ELSEVIER
Discrete Applied Mathematics 51 (1994) 75–83
DISCRETE APPLIED MATHEMATICS

Average distance and independence number

P. Dankelmann
Lehrstuhl II für Mathematik, RWTH Aachen, 51 Aachen, Germany

(Received 1 October 1991; revised 7 July 1992)

Abstract

A sharp upper bound on the average distance of a graph depending on the order and the independence number is given. As a corollary we obtain the maximum average distance of a graph with given order and matching number. All extremal graphs are determined.

1. Introduction

Let $G = (V, E)$ be a finite, simple and undirected graph with vertex set V and edge set E. If G is connected, the *average distance* $\mu(G)$ is defined to be the average of all distances in G:

$$\mu(G) := \binom{|V|}{2}^{-1} \sum_{a,b \in V(G)} d(a,b),$$

where $d(a, b)$ denotes the length of a shortest path joining the vertices a and b. The average distance can be used as a tool in analyzing networks that represent transportation systems. It is a measure on the time needed in the average case, contrary to the diameter, which indicates the maximum transportation time.

The computer program GRAFFITI [6] made the attractive conjecture

$$\mu(G) \leqslant \alpha(G),$$

where $\alpha(G)$ denotes the independence number of G. The weaker inequality $\mu(G) - 1 < \alpha(G)$ was proved by Fajtlowicz and Waller [6]. Chung [3] succeeded in proving the conjecture. So we have a lower bound on the independence number which is computable in polynomial time. She also established that equality holds only for the complete graph, i.e. for $\alpha = 1$. We give an upper bound for μ depending also on n which is sharp for every α. Making use of this bound we are able to answer a question of Erdös (see [6]). He asked for bounds on the independence number of

SSDI 0166-218X(92)E0201-O

a graph with a given average distance. Furthermore, we will give upper and lower bounds on the average distance depending on the matching number.

We need some further notations. The *diameter* of a connected graph G, $\mathrm{diam}(G)$, is the maximum distance between two vertices of G. By K_n we denote the complete graph and by P_n the path of order n, respectively. For disjoint graphs G and H, the *sum* $G + H$ is obtained from G and H by adding all possible edges between vertices of G and H. If $p > 0$ is an integer, pG denotes the disjoint union of p copies of G. The *neighborhood* $N(x)$ of a vertex $x \in V$ is the set of all vertices adjacent to x. The *closed neighborhood* $\bar{N}(x)$ of a vertex $x \in V$ contains also the vertex x itself. $d(x) := |N(x)|$ denotes the *degree* of the vertex x. The *transmission of a vertex* $x \in V$ is the sum of all distances between x and each other vertex of G. The *transmission of the graph* G, $\sigma(G)$, is the sum of all transmissions of the vertices of G:

$$\sigma(x) := \sigma(x, G) := \sum_{y \in V} d(x, y),$$

$$\sigma(G) := \sum_{x \in V} \sigma(x) = \sum_{(x, y) \in V \times V} d(x, y).$$

For the sake of brevity, let $\mathcal{G}(n, \alpha)$ denote the class of all connected graphs of order n and independence number α.

2. Results

We first give a sharp lower bound on the average distance of the members of the class $\mathcal{G}(n, \alpha)$. This bound is an immediate consequence of the following observation, also noted in [5].

Lemma 2.1. *If G is a connected graph with $n \geqslant 2$ vertices and m edges, then*

$$\mu(G) \geqslant 2 - m\binom{n}{2}^{-1}.$$

Equality holds if and only if $\mathrm{diam}(G) \leqslant 2$.

Proposition 2.2. *If G is a connected graph with n vertices and independence number α, then*

$$\mu(G) \geqslant 1 + \frac{\alpha(\alpha - 1)}{n(n - 1)}.$$

Equality holds if and only if $G \cong K_{n-\alpha} + \alpha K_1$.

Proof. Obviously, we have

$$m \leqslant \binom{n}{2} - \binom{\alpha}{2},$$

with equality only for $K_{n-\alpha} + \alpha K_1$. Application of Lemma 2.1 yields the proposition. □

It turns out that the problem of determining the *maximum* average distance of the members of $\mathcal{G}(n, \alpha)$ is much harder. In order to formulate the main result, we need a preliminary definition.

Definition 2.3. (a) For positive integers n, k with $2 \leqslant k \leqslant \frac{1}{2}n$, let $G_{n,k}$ be the graph obtained from a path P_{2k-2} with end vertices v_1, v_2 and two disjoint complete graphs G_1, G_2 of order

$$n_1 := \left\lfloor \frac{n}{2} \right\rfloor - k + 1, \qquad n_2 := \left\lceil \frac{n}{2} \right\rceil - k + 1$$

by joining v_i with each vertex of G_i for $i = 1, 2$.

(b) For positive integers n, k with $\frac{1}{2}n < k \leqslant n - 1$, let $G_{n,k}$ be the graph obtained from a path $P_{2n-2k-1}$ with end vertices v_1, v_2 and two disjoint empty graphs G_1, G_2 of order

$$n_1 := k - \left\lfloor \frac{n-1}{2} \right\rfloor, \qquad n_2 := k - \left\lceil \frac{n-1}{2} \right\rceil$$

by joining v_i with each vertex of G_i for $i = 1, 2$.

Theorem 2.4. *Let n, α be two integers satisfying $2 \leqslant \alpha \leqslant n - 1$ and let $G \in \mathcal{G}(n, \alpha)$. Then we have*

$$\mu(G) \leqslant \mu(G_{n,\alpha}). \tag{1}$$

Equality holds if and only if $G \cong G_{n,\alpha}$.

We only consider the case $2 \leqslant \alpha \leqslant \frac{1}{2}n$, for the proof in the other case is very similar. Preliminarily, we shall state the following two lemmas.

Lemma 2.5. *Let $G \in \mathcal{G}(n, 2)$, $n \geqslant 3$. Then we have*

$$\sigma(G) \leqslant \sigma(G_{n,2}).$$

Equality holds if and only if $G \cong G_{n,2}$.

Proof. Let $G \in \mathcal{G}_{n,2}$ be a graph with maximum transmission. Obviously, the diameter of G equals 2 or 3; otherwise, $\alpha(G) \neq 2$.

Case 1. $\operatorname{diam}(G) = 2$. A Turán-type theorem for connected graphs [4] is of great use. It yields

$$m(G) \geqslant m(G_{n,2}),$$

with equality if and only if $G \cong G_{n,2}$. With $\text{diam}(G) \neq \text{diam}(G_{n,2})$ we have $G \neq G_{n,2}$ and thus

$$\begin{aligned}\sigma(G) &= \sum_{d(a,b)=1} d(a,b) + \sum_{d(a,b)=2} d(a,b)\\ &= 2n(n-1) - 2m(G)\\ &< 2n(n-1) - 2m(G_{n,2})\\ &\leqslant \sigma(G_{n,2}),\end{aligned}$$

a contradiction to the maximality of $\sigma(G)$.

Case 2. $\text{diam}(G) = 3$. There are two vertices $a, b \in V(G)$ with $d(a,b) = 3$. It is easily seen that G consists of two cliques induced by $\bar{N}(a)$ and $\bar{N}(b)$, respectively, and exactly one edge between these cliques. A simple calculation shows that $|d(a) - d(b)| \leqslant 1$ and thus $G \cong G_{n,2}$. □

Lemma 2.6. (a) *For integers* n, k, $2 \leqslant k \leqslant \frac{1}{2}n$, *let* $H_{n,k}$ *denote the connected graph of order* n, *which consists of a path* P_{2k-1} *with an end vertex* v, *a complete graph* K_{n-2k+1}, *and an edge joining* v *to exactly one vertex of the* k-*clique.*

(b) *For positive integers* n, k, $\frac{1}{2}n < k \leqslant n-1$, *let* $H_{n,k}$ *denote the connected graph of order* n, *which consists of a path* P_{2n-2k} *with an end vertex* v, *an empty graph of order* $2k-n$ *and edges joining* v *to all vertices of the empty graph. Let* $H \in \mathscr{G}_{n,k}$ *and* $x \in V(H)$. *Then we have*

$$\sigma_H(x) \leqslant (2k-1)(n-k).$$

Equality holds if and only if $H \cong H_{n,k}$ *and* x *is the unique end vertex of* H.

In order to prove Theorem 2.4 suppose that α is the minimum number for which (1) is false. By Lemma 2.5 we have $\alpha > 2$. Suppose further that, for this α, n is minimum under all values for which (1) does not hold and that $G \in \mathscr{G}_{n,\alpha}$ has maximum transmission.

The idea of the proof is based on the following simple observation: "Shrinking" the extremal graph $G_{n,\alpha}$ by deleting the end vertices of a bridge $xy \in E(G)$ and joining their neighbors yields $G_{n-2,\alpha-1}$. A similar construction will be applied to G and then we shall obtain the inequality (1) by induction.

In the sequel, let $\tilde{E}$ denote the set of all edges of G whose deletion does not increase the independence number.

We first show that $\tilde{E}$ is nonempty.

Assume that $\tilde{E} = \emptyset$, i.e. G is α-critical. Consequently, G is 2-vertex-connected (see e.g. [1, p. 284]). Choose a vertex $a \in V(G)$ with minimum transmission $\sigma_G(a)$. Then

$G - a$ is connected, $\alpha(G - a) = \alpha$, and

$$\sigma_G(a) \leqslant \frac{1}{n-1} \sum_{b \in V - \{a\}} \sigma_G(b)$$

$$\leqslant \frac{1}{n-1} \sum_{b \in V - \{a\}} (\sigma_{G-a}(b) + d_G(a, b))$$

$$= \frac{1}{n-1}\sigma(G - a) + \frac{1}{n-1}\sigma_G(a).$$

Obviously, α cannot equal $\frac{1}{2}n$, because $G_{n, n/2}$ is isomorphic to the path P_n, the graph that has maximum transmission under all connected graphs of order n. Hence, by the minimality of α and n, we have

$$\sigma_G(a) \leqslant \frac{1}{n-2}\sigma(G - a) \leqslant \frac{1}{n-2}\sigma(G_{n-1, \alpha}).$$

Thus, we obtain a contradiction by

$$0 \leqslant \sigma(G) - \sigma(G_{n, \alpha})$$

$$\leqslant \sigma(G - a) + 2\sigma_G(a) - \sigma(G_{n, \alpha})$$

$$\leqslant \sigma(G_{n-1, \alpha}) + \frac{2}{n-2}\sigma(G_{n-1, \alpha}) - \sigma(G_{n, \alpha})$$

$$= \frac{n}{n-2}\sigma(G_{n-1, \alpha}) - \sigma(G_{n, \alpha})$$

$$= \begin{cases} 2n(1 - \alpha) - \frac{8}{3}\alpha(\alpha - 1)(\alpha - 2) & \text{if } n \text{ is even} \\ -\frac{8}{3}\alpha(\alpha^2 - 3\alpha + \frac{11}{4}) + 2 & \text{if } n \text{ is odd} \end{cases}$$

$$< 0,$$

and $\tilde{E}$ cannot be empty.

Next we show that for every $xy \in \tilde{E}$ neither x nor y is an end vertex of G. Let $xy \in \tilde{E}$. Suppose that x is an end vertex of G. Let $(\hat{G})$ be the graph obtained from G by joining x to all other neighbors of y. Using $\alpha(G - x) = \alpha(G) - 1$, it is easily checked that

$$\alpha(\hat{G}) = \alpha(G) - 1.$$

By the minimality of α we have

$$\sigma(\hat{G}) \leqslant \sigma(G_{n, \alpha - 1}).$$

Hence, with $\alpha \leqslant \frac{1}{2}n$ and $n \geqslant 3$, we obtain

$$0 \leqslant \sigma(G) - \sigma(G_{n, \alpha})$$

$$= \sigma(\hat{G}) + 2(n - 2) - \sigma(G_{n, \alpha})$$

$$\leqslant \sigma(G_{n, \alpha - 1}) + 2(n - 2) - \sigma(G_{n, \alpha})$$

$$= \begin{cases} -n^2 + 4n + 4\alpha^2 - 12\alpha + 4, & n \text{ even} \\ -n^2 + 4n + 4\alpha^2 - 12\alpha + 5, & n \text{ odd} \end{cases}$$

$$< 0,$$

a contradiction to the maximality of $\sigma(G)$. Thus, x cannot be an end vertex of G. Similarly, y is not an end vertex of G.

Clearly, every edge in $\tilde{E}$ is a bridge of G, for otherwise the transmission of G could be increased by the deletion of an edge without changing the independence number. Consequently, the induced subgraph $G[\tilde{E}]$ is a nonempty forest. Choose an end vertex x of $G[\tilde{E}]$ and let y denote its unique neighbor. Since x is an end vertex of $\hat{G}$ but not of G, there is an edge $xz \in E(G)$ that fulfils $\alpha(G - xz) = \alpha + 1$. A maximum independent set A' of $G - xz$ contains x and z. Thus, in G exists a maximum independent set $A := A' - \{z\}$ with $x \in A$. So far we have shown the existence of a bridge $xy \in E(G)$ with the following properties:

(i) $\alpha(G - xy) = \alpha(G)$,

(ii) x and y are not end vertices,

(iii) there is a maximum independent set $A \subset V(G)$ that contains x.

Now we are able to complete the proof by constructing a smaller graph G' and by making use of the choice of G as a minimum counterexample to inequality (1) in Theorem 2.4. Choose a bridge $xy \in E(G)$ that fulfils (i)–(iii), join each neighbor of x to each neighbor of y, and delete the vertices x and y.

Obviously, the resulting graph G' is connected and is of order $n - 2$. It is easily checked that

$$\alpha(G') = \alpha(G) - 1.$$

Let G_x and G_y denote the connected component of $G - xy$ that contains x and y, respectively. We first note

$$\sigma_G(x) = \sigma(x, G_x) + \sigma(y, G_y) + |V(G_y)|,$$

$$\sigma_G(y) = \sigma(y, G_y) + \sigma(x, G_x) + |V(G_x)|.$$

Further let $V_1 := V(G_x) - \{x\}$ and $V_2 := V(G_y) - \{y\}$. Then we have

$$\sigma(G') = \left(\sum_{a \neq b \in V_1} + \sum_{a \neq b \in V_2} + 2 \sum_{a \in V_1, b \in V_2} \right) d_{G'}(a, b)$$

$$= \left(\sum_{a \neq b \in V_1} + \sum_{a \neq b \in V_2} + 2 \sum_{a \in V_1, b \in V_2} \right) d_G(a, b) - 4|V_1||V_2|$$

$$= \sigma(G) - 2\sigma_G(x) - 2\sigma_G(y) + 2 - 4|V_1||V_2|$$

$$= \sigma(G) - 2n - 4\sigma(x, G_x) - 4\sigma(y, G_y) + 2 - 4|V_1||V_2|.$$

Applying Lemma 2.6 with $k_1 := \alpha(G_x)$ and $k_2 := \alpha(G_y)$ we derive

$$\begin{aligned}\sigma(G) &\leqslant \sigma(G') + 2n - 2 + 4((2k_1 - 1)(|V_1| - k_1 + 1) \\ &\quad + (2k_2 - 1)(|V_2| - k_2 + 1) + |V_1||V_2|) \\ &=: \sigma(G') + 2n - 2 + F.\end{aligned} \tag{2}$$

A simple calculation using $|V_1| + |V_2| = n - 2$ and $k_1 + k_2 = \alpha$ shows

$$F \leqslant \begin{cases} n^2 + 4n(\alpha - 2) - 4\alpha^2 + 4\alpha + 4, & n \text{ even} \\ n^2 + 4n(\alpha - 2) - 4\alpha^2 + 4\alpha + 3, & n \text{ odd.}\end{cases} \tag{3}$$

By the minimality of α, we have

$$\sigma(G') \leqslant \sigma(G_{n-2,\alpha-1}).$$

Together with (2) and (3) this yields after a simple calculation,

$$\sigma(G) \leqslant \sigma(G_{n,\alpha}),$$

which is equivalent to the inequality (1) of Theorem 2.4.

We prove the uniqueness of the extremal graph by induction on α. The case $\alpha = 2$ is settled by Lemma 2.5. Let $G \in \mathscr{G}(n, \alpha)$, $\alpha > 2$ be a graph with maximum transmission. As we have seen in the proof of (1), G contains a bridge xy having the properties (i)–(iii) stated above. We first note that equality in (1) implies equality in (2) and (3). Lemma 2.6 yields that $G_x \cong H_{n_1+1,k_1}$ and $G_y \cong H_{n_2+1,k_2}$. It is easy to see that F attains its maximum only for $|V_1| = 2k_1 - 1 - \alpha + \frac{1}{2}n$ if n is even and for $|V_1| = 2k_1 - 1 + \frac{1}{2}n \pm \frac{1}{2}$ if n is odd, respectively. This yields $G \cong G_{n,\alpha}$. □

By summing up all distances in $G_{n,\alpha}$ we obtain the following corollary.

Corollary 2.7. (a) *If G is a connected graph of order n and independence number α with $2 \leqslant \alpha \leqslant \frac{1}{2}n$, the following inequality holds*:

$$\mu(G) \leqslant \begin{cases} \alpha\dfrac{n-2}{n-1} + \dfrac{1}{n-1} - 4\dbinom{\alpha}{3}\dbinom{n}{2}^{-1} & \text{if } n \text{ is even,} \\ \alpha\dfrac{n-2}{n-1} + \dfrac{1}{n-1} - \dfrac{\alpha-1}{n(n-1)} - 4\dbinom{\alpha}{3}\dbinom{n}{2}^{-1} & \text{if } n \text{ is odd.}\end{cases}$$

(b) *If G is a connected graph of order n and independence number α with $\frac{1}{2}n < \alpha \leqslant n - 1$, the following inequality holds*:

$$\mu(G) \leqslant \begin{cases} \dfrac{n+1}{3} - \dfrac{(4n-7-2\alpha)(2\alpha+2-n)(2\alpha+2)}{6n(n-1)} & \text{if } n \text{ is even,} \\ \dfrac{n+1}{3} - \dfrac{(4n-7-2\alpha)(2\alpha+2-n)(2\alpha+2) + 6(n+\alpha) - 9}{6n(n-1)} & \text{if } n \text{ is odd.}\end{cases}$$

Obviously, for constant n the upper bound on the average distance given in Corollary 2.7 is strict monotone increasing in the range $2 \leqslant \alpha \leqslant \frac{1}{2}n$ and strict monotone decreasing for $\frac{1}{2}n \leqslant \alpha \leqslant n-1$. So it is easy to see that the inverse function of the bound given in Corollary 2.7(b) yields the minimum independence number of a graph in terms of its order and average distance which Erdös asked for in [6]. The determination of the inverse is an easy exercise in the handling of formula manipulation systems. The derivation of a *lower* bound on α from Proposition 2.2 is much easier, especially with no formula manipulation system.

The problem of determining sharp upper and lower bounds on the average distance of a connected graph whose order and *matching number* are given is less difficult than the problem treated above. In the case of the lower bound this is not surprising, for the unique graph with given order and matching number and maximum size is easy to determine.

Proposition 2.8. *If G is a connected graph with n vertices and matching number β, then we have*

$$\mu(G) \geqslant \begin{cases} 1 & \text{if } \beta \geqslant \left\lfloor \dfrac{n}{2} \right\rfloor, \\ 2 - \dfrac{2\beta}{n-1} + \dfrac{\beta(\beta-1)}{n(n-1)} & \text{otherwise.} \end{cases}$$

Equality holds if and only if G is complete or $G \cong K_\beta + (n-\beta)K_1$, respectively.

Proof. The first inequality is trivial, so we only consider the case $\beta < \lfloor \frac{1}{2}n \rfloor$. It is easily shown that G has at most $\frac{1}{2}\beta(2n - \beta - 1)$ edges. Application of Lemma 2.1 yields the proposition. □

Using an equality due to Buckley [2], we can deduce the maximum average distance of a connected graph with given order and matching number from Theorem 2.4.

Lemma 2.9 (Buckley [2]). *Let T be a tree of order $n > 2$ and $L(T)$ denote its line graph, then the following inequality holds:*

$$\mu(L(T)) = \frac{n}{n-2}(\mu(T) - 1).$$

Definition 2.10. For positive integers n, k with $\frac{1}{2}n \geqslant k$, let $\tilde{G}_{n,k}$ be the graph obtained from a path P_{2k-1} with end vertices v_1, v_2 and two disjoint empty graphs G_1, G_2 of order

$$n_1 := \left\lfloor \frac{n+1}{2} \right\rfloor - k, \qquad n_2 := \left\lfloor \frac{n+1}{2} \right\rfloor - k$$

by joining v_i with all vertices of G_i for $i = 1, 2$.

Theorem 2.11. *Let G be a connected graph of order $n > 4$ and matching number $\beta > 1$. Then we have*

$$\mu(G) \leqslant \mu(\tilde{G}_{n,\beta}).$$

Equality holds if and only if $G \cong \tilde{G}_{n,\beta}$.

Proof. We may assume that G has maximum transmission under all connected graphs of order n and matching number β. We first show that G is a tree.

Suppose that G contains a cycle C. Let M be a maximum matching of G. At most half of the edges of C belong to M. Hence, C contains an edge whose deletion neither decreases the matching number nor destroys the connectivity of G. Thus, $\sigma(G - e) > \sigma(G)$, a contradiction to the maximality of $\sigma(G)$.

Let $L(G)$ denote the line graph of G. Then we have $\alpha(L(G)) = \beta(G)$. Theorem 2.4 yields

$$\sigma(L(G)) \leqslant \sigma(G_{n-1,\beta}),$$

with equality if and only if $L(G) \cong G_{n-1,\beta}$. Lemma 2.9 implies

$$\sigma(T_1) \leqslant \sigma(T_2) \iff \sigma(L(T_1)) \leqslant \sigma(L(T_2)),$$

if T_1 and T_2 are trees of the same order. Together with $L(\tilde{G}_{n,\beta}) \cong G_{n-1,\beta}$ we obtain

$$\sigma(G) \leqslant \sigma(\tilde{G}_{n,\beta}).$$

Equality holds if and only if $L(G) \cong G_{n-1,\beta}$, i.e. for $G \cong \tilde{G}_{n,\beta}$. □

References

[1] C. Berge, Graphs (North-Holland, Amsterdam, 2nd revised ed., 1985) 284.

[2] F. Buckley, Mean distance in line graphs, Congr. Numer. 32 (1981) 153–162.

[3] F.R.K. Chung, The average distance and the independence number, J. Graph Theory 12 (1988) 229–235.

[4] P. Dankelmann, Eckenzusammenhang in Graphen, Diplomarbeit, Aachen (1990).

[5] R.C. Entringer, D.E. Jackson and D.A. Snyder, Distance in graphs, Czechoslvak Math. J. 26 (101) (1976) 283–296.

[6] S. Fajtlowicz and W.A. Waller, On two conjectures of GRAFFITI, Congr. Numer. 55 (1986) 51–56.