Communications in Mathematical and in Computer Chemistry

Average Distance, Connected Hub Number and Connected Domination Number^{*}

Xiao–Lu Gao, Shou–Jun Xu[†]

School of Mathematics and Statistics, Lanzhou University, Lanzhou, Gansu 730000, China

gaoxl14@lzu.edu.cn, shjxu@lzu.edu.cn

(Received April 17, 2018)

Abstract

Let G be a connected graph of given order n and let $\mu(G)$ denote the average of all the distances between any two distinct vertices in G. The connected hub number $h_c(G)$ (resp., the connected domination number $\gamma_c(G)$) of G is the smallest order of a connected subgraph S of G such that each pair of nonadjacent vertices outside S are joined by a path with all internal vertices in S (resp., each vertex outside S is adjacent to one vertex of S). It is easy to see that $h_c(G) \leq \gamma_c(G) \leq h_c(G) + 1$. In view of the close relationship between the two invariants, we can partition connected graphs into two classes and according to this partition, give sharp upper bounds on $\mu(G)$ of the two classes of G in terms of $h_c(G)$, respectively, and further characterize the extremal graphs. As a corollary, we give sharp upper bounds on $\mu(G)$ in terms of $\gamma_c(G)$, and characterize the extremal graphs. Since these graphs are trees, we further address the problem about 2-connected graphs and give some initial properties and results.

1 Introduction

Let G be a simple connected and finite graph with vertex set V(G) and edge set E(G). The *Wiener index* of G, which is introduced originally for approximating the boiling points of alkanes in 1947 by Wiener [25], is the sum of all distances in G:

$$W(G) := \frac{1}{2} \sum_{u, v \in V(G)} d_G(u, v) = \frac{1}{2} \sum_{u \in V(G)} \sigma_G(u),$$

^{*}This work is supported by NSFC (Grants Nos. 11571155, 11201205).

[†]Corresponding author.

where $d_G(x, y)$ denotes the distance between x and y in G and $\sigma_G(u) = \sum_{v \in V(G)} d_G(u, v)$. It is one of the most-known and well-studied topological indices that has numerous applications in analyzing problems in communication networks, geometry and physical chemistry. For details, readers can refer to the recent survey papers [11, 12, 16, 17, 26, 27].

Like the Wiener index, the *average distance* (or *mean distance*) of G is defined to be the average of all distances in G:

$$\mu(G) := \frac{1}{n(n-1)} \sum_{x,y \in V(G)} d_G(x,y).$$

The computer program GRAFFITI [13] made the attractive conjecture: $\mu(G) \leq \alpha(G)$, where $\alpha(G)$ denotes the independence number of G. Chung [3] succeeded in proving the conjecture. In [4] sharp upper and lower bounds for $\mu(G)$, depending as well on the independence number as on the order, were given. Besides these, many efforts have been made by several authors to give several upper or lower bounds on the average distance in terms of other graph parameters, for example, diameter and radius [20], the matching number [4], independence-related invariants [4, 6, 14], (edge-)connectivity [7–9], the girth [1] and the chromatic number [22, 23].

Introduced by Walsh [24], a hub set in a connected graph G with vertex set V is a subset U of V such that any two nonadjacent vertices outside U are connected by a path with all internal vertices in U. If a hub set U induces a connected subgraph, then U is called a connected hub set. The connected hub number $h_c(G)$ (resp. hub number h(G)) is the minimum size of a connected hub set (resp. a hub set) in G. Likewise, a dominating set in G [18] is a subset S of V such that every vertex outside S is adjacent to some vertex in S. If a dominating set S induces a connected subgraph, then S is called a connected dominating set. The connected domination number $\gamma_c(G)$ (resp. domination number $\gamma(G)$) is the minimum size of a connected dominating set (resp. a dominating set) in G. There are close relationships between the parameters $h_c(G)$ and $\gamma_c(G)$. In [19], Johnson, Slater and Walsh gave an inequality about them as follows:

Theorem 1.1. [19] If G is connected, then $h_c(G) \leq \gamma_c(G) \leq h_c(G) + 1$.

Upper bounds on the average distance of a connected graph of given order in terms of domination number and distance domination number are established by Dankelmann [5] and Tian et al. [21], respectively. Here, we introduce Dankelmann's result as follows: **Theorem 1.2.** [5] Let G be a connected graph of order n and domination number γ . Then

$$\mu(G) \leq \begin{cases} \frac{n+1}{3} - \frac{(n-3\gamma)(n-3\gamma+2)(2n+3\gamma-7)}{6n(n-1)}, & \text{if } \gamma \leq \frac{n}{3} \text{ and } n-\gamma \equiv 0 \pmod{2};\\ \frac{n+1}{3} - \frac{(n-3\gamma)(n-3\gamma+2)(2n+3\gamma-7)-9(\gamma-1)}{6n(n-1)}, & \text{if } \gamma \leq \frac{n}{3} \text{ and } n-\gamma \equiv 1 \pmod{2};\\ \frac{n+1}{3} - \frac{(3\gamma-n)(3\gamma-n-2)(5n-6\gamma-4)}{3n(n-1)}, & \text{if } \gamma \geq \frac{n}{3} \text{ and } n-\gamma \equiv 0 \pmod{2};\\ \frac{n+1}{3} - \frac{(3\gamma-n-1)((3\gamma-n-3)(5n-6\gamma-2)+6(2n-3\gamma-1))}{3n(n-1)}, & \text{if } \gamma \geq \frac{n}{3} \text{ and } n-\gamma \equiv 1 \pmod{2}; \end{cases}$$

with equality holding if and only if $G = G_{n,\gamma}$. (The graph $G_{n,\gamma}$ is showed in Figure 1.)



Figure 1. The graph $G_{n,\gamma}$ in Theorem 1.2: (i) when $\gamma \leq \frac{n}{3}$; (ii) when $\gamma \geq \frac{n}{3}$.

In this paper, we are going to establish sharp upper bounds on the average distance $\mu(G)$ of a connected graph G of given order n in terms of the connected domination number $\gamma_c(G)$ and the connected hub number $h_c(G)$, respectively. The rest sections are organized as follows. Section 2 gives preliminary definitions, notations and known results, including the results characterizing the connected graph G satisfying $\gamma_c(G) = h_c(G) + 1$. Section 3 gives sharp upper bounds on the average distance of the two classes of G in terms of $h_c(G)$, respectively, where the classes are partitioned by the close relationship between $h_c(G)$ and $\gamma_c(G)$, and further characterize the extremal graphs. Section 4 shows that the induced subgraph of any minimum connected dominating set U of an edge-minimal (with regard to the two properties "2-connected" and " γ fixed") graph G must be a tree and $G[V \setminus U]$ is an edge-nonempty forest, as well as obtain the sharp upper bound of $\mu(G)$ of 2-connected graphs G with $\gamma_c(G) = 2$ and further characterize the extremal graphs. We end the paper with the concluding section.

2 Preliminaries

In this paper, we consider simple connected undirected graphs G with given order n. For any subset X of V(G) or E(G), let G[X] denote the subgraph of G induced by X. For any $v \in V(G)$, denote by $d_G(v)$, $N_G(v)$ and $N_G[v]$ the degree, open neighborhood and closed neighborhood of v in G, respectively. Here, as elsewhere, we drop the index referring to the underlying graph if the reference is clear. Furthermore, for any $U \subseteq V(G)$, the open and closed neighbourhood of U in G are defined as $N_G(U) = \bigcup_{v \in U} N_G(v) \setminus U$ and $N_G[U] = N_G(U) \cup U$, respectively. For any two graphs G and H, we set their union $G \cup H = (V(G) \cup V(H), E(G) \cup E(H)).$

A vertex v is called a *leaf* of G if $d_G(v) = 1$. We call an edge uv of G as a *chord* of a cycle C if $u, v \in V(C)$ but $uv \notin E(C)$. For any $X, Y \subseteq V(G)$, a path $P = x_1x_2...x_k$ is said to be an (X, Y)-path of G, if $V(P) \cap X = x_1$ and $V(P) \cap Y = x_k$. In particular, if $X = \{x\}$ or $Y = \{y\}$, then we write $(\{x\}, Y)$ -path and $(\{x\}, \{y\})$ -path as (x, Y)-path and xy-path for short, respectively. G is called *edge-minimal* with a given graph property, if G itself has the property but G - e does not, for every edge $e \in E(G)$.

As shown in Theorem 1.1, Johnson et al. [19] gave the upper and lower bounds of $\gamma_c(G)$ by the functions of $h_c(G)$. Meanwhile, they further characterized the graphs attaining the upper bound as follows. For subsets $R, S \subseteq V$, we say that S dominates R if $N_G[S] \supseteq R$.



Figure 2. A schematic illustration of Lemma 2.1.

Lemma 2.1. [19] Let G be a connected graph with vertex set V, U any minimum connected hub set. We denote $T = V - N_G[U]$, $S = N_G(U)$. (Hence $|U| = h_c(G)$ and $V = U \cup S \cup T$. See Figure 2.) Then $\gamma_c(G) = h_c(G) + 1$ if and only if one of the following holds:

0. |U| = 0 and G is a clique.

1. |U| = 1, $\emptyset \neq T$ induces a clique in G, T and S are the parts of a complete bipartite subgraph of G, and no vertex in S has degree |V| - 1.

2. $|U| \ge 2$ and U induces a path with end-vertices, say, u and v, in $G; \emptyset \ne T$ induces a clique in G, and T and S are the parts of a complete bipartite subgraph of G. No vertex

of U other than u, v of the path G[U] has neighbours in S, and the sets $N_G(u) \cap S = A$ and $N_G(v) \cap S = B$ are nonempty and disjoint (see Figure 2).

(a) If |U| = 2 then no vertex of A dominates B, no vertex of B dominates A, and for any edge $ab \in E$, $a \in A$, $b \in B$, $\{a, b\}$ does not dominate $A \cup B$.

(b) If |U| = 3 then for any edge $ab \in E$, $a \in A$, $b \in B$, $\{a, b\}$ dominates neither A nor B.

(c) If $|U| \ge 4$ then there are no edges between A and B.

Given a graph G, we call a subgraph $P \subseteq G$ an *interior path* of G if P is a path and $d_G(v) \ge 2$ for each $v \in P$. Let R(k, t, l) denote the *binary star* of order k + t + l, 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. Especially, denote by R(n, t) the binary star of order $n \ge t+2$, where the maximal interior path has order t and the leaves are as balanced as possible on each side of the binary star (see Figure 3). A *trunk* in a connected graph G is a sub-tree (not necessarily induced) that contains the vertices of a dominating set of G. Obviously, the vertex set of every trunk is a connected dominating set. Conversely, every connected dominating set is exactly the vertex set of some trunk.



Figure 3. Binary star R(n, t)

Lemma 2.2. [10] Let G be a connected graph with a trunk of order $t \ge 1$ (or, equivalently, $\gamma_c(G) \le t$). Then

$$\mu(G) \leqslant \mu(R(n,t)),$$

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

From Lemma 2.2, we readily obtain (note that $\gamma_c(R(n,t)) = t$)

Lemma 2.3. Let G be a connected graph with $\gamma_c(G) \ge 1$. Then

$$\mu(G) \leqslant \mu(R(n, \gamma_c(G))),$$

with equality holding if and only if $G = R(n, \gamma_c(G))$.

3 Upper bounds of $\mu(G)$ in terms of h_c

By Theorem 1.1, we can divide connected graphs G into two classes, one for G with $\gamma_c = h_c + 1$ and the other for G with $\gamma_c = h_c$. In view of these classes, we give sharp upper bounds on $\mu(G)$ of the connected graphs G of given order in terms of h_c , respectively.

Theorem 3.1. Let G be a connected graph of given order n with connected hub number $h_c \geq 1$ and connected domination number γ_c . Then we have

If
$$\gamma_c = h_c + 1$$
, then

$$\mu(G) \leqslant \begin{cases} \frac{-h_c^3 + (2n-1)h_c^2 + (-4n+13)h_c + 8n^2 - 22n + 21}{4n(n-1)}, & \text{if } h_c \text{ is odd;} \\ \frac{-h_c^3 + (2n-1)h_c^2 + (-4n+14)h_c + 8n^2 - 24n + 24}{4n(n-1)}, & \text{if } h_c \text{ is even,} \end{cases}$$
(1)

with equality holding if and only if G is shown as Figure 2 with $|U| \ge 1$, |T| = 1, $|S| \ge 2$ and no edges in G[S].

If
$$\gamma_c = h_c$$
, then

$$\mu(G) \leqslant \begin{cases} \frac{-h_{e}^{3}+3h_{c}^{2}+(3n^{2}-12n+7)h_{c}+9n^{2}-12n+3}{6n(n-1)}, & \text{if } n-h_{c} \text{ is odd;} \\ \frac{-h_{e}^{3}+3h_{c}^{2}+(3n^{2}-12n+10)h_{c}+9n^{2}-12n}{6n(n-1)}, & \text{if } n-h_{c} \text{ is even,} \end{cases}$$
(2)

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

Proof. For convenience, we may assume that G is a graph such that $\mu(G)$ attains the maximum value under the conditions of this theorem. Note that $W(G) = \frac{1}{2}n(n-1)\mu(G)$. We just need to compute the maximum value of the Wiener index W(G) for G.

Case 1. $\gamma_c = h_c + 1$.

Let $a \in S$. Since $\gamma_c = h_c + 1$ and $|U| = h_c \ge 1$, it follows from Results 1 and 2 of Lemma 2.1 that $U \cup \{a\}$ is a minimum connected dominating set of G. If there is an edge $e \in E(G)$ connecting two vertices of S, then G - e is a connected graph with U being a minimum hub set and $U \cup \{a\}$ being a minimum connected dominating set of it. So $\gamma_c(G - e) = h_c(G - e) + 1$. But W(G - e) > W(G), contradicting to our assumption of G. Thus, we assume that G[S] consists of isolated vertices.

Denote by W_U the sum of distances over pairs of vertices in U, W_{US} the sum of distances between a vertex in U and a vertex in S. Similarly, we can define W_S , W_T , W_{ST} , W_{UT} . Then W(G) can be written as

$$W(G) = W_U + W_S + W_T + W_{US} + W_{ST} + W_{UT}.$$
(3)

Note that, from Lemma 2.1, it is deduced that $|S| \ge 2$.

Let |T| = t. It is easily obtained that

$$W_T = \frac{t(t-1)}{2},$$
 (4)

$$W_S = (n - h_c - t)(n - h_c - t - 1),$$
(5)

$$W_{ST} = t(n - h_c - t). \tag{6}$$

It follows from Lemma 2.1 and the assumption for S that in the case $h_c \ge 2$, all vertices in U, a vertex a in S adjacent to u, a vertex $b(\ne a)$ in S adjacent to v and a vertex win T form an induced cycle C of length $h_c + 3$; in the case $h_c = 1$, C consists of the only vertex in U, two vertices in S and a vertex w in T. And the distance between two vertices of C in G is exactly the distance between them in C. Thus the sum W_{aU} (resp. W_{wU}) of distances between a (resp. w) and all vertices in U is

$$W_{aU} = \begin{cases} 2(1+2+\ldots+\frac{h_c+1}{2}) + \frac{h_c+3}{2} - (1+2), & \text{if } h_c \text{ is odd;} \\ 2(1+2+\ldots+\frac{h_c+2}{2}) - (1+2), & \text{if } h_c \text{ is even.} \end{cases}$$
(7)

$$W_{wU} = \begin{cases} 2(1+2+\ldots+\frac{h_c+1}{2}) + \frac{h_c+3}{2} - 2 \times 1, & \text{if } h_c \text{ is odd;} \\ 2(1+2+\ldots+\frac{h_c+2}{2}) - 2 \times 1, & \text{if } h_c \text{ is even.} \end{cases}$$
(8)

Thus,

$$W_{US} = (n - h_c - t)W_{aU} = \begin{cases} (n - h_c - t)((\frac{h_c + 3}{2})^2 - 3), & \text{if } h_c \text{ is odd};\\ (n - h_c - t)((\frac{h_c + 2}{2})^2 + \frac{h_c}{2} - 2), & \text{if } h_c \text{ is even.} \end{cases}$$
(9)

$$W_{UT} = tW_{wU} = \begin{cases} \frac{t(h_c^2 + 6h_c + 1)}{4}, & \text{if } h_c \text{ is odd;} \\ \frac{t(h_c^2 + 6h_c)}{4}, & \text{if } h_c \text{ is even.} \end{cases}$$
(10)

By simple computation, we get

$$W(C_{h_c+3}) = \begin{cases} \frac{1}{8}(h_c+3)^3, & \text{if } h_c \text{ is odd;} \\ \frac{1}{8}(h_c+3)^3 - \frac{h_c+3}{8}, & \text{if } h_c \text{ is even.} \end{cases}$$
(11)

Since

$$W_U = W(C_{h_c+3}) - 2W_{aU} - W_{wU} - 4, (12)$$

substituting Eqs. (11), (7) and (8) for Eq. (12), we get

$$W_U = \begin{cases} \frac{h_c^3 + 3h_c^2 - 9h_c + 5}{8}, & \text{if } h_c \text{ is odd;} \\ \frac{h_c^3 + 3h_c^2 - 10h_c + 8}{8}, & \text{if } h_c \text{ is even.} \end{cases}$$
(13)

Substituting Eqs. (4)-(6), (9), (10) and (13) for (3), and then tidying up, we get

$$W(G) = \begin{cases} \frac{-h_c^3 + (2n-1)h_c^2 + (-4n+8t+5)h_c + 8n^2 - 2(4t+7)n + 4t^2 + 12t+5}{8}, & \text{if } h_c \text{ is odd}; \\ \frac{-h_c^3 + (2n-1)h_c^2 + (-4n+8t+6)h_c + 8n^2 - 8(t+2)n + 4t^2 + 12t+8}{8}, & \text{if } h_c \text{ is even.} \end{cases}$$

As W(G) is a quadratic function in t and a strictly decreasing function for $1 \leq t \leq n - h_c - 2$, W(G) attains the maximum value at t = 1. Therefor

$$W(G) \leqslant \begin{cases} \frac{-h_c^3 + (2n-1)h_c^2 + (-4n+13)h_c + 8n^2 - 22n+21}{8}, & \text{if } h_c \text{ is odd;} \\ \frac{-h_c^3 + (2n-1)h_c^2 + (-4n+14)h_c + 8n^2 - 24n+24}{8}, & \text{if } h_c \text{ is even,} \end{cases}$$
(14)

with equality holding if and only if G is shown as Figure 2 with $|U| \ge 1$, |T| = 1, $|S| \ge 2$ and no edges in G[S].

Case 2. $\gamma_c = h_c$.

Since $h_c(R(n, \gamma_c)) = \gamma_c(R(n, \gamma_c))$ and by Lemma 2.3, W(G) for G of given order n with $\gamma_c = h_c$ attains the maximum value whenever $G = R(n, \gamma_c)$. By simple computation, we have

$$W(G) \leqslant W(R(n,\gamma_c)) = \begin{cases} \frac{-h_c^3 + 3h_c^2 + (3n^2 - 12n + 7)h_c + 9n^2 - 12n + 3}{12}, & \text{if } n - h_c \text{ is odd;} \\ \frac{-h_c^3 + 3h_c^2 + (3n^2 - 12n + 10)h_c + 9n^2 - 12n}{12}, & \text{if } n - h_c \text{ is even,} \end{cases}$$
(15)

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

Again combining $\mu(G) = \frac{2W(G)}{n(n-1)}$ with (14) and (15), we get our proof.

Note that the right-hand side of Ineq. (1) is smaller than that of Ineq. (2). From Theorem 3.1, we can get the upper bound of the average distance of connected graphs Gof given order in terms of $\gamma_c(G)$ or $h_c(G)$ as follows.

Corollary 3.2. Let $m(\ge 1)$ be the connected domination number $\gamma_c(G)$ or the connected hub number $h_c(G)$ of a connected graph G of given order n. Then

$$\mu(G) \leqslant \begin{cases} \frac{-m^3 + 3m^2 + (3n^2 - 12n + 7)m + 9n^2 - 12n + 3}{6n(n-1)}, & \text{if } n - m \text{ is odd;} \\ \frac{-m^3 + 3m^2 + (3n^2 - 12n + 10)m + 3n(3n-4)}{6n(n-1)}, & \text{if } n - m \text{ is even,} \end{cases}$$
(16)

with equality holding if and only if G = R(n, m).

Since the right-hand sides of Ineq. (16) are both strictly increasing functions of n $(n \ge 1)$ and the limits of them as n approaches infinity both equal $\frac{m+3}{2}$, it follows that $\frac{m+3}{2}$ is the supremum of $\mu(G)$. From this, we get the following Corollary 3.3, which is a result obtained by DeLaViña, Pepper and Waller in [10].

Corollary 3.3. Let $m(\geq 1)$ be the connected domination number $\gamma_c(G)$ or the connected hub number $h_c(G)$ of a non-clique connected graph G of given order n. Then

$$m > 2\mu(G) - 3.$$

Moreover, this inequality is best possible.

4 Upper bounds of $\mu(G)$ of 2-connected graphs G

In the previous section, we get the upper bounds of $\mu(G)$ when G is connected. In this section, we further restrict the graph G to be 2-connected, in which way the extremal graphs whose average distances achieve sharp upper bounds will not be trees.

4.1 The analysis for the construction of edge-minimal graphs G

It's easily seen that for a graph G of given order n and connected domination number γ_c , if the average distance $\mu(G)$ is maximum, then G is edge-minimal. Firstly, we show the structure of G where G is edge-minimal with given order and connected domination number.

Lemma 4.1. [2] Let G be a k-connected graph, $x \in V(G)$ and $Y \subseteq V(G) - \{x\}$ with $|Y| \ge k$. Then there exist k internally disjoint (x, Y)-paths in G whose terminal vertices in Y are distinct.

Lemma 4.2. Let G be an edge-minimally 2-connected graph. Then every cycle of G is indeed an induced cycle of G.

Proof. To the contrary suppose that there exists a cycle C of G and $e_0 = uv \in E(G)$ is a chord of C. Since G is an edge-minimally 2-connected, it follows that G - uv is not 2-connected. Thus there exists a cut-vertex x of G - uv. Then u and v are respectively in two distinct components of (G - uv) - x, for otherwise x is also a cut-vertex of G, which contradicts to the 2-connectivity of G. So each uv-path in G - uv must pass x. However, C is also a cycle of G - uv, resulting that there exist two internally disjoint uv-paths in G - uv, a contradiction.

Theorem 4.3. Let G be a 2-connected graph with vertex set V of given order, and U be a minimum connected dominating set of G. If G is 2-connected with $\gamma_c(G) = \gamma_c$ being fixed, and G is edge-minimal with regard to these two properties, then G[U] is a tree and G - U is an edge-nonempty forest. Proof. Firstly, we prove that G[U] is a tree. To the contrary suppose that G[U] contains a cycle C. C must be an induced cycle of G because of Lemma 4.2 as well as that deleting any chord of C does not increase the connected domination number of G. For simplicity, in this proof, we call a non-trivial path P a C-path if $V(P) \cap V(C)$ contains exactly the two end-vertices of P.

Claim 1. For any $u \in V(C)$, $d_G(u) \ge 3$.

Otherwise, if there exists a vertex $v \in V(C)$ such that $d_G(v) = 2$, then $U' = U - \{v\}$ is still a connected dominating set of G, which implies that $\gamma_c(G) \leq |U'| = \gamma_c - 1$, a contradiction.

Claim 2. For any $u \in V(C)$, there exists a C-path P of G connecting u and another vertex v of C such that $v \notin N_C(u)$.

By Claim 1 and C being an induced cycle of G, there exists a vertex $w \in V(G) - V(C)$ such that $uw \in E(G)$. Since C is a cycle, we have $|V(C)| \ge 3$. Since G is 2-connected, there exist two internally disjoint (w, V(C))-paths P_1 and P_2 such that $V(P_1) \cap V(P_2) =$ $\{w\}$ by Lemma 4.1. Without loss of generality, suppose that $u \notin V(P_1)$, and let $V(C) \cap$ $V(P_1) = \{v\}$. Then we must have $v \notin N_C(u)$, for otherwise $P_1 \cup (C - uv) \cup \{uw\}$ is a cycle of G containing a chord uv, resulting that G - uv is still 2-connected by Lemma 4.2 with U its minimum connected dominating set, contradicting to the edge-minimality of G. Let P be the path such that $E(P) = \{uw\} \cup E(P_1)$, then P is a C-path of G satisfying the requirement of Claim 2.

Every C-path P divides the cycle C into two paths P_1 and P_2 (note that both contain the end-vertices of P). Now find a C-path P^* such that one of the two lengths of P_1^*, P_2^* is the shortest among all the paths of divisions of the C-paths. Without loss of generality, say it as P_1^* . That is,

$$|E(P_1^*)| = \min\{|E(P_1)|, |E(P_2)| : P \text{ is a } C\text{-path}\}.$$

Let $V(P^*) \cap V(C) = \{u, v\}$, and $N_C(u) \cap V(P_1^*) = \{u'\}$. By Claim 2, there exists a C-path P' and $v' \in V(C)$ such that $V(C) \cap V(P') = \{u', v'\}$. From the selection of P^* , we know that $v' \notin V(P_1^*)$.



Figure 4. A schematic illustration of Proof of Theorem 4.3.

For any path $P = x_1 x_2 \cdots x_t$, denote by $P(x_i, x_j)$ the sub-path of P from x_i to x_j , where $1 \leq i, j \leq t$. If $P^* \cap P' = \emptyset$, then $P' \cup P_2^*(v', u) \cup P^* \cup P_1^*(v, u')$ is a cycle and uu' is a chord of it (see Figure 4 (i)). However, G' = G - uu' is still 2-connected by Lemma 4.2, with U its minimum connected dominating set, which contradicts to the edge-minimality of G. If $P^* \cap P' \neq \emptyset$, then starting at u, find along P^* the first vertex w such that $w \in P^* \cap P'$. Then $(C - uu') \cup P^*(u, w) \cup P'(w, u')$ is a cycle and uu' is a chord of it. Similarly, G' = G - uu' contradicts to the edge-minimality of G (see Figure 4 (ii)). Thus G[U] contains no cycles, and so it is a tree.

Secondly, we prove that $G[V \setminus U]$ is a forest. To the contrary suppose that $G[V \setminus U]$ contains some cycle C, and pick any edge $e = uv \in C$. Since U is a dominating set of G, there exist a vertex $u' \in U$ and a vertex $v' \in U$ such that $uu', vv' \in E(G)$. Since G[U] is a tree, there exists a unique path P in G[U] connecting u' and v' (note that P might be a single vertex in the case when u' = v'). Then e is a chord of the cycle $(C - e) \cup P \cup \{uu', vv'\}$, and G' = G - e is still a 2-connected graph by Lemma 4.2 with U its connected dominating set. Thus $\gamma_c \geq \gamma_c(G') \geq \gamma_c(G) = \gamma_c$, that is, $\gamma_c(G') = \gamma_c$, contradicting to the edge-minimality of G. Thus $G[V \setminus U]$ is a forest.

Finally, we prove that $G[V \setminus U]$ contains at least one edge. To the contrary suppose that $G[V \setminus U]$ consists of isolated vertices. Pick any leaf $l \in G[U]$, l must dominate some vertex in $V \setminus U$, and any vertex in $V \setminus U$ dominated by l in G is also dominated by some other vertex in U because of the 2-connectivity of G. Thus U - l is still a connected dominating set of G, contradicting to the minimality of U. Thus $G[V \setminus U]$ is an edge-nonempty forest.

4.2 Upper bound of $\mu(G)$ of a 2-connected graph G with $\gamma_c(G) = 2$

Note that it is more difficult for the case that G is 2-connected where the extremal graphs whose average distances achieve sharp upper bounds will not be trees. So we restrict that $\gamma_c(G) = 2$, and then give the sharp upper bound of $\mu(G)$ and characterize the extremal graphs. Before we offer this result, we need to define a graph class $\Gamma_{n,2}$.

Definition 4.4. Define that $G \in \Gamma_{n,2}$ $(n \ge 4)$ if and only if G has the following structure:

- 1. If n is even, then G is the union of $\frac{n-2}{2}C_4$'s, where the C_4 's join exactly one common edge, as shown in Figure 5 (i);
- 2. If n is odd, then G is the union of $\frac{n-3}{2} C_4$'s and one C_3 , where the C_4 's join exactly one common edge, and the C_3 joins one edge with some C_4 satisfying that the edge is adjacent to that common edge, as shown in Figure 5 (ii).



Figure 5. The graph $G \in \Gamma_{n,2}$: (i) when n is even; (ii) when n is odd.

Lemma 4.5. Let G be a 2-connected graph of given order with $\gamma_c(G) = 2$ and $U = \{x, y\}$ be a minimum connected dominating set of G. If $\mu(G)$ attains the maximum value, then $G[V \setminus U]$ is a forest containing at most one isolated vertex.

Proof. Since $\mu(G)$ attains the maximum value, G must be edge-minimal under the conditions of this lemma. By Theorem 4.3, $G[V \setminus U]$ is a forest and can't be an isolated vertex set.

Suppose to the contrary that $G[V \setminus U]$ contains two isolated vertices u and v. Since G is 2-connected, each of u and v is adjacent to both x and y in G. That is, $ux, uy, vx, vy \in E(G)$. Construct a new graph G' satisfying V(G') = V(G) and $E(G') = (E(G) - \{uy, vx\}) \cup \{uv\}$. Then G' is 2-connected and $\gamma_c(G') = \gamma_c(G) = 2$.

Now we compare W(G) and W(G'). Note that $d_G(r,s) = d_{G'}(r,s)$ for any $r,s \in V(G - u - v)$. So it remains to compare $\sigma_G(u) + \sigma_G(v)$ and $\sigma_{G'}(u) + \sigma_{G'}(v)$. For any $w \in V(G - \{x, y, u, v\})$, it holds that

$$d_{G'}(u,w) \ge d_G(u,w);$$

$$d_{G'}(v,w) \ge d_G(v,w);$$

$$7 = \sum_{\substack{m \in \{u,v\}\\n \in \{x,y\}}} d_{G'}(m,n) + d_{G'}(u,v) > \sum_{\substack{m \in \{u,v\}\\n \in \{x,y\}}} d_G(m,n) + d_G(u,v) = 6$$

From the above three inequalities, we can see that W(G') > W(G), contradicting to the maximality of $\mu(G)$.

In fact, the induced subgraph $G[V \setminus U]$ in Lemma 4.5 contains no isolated vertex except for one simple case. In the following theorem, we will show it in two cases according to the parity of n.

Theorem 4.6. Let G be a 2-connected graph with |V(G)| = n and $\gamma_c(G) = 2$. Then

$$\mu(G) \leq \begin{cases} \frac{5n^2 - 16n + 16}{2n(n-1)}, & \text{if } n \text{ is even}; \\ \frac{5n^2 - 16n + 11}{2n(n-1)}, & \text{if } n \text{ is odd}, \end{cases}$$

with equality holding if and only if $G \in \Gamma_{n,2}$.

Proof. For convenience, we may assume that G is such a graph that $\mu(G)$ attains the maximum value under the conditions of this theorem. Let $U = \{x, y\}$ be a minimum connected dominating set of G. By Lemma 4.5, $G[V \setminus U]$ is a forest and contains at most one isolated vertex.

Claim 1. For each $u \in V \setminus U$ with u being not an isolated vertex of $G[V \setminus U]$, u is adjacent to exactly one vertex of $\{x, y\}$ in G.

Otherwise, if there exists a component B of $G[V \setminus U]$ with $|V(B)| \ge 2$ which contains a vertex u such that $ux, uy \in E(G)$. Choose another vertex $v \in V(B)$ such that $uv \in E(G)$. Since $U = \{x, y\}$ is a dominating set of G, v is adjacent to at least one of x and y in G. Without loss of generality, assume that $vx \in E(G)$, then ux is a chord of the cycle xyuv. Then G' = G - ux is still a 2-connected graph by Lemma 4.2 with U its connected dominating set, contradicting to the maximality of $\mu(G)$.

Case 1. n is even.

Firstly, we show that $G[V \setminus U]$ contains no isolated vertex. It's easy to see that G is exactly $\Gamma_{4,2}$ in the case when n = 4. So we suppose that $n \ge 6$. To the contrary suppose that $G[V \setminus U]$ contains an isolated vertex u.

Since n-2 is even, there exists a component C of $G[V \setminus U]$ with odd order at least 3. Since C is a tree, C contains at least 2 leaves, say l and r. Combining these with Claim 1 as well as the condition that G is 2-connected, we must have that one of l and r, say l, satisfies that

$$N_G(x) \cap \{V(C-l)\} \neq \emptyset$$
 and $N_G(y) \cap \{V(C-l)\} \neq \emptyset$.

Suppose without loss of generality that $l \in N_G(x)$, and let $v \in N_C(l)$. Construct a new graph G' from G'' = G - l by removing the edge ux, and adding the vertex l and the edges ul and xl. Observing that G'' is 2-connected by considering the pair x and l (or vand l) as well as using the same method in the proof of Lemma 4.2. Since G' is obtained by increasing a G''-path to G'', G' is also 2-connected (here, the notation "G''-path" is used hereditarily in the proof of Theorem 4.3). It holds also that $\gamma_c(G') = \gamma_c(G) = 2$.

Now we compare W(G) and W(G'). It is not hard to check that $d_G(r,s) = d_{G'}(r,s)$ for any $r, s \in V(G - l - u)$, so we just need to compare $\sigma_G(u) + \sigma_G(l)$ and $\sigma_{G'}(u) + \sigma_{G'}(l)$. For any $w \in V(G) \setminus \{x, y, u, l\}$, it holds that

$$\begin{aligned} d_{G'}(u,w) &\geq d_G(u,w), \\ d_{G'}(l,w) &\geq d_G(l,w), \\ d_{G'}(l,v) &\geq 2 > 1 = d_G(l,v), \\ \sum_{n \in \{x,y\}} d_{G'}(m,n) + d_{G'}(u,l) &= \sum_{\substack{m \in \{u,l\}\\n \in \{x,y\}}} d_G(m,n) + d_G(u,l) = 7. \end{aligned}$$

From the above formulas, we deduce that W(G') > W(G), contradicting to the maximality of $\mu(G)$. Thus $G[V \setminus U]$ contains no isolated vertex.

Next, we continue the calculation for $\mu(G)$. By Claim 1, we have $N_G(x) \cap N_G(y) = \emptyset$. Denote

$$W_G(U) = \sum_{\{u,v\} \subseteq U} d_G(u,v),$$
$$W_G(V \setminus U) = \sum_{\{u,v\} \subseteq V \setminus U} d_G(u,v),$$
$$W_G(U,V \setminus U) = \sum_{u \in U, v \in V \setminus U} d_G(u,v).$$

Then

$$W(G) = W_G(U) + W_G(U, V \setminus U) + W_G(V \setminus U),$$
(17)

where

$$W_G(U) = 1, \tag{18}$$

$$W_G(U, V \setminus U) = d(x) - 1 + 2(d(y) - 1) + d(y) - 1 + 2(d(x) - 1) = 3(n - 2).$$
(19)

Since both $W_G(U)$ and $W_G(U, V \setminus U)$ are constants, we just need to compute $W_G(V \setminus U)$. For $G[V \setminus U]$, denote by c the number of its components, c' the number of components of order at least 3. Consider the distances between vertices of $V \setminus U$ in G.

Since $G[V \setminus U]$ is a forest, the number of pairs whose distance is 1 equals n - 2 - c. In every component in $G[V \setminus U]$, pick the pairs of vertices such that one vertex is in $N(x) \setminus \{y\}$ and the other in $N(y) \setminus \{x\}$. It's easy to see that among these pairs, there exists at least one pair whose distance is 1 in each component of $G[V \setminus U]$, and at least one pair whose distance is 2 in each component of $G[V \setminus U]$ with order at least 3. Besides, the distance between any two vertices in $V \setminus U$ is at most 3, and between any two vertices in $N(x) \setminus \{y\}$ or in $N(y) \setminus \{x\}$ is at most 2, thus

$$W_G(V \setminus U) \le n - 2 - c + 2\left(\binom{n-2}{2} - (n-2-c) - ((d(x)-1)(d(y)-1) - c - c')\right) + 3((d(x)-1)(d(y)-1) - c - c') = 2\binom{n-2}{2} - (n-2) + (d(x)-1)(d(y)-1) - c' \le 2\binom{n-2}{2} - (n-2) + (d(x)-1)(d(y)-1),$$

with equality holding if and only if c' = 0.

Denote
$$n_0 = d(x) - 1$$
. Then $d(y) - 1 = n - d(x) - 1 = n - 2 - n_0$ and

$$n_0(n-2-n_0) = -(n_0 - \frac{n-2}{2})^2 + (\frac{n-2}{2})^2 \le \frac{(n-2)^2}{4},$$

with equality holding if and only if $n_0 = \frac{n-2}{2}$. Thus

$$W_G(V \setminus U) \le 2\binom{n-2}{2} - (n-2) + \frac{(n-2)^2}{4}.$$
 (20)

Combining (17)-(20), we obtain that

$$W(G) = W_G(U) + W_G(U, V \setminus U) + W_G(V \setminus U)$$

$$\leq 1 + 3(n-2) + 2\binom{n-2}{2} - (n-2) + \frac{(n-2)^2}{4}$$

$$= 2\binom{n-2}{2} + 2(n-2) + \frac{(n-2)^2}{4} + 1,$$
(21)

with equality holding if and only if $d(x) - 1 = d(y) - 1 = \frac{n-2}{2}$ and c' = 0. That is, $G \in \Gamma_{n,2}$.

Case 2. n is odd. When n = 5, it's easy to see that the extremal graph achieving the maximum average distance is exactly $\Gamma_{5,2}$, in which case the minimum connected dominating set U can be chosen as the two end-vertices of the common edge joined by the triangle and the C_4 , resulting that $G[V \setminus U]$ contains one isolated vertex. Consider the case when $n \geq 7$ below.

First we show that $G[V \setminus U]$ contains no isolated vertex. By Lemma 4.5, $G[V \setminus U]$ contains at most one isolated vertex. So we can suppose to the contrary that there exists a vertex $w \in V \setminus U$, such that w is an isolated vertex of $G[V \setminus U]$.

Note that the distance between w and each vertex in G is constant:

$$d(w, x) = d(w, y) = 1,$$

$$d(w, u) = 2, \quad \forall u \in V \setminus \{x, y\},$$

and deleting w doesn't change the distances of pairs of other vertices of G. Thus G - wis a graph such that W(G - w) is maximum over all the graphs of order n - 1 satisfying that G - w is 2-connected and $\gamma_c(G - w) = 2$. Therefor $G - w \in \Gamma_{n-1,2}$ by case 1. So G has the construction depicted in Figure 6 (i). However, if we construct a graph G'from G by removing the edge wy as well as adding a new edge incident with w which is shown in Figure 6 (ii)(remark that G' is just the graph depicted in Figure 5 (ii)), then $\sigma_G(w) < \sigma_{G'}(w)$, which implies that W(G) < W(G'), contradicting to the choice of G.



Figure 6. (i) The graph G; (ii) The graph G'.

Now we obtain that $G[V \setminus U]$ contains no isolated vertex. Since n-2 is odd, $G[V \setminus U]$ contains at least one component of order at least 3. Here we still use the notations c, c'

in case 1, where $c' \ge 1$ is restricted. Similarly, we obtain that

$$W_G(V \setminus U) \le n - 2 - c + 2\left(\binom{n-2}{2} - (n-2-c) - ((d(x)-1)(d(y)-1) - c - c')\right) + 3((d(x)-1)(d(y)-1) - c - c') = 2\binom{n-2}{2} - (n-2) + (d(x)-1)(d(y)-1) - c' \le 2\binom{n-2}{2} - n + 1 + (d(x)-1)(d(y)-1),$$

with equality holding if and only if c' = 1, and because of the first inequality, the order of the component of order at least 3 is exactly 3.

Denote $n_0 = d(x) - 1$. Then $d(y) - 1 = n - d(x) - 1 = n - 2 - n_0$ and

$$n_0(n-2-n_0) = -(n_0 - \frac{n-2}{2})^2 + (\frac{n-2}{2})^2 \le \frac{(n-2)^2 - 1}{4}$$

with equality holding if and only if $n_0 = \frac{n-1}{2}$ or $\frac{n-3}{2}$. Thus

$$W_G(V \setminus U) \le 2\binom{n-2}{2} - n + 1 + \frac{(n-2)^2 - 1}{4}.$$
 (22)

Combining (17)-(19) and (22), we obtain that

$$W(G) = W_G(U) + W_G(U, V \setminus U) + W_G(V \setminus U)$$

$$\leq 1 + 3(n-2) + 2\binom{n-2}{2} - n + 1 + \frac{(n-2)^2 - 1}{4}$$

$$= 2\binom{n-2}{2} + 2(n-2) + \frac{(n-2)^2 - 1}{4},$$
(23)

with equality holding if and only if $d(x) - 1 = \frac{n-1}{2}$ or $\frac{n-3}{2}$, and $G[V \setminus U]$ contains exactly one component of order 3 and all the others of order 2. That is, $G \in \Gamma_{n,2}$.

In summary, combining (21), (23) and $\mu(G) = \frac{2W(G)}{n(n-1)}$, we obtain that

$$\mu(G) \le \begin{cases} \frac{5n^2 - 16n + 16}{2n(n-1)}, & \text{if } n \text{ is even};\\ \frac{5n^2 - 16n + 11}{2n(n-1)}, & \text{if } n \text{ is odd}, \end{cases}$$

with equality holding if and only if $G \in \Gamma_{n,2}$, as required.

5 Conclusions

In this paper we focus on the connected hub number h_c , recently introduced by Walsh [24], and the connected domination number γ_c . In view of the close relationship between these two variables, we divide connected graphs into two classes, and, respectively, establish sharp upper bounds on the average distances of each class of graphs of given order in terms of h_c and also characterize the extremal graphs, one class containing cycles and the other class being trees. But from the global point of view, the extremal graphs are trees, like the previous results [5, 21]. Thus we restrict graphs to be 2-connected and obtain that for a 2-connected graph G with given order n and connected domination number γ_c , if G is edge-minimal and U is a minimum connected dominating set of G, then G[U] is a tree and $G[V \setminus U]$ is an edge-nonempty forest. Further, we restrict that $\gamma_c = 2$, and give sharp upper bounds of $\mu(G)$ and characterize the extremal graphs. One open problem is that what the results are if γ_c is not restricted for 2-connected graphs. Of course, another future work can be considered: establishing sharp upper bounds of the average distance of a connected graph of given order in terms of the hub number.

References

- S. Bekkai, M. Kouider, On mean distance and girth, Discr. Appl. Math. 158 (2010) 1888–1893.
- [2] J. A. Bondy, U. S. R. Murty, Graph Theory, Springer, Berlin, 2008.
- [3] F. R. K. Chung, The average distance and the independence number, J. Graph Theory 12 (1988) 229–235.
- [4] P. Dankelmann, Average distance and independence number, *Discr. Appl. Math.* 51 (1994) 75–83.
- [5] P. Dankelmann, Average distance and domination number, *Discr. Appl. Math.* 80 (1997) 21–35.
- [6] P. Dankelmann, Average distance and generalised packing in graphs, *Discr. Math.* 310 (2010) 2334–2344.
- [7] P. Dankelmann, S. Mukwembi, H. C. Swart, Average distance and vertex-connectivity, J. Graph Theory 62 (2009) 157–177.
- [8] P. Dankelmann, S. Mukwembi, H. C. Swart, Average distance and edge-connectivity I, SIAM J. Discr. Math. 22 (2008) 92–101.
- P. Dankelmann, S. Mukwembi, H. C. Swart, Average distance and edge-connectivity II, SIAM J. Discr. Math. 21 (2008) 1035–1052.
- [10] E. DeLaViña, R. Pepper, B. Waller, A note on dominating sets and average distance, *Discr. Math.* **309** (2009) 2615–2619.

- [11] A. A. Dobrynin, R. Entringer, I. Gutman, Wiener index of trees: theory and applications, Acta Appl. Math. 66 (2001) 211–249.
- [12] A. A. Dobrynin, I. Gutman, S. Klavžar, P. Žigert, Wiener index of hexagonal systems, Acta Appl. Math. 72 (2002) 247–294.
- [13] S. Fajtlowicz, W. A. Waller, On two conjectures of GRAFFITI, Congr. Numer. 55 (1986) 51–56.
- [14] P. Firby, J. Haviland, Independence and average distance in graphs, *Discr. Appl. Math.* **75** (1997) 27–37.
- [15] T. Grauman, S. G. Hartke, A. Jobson, B. Kinnersley, D. B. West, L. Wiglesworth, P. Worah, H. Wu, The hub number of a graph, *Inf. Proc. Lett.* **108** (2008) 226–228.
- [16] I. Gutman, S. Klavžar, B. Mohar (Eds.), Fifty years of the Wiener index, MATCH Commun. Math. Comput. Chem. 35 (1997) 1–259.
- [17] I. Gutman, S. Klavžar, B. Mohar (Eds.), Fiftieth anniversary of the Wiener index, Discr. Appl. Math. 80 (1997) 1–113.
- [18] T. W. Haynes, S. T. Hedetniemi, P. J. Slater (Eds.), Fundamentals of Domination in Graphs, Marcel Dekker, New York, 1998.
- [19] P. Johnson, P. Slater, M. Walsh, The connected hub number and the connected domination number, *Networks* 58 (2011) 232–237.
- [20] J. Plesník, On the sum of all distances in a graph or digraph, J. Graph Theory 8 (1984) 1–24.
- [21] F. Tian, J. Xu, Average distances and distance domination numbers, Discr. Appl. Math. 157 (2009) 1113–1127.
- [22] I. Tomescu, On the sum of all distances in chromatic blocks, J. Graph Theory 18 (1994) 83–102.
- [23] I. Tomescu, R. A. Melter, On distances in chromatic graphs, Quart. J. Math. Oxford 40 (1989) 475–480.
- [24] M. Walsh, The hub number of a graph, Int. J. Math. Comput. Sci. 1 (2006) 117-124.
- [25] H. Wiener, Structural determination of paraffin boiling points, J. Amer. Chem. Soc. 69 (1947) 17–20.
- [26] S. Xu, R. Gysel, D. Gusfield, Minimum average distance clique trees, SIAM J. Discr. Math. 29 (2015) 1706–1734.
- [27] H. Zhang, S. Xu, Y. Yang, Wiener index of toroidal polyhexes, MATCH Commun. Math. Comput. Chem. 56 (2006) 153–168.