TAMKANG JOURNAL OF MATHEMATICS
Volume 25, Number 1, Spring 1994

ON THE SUM OF ALL DISTANCES IN GRAPHS

IVAN GUTMAN, YEONG-NAN YEH AND JIANN-CHERNG CHEN

Abstract. The sum W of the distances between all pairs of vertices in a connected graph may be any positive integer, except 2 and 5. We also examine the values that W assumes for connected bipartite graphs and trees.

1. Introduction

In this paper we consider finite, undirected, connected graphs without loops or multiple edges. If G is such a graph, and u and v are its two vertices, then the distance between u and v is the number of edges in the shortest path that connects u and v [1]. The sum of the distances between all pairs of vertices of the graph G will be denoted by $W(G)$.

The quantity W was examined in quite a few mathematical papers (see, for example, [2, 3, 5, 6, 8]). It is also worth noting that W found applications in chemistry [4]; it was first studied in 1947 by the American chemist Harold Wiener [7] and is therefore often referred to as "the Wiener number".

Nevertheless, it seems that an elementary question has not been considered so far, namely which numerical values can W assume. We now provide a solution of this problem.

Let $\mathbf{N}$ be the set of all non-negative integers. Let φ be a set of (connected) graphs and $\mathbf{W}(\varphi) = \{W(G) | G \in \varphi\}$. Then the main result of this paper can be formulated as follows.

Theorem 1. *Let $\mathcal{G}$ be the set of all connected graphs. Then*

$$\mathbf{N} - \mathbf{W}(\mathcal{G}) = \{2, 5\}.$$

In order to prove Theorem 1 we need some preparations.

Received February 18, 1993; revised March 2, 1993.
AMS subject classification (1991): Primary 05C12

2. Some auxiliary results

Lemma 1. *If G is a graph with n vertices and m edaes and if the diameter of G is less than three, then*

$$W(G) = n(n-1) - m. \tag{1}$$

Proof. Bearing in mind that the diameter of a graph is equal to the maximal distance between two vertices of this graph [1], we conclued that m pairs of vertices of G have distance one and the remaining $\binom{n}{2} - m$ pairs are at distance two. Hence $W(G) = m + 2\left[\binom{n}{2} - m\right]$.

Lemma 2. *For every integer value of W, such that $\binom{n}{2} \leq W \leq (n-1)^2$, $n \geq 1$, there exists a graph G with n vertices and diameter less than three, such that $W(G) = W$.*

Proof. The smallest value of m in eq. (1) is $n-1$, because graphs with $m < n-1$ are not connected. The graph with $m = n-1$ and diameter two is unique: this is the star, whose one vertex (of degree $n-1$) is adjacent to all the other vertices (of degree one). If k new edges are introduced into the star, so as to connect k arbitrarily chosen pairs of its vertices of degree one, then an n-vertex graph with $m = n-1+k$ edges and diameter less than three is obtained. The maximal value of k is $\binom{n}{2} - (n-1)$, when the complete graph is obtained. Hence, for every m, $n-1 \leq m \leq \binom{n}{2}$, there exist graphs with n vertices, m edges and diameter less than three.

Lemma 2 follows from Lemma 1.

Theorem 2. *Let $\mathcal{D}_2$ be the set of all graphs whose diameter is less than three. Then*

$$\mathbf{N} - \mathbf{W}(\mathcal{D}_2) = \{2, 5\}.$$

Proof. Theorem 2 follows immediately from Lemma 2. It is sufficient to observe that the inequality $(n-1)^2 \geq \binom{n+1}{2} - 1$ is satisfied for $n \geq 4$. Consequently, $W \in \mathbf{W}(\mathcal{D}_2)$ for all $W \geq 6$, i.e., for $W \geq \binom{n}{2}$, $n = 4$. The fact that the numbers 0, 1, 3 and 4 are also elements of $\mathrm{W}(\mathcal{D}_2)$ whereas the numbers 2 and 5 are not, can bee checked by direct calculation, applying Lemma 2 for $n = 1, 2$ and 3.

3. Proof of Theorem 1 and discussion

In view of Theorem 2 is remains only to verify that there exists no graphs with diameters greater than two, for which $W = 2$ or $W = 5$. This is achieved by the following:

Lemma 3. *If G is a (connected) graph with diameter greater than two, then $W(G) \geq 10$.*

Proof. A graph whose diameter is greater than two contains the path with four vertices as an induced subgraph. Evidently, the sum of the distances in an induced subgraph cannot exceed the sum of the distances in the partent graph. Lemma 3 follows from the fact that $W = 10$ for the path with four vertices.

This completes the proof of Theorem 1.

Theorem 1 is, evidently, just the first and the simplest result in an approach towards the understanding of the number-theoretical properties of W. If, instead of the set of all graphs, we restrict the consideration to some of its subsets, then we may encounter nontrivial problems. First of all, it may well happen that $\mathbf{N} - \mathbf{W}(\varphi)$ is an infinite set, as illustrated by the following elementary example.

Observation 1. Let $\mathcal{B}_{\text{odd}}$ be the set of all connected bipartite graphs with odd numbers of vertices and let $\mathbf{N}_{\text{odd}}$ be the set of all positive odd integers. Then

$$\mathbf{N}_{\text{odd}} \subset \mathbf{N} - \mathbf{W}(\mathcal{B}_{\text{odd}}). \tag{2}$$

Proof. In order to see that relation (2) holds, consider a bipartite graph G with $\underline{a}$ vertices of one color (say, white) and $\underline{b}$ vertices of the other color (say, black). Hence G has $\underline{a} + \underline{b}$ vertices. Observe that the distances between two white or two black vertices are necessarily even. The distance between a white and a black vertex is odd. Hence, the parity of $W(G)$ is equal to the parity of the number of pairs of oppositely colored vertices, i.e. to the parity of the product $\underline{a}\,\underline{b}$. Now, if $\underline{a} + \underline{b}$ is odd, then $\underline{a}\,\underline{b}$ is necessarily even and, consequently, $W(G)$ is even. This implies (2).

If we extend the consideration to bipartite graphs with both even and odd number of vertices, then we arrive at:

Observation 2. Let $\mathcal{B}$ be the set of all connected bipartite graphs. Then $\mathbf{N} - \mathbf{W}(\mathcal{B})$ is finite.

Proof. Construct a bipartite graph $G_{a5}(i, j, k, \ell)$ with $\underline{a}$ white and $\underline{b}$ black vertices, $\underline{a} \geq 1$, $\underline{b} = 5$, $1 \leq \ell \leq k \leq j \leq i \leq \underline{a}$, in the following manner. Denote the white vertices by $v_1, \ldots, v_a$ and the black vertices by $w_1, w_2, \ldots, w_5$. Connect w_1 with $v_1, \ldots, v_a$. Connect w_2 with $v_1, \ldots, v_i$. Connect w_3 with $v_1, \ldots, v_j$. Connect w_4 with $v_1, \ldots, v_k$. Connect w_5 with $v_1, \ldots, v_\ell$.

It can be sjown that

$$W(G_{a5}(i, j, k, \ell)) = a^2 + 12a + 20 - 2(i + j + k + \ell). \tag{3}$$

Now, by an elementary, yet somewhat tedious consideration it can be demonstrated that for $\underline{a} \geq 5$ and $4 \leq i + j + k + \ell \leq 4\,\underline{a}$, the right-hand side of (3) assumes any integer value greater than 78. Consequently, $\mathbf{N} - \mathbf{W}(\mathcal{B})$ has less than 79 elements.

A detailed examination, based on Observation 2, yields a result analogous to, but slightly more complicated than Theorem 1. Its proof is straightforward, but rather lengthy and is therefore omitted.

Theorem 3. *Let $\mathcal{B}$ be the set of all connected bipartite graphs. Then*

$$\mathbf{N} - \mathbf{W}(\mathcal{B}) = \{2,3,5,6,7,11,12,13,15,17,19,33,37,39\}.$$

The case of trees seems to be even more perplexed. By numerical testing we found that

$$\begin{aligned}\mathbf{N} - \mathbf{W}(\mathcal{T}) = \{&2,3,5,6,7,8,11,12,13,14,15,17,19,21,22,23,24,26,27,\\&30,33,34,37,38,39,41,43,45,47,51,53,55,60,61,69,73,\\&77,78,83,85,87,89,91,99,101,\ldots\}\end{aligned}$$

where $\mathcal{T}$ is the set of all trees. Nevertheless, we are inclined to formulate the following:

Conjecture. Let $\mathcal{T}$ be the set of all trees. Then $\mathbf{N} - \mathbf{W}(\mathcal{T})$ is finite.

Acknowledgement

The authors are indebted to the National Science Council of the Republic of China for financial support under the grants NSC-82-0208-M001-042 and VPR-92029. IG thanks also the Mathematical Institute in Belgrade, Yugoslavia, for financial support.

References

[1] F. Buckley and F. Harary, "Distance in Graphs", *Addison-Wesley*, Redwood, 1990.

[2] R. C. Entringer, D. E. Jackson and D. A. Snyder, "Distance in graphs", *Czech. Math. J.*, 26 (1976), 283-296.

[3] I. Gutman, "On distance in some bipartite graphs", *Publ. Inst. Math.* (*Beograd*), 43 (1988), 3-8.

[4] I. Gutman and O.E. Polanský, "Mathematical Concepts in Organic Chemistry", *Springer-Verlag*, Berlin, 1986.

[5] J. Plesnik, "On the sum of all distances in a graph or digraph", *J. Graph Theory*, 8 (1984), 1-21.

[6] L. Šoltés, "Transmission in graphs: a bound and vertex removing", *Math. Slovaca*, 41 (1991), 11-16.

[7] H. Wiener, "Structural determination of paraffin boiling points", *J. Amer. Chem. Soc.* 69 (1947), 17-20.

[8] Y.-N. Yeh and I. Gutman, "On the sum of all distances in composite graphs", *Discrete Math.*, in press.

Faculty of Science, University of Kragujevac, P.O.Box 60, YU-34000 Kragujevac, Yugoslavia.

Institute of Mathematics, Academia Sinica, Taipei 11529, Taiwan, R.O.C.

China Junior College of Industrial and Commercial Mangement, Taipei, Taiwan, R.O.C.