*Combinatorics, Probability and Computing* (1997) 6, 455–463. Printed in the United Kingdom
© 1997 Cambridge University Press

# Maximal Spacing Configurations in Graphs

PETER FIRBY[1] and JULIE HAVILAND[2]†

[1]Department of Mathematics, University of Exeter, North Park Road, Exeter EX4 4QE, UK
(e-mail: firby@maths.exeter.ac.uk)
[2]Department of Mathematical Statistics and Operational Research, University of Exeter, North Park Road, Exeter EX4 4QE, UK
(e-mail: julie@maths.exeter.ac.uk)

*Received 13 July 1995; revised 17 May 1996*

Let $G = (V, E)$ be a simple connected graph of order $|V| = n \geqslant 2$ and minimum degree $\delta$, and let $2 \leqslant s \leqslant n$. We define two parameters, the *s-average distance* $\mu_s(G)$ and the *s-average nearest neighbour distance* $\Lambda_s(G)$, with respect to each of which $V$ contains an extremal subset $X$ of order $s$ with vertices 'as spread out as possible' in $G$. We compute the exact values of both parameters when $G$ is the cycle $C^n$, and show how to obtain the corresponding optimal arrangements of $X$. Sharp upper and lower bounds are then established for $\Lambda_s(G)$, as functions of $s, n$ and $\delta$, and the extremal graphs described.

## 1. Introduction

Let $G = (V, E)$ be a simple connected graph of order $|V| = n \geqslant 2$ and minimum degree $\delta$. Denote the *open neighbourhood* of a vertex $v \in V$ by $\Gamma(v) = \{w : vw \in E\}$. The *distance* between two vertices $x, y \in V$, denoted by $d(x, y)$, is the minimum length of an $x$-$y$ path in $G$. The *average distance* of $G$, which we shall denote by $\mu_n(G)$, is the average value of the distance between all pairs of vertices in $G$, i.e. $\mu_n(G) = \binom{n}{2}^{-1} \sum_{x,y \in V} d(x, y)$. The average (or mean) distance is a graph-theoretical invariant which has been much studied in the literature (see, for example, Bienstock and Győri [1], Chung [4], Dankelmann [5], Firby and Haviland [7], Hendry [9] and Winkler [11]). It has also been a useful tool in the analysis of networks representing transportation systems, where it is a measure of the time, cost, etc. needed in the average case.

In this paper, we are concerned with the general problem of positioning $s \geqslant 2$ vertices in the graph $G$ so that they are as 'spread out as possible'. Such arrangements have been considered, for example, in Hubbard [10], when modelling the distribution of $s$ electrons over the $n$ sites of a chemical chain so as to minimise the interaction energy, and in

† Current address: 19 Florida Drive, Exeter EX4 5EX, UK.

https://doi.org/10.1017/S0963548397003088 Published online by Cambridge University Press

Cocho, Pérez-Pascual and Ruis [3] when modelling coat patterns in animals. To this end, we define two different criteria for selecting a subset of $V$ with this property.

The first parameter, $\mu_s(G)$, is a direct extension of the average distance of $G$ and is defined as follows. Given $X \subseteq V$ of order $|X| = s$ and $x \in X$, the *average distance of $x$ in $X$* is $d_X(x) = \frac{1}{s-1} \sum_{y \in X - \{x\}} d(x, y)$. The *$s$-average distance* of $G$ is defined to be

$$\mu_s(G) = \frac{1}{s} \max_{\substack{X \subseteq V \\ |X| = s}} \sum_{x \in X} d_X(x).$$

The $s$-average distance of $G$ thus represents the best possible average distance between pairs of vertices in $X$, for subsets $X$ of order $s$.

While $\mu_s(G)$ takes into account the distances between all possible pairs of the $s$ vertices, the second parameter, $\Lambda_s(G)$, defined below, is concerned with only nearest neighbour distances. Given $X \subseteq V$ of order $|X| = s$ and $x \in X$, the *nearest neighbour distance of $x$ in $X$* is $p_X(x) = \min_{y \in X - \{x\}} d(x, y)$ and the *$s$-average nearest neighbour distance* of $G$ is defined to be

$$\Lambda_s(G) = \frac{1}{s} \max_{\substack{X \subseteq V \\ |X| = s}} \sum_{x \in X} p_X(x).$$

The quantity $\Lambda_s(G)$ thus represents the best possible average nearest neighbour distance in $X$, for subsets $X$ of order $s$.

Observe that for any $X \subseteq V$ we have $p_X(x) \leqslant d_X(x)$ for all $x \in X$, so if $X_0$ is a subset of $V$ for which $\Lambda_s(G) = \frac{1}{s} \sum_{x \in X_0} p_{X_0}(x)$, then $\Lambda_s(G) \leqslant \frac{1}{s} \sum_{x \in X_0} d_{X_0}(x) \leqslant \mu_s(G)$. Thus for any integer $2 \leqslant s \leqslant n$ and any connected graph $G$, we have $\Lambda_s(G) \leqslant \mu_s(G)$. Note also that $\Lambda_2(G) = \mu_2(G) = \max_{x,y \in V} d(x, y) = \mathrm{diam}(G)$, but that the two parameters are not equal in general. For example, $\Lambda_n(G) = 1$ for all connected graphs $G$, whereas $\mu_n(G)$, the average distance of $G$, is equal to 1 only if $G$ is the complete graph $K_n$.

To gain some idea of the different extremal arrangements of $s$ vertices determined by these two parameters, and therefore to see how well they capture the idea of being 'as spread out as possible', we first consider the special case when $G$ is the cycle $C^n$, in section 2. Then in section 3, we examine the general case, establishing sharp upper and lower bounds for $\Lambda_s(G)$ in terms of $s$ and the order and minimum degree of $G$, and citing the extremal graphs.

In what follows, all notation used is standard, as given in Bollobás [2].

## 2. $\Lambda_s(C^n)$ and $\mu_s(C^n)$

In this section we compare the $s$-average nearest neighbour distance and the $s$-average distance for $C^n$, the cycle on $n \geqslant 3$ vertices. For vertices $x, y \in V(C^n)$ we denote by $\overrightarrow{d}(x, y)$ the length of the clockwise path from $x$ to $y$ in $C^n$.

**Theorem 2.1.** *Let $3 \leqslant n = as + b$ $(0 \leqslant b \leqslant s - 1; a, b \in \mathbb{Z})$.*

(i) *$\Lambda_s(C^n)$ is achieved by choosing $X = \{x_1, \ldots, x_s\}$ labelled clockwise around $C^n$ so that $n_i = \overrightarrow{d}\left(x_i, x_{i(\mathrm{mod}\, s)+1}\right)$ satisfies*

$$n_i = \begin{cases} \lceil \frac{n}{s} \rceil = a + 1 & \text{for } 1 \leqslant i \leqslant b; \\ \lfloor \frac{n}{s} \rfloor = a & \text{for } b + 1 \leqslant i \leqslant s. \end{cases}$$

https://doi.org/10.1017/S0963548397003088 Published online by Cambridge University Press

(ii)

$$\Lambda_s(C^n) = \begin{cases} \frac{n}{s} & \text{if } b = 0; \\ \frac{(n-1)}{s} & \text{if } b \neq 0. \end{cases}$$

**Proof.** Suppose that $X$ is extremal with respect to $\Lambda_s(C^n)$ and let $n_i = \overrightarrow{d}\left(x_i, x_{i(\mathrm{mod}\, s)+1}\right)$ for $1 \leqslant i \leqslant s$. Then $\sum_{i=1}^{s} n_i = n$, so

$$s\Lambda_s(C^n) = \sum_{i=1}^{s} p_X(x_i) = \sum_{i=1}^{s} \min\left\{n_i, n_{i(\mathrm{mod}\ s)+1}\right\} \leqslant n - \varepsilon,$$

where $\varepsilon = 0$ if $n_i = n_{i(\mathrm{mod}\ s)+1}$ for $1 \leqslant i \leqslant s$, and $\varepsilon = 1$ otherwise.

Observe that taking $n_1 = \cdots = n_b = a+1 = \lceil n/s \rceil$ (if $b \neq 0$) and $n_{b+1} = \cdots = n_s = a = \lfloor n/s \rfloor$ gives $\sum_{i=1}^{s} \min\left\{n_i, n_{i(\mathrm{mod}\ s)+1}\right\} = n$ if $b = 0$ and $n-1$ if $b \neq 0$. Therefore $s\Lambda_s(C^n) = n$ for $b = 0$ and $s\Lambda_s(C^n) = n-1$ for $b \neq 0$, which proves (ii), and so the given arrangement of $X$ must be optimal, which proves (i). □

Prior to determining the value of $\mu_s(C^n)$ and the corresponding arrangement of an extremal subset $X$, we require the following technical lemma.

**Lemma 2.2.** *Let* $X = \{x_1, \ldots, x_s\}$ *be labelled clockwise around* $C^n$, *and let* $n_i = \overrightarrow{d}\left(x_i, x_{i(\mathrm{mod}\, s)+1}\right)$ *and* $w_{k,i} = \sum_{\ell=1}^{k} n_{(i+\ell-2)(\mathrm{mod}\, s)+1}$ *for* $1 \leqslant i \leqslant s$, $1 \leqslant k \leqslant s-1$. *Then* $X$ *can be chosen so that for each fixed* $k$ *with* $1 \leqslant k \leqslant s-1$, *no two* $w_{k,i}$ *differ by more than one*, *for* $1 \leqslant i \leqslant s$.

**Proof.** We prove the lemma by induction on $k$. If $k = 1$ then $\sum_{i=1}^{s} w_{1,i} = \sum_{i=1}^{s} n_i$. Writing $n = m_1 s + b$ $(0 \leqslant b < s; m_1, b \in \mathbb{Z})$, we can maximise $\sum_{i=1}^{s} w_{1,i}$ by setting

$$\hat{n}_i = \begin{cases} m_1 + 1 & \text{if } 1 \leqslant i \leqslant b; \\ m_1 & \text{if } b < i \leqslant s. \end{cases}$$

Denote by $U$ (resp. $V$) a term $\hat{n}_i$ of value $m_1$ (resp. $m_1+1$). Initially, we take $U^{s-b}V^b$, considered as a cyclic arrangement of the $\hat{n}_i$'s. Our aim is to progressively rearrange the $U$'s and $V$'s so as eventually to achieve an arrangement with the desired properties. A *p-string* $P$, in the current cyclic arrangement, is a sequence of $p$ consecutive terms, and $\ell(P)$, the *length* of $P$, is the sum of the values of the terms in $P$. Thus 1-strings have length $m_1$ or $m_1+1$. Now let $1 \leqslant k < s-1$ and assume that in the current arrangement, for all $1 \leqslant p \leqslant k$, all $p$-strings have length $m_p$ or $m_p+1$. Furthermore, suppose that the $(k+1)$-strings $A$ and $B$ satisfy $\ell(B) - \ell(A) \geqslant 2$, and that all other pairs of $(k+1)$-strings in the (clockwise or anticlockwise) segment $A \ldots B$ have lengths differing by at most one.

**Claim 1.** $A \cap B = \emptyset$.

**Proof of Claim 1.** If $A \cap B \neq \emptyset$ then $A - (A \cap B)$ and $B - (A \cap B)$ are both $p$-strings for some $p \leqslant k$, so differ in length by at most one. Since $\ell(B) = \ell(B - (A \cap B)) + \ell(A \cap B)$ and $\ell(A) = \ell(A - (A \cap B)) + \ell(A \cap B)$, then $2 \leqslant \ell(B) - \ell(A) = \ell(B - (A \cap B)) - \ell(A - (A \cap B))$, a contradiction.

**Claim 2.** All $(k+1)$-strings in the segment $A \ldots B$ except $A$ and $B$ have length $m_k + m_1 + 1$.

https://doi.org/10.1017/S0963548397003088 Published online by Cambridge University Press

**Proof of Claim 2.** Write $A = A_1a_1 = a_2A_2$ and $B = B_1b_1 = b_2B_2$, where $A_1, B_1, A_2, B_2$ are $k$-strings of length $m_k$ or $m_k + 1$, whilst $a_1, b_1, a_2, b_2 \in \{U, V\}$. Now $\ell(A) = \ell(A_1) + \ell(a_1) = \ell(a_2) + \ell(A_2)$ and $\ell(A) + 2 \leqslant \ell(B) = \ell(B_1) + \ell(b_1) = \ell(b_2) + \ell(B_2)$. Therefore $\ell(a_1) = \ell(a_2) = m_1, \ell(b_1) = \ell(b_2) = m_1 + 1, \ell(A_1) = \ell(A_2) = m_k, \ell(B_1) = \ell(B_2) = m_k + 1$, so $\ell(A) = m_k + m_1$ and $\ell(B) = m_k + m_1 + 2$. Now if $C \neq A, B$ is a $(k+1)$-string in $A \ldots B$ then $\ell(C) > \ell(A)$, for otherwise $\ell(B) - \ell(C) \geqslant 2$, and also $\ell(C) < \ell(B)$, otherwise $\ell(C) - \ell(A) \geqslant 2$. Thus $\ell(C) = m_k + m_1 + 1$.

**Claim 3.** In the segment $A \ldots B$, for $p \leqslant k$ any $p$-string with $a_2$ (resp. $b_2$) as its first term has the same length, $m_p$ (resp. $m_p + 1$), as any $p$-string with $b_1$ (resp. $a_1$) as its last term. Further, the $(p-1)$-string immediately following $a_2$ (resp. $b_2$) has the same length as the $(p-1)$-string immediately preceding $b_1$ (resp. $a_1$).

**Proof of Claim 3.** To prove the first statement, for $p \leqslant k$ let $A'$ be a $p$-string with $a_2$ as its first term and let $B'$ be a $p$-string with $b_1$ as its last term. Then $\ell(A) = \ell(A') + \ell(A - A'), \ell(B) = \ell(B') + \ell(B - B')$ and $\ell(B) = \ell(A) + 2$, which implies $\ell(B') = \ell(A') + 1$ and $\ell(B - B') = \ell(A - A') + 1$, so any $p$-string with $a_2$ as its first term has length $m_p$ and any $p$-string with $b_1$ as its last term has length $m_p + 1$. The second statement follows because $\ell\left(B' - b_1\right) - \ell\left(A' - a_2\right) = \ell(B') - \ell(b_1) - \ell(A') + \ell(a_2) = (\ell(A') + 1) - (m_1 + 1) - \ell(A') + m_1 = 0$. Corresponding arguments apply for $a_1, b_2$.

**Claim 4.** Reversing the ordering of the segment $A_2 \ldots B_1$ affects the current arrangement by converting the $(k+1)$-strings $A$ and $B$ into new $(k+1)$-strings of length $m_k + m_1 + 1$, whilst leaving the length of all other $p$-strings unchanged for $1 \leqslant p \leqslant k + 1$.

**Proof of Claim 4.** After reversal all $(k+1)$-strings between $a_2$ and $b_1$ still have length $m_k + m_1 + 1$, whilst $A^*$ (resp. $B^*$), the new $(k+1)$-string created with first term $a_2$ (resp. last term $b_1$) has length $\ell(a_2) + \ell(B_1) = m_k + m_1 + 1$ (resp. $\ell(b_1) + \ell(A_2) = m_k + m_1 + 1$). For any original $(k+1)$-string $Z \neq A$ with $a_2 \in Z$, we have $\ell(Z) = \ell(Z - A) + \ell(Z \cap A_2) + \ell(a_2)$. If $Z \cap A_2 \neq \emptyset$ then it is a $p$-string immediately following $a_2$, and thus has the same length as the $p$-string immediately preceding $b_1$ which replaces it after reversal. Similarly any original $(k+1)$-string $Z' \neq B$ with $b_1 \in Z'$ will have the same length after the reversal. Clearly for $p \leqslant k$ the length of any $p$-string not containing $a_2$ or $b_1$ is unchanged by reversal. Moreover, for any original $p$-string $Z$ with $a_2 \in Z$, if $Z \cap A_2 = \emptyset$ then $\ell(Z)$ is unchanged by reversal, whilst if $Z \cap A_2 \neq \emptyset$ then for some $p' \geqslant 1$ the $p'$-string $Z \cap A_2$ is replaced by another $p'$-string of the same length, by Claim 3, so again $\ell(Z)$ remains unchanged. A similar conclusion holds for any original $p$-string containing $b_1$, which establishes Claim 4.

The truth of Claim 4 implies that repeating the reversal process moving 'outwards' from the segment $A \ldots B$ eventually results in a cyclic arrangement such that all pairs of $p$-strings, $p \leqslant k + 1$, differ in length by at most one. Since, by Claim 1, any pair of $(k+1)$-strings whose lengths differ by at least two must be disjoint, the process must terminate when $k + 1 \geqslant s/2$.

It follows from this inductive argument that we can construct a cyclic arrangement of the $\hat{n}_i$'s for which all $p$-strings have length $m_p$ or $m_p + 1$, $1 \leqslant p \leqslant s - 1$. Relabel these terms, in order, as $n_1, \ldots, n_s$. If we arrange $x_1, \ldots, x_s$ in $C^n$ so that $n_i = \overrightarrow{d}\left(x_i, x_{i(\mathrm{mod}s)+1}\right)$ then $w_{k,i}$

https://doi.org/10.1017/S0963548397003088 Published online by Cambridge University Press

denotes the length of the $k$-string starting at $n_i$, and for each fixed $k$ with $1 \leqslant k \leqslant s-1$, no two $w_{k,i}$ differ by more than one, for $1 \leqslant i \leqslant s$. □

**Theorem 2.3.**

(i) *$\mu_s(C^n)$ is achieved by choosing $X = \{x_1, \ldots, x_s\}$ labelled clockwise around $C^n$ so that $n_i = \overrightarrow{d}\left(x_i, x_{i(\text{mod}s)+1}\right)$ satisfies the conditions of Lemma 2.2.*

(ii)

$$\mu_s(C^n) = \begin{cases} \frac{n(s+1)}{4s} & \text{if } s \text{ is odd;} \\ \frac{ns}{4(s-1)} & \text{if } s \text{ and } n \text{ are even;} \\ \frac{ns-2}{4(s-1)} & \text{if } s \text{ is even and } n \text{ is odd.} \end{cases}$$

**Proof.** If $n = 3$ then $s = 2$ or 3, and in both cases just one arrangement of $X$ is possible (up to rotation) and $\mu_2(C^3) = \mu_3(C^3) = 1$, which agrees with the condition of (i) and the formulae of (ii).

Henceforth we assume $n \geqslant 4$, and let $X = \{x_1, \ldots, x_s\}$ be a subset of $V$ labelled clockwise around $C^n$. As in Lemma 2.2, write $n_i = \overrightarrow{d}\left(x_i, x_{i(\text{mod}s)+1}\right)$ and $w_{k,i} = \sum_{\ell=1}^{k} n_{(i+\ell-2)(\text{mod}s)+1}$ for $1 \leqslant i \leqslant s,\ 1 \leqslant k \leqslant s-1$. Defining the function $D(i) = \min\{i, n-i\}$ for $1 \leqslant i \leqslant n-1$, we have $D\left(\overrightarrow{d}(x,y)\right) = d(x,y)$ for any $x, y \in V(C^n)$, and $D(w_{k,i}) = d\left(x_i, x_{(i+k-1)(\text{mod}s)+1}\right)$ for $1 \leqslant i \leqslant s,\ 1 \leqslant k \leqslant s-1$.

If we denote by $\overline{D}(X)$ the number $\frac{1}{s(s-1)} \sum_{x,y \in X} d(x,y) = \frac{1}{s} \sum_{x \in X} d_X(x)$, then

$$\mu_s(C^n) = \max_{\substack{X \subseteq V \\ |X|=s}} \overline{D}(X); \tag{1}$$

$$s(s-1)\overline{D}(X) = \sum_{k=1}^{s-1} \sum_{i=1}^{s} D(w_{k,i}). \tag{2}$$

The right-hand side of (2) achieves a maximum if each of the sums $\sum_{i=1}^{s} D(w_{k,i})$ achieves a maximum for $1 \leqslant k \leqslant s-1$. Observe that $D(i)$ is the distance from $v_0$ to $v_i$ in $C^n = v_0 v_1 \ldots v_{n-1} v_0$ and so $D(i+1)+D(j-1) \geqslant D(i)+D(j)$ for $1 \leqslant i < j \leqslant n-1$. It follows that for each fixed $k$ in the range $1 \leqslant k \leqslant s-1$, the maximum value for $\sum_{i=1}^{s} D(w_{k,i})$ is achieved when no two of the $w_{k,i}$ differ by more than one. By Lemma 2.2, there exists an arrangement of $X$ satisfying these conditions, for which, by (1) and (2), $\overline{D}(X)$ achieves its maximum value, $\mu_s(C^n)$. Thus we have established (i).

To prove (ii), first suppose that $s = 2m+1$ is odd. If $1 \leqslant k \leqslant m$ then $w_{k,i} \leqslant w_{m,i}$, whilst if $m < k \leqslant s-1$ then $w_{k,i} > w_{m,i}$. As $|w_{k,j} - w_{k,i}| \leqslant 1$ for $i \neq j$, we must have $w_{k,j} \geqslant w_{m,i}$ for $1 \leqslant i \neq j \leqslant s$. Moreover for $1 \leqslant i \leqslant s$, either $D(w_{k,i}) = \sum_{\ell=1}^{k} n_{(i+\ell-2)(\text{mod}s)+1} = w_{k,i}$ or $D(w_{k,i}) = n - \sum_{\ell=1}^{k} n_{(i+\ell-2)(\text{mod}s)+1} = \sum_{\ell=1}^{s-k} n_{(i+k+\ell-2)(\text{mod}s)+1} = w_{s-k,i+k}$. Now if $1 \leqslant k \leqslant m$ then $m < s-k < s$, so from above $w_{k,i} \leqslant w_{m,i} \leqslant w_{s-k,i+k}$ and thus $D(w_{k,i}) = w_{k,i}$. Similarly if $m < k < s$ then $s-k \leqslant m$, so $w_{s-k,i+k} \leqslant w_{m,i+k} \leqslant w_{k,i}$ and $D(w_{k,i}) = w_{s-k,i+k}$. It follows that

$$\sum_{i=1}^{s} D(w_{k,i}) = \begin{cases} \sum_{i=1}^{s} w_{k,i} = kn & \text{if } 1 \leqslant k \leqslant m; \\ \sum_{i=1}^{s} w_{s-k,i+k} = (s-k)n & \text{if } m < k \leqslant 2m. \end{cases}$$

https://doi.org/10.1017/S0963548397003088 Published online by Cambridge University Press

Thus if $s$ is odd then

$$\overline{D}(X) = \frac{2\sum_{k=1}^{m} kn}{s(s-1)} = \frac{nm(m+1)}{s(s-1)} = \frac{n(s+1)}{4s}.$$

Now consider the case where $s = 2m$ is even. The reasoning above shows that

$$D(w_{k,i}) = \begin{cases} w_{k,i} & \text{if } 1 \leqslant k < m; \\ w_{s-k,i+k} & \text{if } m < k < s. \end{cases}$$

If $k = m$ then $D(w_{k,i}) = D(w_{k,i+k}) = \min\{w_{k,i}, w_{k,i+k}\}$, but $|w_{k,i} - w_{k,i+k}| \leqslant 1$ and so $\min\{w_{k,i}, w_{k,i+k}\} = \lfloor n/2 \rfloor$. It follows that

$$\sum_{i=1}^{s} D(w_{k,i}) = \begin{cases} \sum_{i=1}^{s} w_{k,i} = kn & \text{if } 1 \leqslant k < m; \\ s\lfloor \frac{n}{2} \rfloor & \text{if } k = m; \\ \sum_{i=1}^{s} w_{s-k,i+k} = (s-k)n & \text{if } m < k \leqslant 2m-1. \end{cases}$$

Thus if $s$ is even then

$$\overline{D}(X) = \frac{\left(2\sum_{k=1}^{m-1} kn + s\lfloor \frac{n}{2} \rfloor\right)}{s(s-1)} = \frac{\left(nm(m-1) + s\lfloor \frac{n}{2} \rfloor\right)}{s(s-1)}.$$

If $n$ is even this last expression takes the value $ns/4(s-1)$, and if $n$ is odd it takes the value $(ns-2)/4(s-1)$, which completes the proof of (ii). □

We now present a simple example which demonstrates how to obtain an optimal arrangement of $X$ in $C^n$, and hence the $s$-average distance, by applying the reversal procedure given in the proof of Lemma 2.2.

**Example.** Consider the case $n = 35$ and $s = 8$. As $35 = 4.8 + 3$, we have $k = 1$ and $m_1 = 4$, so starting with the cyclic arrangement 44444555 we apply the procedure of Lemma 2.2 to eliminate any pair of 2-strings differing in length by at least two:

$$444\overset{A}{(44)}\overset{B}{(55)}5 \to 44\overset{A}{(44)}54\overset{B}{(55)} \to 44454545.$$

All 2-strings now have length 8 or 9. Thus $k = 2, m_2 = 8$ and we apply the procedure of Lemma 2.2 to eliminate any pair of 3-strings differing in length by at least two:

$$\overset{A}{(444)}\overset{B}{(545)}45 \to 44544545.$$

All 3-strings now have length 13 or 14. Thus $k = 3, m_3 = 13$ and the process terminates since $4 = k + 1 \geqslant s/2 = 4$. Hence the required arrangement of the clockwise distances between the vertices of $X$ is 44544545, for which the corresponding 8-average distance is 139/14.

Finally, Theorems 2.1 and 2.3 have the following immediate consequences.

**Theorem 2.4.** *Let $G$ be a simple graph of order $n$ and minimum degree $\delta$. If any of the following conditions hold then $\Lambda_s(G) \leqslant \Lambda_s(C^n)$ and $\mu_s(G) \leqslant \mu_s(C^n)$:*
(i) *$G$ is Hamiltonian*;

https://doi.org/10.1017/S0963548397003088 Published online by Cambridge University Press

(ii) $\delta \geqslant n/2$;
(iii) *$G$ is $\ell$-connected and* $2 \leqslant s \leqslant \ell$.

**Proof.** The proof of (i) is trivial. By Dirac [6], if $\delta \geqslant n/2$ then $G$ is Hamiltonian, so (ii) holds. Lastly, it is well-known (e.g. Bollobás [2, p. 10]) that if $G$ is $\ell$-connected with $\ell \geqslant 2$ then for any set of $\ell$ vertices there is a cycle containing all of them; thus if $2 \leqslant s \leqslant \ell$ then any $s$ vertices of $G$ are contained in a cycle of order at most $n$, which implies (iii). □

## 3. Bounds for $\Lambda_s(G)$

In this section we establish general upper and lower bounds for $\Lambda_s(G)$ as functions of $n, s$ and $\delta$. We require the following simple lemma.

**Lemma 3.1.** *For all* $x \in X$,

$$p_X(x) \leqslant \frac{3n - 3s - 1}{\delta + 1} + 1.$$

**Proof.** Choose $x, y \in X$ with $d(x, y) = p_X(x)$ and let $P$ be a minimal $x$-$y$ path. Write $V(P) = \{x_1 = x, x_2, \ldots, x_{k-1}, x_k = y\}$. Then $x_1$ has at least $\delta - 1$ neighbours in $V - X - V(P)$, whilst each of $x_2, \ldots, x_{k-2}$ has at least $\delta - 2$, so there are at least $(k-2)(\delta-2)+1$ edges between $V(P)$ and $V - X - V(P)$. Additionally, by the minimality of $P$, no vertex of $V - X - V(P)$ can have more than three neighbours in $V(P)$, so

$$n - s - k + 2 = |V - X - V(P)| \geqslant \frac{(k-2)(\delta-2)+1}{3};$$

rearranging yields the result. □

**Theorem 3.2.** *If* $m = \left\lceil \frac{n-s+1}{\delta} \right\rceil$ *then*

$$\Lambda_s(G) \leqslant \frac{(3m\delta - \delta - 5)(m-1)}{s(\delta+1)} + 2.$$

**Proof.** Suppose that for some $\ell \geqslant m$ the vertices $v_1, \ldots, v_\ell$ have no neighbours in $X$ and pairwise disjoint neighbourhoods in $V - X$. Then $n - s = |V - X| \geqslant \sum_{i=1}^{\ell} |\Gamma(v_i)| \geqslant m\delta$, contradicting the definition of $m$. Hence at most $m-1$ vertices of $X$ satisfy $p_X(x) > 2$, so by Lemma 3.1 we have

$$\begin{aligned}
s\Lambda_s(G) &\leqslant (m-1)\left(\frac{3n-3s-1}{\delta+1} + 1\right) + 2(s-(m-1)) \\
&= (m-1)\left(\frac{3n-3s-\delta-2}{\delta+1}\right) + 2s \\
&\leqslant (m-1)\left(\frac{3(s+m\delta-1)-3s-\delta-2}{\delta+1}\right) + 2s \quad \text{since } n \leqslant s + m\delta - 1 \\
&= \frac{(3m\delta-\delta-5)(m-1)}{\delta+1} + 2s.
\end{aligned}$$

□

https://doi.org/10.1017/S0963548397003088 Published online by Cambridge University Press

Thus it is relatively simple to establish an upper bound for $\Lambda_s(G)$ of order $3m^2/s$. With considerably more work we can achieve a slight improvement on Theorem 3.2, which we state below, together with a lower bound.

**Theorem 3.3.** *Let* $m = \lceil \frac{n-s+1}{\delta} \rceil$.

*If* $m = 1$ *then* $\Lambda_s(G) = 1$;

*if* $m \geqslant 2$ *then* $1 + \frac{1}{s} \leqslant \Lambda_s(G) \leqslant \max\left\{1 + \frac{(3m-5)(m-1)}{s}, 2 + \frac{3(m-2)^2}{s}\right\}$.

**Proof.** For detailed proofs of the upper bounds the reader is referred to Firby and Haviland [8]. The lower bound is obvious because if $m \geqslant 2$, i.e. $s \leqslant n - \delta$, then we can find a set $X \subseteq V$ of order $s$ with $\Gamma(x) \cap X = \emptyset$ for some $x \in X$, implying $p_X(x) \geqslant 2$ and $\Lambda_s(G) \geqslant (s+1)/s$. □

All bounds stated for the case $m \geqslant 2$ are achieved by certain graphs. For the lower bound, consider the graph $G$ comprising a $K_{n-1}$ with a vertex $x$ of degree $\delta$ attached, where $1 \leqslant \delta \leqslant n-2$. If $X = \{x\} \cup Y$, where $Y \subseteq V - \Gamma(x), Y \neq \emptyset$, then $2 \leqslant s \leqslant n - \delta$ and $\Lambda_s(G) = (s+1)/s$.

The first upper bound is achieved by the graph $G$ with

$$\begin{aligned} V &= \{x_i : 0 \leqslant i \leqslant 4\} \cup \{y_i : 0 \leqslant i \leqslant 4\}, \quad X = \{x_1, y_1\}, \\ E &= \{x_3y_3\} \cup \{x_0x_i, x_ix_{i(\text{mod } 4)+1}, y_0y_i, y_iy_{i(\text{mod } 4)+1} : 1 \leqslant i \leqslant 4\}. \end{aligned}$$

We have $n = 10, \delta = 3, s = 2$ and $d(x_1, y_1) = \text{diam}(G) = 5$, so $m = \lceil (n-s+1)/\delta \rceil = 3$ and $\Lambda_s(G) = 1 + (3m-5)(m-1)/s = 5$.

The second upper bound is attained if $G$ is a complete graph of order $n$ with the edges of a $K_t$ removed, and $X$ comprises $s$ vertices of the $K_t$, where $\max\{2, n-2\delta+1\} \leqslant s \leqslant t = n-\delta$. Thus $m = \lceil (n-s+1)/\delta \rceil = 2$ and $\Lambda_s(G) = 2 + (m-2)^2/s = 2$.

However, it is possible to improve upon the upper bounds for some values of $\delta$ and $s$. The next theorem provides an upper bound for $\Lambda_s(G)$ which is independent of the minimum degree $\delta$, and is sharper than Theorems 3.2 and 3.3 when $\delta = o(n)$ and $s = O(n)$.

**Theorem 3.4.**

$$\Lambda_s(G) \leqslant \frac{2}{s}(n-1).$$

**Proof.** Choose $X$ with $\Lambda_s(G) = \frac{1}{s}\sum_{x \in X} p_X(x)$ and suppose $k = \min_{x \in X} p_X(x)$ and $K = \max_{x \in X} p_X(x)$. Writing $V_i = \{x \in X : p_X(x) = i\}$ we have

$$s\Lambda_s(G) = \sum_{i=k}^{K} i|V_i|. \tag{1}$$

For each $x \in X$ choose a shortest path $x_1x_2 \ldots x_{i+1}$, where $x = x_1$, joining it to $X - \{x\}$. Let $Q_x$ be the subpath $x_1x_2 \ldots x_{\lceil i/2 \rceil}$. Then the $Q_x$s, being the first half of the shortest

https://doi.org/10.1017/S0963548397003088 Published online by Cambridge University Press

paths from $x$ to $X - \{x\}$ for all $x \in X$, are clearly disjoint. Therefore

$$\sum_{x \in X} |V(Q_x)| = \sum_{i=k}^{K} |V_i| \left\lceil \frac{i}{2} \right\rceil . \tag{2}$$

Suppose that for some even $i$ there exists $x \in V_i$; then $d(x, x_{i/2+1}) = i/2$. Now if $x_{i/2+1} \in Q_y$ for some $y \in V_j$, $y \neq x$, then $d(y, x_{i/2+1}) \leqslant \lceil j/2 \rceil - 1$, so $d(x, y) \leqslant i/2 + \lceil j/2 \rceil - 1$. If $j \leqslant i$ this implies $d(x, y) < i$, contradicting the fact that $x \in V_i$, whilst if $i < j$ this implies $d(x, y) < j$, contradicting the fact that $y \in V_j$. Thus if $V_i \neq \emptyset$ for some even $i$ then $V - \cup_{x \in X} V(Q_x) \neq \emptyset$. Counting vertices of $G$ and using (1) and (2), we have

$$n \geqslant \sum_{x \in X} |V(Q_x)| + 1 \geqslant \sum_{i=k}^{K} |V_i| \frac{i}{2} + 1 = \frac{s\Lambda_s(G)}{2} + 1,$$

so $\Lambda_s(G) \leqslant 2(n-1)/s$. Otherwise if all $i$ with $V_i \neq \emptyset$ are odd, then counting vertices of $G$ and using (1) and (2), we have

$$n \geqslant \sum_{x \in X} |V(Q_x)| = \sum_{i=k}^{K} |V_i| \frac{(i+1)}{2} \geqslant \frac{s(\Lambda_s(G)+1)}{2}.$$

Hence $\Lambda_s(G) \leqslant 2n/s - 1 \leqslant 2(n-1)/s$ for $s \geqslant 2$, which proves the theorem. □

Note that Theorem 3.4 is best possible, since the upper bound is attained by the generalized star $S\left(2(n-1)/k, k/2\right)$, which comprises $2(n-1)/k$ paths of length $k/2$ all with one endvertex in common.

## References


[1] Bienstock, D. and Győri, E. (1988) Average distance in graphs with removed elements. *J. Graph Theory* **12** 375–390.

[2] Bollobás, B. (1978) *Extremal Graph Theory*. Academic Press.

[3] Cocho, G., Pérez-Pascual, R. and Ruis, J. L. (1987) Discrete systems, cell-cell interactions and color pattern of animals. I. Conflicting dynamics and pattern formation. *J. Theor. Biol.* **125** 419–435.

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

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

[6] Dirac, G. A. (1952) Some theorems on abstract graphs. *Proc. London Math. Soc.* **2** 69–81.

[7] Firby, P. and Haviland, J. (1997) Independence and average distance in graphs. *Discrete Appl. Math.* (To appear.)

[8] Firby, P. and Haviland, J. (1995) Minimal energy configurations in graphs. University of Exeter, Department of Mathematics Preprint M95/17.

[9] Hendry, G. R. T. (1986) Existence of graphs with prescribed mean distance. *J. Graph Theory* **10** 173–175.

[10] Hubbard, J. (1978) Generalized Wigner lattices in one dimension and some applications to tetracyanoquinodimethane (TCNQ) salts. *Phys. Rev. B* **17** 494–505.

[11] Winkler, P. M. (1986) Mean distance and the 4/3 conjecture. *Proc. 17th S.E. Conference on Combinatorics, Graph Theory and Computing: Congr. Numer.* **54** 63–72.


https://doi.org/10.1017/S0963548397003088 Published online by Cambridge University Press