Average distance, minimum degree, and spanning trees

Article in Journal of Graph Theory · January 2000
DOI: 10.1002/(SICI)1097-0118(200001)33:13.0.CO;2-L

Peter Dankelmann
University of Johannesburg

Average Distance, Minimum Degree, and Spanning Trees

Peter Dankelmann and Roger Entringer

Abstract

The average distance $\mu(G)$ of a connected graph G of order n is the average of the distances between all pairs of vertices of G, i.e. $\mu(G) = \binom{n}{2}^{-1} \sum_{\{x,y\} \subset V(G)} d_G(x,y)$ where $V(G)$ denotes the vertex set of G and $d_G(x,y)$ is the distance between x and y. We prove that every connected graph of order n and minimum degree δ has a spanning tree T with average distance at most $\frac{n}{\delta+1} + 5$. We give improved bounds for K_3-free graphs, C_4-free graphs, and for graphs of given girth.

Let G be a connected graph with vertex set $V(G)$ of order n. The *average distance* $\mu(G)$ of G is defined as $\mu(G) = \binom{n}{2}^{-1} \sum_{\{x,y\} \subset V(G)} d_G(x,y)$, where $d_G(x,y)$ denotes the distance between the vertices x and y in G. In 1987, the computer program GRAFFITI [14] made the attractive conjecture that for every δ-regular connected graph G of order n

$$\mu(G) \le \frac{n}{\delta}.$$

An asymptotically slightly stronger form of this conjecture, stating that $\mu(G) \le n/(\delta + 1) + 2$ for every connected graph of order n and minimum degree δ, was recently proved by Kouider and Winkler [16].

In this paper we consider the problem of finding an upper bound on the average distance for graphs of given minimum degree in connection with the problem of finding a spanning tree of a given graph with small average distance. Johnson, Lenstra, and Rinnooi-Kan [15] showed that the problem of finding a spanning tree T of a given graph G with minimum average distance among all spanning trees of G is NP-hard. Entringer, Kleitman, and Szekely [11] proved that, given a graph G, one can find a spanning tree of G with average distance less than $2\mu(G)$. This, in connection with the result by Kouider and Winkler, implies that every connected graph of order n and minimum degree δ has a spanning tree T with $\mu(T) \le 2n/(\delta+1) + O(1)$. We prove that this bound can be improved to

$$\mu(T) \le \frac{n}{\delta+1} + 5.$$

This gives an alternative proof of an inequality only slightly weaker than the result by Kouider and Winkler. The proof rests on a generalization for vertex weighted graphs of the well known result that among all connected graphs of given order the path has the largest average distance. Applying a similar proof technique, we prove that if G is triangle-free this bound can be improved to

$$\mu(T) \le \frac{2}{3}\frac{n}{\delta} + \frac{25}{3}.$$

and for C_4-free graphs to

$$\mu(T) \le \frac{5}{3}\frac{n}{\delta^2 - 2\lfloor \delta/2 \rfloor + 1} + \frac{29}{3}.$$

Moreover we give similar bounds for graphs of given girth.

For a survey of results on the average distance of graphs before 1984 see Plesník's paper [19]. More recent results can be found in [4, 5, 8, 6, 7, 12, 18, 20, 21].

We use the notation of [2]. In particular we use n, $\delta(G)$ and $g(G)$ for the *order*, the *minimum degree* and the *girth* of a given graph G. If no confusion can occur, we drop the argument G. The kth *neighbourhood* of a given subset $A \subset V(G)$, denoted by $\overline{N}_G^k(A)$ is the set of all vertices x of G with $d_G(x,a) \leq k$ for some $a \in A$. For $\overline{N}_G^k(\{a\})$ we write simply $\overline{N}_G^k(a)$. The *k-th power of G*, denoted by G^k, is the graph with the same vertex set as G, in which two vertices $u \neq v \in V(G)$ are adjacent if $d_G(u,v) \leq k$. For a subset $A \subset V(G)$, the subgraph of G^k induced by A is denoted by $G^k[A]$. For a positive integer k, a *k-packing* of G is a subset $A \subset V(G)$ with $d_G(a,b) > k$ for all $a, b \in A$. If M is a subset of the edge set of G, then $V(M)$ denotes the set of vertices incident with at least one edge of M. For the set of all inegers we use Z.

The following definition is motivated as follows. If the vertices of a graph G stand for sites of facilities, where in each vertex exactly one facility is located, then the average distance $\mu(G)$ gives the expected distance between two randomly selected distinct facilities. Assume now that some vertices host more than one facility and that the distance between facilities located in the same vertex is zero. Let $c(x)$ be the number of facilities located in vertex x and let $N = \sum_{x \in V(G)} c(x)$ be the total number of facilities. Then the expected distance between two randomly selected distinct facilities equals $\binom{N}{2}^{-1} \sum_{x,y \in V(G)} c(x)c(y)d(x,y)$. The weighted total distance (or short distance) was introduced in [3].

Definition 1 *For a weighted graph G with weight function $c : V(G) \to Z$ define the distance of G with respect to c by*

$$\sigma_c(G) = \sum_{\{x,y\} \subset V(G)} c(x)c(y)d_G(x,y)$$

and the average distance of G with respect to c by

$$\mu_c(G) = \binom{N}{2}^{-1} \sigma_c(G),$$

where $N = \sum_{x \in V(G)} c(x)$ is the total weight of the vertices in G.
The distance of a vertex $v \in V(G)$ with respect to c is defined by

$$\sigma_c(v,G) = \sum_{x \in V(G)} c(x)d_G(v,x).$$

It has been proved by several authors [10, 9, 17] that the average distance of a connected graph with n vertices is at most $(n+1)/3$ and that this bound is achieved only by the path of order n. The following lemma generalizes this result.

Lemma 1 *Let G be a weighted graph with weight function c and let k, N be positive integers, N a multiple of k, such that $c(v) \geq k$ for every vertex v of G and $\sum_{v \in V(G)} c(v) \leq N$. Then*

$$\mu_c(G) \leq \frac{N-k}{N-1}\frac{N+k}{3k}.$$

Equality holds if and only if G is a path and $c(v) = k$ for every $v \in V(G)$.

Proof. We prove the equivalent statement

$$\sigma_c(G) \leq \frac{N(N-k)(N+k)}{6k}.$$

The proof is by induction on N/k. If $N/k = 1$ then G has only one vertex of weight k and the average distance of G equals zero. Hence let $N > k$.

Let G be a graph satisfying the hypothesis of the lemma such that $\sigma_c(G)$ is maximum. Then G is a tree. Denote the order of G by n. Note that n is not specified in the statement of the lemma.

CLAIM 1: G is a path.

Suppose that G is not a path and thus has a vertex u of degree at least three. Let $v_1, v_2, \ldots, v_d$ be the neighbours of u and denote the component of $G - u$ containing v_i by G_i and the total weight of G_i by C_i. Without loss of generality we can assume that $C_1 \geq C_2 \geq \ldots \geq C_d$. Consider the graph $H = G - uv_d + v_d v_{d-1}$ with the same vertex weight function c. Since now the distances between the vertices of $G_1 \cup G_2 \cup \ldots \cup G_{d-2} \cup \{u\}$ and the vertices of G_d have increased by one, and the distances between the vertices of G_{d-1} and G_d have decreased by one, we have

$$\sigma_c(H) = \sigma_c(G) + C_d(C_1 + C_2 + \ldots + C_{d-2} + c(u) - C_{d-1}) \geq \sigma_c(G) + C_d c(u) > \sigma_c(G),$$

contradicting the choice of G. Hence G is a path. Denote it by $G = v_1, v_2, \ldots, v_n$.

CLAIM 2: $c(v_1) = k$ or $c(v_n) = k$.

Suppose that $c(v_1), c(v_n) \geq k + 1$. First we note that $c(v_1) \leq 2k - 1$, since otherwise adding a new vertex v_0, joining it to v_1 and splitting the weight of v_1 equally between v_1 and v_0 yielded a graph with larger average distance, contradicting our choice of G.

Let $v_r \in V(G) - \{v_1\}$ be the vertex closest to v_1 with $c(v_r) > k$. Define the weight function $\overline{c}$ by

$$\overline{c}(v_i) = \begin{cases} c(v_1) + 1 & \text{if } i = 1, \\ c(v_r) - 1 & \text{if } i = r, \\ c(v_i) & \text{otherwise.} \end{cases}$$

A simple calculation yields

$$\sigma_{\overline{c}}(G) = \sigma_c(G) + (r-1)\Big(N - (r-1)k - 2c(v_1) - 1\Big).$$

By the maximality of $\sigma_c(G)$ we have

$$0 \geq N - (r-1)k - 2c(v_1) - 1 \geq N - (r+3)k + 1,$$

and thus, since N is a multiple of k,

$$N \leq (r+2)k.$$

Since there are vertices of weight greater than k, this implies

$$n \leq r + 1.$$

Analogously, if $v_s \in V(G) - \{v_n\}$ is the vertex closest to v_n with $c(v_s) > k$, we obtain

$$s \leq 2.$$

We consider two cases: If there is a vertex v_i with $1 < i < n$ and $c(v_i) > k$, then we have $n-1 \le r \le i \le s \le 2$, and thus $n = 3$. Moreover, each vertex has weight greater than k. Using elementary calculus one can show that in this case the statement of the lemma holds.
In the other case, if each vertex except v_1 and v_n has weight k, simple calculus shows that introducing a new vertex v_0, reducing the weight of v_1 and v_n to k and shifting the difference to v_0 yields a graph of larger average distance than G, a contradiction. Note that $c(v_0) = c(v_1) + c(v_2) - 2k \ge k$ since N is a multiple of k. Hence G has at least one end vertex, without loss of generality v_1, of weight k.

Let $\overline{c}$ be the vertex weight function of G restricted to $V(G) - \{v_1\}$. By induction we have

$$\sigma_{\overline{c}}(G - v_1) \le \frac{(N-k)(N-2k)N}{6k}.$$

Since $\sigma_c(v_1, G)$ is maximised if $c(v_2), c(v_3), \ldots = k$, we have

$$\sigma_c(v_1, G) \le k + 2k + 3k + \ldots + \frac{N-k}{k}k = \frac{N}{2}\frac{N-k}{k}. \tag{1}$$

This yields in total

$$\begin{aligned} \sigma_c(G) &= \sigma_{\overline{c}}(G - v_1) + k\sigma_c(v_1, G) \\ &\le \frac{(N-k)(N-2k)N}{6k} + \frac{1}{2}N(N-k) \\ &= \frac{N(N-k)(N+k)}{6k}, \end{aligned}$$

with equality only if equality holds in (1), which implies $c(v_1) = c(v_2) = \ldots = c(v_n) = k$. $\square$

Theorem 1 *Let G be a connected graph with n vertices and minimum degree δ. Then G has a spanning tree T with*

$$\mu(T) \le \frac{n}{\delta + 1} + 5. \tag{2}$$

Apart from the additive constant, this inequality is best possible.

Proof. First we find a maximal 2-packing A of G using the following procedure. Choose an arbitrary vertex v of G and let $A = \{v\}$. If there exists a vertex u in G with $d_G(u, A) = 3$ add u to A. Add vertices u with $d_G(u, A) = 3$ to A until each of the vertices not in A is within distance two of A. Note that $G^3[A]$ is connected.
Let $T_1 \le G$ be the forest with vertex set $\overline{N}_G(A)$ and whose edge set consists of all edges incident with a vertex in A. By our construction of A, there exist $|A| - 1$ edges in G, each of them joining two neighbours of distinct elements of A, whose addition to T_1 yields a tree $T_2 \le G$.
Now each vertex $u \in V(G) - V(T_2)$ is adjacent to some vertex $u' \in V(T_2)$. Let T be the spanning tree of G with edge set $E(T_2) \cup \{uu' | u \in V(G) - V(T_2)\}$.
We now prove that

$$\mu(T) \le \frac{n}{\delta + 1} + 5.$$

For every vertex $u \in V(T)$ let u_A be the unique vertex in A closest to u in T. We now move the weight of every vertex to the closest vertex in A, i.e., we define a weight function $c : V(T) \to Z$ by

$$c(u) = \Big|\{x \in V(T) | x_A = u\}\Big| \quad \text{for } u \in V(T).$$

Note that $c(u) = 0$ if $u \notin A$ and that

$$c(u) \geq \delta + 1 \quad \text{for each } u \in A.$$

Since the weight of each vertex was moved over a distance not exceeding two, no distance between two weights has changed by more than 4 and thus

$$\mu(T) \leq \mu_c(T) + 4.$$

Since the weight c is concentrated exclusively on the vertices of A and since $T^3[A]$ is connected, we have

$$\mu_c(T) \leq 3\mu_c(T^3[A]).$$

Let N be the least multiple of $\delta + 1$ with $N \geq n$. By Lemma 1 we have

$$\mu_c(T^3[A]) \leq \frac{N - \delta - 1}{N - 1} \frac{N + \delta + 1}{3(\delta + 1)} \leq \frac{N + 1}{3(\delta + 1)}.$$

Combining these inequalities, in conjunction with $N \leq n + \delta$, yields

$$\begin{aligned} \mu(T) &\leq 3\frac{N + 1}{3(\delta + 1)} + 4 \\ &\leq \frac{n + \delta + 1}{\delta + 1} + 4 \\ &= \frac{n}{\delta + 1} + 5, \end{aligned}$$

as desired.

It remains to show that, apart from the value of the additive constant, (2) is best possible. For given integers n, δ, k with $n = k(\delta + 1)$, let $G_1, G_2, \ldots, G_k$ be disjoint copies of the complete graph $K_{\delta+1}$ and let $a_i b_i \in E(G_i)$. Let $G_{n,\delta}$ be the graph obtained from the union of $G_1, G_2, \ldots, G_k$ by deleting the edges $a_i b_i$ for $i = 2, 3, \ldots, k - 1$ and adding the edges $a_{i+1} b_i$ for $i = 1, 2, \ldots, k - 1$. As stated in [16], $G_{n,\delta}$ has order n, minimum degree δ and

$$\mu(G_{n,\delta}) > \frac{n}{\delta + 1}.$$

Hence every spanning tree T of $G_{n,\delta}$ has average distance greater than $n/(\delta + 1)$. □

Theorem 2 *Let G be a connected triangle-free graph with n vertices and minimum degree δ. Then G has a spanning tree T with*

$$\mu(T) \leq \frac{2}{3}\frac{n}{\delta} + \frac{25}{3}. \tag{3}$$

Apart from the additive constant, this inequality is best possible.

Proof. The proof is similar to the proof of Theorem 1. First we find a matching M of G using the following procedure. Choose an arbitrary edge $e \in E(G)$ and let $M = \{e\}$. If there exists an edge f in G with $d_G(f, V(M)) = 3$ add f to M. Add edges with $d_G(f, V(M)) = 3$ to M until each of the edges not in M is within distance two of M.
Let $T_1 \leq G$ be the forest with vertex set $\overline{N}_G(V(M))$ and whose edge set consists of all edges incident with a vertex in $V(M)$. By our construction of M, there exist $|M|-1$ edges in G, each of them joining two distinct components of T_1, whose addition to T_1 yields a tree $T_2 \leq G$.
Now each vertex in $V(G)-V(T_2)$ is within distance three of some vertex of T_2. Let $T \geq T_2$ be a spanning tree of G in which $d_T(x, V(M)) = d_G(x, V(M))$ for every vertex $x \in V(G)$. We now prove that

$$\mu(T) \leq \frac{2}{3}\frac{n}{\delta} + \frac{25}{3}.$$

For every vertex $u \in V(T)$ let u_M be the unique vertex in $V(M)$ closest to u in T. We now move the weight of every vertex to the closest vertex in $V(M)$, i.e., we define a weight function $c : V(T) \to Z$ by

$$c(u) = \Big|\{x | x_M = u\}\Big| \quad \text{for } u \in V(T).$$

Note that $c(u) = 0$ if $u \notin V(M)$. Let $uv \in M$. Since G is triangle free u and v do not have any neighbours in common. Hence $deg_T(u) = deg_G(u) \geq \delta$, which implies that

$$c(u) \geq \delta \quad \text{for each } u \in V(M).$$

Since the weight of each vertex was moved over a distance not exceeding three, no distance between two weights has changed by more than 6 and thus

$$\mu(T) \leq \mu_c(T) + 6.$$

Now the weight c is concentrated exclusively on the vertices of $V(M)$. Consider the line graph $L = L(T)$. Define the weight function $\overline{c}$ on $V(L) = E(T)$:

$$\overline{c}(uv) = \begin{cases} c(u) + c(v) & \text{if } uv \in M, \\ 0 & \text{if } uv \notin M. \end{cases}$$

Note that $\overline{c}(uv) \geq 2\delta$ for $uv \in M$.
If $e_1, e_2 \in E(T)$ are edges incident with vertices $v_1, v_2 \in V(T)$, respectively, then $d_T(v_1, v_2) \leq d_L(e_1, e_2) + 1$. Hence no distance between weights has increased by more than one and thus

$$\mu_c(T) \leq \mu_{\overline{c}}(L) + 1.$$

If the distance $d_T(e, f)$ between two matching edges $e, f \in M$ equals three, then $d_L(e, f) \leq 4$. Hence $L^4[M]$ is connected and we have

$$\mu_{\overline{c}}(L) \leq 4\mu_{\overline{c}}(L^4[M]).$$

Let N be the least multiple of 2δ with $N \geq n$. By Lemma 1 we have

$$\mu_{\overline{c}}(L^4[M]) \leq \frac{N - 2\delta}{N - 1}\frac{N + 2\delta}{6\delta} \leq \frac{N + 1}{6\delta}.$$

Combining these inequalities, in conjunction with $N \leq n+2\delta-1$, yields

$$\begin{aligned}
\mu(T) &\leq 4\mu_{\overline{c}}(L^4)+7 \\
&\leq \frac{2}{3}\frac{N+1}{\delta}+7 \\
&\leq \frac{2}{3}\frac{n+2\delta}{\delta}+7 \\
&\leq \frac{2}{3}\frac{n}{\delta}+\frac{25}{3},
\end{aligned}$$

as desired.

It remains to show that, apart from the value of the additive constant, (3) is best possible. Given integers n, δ, k with $n = k\delta/2$, let H_i be a disjoint copy of the empty graph $(\delta/2)K_1$ for $1 \leq i \leq k-2$, $i \neq 2, k-3$. For $i = 2, k-3$ let H_i be a copy of the empty graph δK_1. Let $G'_{n,\delta}$ be the graph obtained from the union of $H_1, H_2, \ldots, H_{k-2}$ by joining each vertex in H_i to each vertex in H_{i+1} for $i = 1, 2, \ldots, k-3$. It is easy to verify that $G'_{n,\delta}$ is bipartite and therefore triangle free, that it has order n, minimum degree δ and that for large n

$$\mu(G'_{n,\delta}) = \frac{2n}{3\delta} + O(1).$$

If δ is odd, a similar construction can be devised. Hence every spanning tree T of $G'_{n,\delta}$ has average distance greater than $2n/(3\delta)+c$ for some constant c. □

Corollary 1 *Let G be a connected triangle-free graph with n vertices and minimum degree δ. Then*

$$\mu(G) \leq \frac{2}{3}\frac{n}{\delta}+\frac{25}{3}.$$

Apart from the additive constant, this inequality is best possible. □

Theorem 3 *(i) Let G be a connected C_4-free graph with n vertices and minimum degree δ. Then G has a spanning tree T with*

$$\mu(T) \leq \frac{5}{3}\frac{n}{\delta^2-2\lfloor\delta/2\rfloor+1}+\frac{29}{3}.$$

(ii) There exists an infinite number of C_4-free graphs with n vertices and minimum degree δ such that for every spanning tree T of G

$$\mu(T) \geq \frac{5}{3}\frac{n}{\delta^2+3\delta+2}+O(1).$$

Proof. (i) First we find a maximal 4-packing A of G using the following procedure. Choose an arbitrary vertex v of G and let $A = \{v\}$. If there exists a vertex u in G with $d_G(u, A) = 5$ add u to A. Add vertices with $d_G(u, A) = 5$ to A until each of the vertices not in A is within distance four of A. Note that $G^5[A]$ is connected.
For $u \in A$ let $T_1(u)$ be a tree with vertex set $\overline{N}^2_G(u)$ which is distance preserving to u. Then $T_1 = \cup_{u\in A} T_1(u)$ is a subforest of G. By our construction of A, there exist $|A|-1$ edges in G, each of them joining two components of T_1, whose addition to T_1 yields a tree $T_2 \leq G$.

Extend the tree T_2 to a spanning tree T of G with $d_T(x,A) = d_G(x,A)$ for each $x \in V(G)$. We now prove that

$$\mu(T) \le \frac{5}{3}\frac{n}{\delta^2 - 2\lfloor\delta/2\rfloor + 1} + \frac{29}{3}.$$

For every vertex $u \in V(T)$ let u_A be the unique vertex in A closest to A in T. We now move the weight of every vertex to the closest vertex in A, i.e., we define a weight function $c : V(T) \to Z$ by

$$c(u) = \Big|\{x | x_A = u\}\Big| \quad \text{for } u \in V(T).$$

Note that $c(u) = 0$ if $u \notin A$. Now every vertex $v \in V(G)$ has at least δ neighbours. Since G is C_4-free no two neighbours of v have a common neighbour apart from v. Hence $|\overline{N}^2_G(v)| \ge \delta^2 - 2\lfloor\delta/2\rfloor + 1$ and thus

$$c(u) \ge \delta^2 - 2\lfloor\delta/2\rfloor + 1 \quad \text{for each } u \in A.$$

Since the weight of each vertex was moved over a distance not exceeding four, no distance between two weights has changed by more than 8 and thus

$$\mu(T) \le \mu_c(T) + 8.$$

With similar arguments as in the proof of Theorem 1 we obtain

$$\begin{aligned} \mu(T) &\le \frac{5}{3}\frac{N+1}{\delta^2 - 2\lfloor\delta/2\rfloor + 1} + 8, \\ &\le \frac{5}{3}\frac{n}{\delta^2 - 2\lfloor\delta/2\rfloor + 1} + \frac{29}{3}, \end{aligned}$$

as desired.

(ii) To prove the second part of the theorem consider the following graph $G''_{n,\delta}$, first described in [13]. Let H be the graph whose vertices are the triples $\underline{x} = (x_1, x_2, x_3) \ne 0$ where $x_1, x_2, x_3 \in GF(q)$, the finite field of order q. We consider $\underline{x}$ and $\underline{y}$ identical if $\underline{x} = \lambda \underline{y}$ for some $\lambda \in GF(q)$, $\lambda \ne 0$. Let $\underline{x}$ and $\underline{y}$ be adjacent in H if $\underline{xy} = 0$. Clearly H is C_4-free and has $q^2 + q + 1$ vertices, each of degree q or $q+1$. Let $\underline{u}, \underline{v}, \underline{z} \in V(H)$ be fixed vertices satisfying $\underline{uz} = \underline{vz} = \underline{zz} = 0$. Let $\underline{z} = \underline{u_0}, \underline{u_1}, \ldots, \underline{u_{q+1}}$ and let $\underline{z} = \underline{v_0}, \underline{v_1}, \ldots, \underline{v_{q+1}}$ denote the neighbours of $\underline{u}$ and $\underline{v}$, respectively. For every i with $1 \le i \le q$ there exists a unique $j(i)$ with $1 \le j(i) \le q$ such that $\underline{u_i v_{j(i)}} \in E(H)$. On the other hand no u_i or v_i is adjacent to z in H.

Let H' denote the graph obtained from H by removing the vertex $\underline{z}$ and all edges of the form $\underline{u_i v_{j(i)}}$, $1 \le i \le q$. Then $d_{H'}(\underline{u}, \underline{v}) = 4$ and every vertex has degree at least $q - 1 = \delta$ in H'.

For n a multiple of $q^2 + q + 1 = \delta^2 + 3\delta + 2$ let $G''_{n,\delta}$ be the graph obtained from the union of $k = n/(\delta^2 + 3\delta + 2)$ disjoint copies $H'_1, H'_2, \ldots, H'_k$ of H' by adding the edges $u^t v^{t+1}$ for $1 \le t \le k-1$, where u^t and v^t are the vertices in H'_t corresponding to $\underline{u}$ and $\underline{v}$ in H'. A simple calculation shows that $G''_{n,\delta}$ has $k(\delta^2 + 3\delta + 2)$ vertices and that

$$\mu(G''_{n,\delta}) = \frac{5}{3}\frac{n}{\delta^2 + 3\delta + 2} + O(1).$$

Since every spanning tree of $G''_{n,\delta}$ has average distance at least $\mu(G''_{n,\delta})$, the theorem follows. □

Corollary 2 *Let G be a connected C_4-free graph with n vertices and minimum degree δ. Then*

$$\mu(G) < \frac{5}{3}\frac{n}{\delta^2 - 2\lfloor \delta/2 \rfloor + 1} + \frac{29}{3}.$$

$\square$

Theorem 4 *Let G be a connected graph with n vertices, minimum degree $\delta \geq 3$ and girth g.*
(a) If g is odd then G has a spanning tree T with

$$\mu(T) \leq \frac{gn}{3K} + \frac{7}{3}g - 2,$$

where

$$K = 1 + \delta \frac{(\delta-1)^{(g-1)/2} - 1}{\delta - 2}.$$

(b) If g is even then G has a spanning tree T with

$$\mu(T) \leq \frac{gn}{3L} + \frac{7}{3}g - 2,$$

where

$$L = 2\frac{(\delta-1)^{g/2} - 1}{\delta - 2}.$$

Proof. (a) First we find a maximal $(g-1)$-packing A of G using the following procedure. Choose an arbitrary vertex v of G and let $A = \{v\}$. If there exists a vertex u in G with $d_G(u, A) = g$ add u to A. Add vertices u with $d_G(u, A) = g$ to A until each of the vertices not in A is within distance at most $g-1$ of A.
Let $T_1 \leq G$ be the forest with vertex set $\overline{N}_G^{(g-1)/2}(A)$ and edges such that the distance from every vertex to A is preserved. By our construction of A, there exist $|A|-1$ edges in G, each of them joining two vertices of distinct components of T_1, whose addition to T_1 yields a tree $T_2 \leq G$.
Now each vertex $u \in V(G) - V(T_2)$ is within distance $(g-1)/2$ to some vertex $u' \in V(T_2)$. Let T be a spanning tree of G containing T_2, such that the distance of every vertex to A is preserved.
We now prove that

$$\mu(T) \leq \frac{g}{3}\frac{n-1}{K} + 2g - \frac{5}{3},$$

For every vertex $u \in V(T)$ let u_A be the unique vertex in A closest to u in T. We now move the weight of every vertex to the closest vertex in A, i.e., we define a weight function $c : V(T) \to Z$ by

$$c(u) = \Big| \{x \in V(T) | x_A = u\} \Big| \quad \text{for } u \in V(T).$$

Note that $c(u) = 0$ if $u \notin A$. Since G has girth g we have, for each $u \in A$,

$$\begin{aligned} c(u) &\geq |\overline{N}_G^{(g-1)/2}(u)| \\ &\geq 1 + \delta + \delta(\delta-1) + \delta(\delta-1)^2 + \ldots + \delta(\delta-1)^{(g-3)/2} \\ &= 1 + \delta\frac{(\delta-1)^{(g-1)/2} - 1}{\delta - 2} \\ &= K. \end{aligned}$$

The remainder of the proof is analogous to the proof of Theorem 1.

(b) The proof of part (b) is analogous to the proof of Theorem 2. $\square$

We remark that, in general, the bounds given in Theorem 4 are not sharp. If, however, δ and g are such that there exists a Moore graph of minimum degree δ and girth g, i.e. a graph with minimum degree δ, girth g and order $1+\delta+\delta(\delta-1)+\delta(\delta-1)^2+\ldots+\delta(\delta-1)^{(g-3)/2}$ (if g is odd) or $2(1+(\delta-1)+(\delta-1)^2+(\delta-1)^3+\ldots+(\delta-1)^{(g-2)/2}$ (if g is even), then the bounds are asymptotically sharp. This is shown by the graph $G_{k,\delta,g}$ constructed below. For more on Moore graphs we refer the reader to Biggs' book [1]).

The graph $G_{k,\delta,g}$ is constructed as follows. Let $G_1, G_2, \ldots, G_k$ be disjoint copies of the (δ,g)-Moore graph and let a_ib_i be an edge of G_i. Let $G_{k,\delta,g}$ be the graph obtained from the union of $G_1, G_2, \ldots, G_k$ by deleting the edges a_ib_i for $i=2,3,\ldots,k-1$ and adding the edges $a_{i+1}b_i$ for $i=1,2,\ldots,k-1$. It is easy to verify that, for large k,

$$\mu(G_{k,\delta,g}) = \begin{cases} \frac{g}{3}\frac{n-1}{K} + O(1) & \text{for odd } g, \\ \frac{g}{3}\frac{n-1}{L} + O(1) & \text{for even } g. \end{cases}$$

Hence no spanning tree of $G_{k,\delta,g}$ can have a smaller average distance.

References

[1] Biggs, N., "Algebraic Graph Theory", Second Edition, Cambridge University Press, Cambridge, 1993.

[2] Chartrand, G. and L. Lesniak, "Graphs and Digraphs", Second Edition, Wadsworth & Brooks/Cole, Monterey, 1986.

[3] Dankelmann, P., Computing the average distance of an interval graph. Inform. Process. Lett. **48** (1993), 311 – 314.

[4] Dankelmann, P., Average distance and independence number. Discrete Appl. Math. **51** (1994), 75 – 83.

[5] Dankelmann, P., Average distance and domination number, Discrete Appl. Math. **80** (1997), 21 – 35.

[6] Dankelmann, P., O.R. Oellermann and H.C. Swart, The average Steiner distance of a graph. J. Graph Theory **22** (1996), 15 – 22.

[7] Dankelmann, P., O.R. Oellermann and H.C. Swart, On the average Steiner distance of graphs with prescribed properties, Discrete Appl. Math. **79** (1997), 91 – 103.

[8] Dankelmann, P., A note on MAD spanning trees. J. Combin. Math. Combin. Comput. (to appear).

[9] Doyle, J.K. and J.E. Graver, Mean distance in a graph. Discrete Math. **7** (1977), 147 – 154.

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

[11] Entringer, R.C., D.J. Kleitman and L.A. Székely, A note on spanning trees with minimum average distance. Bull. Inst. Combin. Appl. **17** (1996), 71–78.

[12] Entringer, R.C., Distance in graphs: Trees. J. Combin. Math. Combin. Comput. **24** (1997), 65 – 84.

[13] Erdös, P., J. Pach, R. Pollack, Z. Tuza, Radius, diameter, and minimum degree. J. Combin. Theory B **47** (1989), 73–79.

[14] Fajtlowicz, S. and W.A. Waller, On conjectures of GRAFFITI II. Congr. Numer. **60** (1987), 187–197.

[15] Johnson, D.S., J.K. Lenstra, and A.H.G. Rinnooy-Kan, The complexity of the network design problem. Networks **8** (1978), 279–285.

[16] Kouider, M. and P. Winkler, Mean distance and minimum degree. J. Graph Theory **25** (1997), 95–99.

[17] Lovász, L., Combinatorial Problems and Exercises. Akadémiai Kiadó, Budapest (1979).

[18] Oellermann, O.R., Computing the average distance of a distance hereditary graph in linear time. Congr. Numer. **103** (1994), 219–223.

[19] Plesník, J., On the sum of all distances in a graph or digraph. J. Graph Theory **8** (1984), 1–21.

[20] Tomescu, I., On the sum of all distances in chromatic blocks. J. Graph Theory **18** (1994), 83–102.

[21] Yeh, Y.-N. and I. Gutman, On the sum of all distances in composite graphs. Discrete Math. **135** (1994), 359–365.

Peter Dankelmann, School of Mathematics and Statistics, University of Natal, Durban 4041, South Africa. Email dankelma@scifs1.und.ac.za
Roger Entringer, Department of Mathematics, University of New Mexico, Albuquerque, NM, USA 87131. Email entring@math.unm.edu