Disproof of a conjecture about average Steiner distance

Tao Jiang *

April 29, 2004

Abstract

Given a connected graph G and a subset S of vertices, the *Steiner distance* of S in G is the minimum number of edges in a tree in G that contains all of S. Given a positive integer m, let $\mu_m(G)$ denote the average Steiner distance over all sets S of m vertices in G. In particular, $\mu_2(G)$ is just the average distance of G, often denoted by $\mu(G)$. Dankelmann, Oellermann, and Swart [1] conjectured that if G is a connected graph of order n and $3 \leq m \leq n$, then $\frac{\mu_m(G)}{\mu(G)} \geq 3(\frac{m-1}{m+1})$. In this note, we disprove their conjecture by showing that

$$\lim_{m\to\infty} \inf \left\{ \frac{\mu_m(G)}{\mu(G)} : G \text{ is connected and } n(G) \geq m \right\} = 2.$$

1 Introduction

We only consider connected simple graphs. Given two vertices u, v in a graph G, the distance between u and v, denoted by $dist_G(u, v)$, is the number of edges in a shortest u, v-path. The distance notion can

*Department of Mathematics and Statistics, Miami University, Oxford, OH 45056, USA, jiangt@muohio.edu.

be generalized for a set S of more than two vertices by considering a smallest connected subgraph (tree) of G that contains S. More specifically, given a set S of vertices in a connected simple graph G, we call a smallest tree in G that contains S a *Steiner tree* for S. The *Steiner distance* of S in G, denoted by $d_G(S)$, is defined to be the number of edges in a Steiner tree for S. In particular, when $|S| = 2$, the Steiner distance of S is the same as the distance between the two vertices in S.

The Steiner distance was introduced by Dankelmann, Oellermann, and Swart in [1]. Motivated by the extensive study of average distance of a graph in the literature, the authors introduced the average Steiner distance of a graph as follows. Given a positive integer m, the average Steiner m-distance of G, denoted by $\mu_m(G)$ is the average of $d_G(S)$ over all sets S of m vertices in G. Note that $\mu_2(G)$ is just the average distance of G, which is often denoted by $\mu(G)$ in the literature.

In [1], Dankelmann, Oellermann, and Swart quite thoroughly studied the function $\mu_m(G)$, obtaining various bounds (most of which are sharp) on $\mu_m(G)$ in terms of m and the order of G. Because of the interest in average distance $\mu(G)$ of a graph G, it is natural to wonder how $\mu_m(G)$ is generally related to $\mu(G)$. In [1] (Corollary 2.1), it is shown that $\mu_m(G) \leq (m-1)\mu(G)$ always, with equality achieved by complete graphs. In other words, one has $\frac{\mu_m(G)}{\mu(G)} \leq m - 1$. For a lower bound, Danklemann, Oellerman, and Swart [1] conjectured the following.

Conjecture 1.1 *[1] If G is a connected graph with n vertices and $3 \leq m \leq n$, then*

$$\frac{\mu_m(G)}{\mu(G)} \geq 3\frac{m-1}{m+1}.$$

This conjecture was verified for $m = 2, 3$ in the same paper. The authors also noted that the conjectured bound would be sharp since equality can be achieved for paths.

In this note, we disprove Conjecture 1.1 for all $m \geq 4$ by proving

Theorem 1.2 *For all positive integers $m \geq 2$, we have for even m*

$$\begin{aligned} 2-\frac{2}{m} &\leq \inf\left\{\frac{\mu_m(G)}{\mu(G)} : G \text{ is connected and } n(G) \geq m\}\right\} \\ &\leq 2-\frac{1}{2^{m-2}}, \end{aligned}$$

and for odd m

$$\begin{aligned} 2-\frac{2}{m+1} &\leq \inf\left\{\frac{\mu_m(G)}{\mu(G)} : G \text{ is connected and } n(G) \geq m\}\right\} \\ &\leq 2-\frac{1}{2^{m-2}}. \end{aligned}$$

Hence, in particular,

$\lim_{m\to\infty} \inf\left\{\frac{\mu_m(G)}{\mu(G)} : G \text{ is connected and } n(G) \geq m\right\} = 2.$

Note that while Theorem 1.2 disproves Conjecture 1.1 for $m \geq 4$, it does confirm the fact that Conjecture 1.1 holds for $m = 2, 3$.

2 Main results

For convenience, we split Theorem 1.2 into two propositions as follows. As usual, $[n]_k = n(n-1)\cdots(n-k+1)$. The *length* of a path is the number of edges in it.

Proposition 2.1 *For all $m \geq 2$, we have*

$\inf\left\{\frac{\mu_m(G)}{\mu(G)} : G \text{ is connected and } n(G) \geq m\right\} \leq 2 - \frac{1}{2^{m-2}}.$

Proof. Let $M = 2 - \frac{1}{2^{m-2}}$. It suffices to prove that given any $\epsilon > 0$, there exists a graph G with $\frac{\mu_m(G)}{\mu(G)} \leq M + \epsilon$. Let p be a large enough positive integer and $\gamma < \frac{1}{4}$ a small enough positive number such that $\frac{4m}{p+1} \leq \frac{\epsilon}{2}$ and $\frac{M+2\gamma}{1-2\gamma} \cdot \frac{m+p-1}{p+1} \leq M + \frac{\epsilon}{2}$. Such p and γ clearly exist. For the now fixed p, let q be a positive integer such that $\frac{2q^2}{[p+2q]_2} \geq \frac{1}{2} - \gamma$ and $\frac{2[q]_m}{[p+2q]_m} \geq \frac{1}{2^{m-1}} - \gamma$. It is easy to see that such q exists.

Let G denote the tree obtained from two vertex disjoint stars with q leaves, say with centers u, v, respectively, by joining u, v with a path of length $p-1$. Thus, G is a tree on $p+2q$ vertices in which u, v each has q leaf neighbors. For convenience, let A denote the set of q leaf neighbors of u, and let B denote the set of q leaf neighbors of v. Note that removing A and B from G leaves a path on p vertices with u, v as endpoints.

We now estimate $\mu_m(G)$ and $\mu(G)$. Given a subset S of vertices in G, let $d(S)$ denote the Steiner distance of S. Suppose we randomly choose a subset S of $V(G)$ of size t, with each subset of $V(G)$ of size t being chosen with equal probability. Then $\mu_t(G)$ is just the expected value of $d(S)$, which we denote by $E(d(S))$.

Let S be a random 2-subset of $V(G)$. Clearly, we have $d(S) = p+1$ if $|S \cap A| = |S \cap B| = 1$. Also, $Prob(|S \cap A| = |S \cap B| = 1) = \frac{q^2}{\binom{p+2q}{2}} = \frac{2q^2}{[p+2q]_2} \geq \frac{1}{2} - \gamma$ by our choice of p and q. Thus,

$$\begin{aligned} \mu(G) = \mu_2(G) = E(d(S)) &\geq Prob(|S \cap A| = |S \cap B| = 1) \cdot (p+1) \\ &\geq (\frac{1}{2} - \gamma)(p+1). \end{aligned} \tag{1}$$

Let S be a random m-subset of $V(G)$. We have $d(S) = m$ if $S \subseteq A$ or $S \subseteq B$ and $d(S) \leq m+p-1$ otherwise. Furthermore, $Prob((S \subseteq A)\text{or}(S \subseteq B)) = \frac{2\binom{q}{m}}{\binom{p+2q}{m}} = \frac{2[q]_m}{[p+2q]_m} \geq \frac{1}{2^{m-1}} - \gamma$, by our choice of p and q. Hence,

$$\begin{aligned} \mu_m(G) = E(d(S)) &\leq [1 - (\frac{1}{2^{m-1}} - \gamma)](m+p-1) + (\frac{1}{2^{m-1}} - \gamma)m \\ &\leq (1 - \frac{1}{2^{m-1}} + \gamma)(m+p-1) + m. \end{aligned} \tag{2}$$

By Equations (1) and (2) and our choice of p, q, γ, we have

$$\begin{aligned} \frac{\mu_m(G)}{\mu(G)} &\leq \frac{1 - \frac{1}{2^{m-1}} + \gamma}{\frac{1}{2} - \gamma} \cdot \frac{m+p-1}{p+1} + \frac{m}{(\frac{1}{2} - \gamma)(p+1)} \\ &\leq \frac{M + 2\gamma}{1 - 2\gamma} \cdot \frac{m+p-1}{p+1} + \frac{4m}{p+1} \\ &\leq (M + \frac{\epsilon}{2}) + \frac{\epsilon}{2} = M + \epsilon. \end{aligned}$$

■

The following proposition implies the lower bounds in Theorem 1.2.

Proposition 2.2 *Let $m \geq 2$ be a positive integer. Let G be a connected graph on at least m vertices. If m is even, then $\frac{\mu_m(G)}{\mu(G)} \geq 2-\frac{2}{m}$. If m is odd, then $\frac{\mu_m(G)}{\mu(G)} \geq 2-\frac{2}{m+1}$.*

Proof. Let W be any set of m vertices in G. Let $d(W)$ denote the Steiner distance of W. Let T denote a Steiner tree for W. Then $e(T) = d(W)$, where $e(T)$ denotes the number of edges in T. For any pair u, v of vertices in W, let $P_{u,v}$ denote the unique u,v-path in T. Let $l(P_{u,v})$ denote the length of $P_{u,v}$. Let $\Pi = \{P_{u,v} : u, v \in W\}$, i.e. Π is the collection of all paths $P_{u,v}$ in T connecting two vertices u, v of W.

For any edge e in T, let $f(e)$ denote the number of paths in Π that contain e. Let F and F' denote the two components of $T - e$. Suppose F contains a vertices of W and F' contains b vertices of W. Then $a + b = m$ and $f(e) = ab \leq \beta_m$, where $\beta_m = \frac{m^2}{4}$ if m is even and $\beta_m = \frac{(m-1)(m+1)}{4}$ if m is odd. Hence, $\Sigma_{u,v\in W} l(P_{u,v})$ counts each edge of T at most β_m times. This yields

$$\Sigma_{u,v\in W} dist_G(u,v) \leq \Sigma_{u,v\in W} l(P_{u,v}) \leq \beta_m e(T) = \beta_m d(W)$$

Summing over all m-subsets W of $V(G)$, we have

$$\Sigma_W \Sigma_{u,v\in W} dist_G(u,v) \leq \beta_m \Sigma_W d(W). \tag{3}$$

Now, since each pair u, v is contained in $\binom{n-2}{m-2}$ m-subsets W,

$$\begin{aligned} \Sigma_W \Sigma_{u,v\in W} dist_G(u,v) &= \binom{n-2}{m-2} \Sigma_{u,v\in V(G)} dist_G(u,v) \\ &= \binom{n-2}{m-2}\binom{n}{2}\mu(G), \end{aligned}$$

where the last equality follows from the definition of $\mu(G)$. Also, $\Sigma_W d(W) = \binom{n}{m}\mu_m(G)$ by the definition of $\mu_m(G)$. Hence inequality

(3) becomes

$$\binom{n-2}{m-2}\binom{n}{2}\mu(G) \le \beta_m \binom{n}{m}\mu_m(G),$$

from which we have $\frac{\mu_m(G)}{\mu(G)} \ge \frac{m(m-1)}{\beta_m}$. Using $\beta_m = \frac{m^2}{4}$ for even m and $\beta_m = \frac{(m-1)(m+1)}{4}$ for odd m, we have $\frac{\mu_m(G)}{\mu(G)} \ge 2 - \frac{2}{m}$ for even m and $\frac{\mu_m(G)}{\mu(G)} \ge 2 - \frac{2}{m+1}$ for odd m. ■

References

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

[2] P. Dankelmann, H. Swart, and O. Oellermann, *The average Steiner distance of graphs with prescribed properties*, Discr. Appl. Math. **79** (1997), 91–103.

[3] J. Plesník, *On the sum of all distances in a graph or digraph*, J. of Graph Theory **8** (1984), 1–24.