Applications in graph theory

Yuan Liu*

International curriculum center, High school affiliated to Renmin University of China, Beijing13080, China

*Corresponding author: 18407239@masu.edu.cn

Keywords: the graph theory, applications, development.

Abstract: This paper is talking about the applications of graph theory. The graph theory was introduced by Euler in 1736, which made more and more people pay attention to this theory. The graph theory has close connections with matrix theory, probability theory, topology, numerical analysis, etc. This paper discusses the applications of graph theory in a few areas, such as mathematics, computer science, chemistry, and biology, and it introduces and goes through the applications by giving some knowledge supplement and reasoning. It suggests that the graph theory has been developed quickly and spread into different areas and subjects because of the advancement of computer science and technology. This paper may help you to have a deeper understanding of the graph theory, and this review will try to highlight the most important applications in these different areas. By first presenting the basic knowledge of these fields that are applied to different subjects, this paper aims to help you better understand the importance and how it works.

1. Introduction

1.1 Background

In 1736, Euler published the first written report on graph theory in his publications. The early issues he considered have a solid practical foundation. Later, studies of electrical networks by German naturalist Kirchhoff led to the discovery of the basic concepts and theorems of graph theory; British mathematician Kelly used trees to calculate isomerism in organic chemistry; Hamilton, an English mathematician, posed a difficult problem related to graph theory; and, finally, the famous four-color hypothesis appeared, contributing to the emergence and development of graph theory. A vast number of new findings in graph theory were utilized in a variety of fields over the twentieth century. Current applications of graph theory are applied in chemistry, problem physics, biology, operations research, network theory, information theory, cybernetics, economics, social sciences, etc [1,2].

Most people think that the graph theory origins from the problem of the Seven Bridges of Königsberg. This question is based on the following scenario: The city of Konigsberg in East Prussia straddled both sides of the Pregolya River at the time, and the river had two small islands in the middle. Seven bridges connect the island to both sides of the river. What strategies could individuals use to walk through all of the bridges at the same time? Euler gave the answer in his second paper that this question did not have any solution. Each bridge is viewed as a line, and the area connected by the bridge is regarded as a point since Euler simplifies the actual abstract problem as a mix of points and lines on a plane. In this method, if you start at one location and return there, the number of lines at that position must be an even number, and that site is known as an even vertex. An odd vertex, on the other hand, is a point connected by an odd number of lines. Euler said that because the Königsberg problem has four odd vertices, it cannot be traversed in a way that answers the query [1].

Another famous problem in the history of graph theory is the four-color problem. One of the three great mathematical issues in the modern world is the four-color problem. This means that any map can utilize four colors to create countries with various colored borders. It was designed to travel to the United Kingdom [3]. In 1852, Fernandez Ghenry, a University of London graduate, found an intriguing phenomenon: each map can be colored in four different hues, resulting in countries sharing a shared border being colored in various colors [4-7].

1.2 Definition and basic knowledge

The study of graphs, which are mathematical structures used to model pairwise relationships between things, is known as graph theory. In this context, a graph is made up of vertices connected by edges. There are two kinds of graphs, the first one is the undirected graph, and the other is the directed graph. The undirected graph is the graph whose edges link two vertices symmetrically. However, the directed graph is the graph whose edges link two vertices asymmetrically [8].

A graph can be defined as G (V, E). The V represents the set of vertices and the E represents the set of edges. Graph theory nomenclature is not yet standardized. Some authors prefer to use the terms instead of the terms "Rather than "vertex" and "edge," use "point" and "line." In issues involving both graphs and geometrical or topological structures, its usage may be inconvenient. In some of the older papers we may find "branch" used for "edge, " and "node" for "vertex. When the number of ends on an edge is two or one, it is called a link or a loop. We will, however, start referring to each edge as having two ends, with the caveat that in the instance of a loop, the two ends will be coincident. The two ends of an edge are said to be adjacent when they are linked by that edge. As a result, if and only if a vertex is an incident with a loop, we say it is linked to itself or next to itself. A multiple join graph is defined as two or more links with the same pair of ends. A strict graph is one that has no loops or multiple joins. Only strict graphs are of interest in a number of graph theory topics. As a result, some authors limit the term "graph" to what we refer to as a strict graph. They refer to "multigraphs" when they need to add loops or numerous joins to their structures [9].

1.3 The development of Graph theory

After the concept of graph theory was first announced, the most popular applications of the graph theory are some game problems and game problems, such as the problem of Seven Bridges of Königsberg. With the development of the scientists who were drilling the graph theory, it was not limited to the simple game problems and research questions, and it develops a new set of problems, such as coloring problems and matrix representation problems, as well as the shortest distance problem. In 1936, the mathematician König wrote the first monograph on graph theory, and graph theory became an independent discipline. Later, with the development of computers, graph theory was applied to various practical problems. Recently, the graph theory has applications in physics, chemistry, operations research, computer science, electronics, cybernetics, information theory, and social sciences.

2. Applications in mathematics

2.1 The short path problem

This problem is about getting the shortest path when traveling through a lot of different places, which is really practical. The problem of urban public transit, tourist route selection, and the building of various key transmission networks, such as mining areas, may all be solved using the shortest path. Both businesses and individuals will benefit from it. As a result, the shortest path problem is essential not only in terms of theory, but also in terms of application [10].

The Floyd algorithm is a basic method of solving this problem, which is stretched from the graph theory. It's a method for determining the shortest pathways in a graph with positive or negative edge weights. The technique will find the lengths of shortest paths between all pairs of vertices in a single run. It is possible to reconstruct the paths with simple tweaks to the method, even if it does not return details of the paths themselves [11]. When solving the short path problem using the Floyd algorithm, we need to first define $n \times n$ as the sequence matrix sequence of $D_{-1}, D_0, \ldots D_{n-1}$. Then Initialize the equation of $D_{-1} = C$. Take $D_{-1}[i][j]$ as the length of edge$< i, j >$, which represents the initial length of the shortest path from $i\ to\ j$, and it is the shortest path from $i\ to\ j$ without passing through other intermediate points. Through Iteration, we need to think that if D_{k-1} has been solved, how to get $D_k (0 \le K \le n-1)$. $D_{k-1}[i][j]$ represents the shortest path, whose intermediate point from I to J is not greater than $K-1$. Consider adding the vertex K to the path P to obtain vertex sequence q,

which is $i,\ k\ and\ j$. If q is not a path, the current shortest path is still the previous result, which is $D_k[i][j]$; Otherwise, if the length of q is less than the length of P, q would replace P as the shortest path from $i\ to\ j$. Since the two subpath $i \ldots j\ and\ k \ldots j$ of q are the shortest paths, whose intermediate points are not greater than $k-1$, the shortest path length from i to j whose intermediate points are not greater than k is:

$$D_k[i][j] = min\{D_k[i][j], D_k[i][k] + [k][j]\} \tag{1}$$

2.2 The minimum weight spanning tree problem

If each edge in a graph is associated with a label from a finite label set instead of a weight, the minimum labeling spanning tree challenge is to create a spanning-tree with the fewest types of labels. In a spanning tree, the highest weighted edge is called a bottleneck edge. For example, if we are constructing a new railway that needs to connect a lot of cities, it is easy to connect all the cities, but we need to minimize the cost of construction. Take R as a graph, representing a tree connecting all the cities. Although there are a lot of different trees, the problem is to find the tree that has the minimum constructing cost [8].

Graphs can be used to model many applications that are based on real-world networks. Distances, travel expenses, building costs, and times are all common properties of the edges of such graphs. These graphs are known as weighted graphs. Finding a connected spanning subgraph with the minimum possible weight is a common challenge in weighted graphs. As previously stated, such a subgraph is a tree because it lacks a cycle. In the construction of electric power grids, railway networks, communication networks, water supply networks, natural gas networks, and other networks, the difficulty of constructing a spanning tree with the shortest possible weight can emerge. Kruskal's or Prim's algorithms can be used to solve this problem quickly [9,12].

Using Kruskal's method to solve the minimum weight spanning tree problem. First, making the initial state of the least spanning tree be a non-connected graph T with just n vertices and no edges, and each vertex in the graph becomes a connected component by itself, assuming a connected network $G = (V, E)$. In E, choose the edge with the lowest cost. Add this edge to T if the vertices related to it are on different connected components in T; else, ignore it and choose the next edge with the lowest cost. The process continues until all of T's vertices form a connected component [13].

2.3 The four-color problem

According to the four-color theorem, every map in a plane can be colored with four colors in such a way that regions having a common boundary do not have the same color. Although maps are mentioned in the presentation of the four-color theorem, map painting does not require it: it only needs to color, and it does not need to use the least color. Four colors are rarely utilized when designing a map. The four-color dilemma is most commonly applied to scheduling and allocation issues. There are various tasks, for example, each of which takes a day. Several of these jobs are incompatible, thus, therefore, cannot all be accomplished on the same day. Now, in four days, hopefully. Utilize the graph with tasks as vertices, link edges between conflicting tasks, use dates as colors, and color the graph to solve the four-color issue.

3. Applications in computer science

3.1 The computer drum design problem

The rotating drum is designed to divide the drum's surface into 16 sectors, as shown in Figure.1, with each sector consisting of either a conductor (shaded region) or an insulator (blank area), and four contacts, as shown in Figure.2. The output is 1 when the conductor is in touch with the sector, and 0 when the insulator is in contact. The drum revolves clockwise, and every time it turns one sector, the contact outputs a binary signal. The key issue is figuring out how to arrange the conductors or

insulators on the drum's 16 sectors so that the contacts produce a distinct set of binary signals after each round [9,14].

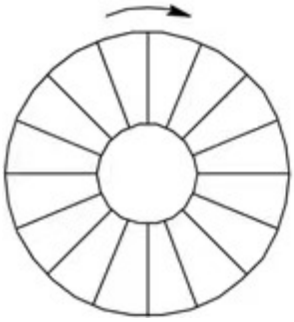

Figure 1. Step 1

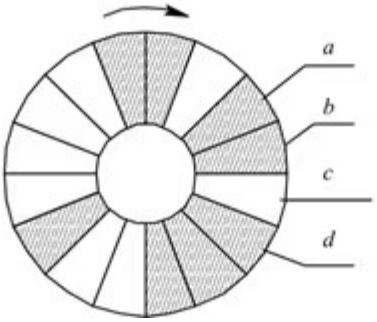

Figure 2. Step 2

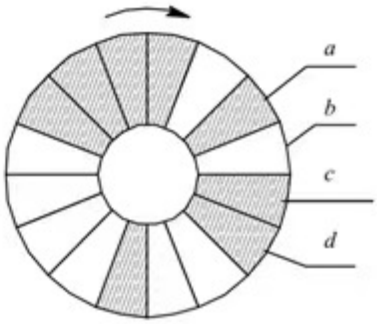

Figure 3. Step 3

Obviously, the signals obtained during rotation are 0010, 1001, 0100, 0010, ..., as shown in Figure.2; however, 0010 appears twice, indicating that this drum does not match the design requirements. The 16 positions of the drum wheel and the 16 four-bit binary signals output by the contact should be in one-to-one correspondence, that is, the 16 binary numbers should be arranged in a cyclic sequence, so that the 16 four-digit numbers formed by every four consecutive numbers are in one-to-one correspondence, according to the requirements of the question. All of the bit binary subsequences are unique. The de Bruijn sequence is the name given to this cyclic sequence. The binary cyclic sequence corresponding to the 16 sectors is exactly the de Bruijn sequence, as shown in Figure.3. Let $V = 000,001,010,011.100,101,110,111$ be the item point of c; the association between vertices and directed edges, as well as the adjacency between vertices, are as follows: Where a_i=0 or $a_i = 1$ $(j = 1,2,3)$, the vertex $v_i = a_{i_1}a_{i_2}a_{i_3}(i = 0,1,2,\ldots,7)$. $e_{i0} = a_{i_1}a_{i_2}a_{i_3}0$and $e_{i1} = a_{i_1}a_{i_2}a_{i_3}1$ are the two directed edges produced from v_i, with e_{ij}associated with the e_{ij} term point $a_{i_2}a_{i_3}0$ associated with the vertex $a_{i_2}a_{i_3}0$. The 16 edges of the loop correspond to a graphic arrangement of 16 binary bits, according to adjacency notation [4,15].

3.2 de Bruijn sequence

$B(k,n)$ is a k-element cyclic sequence. All k-element constituent sequences of length n appear just once in its subsequences and in a circular form. That is why the Figure.2 does not match the requirement since the 0010 repeat again, but in the de Bruijn sequence, all the element can only appear once [16].

3.3 Transitive closure

Let A be the Boolean adjacency matrix of a directed graph G with N vertices, with element $a_{ij} = 1$ if and only if a directed edge exists between vertex I and j. The so-called transitive closure of A^+, also known as vertex i to j, is a NxN Boolean matrix whose element b_{ij}=1 if and only if: 1. $i = j$: or 2. a directed path from $i\ to\ j$, also known as vertex I to j, is reachable. The transitive closure issue is the task of determining G's transitive closure from its Boolean adjacency matrix A. The transitive closure problem is well-known in scientific computing and has a long history of use. The use of multiplication of Boolean matrices to solve the transitive closure problem is a classic technique. This section will show how to build the method on both serial and parallel platforms [17].

The following is the principle of utilizing Boolean matrix multiplication to solve the transitive closure problem: $b_{ij} = 1$ signifies that there is an accessible path of length less than or equal to k from $i\ to\ j$, otherwise, $b_{ij} = 0$. Clearly, $b_{ij} = 1\ in\ ((A + I)^2)^2\ for\ k = 1$ if and only if the path length from $i\ to\ j$ is $0(i = j)\ or\ 1$. In $(A + I)^2$, $b_{ij} = 1$ if and only if the path length from $i\ to\ j$ is less than or equal to 2; In $((A + I)^2)^2$, $b_{ij} = 1$ if and only if the path length from $i\ to\ j$ is less than or equal to 4, and so on in the fourth power of. Because the length of each attainable path between any two places is at most $N - 1$, $(A + I)^k$, and is the needed transitive closure A^+ when $k \geq N -$

1. As a result, so the matrix obtained after $log\ n$ times of self-multiplication of the matrix of (A+I) is the required transitive closure.

3.4 Connected component

If there is a graph G, the connected component of the graph G is the maximum connected subgraph. In this subgraph, there is a path connecting each pair of vertices. The linked component problem refers to the difficulty of locating all of the graph G's connected components. One of the common ways of solving this problem is using the vertex collapse algorithm.

The N vertices in the network are first treated as N isolated super vertex in the vertex collapse algorithm. The edge-connected super vertex is merged one by one until the final connected component is generated during the algorithm's operation. Each vertex belongs to only one super vertex, and the root of the super vertex is the super vertex with the smallest label. The algorithm's flow is made up of a sequence of loops. Each loop is broken down into three sections: 1. Find the smallest label of each vertex next to the super vertex; 2. Connect the root of each super vertex to the root of the smallest label adjacent to the super vertex; 3. Collapse and merge all of the super vertices connected together in step 2 [18].

3.5 Single source shortest path

It is the distance between a single vertex s and all other vertices i the distance between a single vertex s and all other vertices i is called a single source. Let $G(V,E)$ be a directed weighted network, with V and E representing the vertex and edge sets, W representing the edge weight adjacency matrix, and $w(ij) > 0$ representing the edge weight. i,j belongs to V, so V is the integers between 0 and $N-1$ [19].

4. Applications in Chemistry

The study of graph energy in the field of chemistry. The total electron energy E of the Huckel molecular orbital (HMO) is a well-known topological indicator and plays a very important role in theoretical chemistry. The total π electron energy of a conjugated molecule is the difference between the chemical structure and the thermodynamic stability of the conjugated molecule. It can also explain the relationship between the structure and properties of molecules. Through graph theory, the relationship between energy and graph can be explored [20].

4.1 Sachs's formula

Sachs' formula [1] can be used to connect the structure of a non-rooted graph G (for example, a graph describing the topology of carbon atoms in a conjugated hydrocarbon) to the characteristic polynomial $[P(G;x)]$ of G's vertex adjacent matrix. [21] Sachs' formula can be expressed as follows:

$$P(G;x) = \sum_{n=0}^{N} \sum_{s\varepsilon S_n} (-)^{c(s)} 2^{r(s)} x^{N-n} \tag{2}$$

Where s is a Sachs graph, S_n is the set of all Sachs graphs with n vertices, $c(s)$ is the number of components [22], and $r(s)$ is the number of cycles in S, and Sn is the set of all Sachs graphs with n vertices. The number of vertices in G is N. As demonstrated by Graovac et al., this formula may be used to the graphs associated with conjugated hydrocarbons in a simple way. Several alternative (ostensibly different) methods to this problem have been demonstrated to be equal [23].

4.2 Calculating the energy in the graph

Consider a graph G with many vertices and m edges, sometimes known as a (n,m) graph. T polynomial of the graph G is:

$$\phi(G,\lambda) = \det(\lambda I - A(G)) \tag{3}$$

The root of $\phi(G,\lambda)=0$, which is $\lambda_1,\lambda_2\ldots,\lambda_n$, is the eigenvalue of the graph G, and the adjacency matrix of G is $A(G)$. Because $A(G)$ is a real symmetric matrix, graph G's eigenvalues are all real numbers, and graph G's energy is best characterized as:

$$E(G)=|\lambda_1|+|\lambda_2|+\cdots+|\lambda_n| \tag{4}$$

The total π-electron energy of the molecule represented by the graph is closely related to the energy of a molecular graph in chemistry. The total π-electron energy of the molecule is nearly equal to the thermal energy generated by the synthesis of the conjugated molecule, and the calculation of the total π-electron energy of the molecule may be attributed to the calculation of [24]:

$$E(G)=\sum_{i=1}^{i=n}|\lambda_i| \tag{5}$$

$G(n,2)$ is a set of unicyclic graphs with n vertices and a cycle C $(3\le l\le n)$; in the unicycle graphs of order 7 and below, there are only a pair of equal energy graphs, and they are homophonic [25].

Proof: For the calculation of unicyclic graph energy, unicyclic graphs can be divided into two categories: One is a cyclic graph; the other is a unicyclic graph with dangling edges; it can be known that if G is a cyclic graph of order n, and in the first row of A(G), the number 1 is in a_i+1, and the number of other positions is 0, then its eigenvalue is:

$$\{\omega^{j\alpha_1}+\omega^{j\alpha_2}+\cdots+\omega^{j\alpha_k}: 0\le j\le n-1, \omega\ is\ 1's\ nth\ root\ of\ unit\} \tag{6}$$

5. Applications in biomedical

The inverse problem of the topological index of molecular graphs has attracted interest in recent years due to the necessity to find novel medications in biomedicine. People frequently require new medications with specific chemical or physical qualities. To synthesize a desired new drug or substance, empirical formulas are used to determine the topological index value that the molecular pattern of this substance should have, and then a computer search is used to create a database of all possible molecular patterns with this index value. Finally, the most ideal and synthetic graphics in the library are selected to synthesize them. The creation of a database of molecular graphs with a specific topological index value is a crucial stage in this procedure. The problem of molecular graphs with a certain topological index value must be addressed in order to create this database. As a result, doing an in-depth study on this topic is critical for the deliberate synthesis of novel medications [26].

5.1 Molecular Topology Index

A molecular graph is essentially a non-numerical mathematical object. Molecules' various quantifiable properties are frequently stated numerically. As a result, the information gained from the molecular diagram must first be translated into a quantity that can be stated numerically in order to relate the topological features of molecules to the quantifiable aspects of molecules. The quantities that can play this role are the numerous invariants of molecular graphs. That is to say, the invariant of the molecular map may be utilized to not only express the molecule's structure quantitatively, but also to correlate the relationship between the molecule's structure and attributes. The invariant of a molecular graph with this effect is commonly referred to as the molecular topological index [27]. Nowadays, there are hundreds of molecular topological indices [28].

5.2 Randic index

Randic proposed a topological invariant of a molecular graph G [29], which is:

$$\omega_{-\frac{1}{2}}(G)=\sum_{u,v}(d(u)d(v))^{-1/2} \tag{7}$$

The aim was to use it as a measure of the branching nature of the skeleton of carbon atoms in organic molecules, which he found to be related to a number of physicochemical properties, such as boiling points, chromatographic retention times, enthalpies of formation, surface areas and solubility in water. Later, other scientists use the arbitrary real numbers to substitute $-\frac{1}{2}$ [30].

$$\omega_{\alpha}(G) = \sum_{(u,v)\in E(G)} (d(u)d(v))^{\alpha} \tag{8}$$

Since 1975, the Randic index, and later the generalized Randic index, has been one of the most widely used molecular structure descriptors for predicting the physical and chemical properties of organic compounds in quantitative structure-activity correlations and quantitative structure-property relationships. Pharmacological qualities are particularly important.

6. Conclusion and future work

Graph theory has a long history, and its progress has been accelerated by the rapid advancement of computer technology. Graph theory is everywhere; it can be found in many aspects of our life, but because most people don't comprehend it, it has received little attention. Despite the fact that it is an active subject with a wide range of applications, the approach still has limits and requires further research. Through the application of graph theory in various fields above, we can see the importance of graph theory in our lives. A graph isomorphism means that, despite the fact that there are many distinct types of graphs, they all have the same term points and edges, as well as the same continuity. This has a bright future in terms of applications, with applications in chemistry, computer science, operations research, electronics, and other domains. However, solving the isomorphism of complicated graphs is challenging since the time complexity technique has constraints when it comes to employing objects.

References

[1] H. Sachs, Publ. Math. (Debrecen) 11, 199 [1963].

[2] West D B. Introduction to graph theory [M]. Upper Saddle River: Prentice Hall, 2001.

[3] Ore O. The four-color problem [M]. Academic Press, 2011.

[4] Bondy J A, Murty U S R. Graph theory with applications [M]. London: Macmillan, 1976.

[5] Biggs N, Lloyd E K, Wilson R J. Graph Theory, 1736 - 1936 [M]. Oxford University Press, 1986.

[6] Gross J L, Yellen J. Handbook of graph theory [M]. CRC press, 2003.

[7] Chen W K. Applied graph theory [M]. Elsevier, 2012.

[8] Blanco R, Lioma C. Graph-based term weighting for information retrieval [J]. Information retrieval, 2012, 15 (1): 54 - 92.

[9] Tutte W T, Tutte W T. Graph theory [M]. Cambridge university press, 2001.

[10] Newman M E J. Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality [J]. Physical review E, 2001, 64 (1): 016132.

[11] eisstein E W. Floyd-warshall algorithm [J]. https://mathworld. wolfram. com/, 2008.

[12] Zverovich V. Modern Applications of Graph Theory [M]. Oxford University Press, 2021.

[13] Gallager R G, Humblet P A, Spira P M. A distributed algorithm for minimum-weight spanning trees[J]. ACM Transactions on Programming Languages and systems (TOPLAS), 1983, 5 (1): 66 - 77.

[14] Van Steen M. Graph theory and complex networks[J]. An introduction, 2010, 144.

[15] Riaz, F., & Ali, K. M. (2011, July). Applications of graph theory in computer science. In 2011 Third International Conference on Computational Intelligence, Communication Systems and Networks (pp. 142-145). IEEE.

[16] Chan A H, Games R A, Key E L. On the complexities of de Bruijn sequences[J]. Journal of Combinatorial Theory, Series A, 1982, 33 (3): 233 - 246.

[17] Hirschberg D S. Parallel algorithms for the transitive closure and the connected component problems[C]//Proceedings of the eighth annual ACM symposium on Theory of computing. 1976: 55-57.

[18] Hirschberg D S, Chandra A K, Sarwate D V. Computing connected components on parallel computers [J]. Communications of the ACM, 1979, 22 (8): 461 - 464.

[19] Dijkstra E. A Note on Two Problems in Connection with Graphs. Numerische Mathmatic, 1858, 1: 269 - 271.

[20] Balaban A T. Applications of graph theory in chemistry [J]. Journal of chemical information and computer sciences, 1985, 25 (3): 334 - 343.

[21] F. Harary, Graph Theory, Addison-Wesley, Reading, Mass.1961.

[22] Graovac A, Gutman I, Trinajstić N, et al. Graph theory and molecular orbitals [J]. Theoretica chimica acta, 1972, 26 (1): 67 - 78.

[23] A. Moskowitz, J. Combinatorial Theory 12 (B), 177 [1971].

[24] Hou Y. Unicyclic graphs with minimal energy [J]. Journal of Mathematical Chemistry, 2001, 29 (3): 163 - 168.

[25] R. bkrishnan. The energy of a graph. Linear Algebra and its Appli- cation.2004, 387: 287 - 295.

[26] Chao S Y. Graph theory and analysis of biological data in computational biology [M]//Advanced Technologies. IntechOpen, 2009.

[27] Amigó J M, Gálvez J, Villar V M. A review on molecular topology: applying graph theory to drug discovery and design[J]. Naturwissenschaften, 2009, 96 (7): 749 - 761.

[28] Türker L. Contemplation on the Hosoya indices [J]. Journal of Molecular Structure: THEOCHEM, 2003, 623 (1-3): 75 - 77.

[29] Araujo O, Rada J. Randić index and lexicographic order [J]. Journal of Mathematical Chemistry, 2000, 27 (3): 201 - 212.

[30] Bollobás B, Erdos P. Graphs of extremal weights [J]. Ars combinatoria, 1998, 50: 225 - 233.