degree, betweenness, and closeness centralities. A resident cluster's contact network is described in this study using the SNA graph. You can learn about the cluster's contact relationship structure by looking at the created SNA graph. Using the SNA method's centrality assessment, we can see who's really connected and close to one other in the cluster. The node with the highest degree of centrality is Node-F391, which has 3,203 interactions. This suggests that the node is most active in its social interactions with nearby residents in the 14 days preceding and following the day of the COVID-19 swab test. With a Betweenness Centrality score of 0.006, nodes B041 and C371 stand out as key players. An significant link is formed between the two nodes. So that contact tracing based on illness path mapping may be carried out, graphs on SNA allow visualisation of the contact network to differentiate potential links between close contacts. Healthcare providers can benefit from graph visualisation on the SNA as an auxiliary tool in contact tracing investigations, as opposed to the traditional method of evaluating individual contact records.

Dr. P K Srimani *et al*. (2012) [5] In this research, we explore the graph theory method for building and proving a learning model that optimises elementary school students' mathematical pathways. Data pertaining to the mathematical concepts a child needs to learn at the elementary level (Class I to VII) can be represented using Concept Flow Graphs by rearranging nodes according to the learning progression, partitioning the graphs into subgraphs to represent levels of learning, optimising the sub-graphs using merging and elimination technique, and identifying / marking the optional nodes. The next step is to optimise these graphs using graph theory algorithms and techniques. The architecture of the framework is based on the graph theoretical method, which is proven by applying the Networks methodology. This form serves as the basic Mathematical Pathway driver for the Learning model. The approach is novel, and the taught model is highly accurate.

Network Theory

A network's nodes and edges are its building blocks, therefore let's start with a quick overview of them.

Nodes Nodes A, B, C, D, and E in the network may possess both network-based and self-properties, such as size, location, weight, and cluster. These properties provide information about the nodes and their relationships within the network.

Edges Illustrate the links between the nodes, which may also include additional attributes (such as weight to indicate the connection's strength, direction for asymmetric relationships, or time if relevant).

These two building blocks can characterise a vast array of phenomena, such as biological linkages, road networks, social connections, virtual routing networks, physical electricity networks, and countless more.

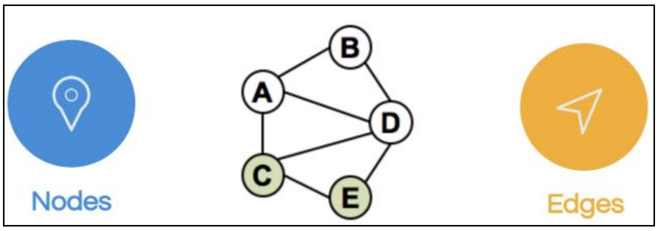

Fig 1: Example network and components

Real-world networks

A real-world network, particularly a social network, differs structurally from a random mathematical network. Figure 2, taken from, shows elaborate networks.

- **Small World phenomenon:** Asserts that in practice, most networks do not have very long pathways connecting any two nodes. This holds true for both online and offline social networks (the six handshakes idea), including physical networks such as airports or the efficacy of web-traffic routing.
- **Scale Free:** In networks with a power-law degree distribution, there is an imbalance between the number of nodes with strong connections (such as social influences) and the number of nodes with weak connections.
- **Homophily**: The propensity for people to form relationships with those who have their characteristics, leading to a proliferation of identical homes in the surrounding area.

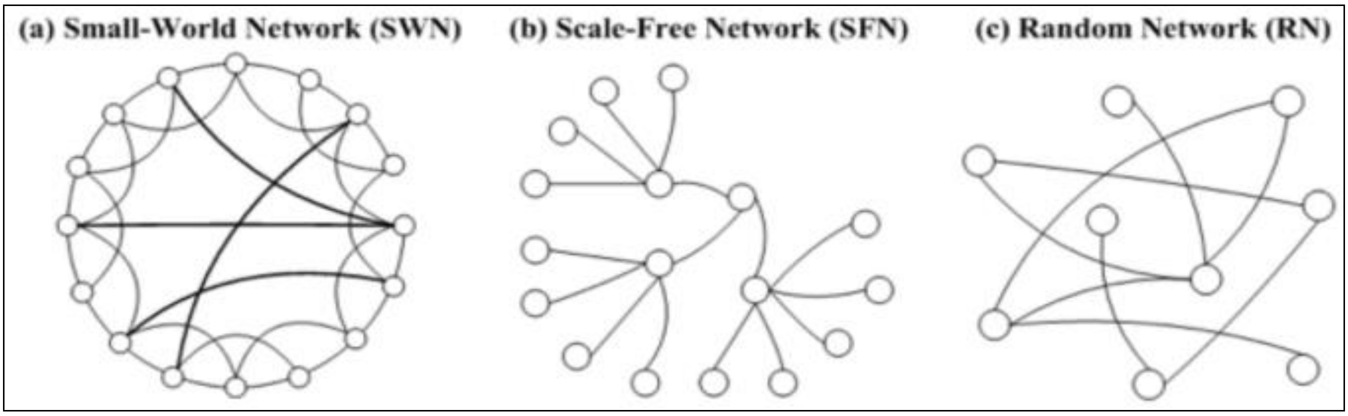

Fig 2: Complex networks

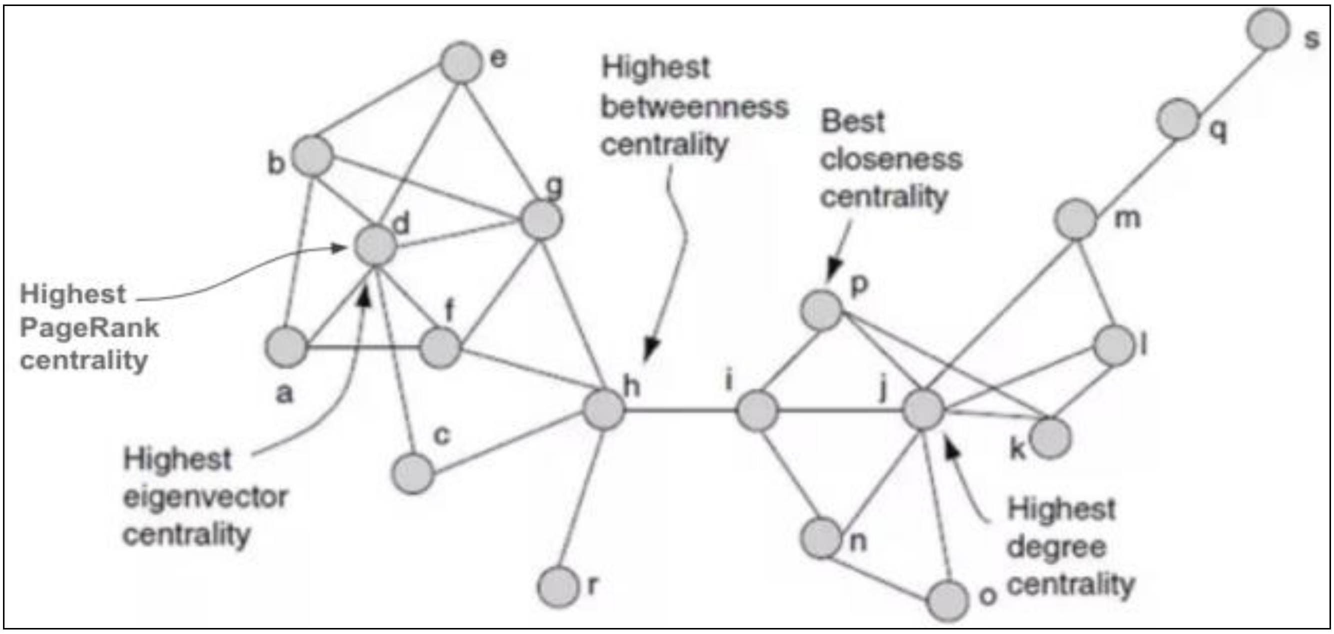

Fig 3: Illustration of various centrality measures

Centrality Measures

Many of a network's dynamics revolve around its most important nodes. The relevance and definition of centrality can differ from one instance to the next, as illustrated in Figure 3, and there are numerous centrality measures that can be employed.

- Degree - the amount of neighbors of the node
- Eigen Vector / PageRank - iterative circles of neighbors
- Closeness - the level of closeness to all of the nodes
- Betweenness - the amount of short path going through the node

Diverse metrics have their uses; for example, they can improve search engine optimisation (page rank), identify crucial points (betweenness), and locate transit hubs (closeness).

Building a Network

With the ability to specify the relations between nodes, networks can be built from a variety of datasets. Here we will demonstrate how to use the Python network x module to construct and display a network representing the 2018 Eurovision votes (using official data). Using a Pandas data frame, we can extract the vote tallies from an Excel spreadsheet. We will melt the dataset such that each row contains all of a country's votes in order to guarantee that each row represents a single vote (edge) between two countries (nodes). The next step is to use the edge list as a pandas data frame and network to construct a directed graph. Then, as demonstrated in code 1 (the complete code is available in the Github repository), we will attempt the generic way of visualisation.

Visualization

As you can see in image 4, the built-in draw method produces a highly illegible figure. Without any helpful "hints," the approach is unable to deduce much meaning from the data, despite its attempts to draw a highly connected network. By using what we know about the entities to divide and conquer various visual parts of the plot, we can improve the figure:

- **Position -** each country is assigned according to its geo-position
- **Style -** each country is recognized by its flag and flag colors
- **Size -** the size of nodes and edges represents the amount of points Plotting of the network components in parts is shown in code 2.

Listing 2: Network components plotting

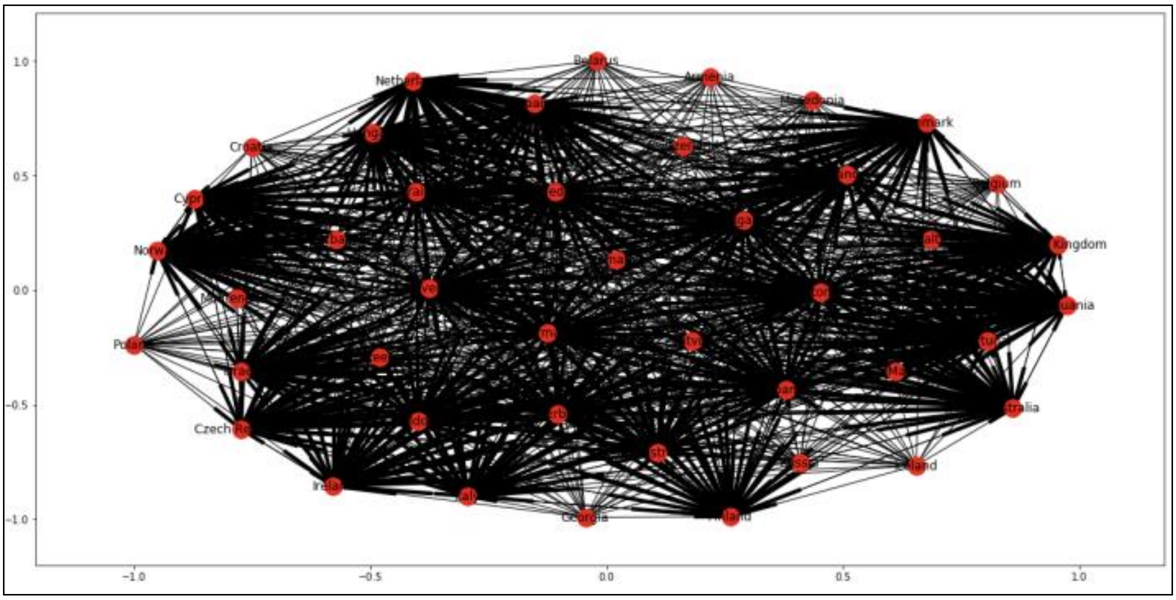

Fig 4: NX. draw network (G) outcome on Eurovision 2018 votes network

The updated figure 5 provides a concise summary of the votes and is slightly easier to read. In general, it's not easy to plot networks; you have to carefully balance the amount of data you show with the message you're trying to convey. (Another option is to look into GraphChi, Pyvis, or Gephi, which are all network visualisation tools).

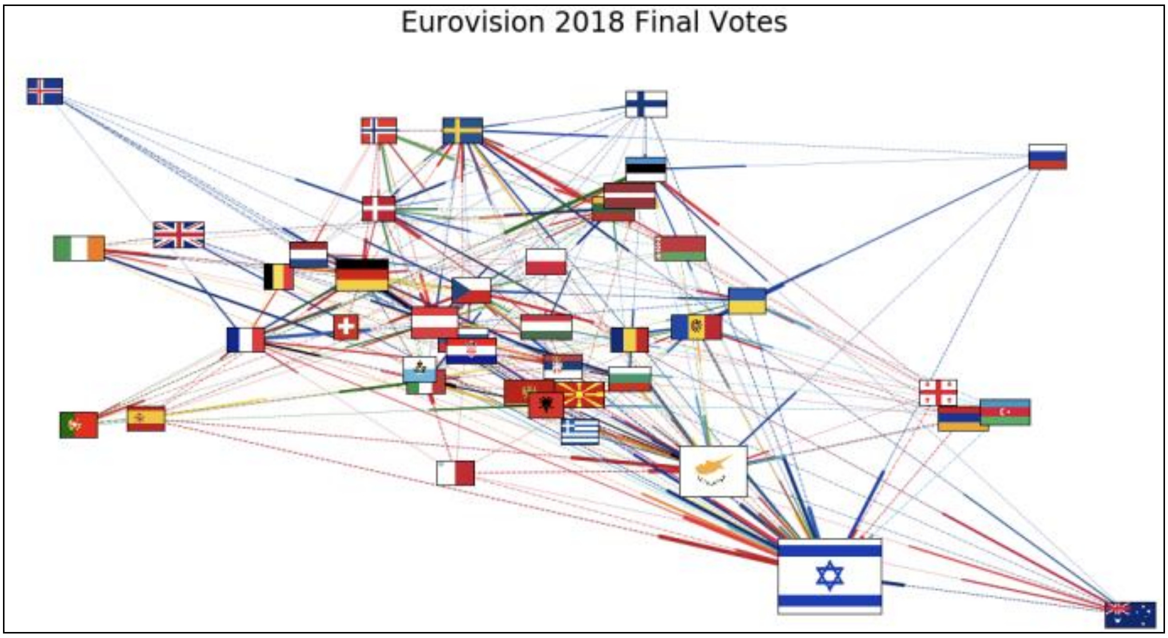

Fig 5: Step-by-step Eurovision network plot

Information Flow

There are certain similarities between the dynamics of information dissemination and the transmission of a virus, which involves the infectious hopping from one person to their social neighbours. When describing the process, two prominent basic models are:

The cumulative influence from the node's many neighbours builds up and becomes active only when it surpasses a certain threshold; this threshold-based behaviour is specified by Linear Threshold and is expressed in the following formula.

$$\sum_{active_u} W_{uv} \geq \theta_v$$

When you hear a lot of good things about a movie from your friends, you might be tempted to see it. This is a common occurrence in movie recommendations.

All of the node's neighbour nodes that are now active have an equal and random chance of activating the node under the Independent Cascade model. All of these social contacts could potentially set off the infection, much as the COVID-19 pandemic.

Information Flow Example

The Storm of Swords network, inspired by the characters in Game of Thrones, will serve as our example of how news travels. The network was built around characters that appeared in both the "Song of Ice and Fire books" and separately.

We will aim to uncover the dynamics of rumour spreading, which are prevalent in this show, by utilising the independent cascade model. To illustrate the point, let's pretend that Jon Snow starts off in the dark, but Bran Stark and Samwell Tarly, two of his most faithful companions, are privy to a crucial life secret. Let's see how the word gets around using the Independent Cascade model:

Table 1: Leading characters per centrality measure

Degree		Weighted Degree		Page Rank		Between ness	
Name	Score	Name	Score	Name	Score	Name	Score
Tyrion	40	Tyrion	1842	Tyrion	0.052	Robert	0.22
Robert	37	Cersei	1627	Jon Snow	0.048	Brienne	0.12
Joffrey	35	Joffrey	1518	Cersei	0.046	Rodrik	0.11

Figure 6 shows how the gossip spreads quickly across the network and eventually becomes public knowledge. At t=1, it reaches Jon, and in the time-steps that follow, it reaches his neighbours. Controlling the diffusion process using model parameters has a significant impact on these dynamics.

Influence maximization

To maximise influence, one must first identify a small subset of nodes in a network, called the seed set, and then choose those nodes in such a way that their influence spreads organically to the largest possible subset of nodes. This setup is applicable outside of marketing as well. Consider inviting a select group of influential people to a high-profile product launch party so that they can inform their respective networks about the occasion. One of the many ways to find these influential people is to use the centrality measures that we discussed before.

However, Maximising Influence is Not an Easy Task. It is actually an NP-Hard problem. The optimal seeding set for an efficient calculation can be found using a number of approaches. Making an effort to discover the most effective.

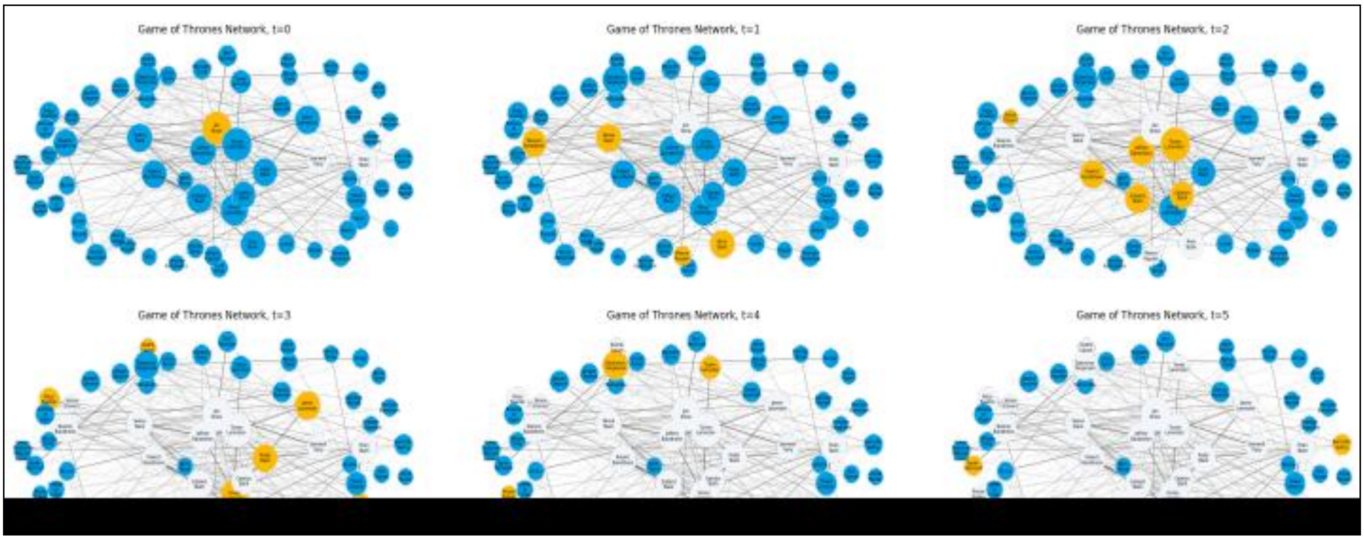

Fig 6: Independent Cascade diffusion simulation on Game of Thrones network

Choosing Khal Drogo and Robert Baratheon as our network's seeded couple took 41 minutes and yielded 56% coverage, an accomplishment that would have been difficult to achieve with centrality techniques.

Graph theory has various applications in social networks. It can be used to explore and visualize social networks by representing them as graphs. Graph theory concepts are helpful in formulating and studying statistical parameters that analyze social patterns in these networks. Clustering online documents, analysing algorithm execution, cryptography, and simplifying and analysing electronic circuits are all applications of graph theory in computer science. Social network analysis combines graph theory with other techniques to investigate social structures, explain underlying dynamics, and analyze social network properties. Graph theory is also used in the analysis of criminal networks, studying weight distribution and shortest path length in interactions among suspects. Additionally, graph theory has potential applications in agriculture, marketing networks, fraud detection, and recommendation systems.

Applications of graph theory in social network analysis

The field known as "social network analysis" examines the interconnections, friendships, and information flows that make up various social networks. Graph theory offers methods for studying social phenomena, analyzing network features, and identifying important nodes or groups.

Here’s how SNA and graph theory work together:

1. Grasping Interconnections: Graph theory enables us to depict social networks as graphs, with every individual serving as a node and their connections as edges. Because of this, it is much simpler to see and examine the relationships between people.
2. Graph theory algorithms allow us to identify important people or groups in a social network, which is crucial for identifying key players. "Hubs" or "influencers" are terms used to describe persons with a large number of

connections or who bring together various groupings.

Behaviour Prediction: By applying graph theory to the structure of social networks, we can foretell how information will circulate and how behaviours will spread. Pattern recognition, prediction, and therapy development are all possible outcomes of this.
All things considered, graph-theory-based social network analysis sheds light on the intricate dynamics of human connections, including friendships, partnerships, information exchange, and influence. Sociology, psychology, marketing, and cybersecurity are just a few of the areas that may benefit from it.

Graph Theory and Communication Network
A group of interconnected devices called "nodes" connections, and terminals that allow for the connection of communication between various users; graph theory is mostly used to depict such networks. It is easy for messages to reach their destinations since each terminal has an appropriate address. Communication networks allow equipment like computers, phones, and processors to exchange data packets, and address space is just a database of those addresses. The function of graphs in communication networks is crucial. A mathematical framework for creating model-pair relations between things, graph theory is the study of graphs. Connected by edges, lines, or arcs, vertices, nodes, or points comprise a typical graph. When any two nodes linked by an edge can be perceived as being the same, or when any two nodes can be linked by an infinite number of edges, we say that the graph is undirected. Graph terminologies have a lot of potential applications since computer science is not a strictly defined field.

Graph Theory in Electrical Network Analysis
Introduction
G. Kirchhoff first proposed using graph theory to electrical networks in 1847; I.C. Mon well refined the concept in 1892; and W.S. Percieval reached a watershed point in graph theoretic analysis of electrical networks. "Branch" denotes an edge in an electrical network, "Node" a vertex, and "Loop" a circuit. In electrical networks, directed edge trees are very important. Using ideas from graph theory, this paper will provide significantly better answers to circuit analysis difficulties. Components like capacitors, inductors, and resistors make up what is known as an electrical network. The study of electrical network topology is known as electric network analysis and synthesis. Graph theory and electrical networks are the subjects of this chapter's theorems and results.

Types of electrical networks
According to the characteristics and behaviors of the electrical network, it's classified into several types. They are mentioned below.

1. **Linear network:** The parameters or the elements of the circuit or network are constant respective to the change in time, time, temperature etc. is called the linear network. The resistors, inductors and capacitors are the elements of the electrical network. For solving these linear networks mainly superposition theorem is used.
2. **Nonlinear networks:** In nonlinear networks, the values of the circuit's parameters vary in relation to variables like voltage, temperature, and time.
3. **Bilateral network:** Switching the direction of current through the network's elements has no effect on the circuit's properties or behaviour. Example of bilateral network is the electrical circuit which has only resistors.
4. **Unilateral network:** Unilateral networks are those in which the direction of current flow through the network's elements causes a change in the circuit's properties and behaviour.
5. **Active network:** A circuit has a source of energy is called active network. The voltage and current are the energy sources of the electrical circuit.
6. **Passive network:** A circuit has no source of energy is called passive network. It contains only resistor, capacitor and inductors.
7. **Lumped network:** All the elements in the circuit are physically separable for analysis purpose is called lumped network. Most electrical networks are lumped type.
8. **Distributed network:** All the elements in the circuit are distributed across the entire length of transmission line and cannot be separated.

Conclusion
Graph theory's special capacity to aid in the modelling of real-world issues has tremendous practical implications in many domains, including electrical and communication networks, interpersonal interactions, social networking, and many more. These networks include social media in technology, GSM mobile phone networks, rapid communication in sensor networks, traffic lights, and so on. When it comes to designing a network's defenses, this paper offers some fundamental answers.

Reference

1. Adadi, Berrada M. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI), in IEEE Access. 2018;6:52138-52160. DOI: 10.1109/ACCESS.2018.2870052. keywords: {Conferences; Machine learning; Market research; Prediction algorithms; Machine learning algorithms; Biological system modeling; Explainable artificial intelligence; interpretable machine learning; black-box models},
2. Popereshnyak S, *et al*. Social Networks: Analysis, Algorithms and Their Implementation; c2021.
3. Lanel GHJ, Jayawardena H. A Study on Graph Theory Properties of On-line Social Networks. International Journal of Scientific and Research Publications (IJSRP). 2020;10:9929. 10.29322/IJSRP.10.03.2020.p9929.
4. Rukmi AM, Zakky W, Shahab ML. Contact Tracking with Social Network Analysis Graph. In International Conference on Mathematics, Geometry, Statistics, and Computation (IC-MaGeStiC 2021); c2022. p. 120-124. Atlantis Press.
5. Srimani, Dr. Chairman, Kamath F, Annapurna. Graph Theoretical and Networks Approach for The Development of A Learning Model – A Case Study.

International Journal of Engineering Science and Technology. 2012, 4.
6. Dr. Velammal S, *et al*. Applications of Graph Theory in Network Analysis IJSTE - International Journal of Science Technology & Engineering, 2018, 5(1).
7. Sadavare AB, *et al*. A Review of Application of Graph Theory for Network International Journal of Computer Science and Information Technologies. 2012;3(6):5296-5300.
8. Arul SM, Senthil G, Jayasudha S, Alkhayyat A, Azam K, Elangovan R. Graph Theory and Algorithms for Network Analysis. InE3S Web of Conferences. 2023;399:08002. EDP Science. DOI: 10.1051/e3sconf/202339908002.
9. Sporns O. Graph theory methods: applications in brain networks. Dialogues Clin Neurosci. 2018;20(2):111-121. DOI: 10.31887/DCNS.2018.20.2/osporns. PMID: 30250388; PMCID: PMC6136126.
10. Chambers J, Cozza S, Patel MK. Applications of graph theory in district heat network analysis at national scale. In Journal of Physics: Conference Series. 2019;1343(1):012045. IOP Publishing.
11. Kumar S, Yadav RBS. Application of Graph Theory in Network Analysis, International Journal of Mathematics Trends and Technology (IJMTT). 2016;31(2):44-45. Crossref, https://doi.org/10.14445/22315373/IJMTT-V31P511
12. Rangaswamy KD, Gurusamy M. Application of Graph Theory Concepts in Computer Networks and its Suitability for the Resource Provisioning Issues in Cloud Computing-A Review. J. Comput. Sci. 2018;14(2):163-72. DOI: 10.3844/jcssp.2018.163.172
13. Berdewad OK, Deo SD. Application of graph theory in electrical network. International Journal of Science and Research. 2016;5(3):981-982.
14. Elumalai A. Graph theory applications in computer science and engineering Malaya Journal of Matematik, 2020;2:4025-4027.
15. Barabási AL, Albert R, Jeong H. Scale-free characteristics of random networks: the topology of the world-wide web. Physica A: statistical mechanics and its applications. 2000;281(1):69-77.