WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Metanalysis of Algebraic Graph Theory

Manuel Baro [a]*; Judith Valdiviezo[b]; Mayra Perez[c]; Rosa Ma. Toral[d]

[ac]*Tecnológico Nacional de México Campus Nuevo Casas Grandes, Nuevo Casas Grandes, Chih., México; mbaro@itsncg.edu.mx*

[bd] *Tecnológico Nacional de México Campus Chihuahua II, Chihuahua, México*

DOI : https://doi.org/10.55248/gengpi.6.0225.0721

A B S T R A C T

Recent advancements in Algebraic Graph Theory (AGT) emphasize community detection, graph clustering, and biological networks, such as Sequence Similarity Networks, Protein-Protein Interaction Networks (PPIs), and Gene Regulatory Networks (GRNs), which help researchers understand the complexity of biological systems. Environmental and Ecosystem Networks, including Ecological networks, Food webs, and Phylogenetic networks, contribute to understanding the mechanisms affecting fragile ecosystems and disease transmission. Novel graph networks, such as knowledge networks, graph neural networks, and brain networks, enable researchers to comprehend complex relationships in various systems. This article presents a comprehensive analysis of the recent advancements, trends, impact, and collaborative patterns within the field of AGT, with a focus on networks and graphs, biological networks, environmental and ecosystem networks, and several interconnected disciplines. A total of 772 scholarly publications from 1980 to 2022 were examined, revealing substantial growth in the field since 1990, along with an increasing number of publications, authors, and institutions involved. The United States and China have dominated AGT research, with 18,998 and 14,503 total citations, respectively, and high H-index scores of 58 for China and 35 for the United States. Countries such as Australia, with an average citation per publication of 74.8, demonstrate the quality and impact of their research, while Saudi Arabia, Qatar, Denmark, and South Africa exhibit high collaboration rates. The analysis of the top twenty institutions in AGT showed that 35% were from China and 30% from the United States.

Keywords: Algebraic Graph Theory; Multi-Agent Systems; Community Detection.

1. Introduction

Algebraic Graph Theory (AGT) is a fast-developing topic of study with major applications in a wide range of disciplines, including computer science, engineering, social network analysis, and data mining. As the area continues to grow, the necessity for quantitative analysis of the literature and trends of research output grows. Bibliometrics, the study of publishing trends and their relevance, is an efficient method for getting insight into the history and development of a research topic. With bibliometric analysis, one may discover major trends, notable contributions, and new study topics, as well as quantify the effect of research and monitor the spread of information. The purpose of this study is to use bibliometric techniques in AGT research output to provide a full picture of the field's development, significant contributors, and upcoming trends.

AGT is the area of mathematics that uses algebraic techniques to study graphs, which are mathematical structures composed of nodes and edges that reflect their interactions. Using algebraic structures like groups, rings, and matrices, it seeks to investigate the characteristics of graphs. The core subfields of AGT are spectrum, topological, and extremal graph theory. Spectral graph theory examines the link between the eigenvalues and eigenvectors of graph-representing matrices and their characteristics. In contrast, topological graph theory focuses on the topological features of graphs, including embeddings, immersions, and knots. Finally, extremal graph theory aims to discover the greatest and least number of edges and nodes in a graph that meets particular constraints. Among other uses, it has contributed to creating algorithms for network analysis, investigating the characteristics of molecules and crystals, and analyzing electrical networks [1].

Several recent researchers reviewed some aspects of graph theory. For many years, graph theory has been used efficiently in gear transmission design. Based on over 100 references, a study of graph-based techniques for kinematic and static force analysis, power flow, and mechanical efficiency calculation was done [2]. The approach is based on the basic circuit notion that correlates to a fundamental epicyclic gear train. A 1-dof epicyclic gear train and a two-stage planetary gear train are used to show the implementation of this approach. The paper also examines the determination of isomorphism and the enumeration of 1-dof epicyclic gear train graphs. Also presented are automated approaches for spotting superfluous gears and degenerate structures in epicyclic gear systems.

In the discipline of computer science, graph theory offers the potential for several applications. Clustering online content, cryptography, and algorithm analysis are promising applications. Moreover, graph theory may be used to analyze and simplify electrical circuits. Lately, graphs have been extensively utilized to represent and study the architecture, processes, and users of social networks. The exhaustive review of Majeed and Rauf [1] outlines the use of graph theory in social networks. It shows the relevance of graph theory in social network modeling and analysis using principles and pertinent examples. The authors demonstrated graph theory's pervasive and potential applications in social networks and computer science.

From the mid-1990s, graph theory has emerged as a crucial tool for comprehending brain connection architecture, notably in fMRI studies of the human connectome. examined how brain characteristics may emerge through the interactions of distinct neural units, as well as evaluating the existing functional and effective connection methodologies used to form the brain network. Since the Human Connectome Project started in 2009, graph theory has gained growing importance. This article investigated the use of theoretical graph metrics to derive neurobiological conclusions about the mechanisms underlying human cognition, behavior, and different brain disorders. Increasing volumes of empirical data on brain networks of various sizes contribute to network neuroscience's fast expansion. To evaluate and represent this data effectively, relevant tools and methodologies, such as those offered by graph theory, are necessary. Sporns focused on the most prevalent and informative graph measurements and methodologies used for neurobiological reasons. Identification of network modules and essential parts that permit communication and signal transmission is crucial. Among the growing themes are generative models, dynamic and multilayer networks, and algebraic topology. Methods based on graph theory are vital for comprehending the structure, growth, and evolution of brain networks.

Integration of network-based machine learning and graph theory methods has gained momentum in precision oncology [12]. Analyzing dysregulated or mutant pathways and networks is more successful than examining individual mutations for understanding cancer, according to research, and disease modules in molecular networks may predict the efficiency of repositioning medications. In their research, Zhang et al. [12] scrutinized the development and mathematical underpinnings of network-oriented machine learning and graph theory techniques for analyzing individual genetic information and biomedical databases. Their aim was to pinpoint tumor-related molecular processes, potential targets, and repurposed drugs for customized therapies. They evaluated methods employed in three different situations for integrating genetic data with network models and assessed three distinct network-driven approaches for repositioning medications in the context of drug-disease-gene networks.

Electrical networks, from microscopic integrated circuits to continental-scale power systems, have been studied using AGT. Dörfler et al. [13] studied algebraic graph theory's historical and contemporary applications in electrical network analysis, dynamics, and design. The authors explored the algebraic and spectral features of graph matrices, as well as their relationship to the network analysis and dynamics of specific electrical networks. They examined these relationships for more sophisticated models, from direct current to alternating current power flow. They also discussed unresolved issues at the confluence of algebraic graph theory and electrical networks. Besides, the study of Easttom highlights the potential of algebraic graph theory in network forensics, which may dramatically enhance intrusion detection and prevention [3].

2. AGT and its recent applications

2.1. Exploring Networks and Graphs

Networks and graphs are mathematical representations that show the connections between entities or objects. Comprised of vertices (or nodes) connected by edges (or links), they represent various real-world networks, such as electricity infrastructure, the Internet, social networks, and more. Graph theory is the study of these networks and is applicable to many disciplines, such as biology and medicine.

Subgraphs are subsets of both vertices and edges in a graph G = (V, E). Graph categories include trees, hypergraphs, multi-edge, bipartite, weighted, directed, undirected, and undirected graphs, such as gene co-expression networks, have a single connection between vertices, whereas directed graphs have arrows indicating direction (e.g., pathway). The nodes of weighted graphs are assigned a weight function (e.g., sequence similarity network). Vertices in bipartite graphs are divided into two distinct groupings (e.g., gene-disease networks). Multiple edges connect vertices in multi-edge graphs (e.g., knowledge/integration networks). Hypergraphs, such as biochemical networks, have edges that connecting many vertices. Trees are connected acyclic undirected graphs (e.g., ontologies, phylogenies). Graph connectivity is predicated on paths between points, with complete graphs having unique edges between distinct vertices. The disjoint union of complete graphs forms clusters, and cliques are subsets of vertices with all pairs connected.

Adjacency matrices, adjacency lists, and sparse matrices can all be used to represent networks. Adjacency matrices represent graphs as square matrices of size N. The matrix is symmetric for undirected graphs but not for directed graphs. Bipartite graphs can be represented as two biadjacency matrices and have distinctive properties.

Adjacency matrices can be resource-intensive for larger sparse networks, requiring $O(V)^2$ memory. O notation evaluates how an algorithm's requirements increase with input size. By storing the upper triangular portion of the adjacency matrix linearly, memory consumption can be reduced to $V(V-1)^2$. Adjacency lists are a more efficient data structure option because they require less space $(O(V+E))$ and allow for quicker access to adjacent vertices. Sparse matrix structures, which store only non-zero elements and their positions, can further reduce memory requirements.

The development of large-scale networks, such as social networks, communication networks, biological networks, and other recent networks in knowledge, deep learning, the brain, etc., are significant factors contributing to the revival of algebraic graph theory. Researchers have recognized the need for effective modeling, analysis, and comprehension tools for these complex systems. Algebraic graph theory provides a unique set of instruments for this purpose, such as adjacency matrices, Laplacian matrices, and spectral graph theory, which facilitate the study of graph properties and the design of efficient algorithms. The following section briefly addresses some significant advancements of AGT in recent scientific fields [4].

2.2. Recent types of graphs

2.2.1. Community Detection and Graph Clustering

Community detection in social networks assists in the identification of groups of individuals with similar interests or behaviors. Using algebraic graph theory, particularly spectral clustering techniques, is an efficient method for detecting communities. Due to the abundance of information in these networks, social network analysis (SNA) has acquired popularity. SNA employs graph theory to comprehend social structures and predict future linkages based on existing connections. Applications of link prediction and community detection include recommendations, friend suggestions, criminal associations, biological network inferences, and marketing trend analysis [25].

Graph clustering has advanced, but the discipline lacks a theoretical framework and direction. The primary responsibility of the scientific community is to define a set of trustworthy benchmark graphs for testing algorithms. Establishing criteria necessitates agreement on the fundamental concepts of community and partition, which will result in more cohesion in the field's development. The planted '-partition model [27] and the benchmark graphs presented by Lancichinetti et al. and Sawardecker et al. are significant steps in this direction. Another crucial area of research is defining null models or graphs without community structure. A precise definition and classification of null models are required to distinguish "true" communities from random fluctuations and test clustering algorithms' dependability. Currently, there is no ideal method for clustering graphs. Undirected and unweighted edges are the main focus of most algorithms. However, actual networks may have directed, weighted, and bipartite connections, necessitating the development of new methods. As timestamped network datasets become available, the study of graphs that change over time is also gathering traction. Tracing the evolution of community structure reveals the formation and interaction of communities.

The computational complexity of graph clustering algorithms has significantly improved, but further progress is required to attain linear or sublinear complexity. Local methods may increase efficiency and parallelism. Finally, it remains difficult to comprehend the implications of community detection results. The primary concern is how vertex classifications provided by algorithms relate to actual classifications. It is essential to ascertain the relationships between communities and the underlying system, as well as potentially uncover hidden relationships between vertices [5].

2.2.2. Biological Networks

Biomedical research benefits from the use of graphs to represent the relationships between various biological entities, such as proteins, genes, molecules, diseases, drugs, and database records. These networks help researchers understand the complexity of biological systems and have various applications. Sequences are represented by nodes in Figure 2(a), with an edge connecting two nodes if their similarity exceeds a predetermined threshold (A). Separate subnetworks, also known as connected components, make up the network diagrams (B). A global articulation point is a node that, when removed, disconnects a connected component (C). The quantity of a node's immediate neighbors determines its degree (D). In cases of partial assembly, a significant proportion of transcripts may have a degree of 1 (illustrated in blue) relative to reference transcripts (red). The Jaccard Index (JI) is used to measure the proximity of two nodes (E). For instance, the red and purple nodes have a total of five neighbors (black/gray nodes), but only the black nodes are shared, resulting in a JI of 3/5. A JI value of 1 indicates that two nodes share identical neighbors, whereas a JI value of 0 indicates that there are no common neighbors. When assembly techniques create chimeras (blue nodes) of unrelated proteins connecting unrelated reference proteins (red, purple, black, and gray nodes), a star-like pattern appears (F). When assembly methods generate non-overlapping, short partial assemblies (blue nodes) that do not connect to the majority of the complete reference sequence (red node), a similar star-like pattern appears (G). A chain-like pattern is produced by fragmentary assemblies that overlap (blue nodes) (H) [6].

3. A bibliometrics overview of AGT

3.1. AGT bibliometrics research scope

The present section aims to explore and understand the trends, impact, and collaborative patterns within the field of Algebraic Graph Theory (AGT) by examining the scholarly publications available. The Web of Science databases were searched for the period from 1980 to 2022 using the query "TS=('Algebraic Graph Theory')". The search resulted in 784 records, of which 12 were related to 2023. Since the analysis is being conducted at the beginning of 2023, these 12 records were excluded, leaving 772 records for further investigation.

Figure 1 illustrates the types and percentages of these 772 publications, highlighting the total number and percentage of each publication type. It is worth noting that Editorial Material, Articles; Book Chapter, Letter, Book Review, Book, Editorial Material; Book Chapter, Correction, and Note collectively make up only 18 articles, accounting for 2.3% of the total publications. Additionally, there are six Review articles, comprising 0.8% of the dataset. Due to the small proportion of these publication types, the research will focus on the more prevalent types: Articles, Proceedings Paper, and Articles; Proceedings Paper. These publication types collectively represent almost 97% of the dataset, with 748 publications related to AGT [7].

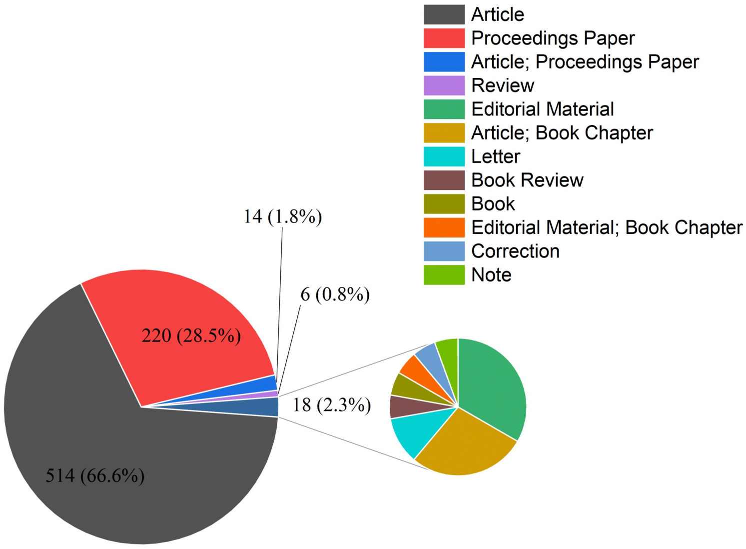

Figure 1. The details of 772 publication records related to AGT

Table 1. Languages of the publications

Language	TP	PR%
English	742	99.2
Chinese	5	0.7
Russian	1	0.1

3.2. Data analysis

This research study seeks to evaluate the productivity, influence, and collaboration patterns in a collection of articles using an array of variables. By examining these variables, the study aims to provide valuable insights into the development and trends within a particular research area.

The variables assessed in this study include the total number of publications in a dataset, the percentage of each publication type within the dataset, and the total number of authors who have published articles in the dataset. Additionally, the research considers the number of institutions associated with the authors, the number of countries represented by the authors, and the number of times articles in the dataset have been cited by other publications.

Furthermore, the study examines the average number of citations per publication, the average number of authors per publication, and the H-index, which measures the productivity and influence of researchers or groups of scholars by taking into account the number of publications and their respective citation counts. The research also investigates the number of publications with the same first author and the number of publications with the same corresponding author.

The study also explores the number of independent and collaborative publications, as well as the proportion of collaborative publications within the dataset. The impact factor, a statistic that quantifies the annual average number of citations received by journal articles and is often used to assess a journal's significance within its field, is another variable analyzed in this research.

Through a comprehensive analysis of these variables, the research aims to provide an understanding of the productivity and impact of institutions and countries participated in a given research area. Additionally, the examination of collaboration patterns will offer insights into how researchers and organizations work together to advance the field [8].

4. Results and discussions

4.1. Annual analysis

The field of AGT has experienced significant growth in publications over the years. In the early years (from 1990 to 2002), the number of publications was low, with only 1 to 4 publications per year. In 2003, there was a slight increase, but the real surge in publications began in 2005. From 2005 to 2022, the number of publications per year increased steadily, suggesting a growing interest in the field. In 2010, there was a notable increase in the total publications, reaching 28, which is 3.7% of the total publications. The growth continued in the following years, with a peak in 2016, when the number of publications reached 77, accounting for 10.3% of the total publications. After this peak, there was a slight decline in the number of publications, but the overall trend remained upward.

The number of authors and institutions involved in AGT research have followed an upward trend over the years. In terms of authors (AU) the early years (1990 to 2002) saw relatively low involvement, ranging from 1 to 6 authors per year. From 2003 onwards, there has been a noticeable increase in the number of authors contributing to AGT research. The growth accelerates from 2005, with a significant jump from 11 authors in 2004 to 26 authors in 2005. This upward trajectory continued in the subsequent years, with a peak of 210 authors in 2016. After this peak, there is a slight decline, but the overall trend remains positive, with the number of authors still considerably higher than in the early years.

Similarly, the number of institutions (Inst) engaged in AGT research has also exhibited growth over time. In the initial years (1990 to 2002), the number of institutions per year was quite low, with only 1 to 4 institutions involved in the research. From 2003 onwards, the number of institutions began to rise, reaching 14 institutions in 2007. The growth further accelerated from 2008, with the number of institutions involved in AGT research almost doubling in just a year, from 14 in 2007 to 25 in 2008. The upward trend continued in the following years, reaching a peak of 99 institutions in 2016 and 2022. Although there are slight fluctuations in the number of institutions, the overall trend indicates increased involvement of various institutions in AGT research. The data also reveals a trend of increasing international collaboration in the field. In the initial years, the number of countries involved in AGT research was quite limited, with only 1 to 3 countries per year. However, starting from 2008, the number of countries engaged in AGT research began to increase, reaching a peak of 19 countries in 2015 and 2016. This indicates a growing global interest in AGT and a higher degree of international cooperation among researchers. Another noticeable trend is the increase in the number of authors and institutions involved in AGT research. The average number of authors per publication (AU/TP) increased from 1.0 in the early 1990s to around 3.4 in 2019 and 2022. This suggests a rise in collaborative efforts among researchers, which might be a result of the increasing complexity of AGT problems and the need for interdisciplinary expertise. Moreover, the total number of citations (TC) shows a considerable variation over the years, with some years witnessing significantly higher citations than others. For instance, in 2004, the total citations reached 11,745, with an average of 1957.5 citations per publication, which is an exceptionally high number compared to other years. It is possible that a highly influential paper or a groundbreaking discovery was published during that year, contributing to the high citation count. In general, the average number of citations per publication (TC/TP) fluctuates over the years but appears to be decreasing in recent years, possibly due to the growing number of publications and the time it takes for new research to accumulate citations see table 2.

Table. 2. The features of the published records for each year in AGT

Year	TP	TP%	AU	Inst	Count	TC	TC/TP	AU/TP
1990	1	0.1	1	1	1	5	5.0	1.0
1993	1	0.1	1	1	1	15	15.0	1.0
1994	1	0.1	3	1	1	32	32.0	3.0
1995	1	0.1	3	2	2	2	2.0	3.0
1997	1	0.1	2	1	1	0	0.0	2.0
1998	3	0.4	1	1	1	20	6.7	0.3
1999	2	0.3	3	3	2	6	3.0	1.5
2000	4	0.5	6	4	2	26	6.5	1.5
2001	2	0.3	4	2	2	1	0.5	2.0
2002	3	0.4	3	3	2	60	20.0	1.0
2003	2	0.3	6	4	3	20	10.0	3.0
2004	6	0.8	11	6	2	11745	1957.5	1.8
2005	12	1.6	26	12	5	845	70.4	2.2
2006	6	0.8	9	5	4	527	87.8	1.5
2007	10	1.3	22	14	4	1487	148.7	2.2
2008	18	2.4	38	25	12	748	41.6	2.1
2009	24	3.2	51	36	12	1421	59.2	2.1
2010	28	3.7	57	33	10	3157	112.8	2.0
2011	24	3.2	56	38	10	1552	64.7	2.3
2012	41	5.5	93	47	10	1323	32.3	2.3
2013	48	6.4	104	58	14	1979	41.2	2.2

2014	45	6	128	66	16	1136	25.2	2.8
2015	67	9	162	90	19	1555	23.2	2.4
2016	77	10.3	210	99	19	2022	26.3	2.7
2017	64	8.6	177	82	14	1540	24.1	2.8
2018	47	6.3	130	73	16	1020	21.7	2.8
2019	53	7.1	181	82	16	738	13.9	3.4
2020	59	7.9	184	98	18	548	9.3	3.1
2021	43	5.7	141	69	16	271	6.3	3.3
2022	55	7.4	187	99	18	122	2.2	3.4

5. Conclusions

In recent years, the field of Algebraic Graph Theory (AGT) has witnessed considerable advancements in various areas, such as community detection, graph clustering, biological networks, and environmental and ecosystem networks. This study aimed to explore trends, impact, and collaborative patterns within AGT by examining scholarly publications from 1980 to 2022. The analysis revealed significant growth in publications, authors, and institutions involved in AGT research since 1990, with increasing international collaboration and the dominance of countries like the United States and China in terms of total citations and high H-index scores. The research also emphasized the importance of international collaborations and diverse expertise in driving advancements in AGT. The main outcomes of the study can be summarized as follows:

Recent advancements in graph theory emphasize community detection, graph clustering, and biological networks. Community detection identifies groups in social networks with shared interests or behaviors. Graph clustering, a developing field, requires more research on null models and algorithm efficiency, with time-evolving graphs also gaining traction. Biological networks, such as Sequence Similarity Networks, Protein-Protein Interaction Networks (PPIs), and Gene Regulatory Networks (GRNs), help researchers understand the complexity of biological systems. PPIs provide information about protein interactions, while GRNs represent the underlying system governing transcriptional regulation. Expression Quantitative Trait Loci (eQTL) Networks explain genetic variance in gene expression phenotypes and reveal genotype-phenotype relationships. Environmental and Ecosystem Networks include Ecological networks, Food webs, and Phylogenetic networks. Ecological networks identify mechanisms affecting fragile ecosystems, while Epidemiological networks study disease transmission. Between-species networks characterize pairwise interactions between species, and within-species networks quantify associations between individuals. Biomedical diagrams and novel graph networks, such as knowledge networks, graph neural networks, and brain networks, help researchers understand complex relationships in various systems. These networks can be integrated to create knowledge networks, providing valuable insights across scientific fields.

In conclusion, Algebraic Graph Theory represents a leading area of mathematical research, offering insights into the structural properties of graphs and their applications across various domains. As researchers continue to explore new techniques in spectral graph theory, develop novel algebraic graph invariants, and foster interdisciplinary collaborations, the field is poised for significant advancements. The future of Algebraic Graph Theory holds promise for addressing complex network problems, from understanding the dynamics of social networks to optimizing communication networks and analyzing biological systems. By employing the power of algebraic techniques, researchers can elucidate the intricate relationships within networks, paving the way for innovative solutions to real-world challenges. As we embark on this journey of discovery, Algebraic Graph Theory remains a vibrant and dynamic field, offering boundless opportunities for exploration and discovery.

6. Discussion and Future Works

Another promising avenue for future research is the study of algebraic graph invariants. Graph invariants are properties of graphs that remain unchanged under certain graph transformations, such as isomorphism or graph automorphisms [2] . Algebraic graph invariants, derived from the algebraic properties of graph representations, provide insights into the structural properties of graphs and often have applications in graph classification, graph recognition, and graph drawing. Future research may seek to develop novel algebraic graph invariants, examine their interrelationships with existing graph invariants, and investigate their potential applications in addressing practical graph-related issues [5].

References

1. Algebraic graph theory. In: Graph Spectra for Complex Networks. Cambridge University Press; 2023. p. 15–50.
2. Lin WWK. On Strongest Algebraic Program Invariants [Internet]. Available from: https://www.researchgate.net/publication/375895164
3. Dalfó C, Fiol MA. On the algebraic connectivity of some token graphs. Journal of Algebraic Combinatorics. 2024 May 7;
4. O. A. S. Rapidah et al. Algebraic Structures and Graph Theory II Guest. Journal of Fundamental and Applied Sciences. 2018;4(1):9–10.

5. Shaukat B, Ishaq M, Ul Haq A. Algebraic invariants of edge ideals of some circulant graphs. AIMS Mathematics. 2024;9(1):868–95.

6. Jiang Q, Wang R, Zhang W, Jiao L, Li W, Wu C, et al. Monitoring Dynamically Changing Migratory Flocks Using an Algebraic Graph Theory-Based Clustering Algorithm. Remote Sens (Basel). 2024 Mar 29;16(7):1215.

7. Wróblewska A, Ustymenko V, Pustovit O. Extremal graph theory and generation of quadratic multivariate transformations of Algebraic Post-Quantum Cryptography. Theoretical and Applied Cybersecurity. 2023 Sep 19;5(1).

8. Alali AS, Ali S, Hassan N, Mahnashi AM, Shang Y, Assiry A. Algebraic Structure Graphs over the Commutative Ring Zm: Exploring Topological Indices and Entropies Using M-Polynomials. Mathematics. 2023 Sep 7;11(18):3833.