# A NOTE ON DOMINATING SETS AND AVERAGE DISTANCE

ERMELINDA DELAVIÑA, RYAN PEPPER, AND BILL WALLER
UNIVERSITY OF HOUSTON – DOWNTOWN, HOUSTON, TX, 77002

Abstract. We show that the total domination number of a simple connected graph is greater than the average distance of the graph minus one-half, and that this inequality is best possible. In addition, we show that the domination number of the graph is greater than two-thirds of the average distance minus one-third, and that this inequality is best possible. Although the latter inequality is a corollary to a result of P. Dankelmann, we present a short and direct proof.

## 1. Introduction and Key Definitions

Let $G = (V, E)$ be a simple connected graph of finite of order $|V| = n$. Although we may identify a graph $G$ with its set of vertices, in cases where we need to be explicit we write $V(G)$ to denote the vertex set of $G$. A set $D$ of vertices of a graph $G$ is called a *dominating set* provided each vertex of $V - D$ is adjacent to a member of $D$. The *domination number* of $G$, denoted $\gamma = \gamma(G)$, is the cardinality of a smallest dominating set in $G$. Likewise, a set $D$ of vertices is called a *total dominating set* provided each vertex of $V$ is adjacent to a member of $D$. The *total domination number* of $G$, denoted $\gamma_t = \gamma_t(G)$, is the cardinality of a smallest total dominating set in $G$. The *distance* between two vertices $u$ and $v$ in $G$ is the length of a shortest path in $G$ connecting $u$ and $v$. The *Wiener index* or *total distance* of $G$, denoted by $W = W(G)$, is the sum of all distances between unordered pairs of distinct vertices of $G$ [5]. The *average distance* of $G$, denoted by $\bar{D} = \bar{D}(G)$, is $2W/[n(n-1)]$. Put another way, this number gives, on average, the distance between a pair of vertices of $G$. Unless stated otherwise, when we refer to a subgraph of $G$, we mean an induced subgraph.

The total domination number of a graph was first introduced in [2]. This invariant remains of interest to researchers as evidenced by numerous recent papers. Various upper and lower bounds on $\gamma_t$ have been discovered. The domination number has, of course, been well studied [8,9].

The average distance of a graph has sometimes been used to provide lower bounds for domination-related invariants, including the domination number itself [4]. One of the first results along these lines is the following theorem due to F. Chung in [1], which originated as a conjecture of the computer program Graffiti [6]. The *independence number* of $G$, denoted by $\alpha = \alpha(G)$, is the cardinality of a largest set of mutually non-adjacent vertices.

1991 *Mathematics Subject Classification.* 05C35.

*Key words and phrases.* average distance, domination number, total domination number, Wiener index.

**Theorem 1** (Chung)**.** *Let $G$ be a graph. Then*

$$\alpha \geq \bar{D},$$

*with equality holding if an only if $G$ is complete.*

Recently, this theorem has been generalized by Hansen et al. as a result about the *forest number* $f = f(G)$ of a graph $G$ [7]. This is the maximum order of an induced forest of $G$.

**Theorem 2** (Hansen et al.)**.** *Let $G$ be a graph. Then*

$$f \geq 2\bar{D}.$$

This theorem was also motivated by a conjecture of Graffiti [10]. Its proof is based on techniques introduced by Dankelmann in [3]. Dankelmann uses similar techniques in [4] to characterize graphs with fixed order and domination number that have maximum average distance. One can derive the following theorem as a corollary of this characterization (although this is not stated in [4]).

**Theorem 3.** *Let $G$ be a graph. Then*

$$\gamma > \frac{2}{3}\bar{D} - \frac{1}{3}.$$

*Moreover, this inequality is best possible.*

The proof of Danklelmann's characterization result is lengthy and technical. We give a short direct proof of Theorem 3, as well as the following Theorem 4, which is the main result of our paper. We defer the proofs to a later section.

**Theorem 4** (Main Theorem)**.** *Let $G$ be a graph. Then*

$$\gamma_t > \bar{D} - \frac{1}{2}.$$

*Moreover, this inequality is best possible.*

## 2. Other Definitions

Let $R(k, t, l)$ denote the *binary star* on $k + t + l$ vertices, where the maximal interior path has order $t$ and there are $k$ leaves on one side of the binary star and $l$ leaves on the other. See Figure 1.

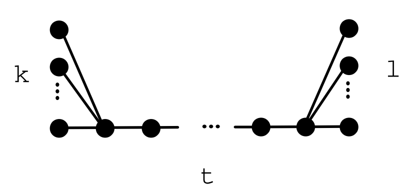


Figure 1. Binary star $R(k, t, l)$.

Now let $R(n, t)$ denote the binary star of order $n$ where the maximal interior path has order $t$ and the leaves are as balanced as possible on each side of the binary star.

A set $D$ of vertices of a graph $G$ is called a *connected dominating set* provided $D$ is a dominating set that induces a connected subgraph of $G$. The *connected* *domination number* of $G$, denoted $\gamma_c = \gamma_c(G)$, is the cardinality of a smallest connected dominating set in $G$. A *trunk* for a graph $G$ is a sub-tree (not necessarily induced) that contains the vertices of a dominating set of $G$. Hence, every spanning tree of $G$ is a trunk for $G$, and every connected dominating set is the vertex set of some trunk. Standard graph theoretical terms not defined in this paper can be found in [11], for instance.

## 3. Lemmas

The proof of Lemma 5 involves elementary algebra, counting, and limit arguments; we therefore omit it.

**Lemma 5.** *For integers $k \geq 0$ and $t \geq 1$,*

$$W(R(k,t,k)) = (t+3)k^2 + (t+2)(t-1)k + \frac{t(t+1)(t-1)}{6}, \text{ and}$$

$$W(R(k,t,k+1)) = (t+3)k^2 + (t+1)^2k + \frac{t(t+1)(t+2)}{6}.$$

*Moreover,*

$$W(R(k,t,k)) < W(R(k,t,k+1)) < W(R(k+1,t,k+1)), \text{ and}$$

$$\lim_{k\to\infty} \bar{D}(R(k,t,k)) = \frac{t+3}{2}.$$

The following lemma is proven in [6, Theorem 2].

**Lemma 6.** *Let $G$ be a graph with a trunk of order $t \geq 1$. Then*

$$\bar{D}(G) \leq \bar{D}(R(n,t)),$$

*with equality holding if and only if $G = R(n,t)$.*

The next lemma follows by combining the two previous lemmas.

**Lemma 7.** *Let $G$ be a graph with a trunk of order $t \geq 1$. Then*

$$\bar{D}(G) < \frac{t+3}{2}.$$

An immediate consequence of Lemmas 5 and 7 is the following corollary, which defines the relationship between the minimum order of a connected dominating set of a graph $G$, denoted $\gamma_c = \gamma_c(G)$, and its average distance.

**Corollary 8.** *Let $G$ be a graph. Then*

$$\gamma_c > 2\bar{D} - 3.$$

*Moreover, this inequality is best possible.*

*Proof.* Let $D$ be a minimum connected dominating set. Then any spanning tree of the subgraph induced by $D$ is a trunk for $G$. Hence, by Lemma 7,

$$\bar{D}(G) < \frac{\gamma_c + 3}{2}.$$

To show this inequality is best possible, consider $R(j,t,j)$, where $t \geq 1$ and $j \geq 0$. It is easy to see that $\gamma_c(R(j,t,j)) = t$. But by Lemma 5,

$$\lim_{j\to\infty} \bar{D}(R(j,t,j)) = \frac{t+3}{2} = \frac{\gamma_c + 3}{2}.$$

□

One final lemma is needed. The next simple lemma provides some relations that hold for the number of edges induced by dominating sets and their complements. Given a graph $G$ with dominating set $D$, a vertex $v \notin D$ is *over-dominated* by $D$ if it has two or more neighbors in $D$. The *over-domination number of $v$ with respect to $D$*, denoted by $O_D(v)$, is one less than the number of neighbors $v$ has in $D$.

**Lemma 9.** *Let $T$ be a tree with minimum dominating set $D$ such that the number of components of $D$ is $k$. Denote the number of edges with both endpoints in $D$ by $e_1$, the number of edges with both endpoints in $H = T - D$ by $e_2$, and the number of edges with one endpoint in $D$ and the other endpoint in $H$ by $e_3$. Moreover, let $j$ be the number of non-trivial components of $H$ with at least two neighbors in $D$ and let $l_H$ be the number of components of $H$ with exactly one neighbor in $D$ (the leaves of $H$). Then*

a) $e_1 = |D| - k$
b) $e_2 = k - 1 - \sum_{v \in H} O_D(v)$
c) $e_3 = n - |D| + \sum_{v \in H} O_D(v)$
d) $2j + l_H \leq e_3 = k + j + l_H - 1$
e) $n - l_H + 2 + \sum_{v \in H} O_D(v) \leq 2k + |D|$.

*Proof.* Part a) holds because $D$ induces a forest with $k$ trees. Part c) is true because every vertex in $H$ has a neighbor in $D$, giving the $n - |D|$, and because the summation contributes the extra edges that have one endpoint $D$ and one in $H$. Part b) follows immediately from parts a) and c), since $n - 1 = e_1 + e_2 + e_3$ for a tree.

The left hand side of d) comes from the fact that, when counting the edges between $D$ and $H$, each of the $l_H$ leaves in $H$ contributes exactly one edge while each of the $j$ non-trivial components of $H$ contributes at least two edges. The right hand side of d) follows easily by viewing the components of $D$ together with the components of $H$ as the vertices of a new tree with $e_3$ edges and $k + j + l_H$ vertices.

From d) we deduce that there are at most $k - 1$ non-trivial components of $H$, that is, $j \leq k - 1$. Combining this with the right hand side of d) and part c), we arrive at inequality e). □

## 4. Theorem Proofs

Our strategy for proving Theorem 4 is as follows. Given a minimum total dominating set $D$ of a graph $G$, we form a particular spanning tree $T$ of $G$ so that $D$ is also a minimum total dominating set of $T$. Then we apply the lemmas from the previous section to obtain the desired result.

**Theorem 4** (Main Theorem) *Let $G$ be a graph. Then*

$$\gamma_t > \bar{D} - \frac{1}{2}.$$

*Moreover, this inequality is best possible.*

*Proof.* Let $D$ be a minimum total dominating set of $G$. Suppose that $D$ has $k$ components. We form a spanning tree $T$ of $G$ such that $D$ is also a minimum total dominating set of $T$. If $G$ is a tree, then put $T = G$ and we are done. Otherwise, let $C$ be a cycle in $G$. We delete an edge from $C$ as follows.

i) If $C$ has two consecutive vertices $x$ and $y$ such that $x \notin D$ and $y \notin D$, then delete the edge between them. The set $D$ is still total dominating set for the resulting graph.
ii) Suppose the first case does not apply. If $C$ has two consecutive vertices $x$ and $y$ such that $x \in D$ and $y \notin D$, then delete the edge between them. Since the other neighbor of $y$ on $C$ is necessarily in $D$ (or else the first case applies), the set $D$ is still a total dominating set for the resulting graph.
iii) If neither of the first two cases apply, then all of the vertices of $C$ are in $D$. Delete any edge of $C$ and the set $D$ is still a total dominating set for the resulting graph.

Repeat this process until all cycles are removed. Call the resulting spanning tree $T$. Since $D$ is a total dominating set of $T$, $\gamma_t(T) \leq |D| = \gamma_t(G)$. Since the total domination number of a graph is at most the total domination number of any of its spanning trees, $\gamma_t(G) \leq \gamma_t(T)$. Thus, $\gamma_t(T) = |D|$ and $D$ is a minimum total dominating set of $T$.

Now, let $L_H$, of cardinality $l_H$, denote the leaves of $T$ that are in $H = T - D$ (the leaves of $T$ that are not in $D$). Observe that the sub-tree $T - L_H$ contains the total dominating set $D$ of $G$ and is thereby a trunk for $G$. From Lemma 7,

$$2\bar{D} - 3 < |T - L_H| = n - l_H.$$

Hence by Lemma 9 part e), and since $2k \leq \gamma_t$,

$$2\bar{D} - 3 < 2k + \gamma_t - 2 - \sum_{v \in H} O_D(v) \leq 2\gamma_t - 2 - \sum_{v \in H} O_D(v) \leq 2\gamma_t - 2.$$

Rearranging yields the desired inequality.

To show the inequality is best possible, consider $R(j,t,j)$, where $t \equiv 2 \pmod 4$ and $j \geq 0$. It is easy to see that $\gamma_t(R(j,t,j)) = \frac{t}{2} + 1$. But by Lemma 5,

$$\lim_{j \to \infty} \bar{D}(R(j,t,j)) = \frac{t}{2} + \frac{3}{2} = \gamma_t + \frac{1}{2}.$$

□

The proof of the theorem provides a necessary condition for $\gamma_t = \lceil \bar{D} - \frac{1}{2} \rceil$. In the proof we found a spanning tree $T$ of a connected graph $G$ such that a minimum total dominating set of $G$ was also a total dominating set for $T$. We let $H = T - D$ and found that

$$\gamma_t > \bar{D} - \frac{1}{2} + \frac{1}{2} \sum_{v \in H} O_D(v).$$

Now if $\gamma_t = \lceil \bar{D} - \frac{1}{2} \rceil$, then

$$\lceil \bar{D} - \frac{1}{2} \rceil = \gamma_t \geq \lceil \bar{D} - \frac{1}{2} + \frac{1}{2} \sum_{v \in H} O_D(v) \rceil,$$

which immediately suggests that $D$ may over-dominate at most one vertex of $H$, and if there is an over-dominated vertex of $H$, its over-domination number is 1.

To see that there exist graphs in which any spanning tree containing a minimum total dominating set of the graph (as a total dominating set for the spanning tree) over-dominates exacly one vertex (with over-domination number 1) of $H$ and $\gamma_t = \lceil \bar{D} - \frac{1}{2} \rceil$, consider $R(j,t,j)$, where $t > 1$, $t \equiv 1 \pmod 4$ and $j \geq t$. On the other hand, that this condition is not sufficient for equality is seen in $P_{4k+3}$ (the path on $4k + 3$ vertices) for $k \geq 1$. Any minimum total dominating set $D$ in $P_{4k+3}$ over-dominates exactly one vertex $v$ of $V - D$, and $v$ has over-domination number 1, but $\gamma_t$ is about one half the number of vertices and $\bar{D}$ is about one third of the number vertices.

Next we present a short and direct proof of Theorem 3. As mentioned previously, this result can be deduced from a result of Dankelmann in [4].

**Theorem 3** *Let $G$ be a graph. Then*

$$\gamma > \frac{2}{3}\bar{D} - \frac{1}{3}.$$

*Moreover, this inequality is best possible.*

*Proof.* Let $D$ be a minimum dominating set of $G$. Suppose that $D$ has $k$ components. We will form a spanning tree $T$ of $G$ such that $D$ is also a minimum dominating set of $T$. If $G$ is a tree, then put $T = G$ and we are done. Otherwise, let $C$ be a cycle in $G$. We delete an edge from $C$ as follows.

i) If $C$ has two consecutive vertices $x$ and $y$ such that $x \notin D$ and $y \notin D$, then delete the edge between them. The set $D$ still dominates the resulting graph.
ii) Suppose the first case does not apply. If $C$ has two consecutive vertices $x$ and $y$ such that $x \in D$ and $y \notin D$, then delete the edge between them. Since the other neighbor of $y$ on $C$ is necessarily in $D$ (or else the first case applies), the set $D$ still dominates the resulting graph.
iii) If neither of the first two cases apply, then all of the vertices of $C$ are in $D$. Delete any edge of $C$ and the set $D$ still dominates the resulting graph.

Repeat this process until all cycles are removed. Call the resulting spanning tree $T$. Since $D$ is a dominating set of $T$, $\gamma(T) \leq |D| = \gamma(G)$. Since the domination number of a graph is at most the domination number of any of its spanning trees, $\gamma(G) \leq \gamma(T)$. Thus, $\gamma(T) = |D|$ and $D$ is a minimum dominating set of $T$.

Now, let $L_H$, of cardinality $l_H$, denote the leaves of $T$ that are in $H = T - D$ (the leaves of $T$ that are not in $D$). Observe that the sub-tree $T - L_H$ contains the dominating set $D$ of $G$ and is thereby a trunk for $G$. From Lemma 7,

$$2\bar{D} - 3 < |T - L_H| = n - l_H.$$

Hence by Lemma 9 part e), and since $2k \leq 2\gamma$,

$$2\bar{D} - 3 < 2k + \gamma - 2 - \sum_{v \in H} O_D(v) \leq 3\gamma - 2 - \sum_{v \in H} O_D(v) \leq 3\gamma - 2.$$

Rearranging yields the desired inequality.

To show the inequality is best possible, consider the family of stars $S_n$. Since the average distance in stars can be made arbitrarily close to 2, $\frac{2}{3}\bar{D}(S_n) - \frac{1}{3}$ can be made arbitrarily close to $\gamma(S_n) = 1$. □

As was the case for total domination number and average distance, one can deduce from the proof a similar necessary condition for equality in $\gamma = \lceil \frac{2}{3}\bar{D} - \frac{1}{3} \rceil$.

## References

[1] F. Chung, *The average distance is not more than the independence number*, J. Graph Theory, 12 (1988), p. 229-235.

[2] E. Cockayne, R. Dawes and S. Hedetniemi, *Total domination in graphs*, Networks, 10 (1980), p. 211-219.

[3] P. Dankelmann, *Average distance and the independence number*, Discrete Applied Mathematics, 51 (1994), p. 73-83.

[4] P. Dankelmann, *Average distance and the domination number*, Discrete Applied Mathematics, 80 (1997), p. 21-35.

[5] A. Dobrynin, R. Entringer and I. Gutman, *Wiener index of trees: Theory and applications*, Acta Applicandae Mathematicae, 66 (2001), p. 211-249.

[6] S. Fajtlowicz and W. Waller, *On two conjectures of Graffiti*, Congressus Numerantium, 55 (1986), p. 51-56.

[7] P. Hansen, A. Hertz, R. Kilani, O. Marcotte and D. Schindl, *Average distance and maximum induced forest*, pre-print, 2007.

[8] T. Haynes, S.T. Hedetniemi and P.J. Slater, "Fundamentals of Domination in Graphs," Marcel Decker, Inc., NY, 1998.

[9] T. Haynes, S.T. Hedetniemi and P.J. Slater, "Domination in Graphs: Advanced Topics," Marcel Decker, Inc., NY, 1998.

[10] D.B. West, *Open problems column #23*, SIAM Activity Group Newsletter in Discrete Mathematics, 1996.

[11] D.B. West, "Introduction to Graph Theory (2nd ed.)," Prentice-Hall, NJ, 2001.