Distance in Graphs

Wayne Goddard[1] and Ortrud R. Oellermann[2]

[1] Clemson University, Clemson SC USA, goddard@clemson.edu
[2] University of Winnipeg, Winnipeg MN Canada, o.oellermann@uwinnipeg.ca

Summary. The distance between two vertices is the basis of the definition of several graph parameters including diameter, radius, average distance and metric dimension. These invariants are examined, especially how they relate to one another and to other graph invariants and their behaviour in certain graph classes. We also discuss characterizations of graph classes described in terms of distance or shortest paths. Finally, generalizations are considered.

1 Overview of Chapter

The distance between two vertices in a graph is a simple but surprisingly useful notion. It has led to the definition of several graph parameters such as the diameter, the radius, the average distance and the metric dimension. In this chapter we examine these invariants; how they relate to one another and other graph invariants and their behaviour in certain graph classes. We also discuss characterizations of graph classes that have properties that are described in terms of distance or shortest paths. We later consider generalizations of shortest paths connecting pairs of vertices to shortest trees, called Steiner trees, that connect three or more vertices.

2 Distance, Diameter and Radius

A path in a graph is a sequence of distinct vertices, such that adjacent vertices in the sequence are adjacent in the graph. For an unweighted graph, the *length* of a path is the number of edges on the path. For an (edge) weighted graph, the length of a path is the sum of the weights of the edges on the path. We assume that all weights are nonnegative and that all graphs are connected. We start with undirected graphs.

The *distance* between two vertices u and v, denoted $d(u, v)$, is the length of a shortest $u - v$ path, also called a $u - v$ *geodesic*. The distance function is a metric on the vertex set of a (weighted) graph G. In particular, it satisfies the *triangle inequality*:

$$d(a,b) \leq d(a,c) + d(c,b)$$

for all vertices a, b, c of G. This follows from the fact that, if you want to get from a to b, then one possibility is to go via vertex c.

2.1 Diameter and Radius

Two of the most commonly observed parameters of a graph are its radius and diameter. The *diameter* of a connected graph G, denoted $\mathrm{diam}(G)$, is the maximum distance between two vertices. The *eccentricity* of a vertex is the maximum distance from it to any other vertex. The *radius*, denoted $\mathrm{rad}(G)$, is the minimum eccentricity among all vertices of G. Of course the diameter is the maximum eccentricity among all vertices.

For a (weighted) undirected graph G:

$$\mathrm{rad}(G) \leq \mathrm{diam}(G) \leq 2\,\mathrm{rad}(G).$$

The upper bound follows from the triangle inequality, where c is a vertex of minimum eccentricity.

The radius and diameter are easily computed for simple graphs:

Fact 1 *1. Complete graphs:* $\mathrm{diam}(K_n) = \mathrm{rad}(K_n) = 1$ *(for $n \geq 2$).*
2. *Complete bipartite graphs:* $\mathrm{diam}(K_{m,n}) = \mathrm{rad}(K_{m,n}) = 2$ *(if n or m is at least 2).*
3. *Path on n vertices:* $\mathrm{diam}(P_n) = n-1$; $\mathrm{rad}(P_n) = \lceil (n-1)/2 \rceil$.
4. *Cycle on n vertices:* $\mathrm{diam}(C_n) = \mathrm{rad}(C_n) = \lfloor n/2 \rfloor$.

Note that cycles and complete graphs are vertex-transitive, so the radius and diameter are automatically the same (every vertex has the same eccentricity).

The *centre* is the subgraph induced by the set of vertices of minimum eccentricity. Graphs G where $\mathrm{rad}(G) = \mathrm{diam}(G)$ are called *self-centred*.

A famous result, due originally to Jordan [44], is that:

Fact 2 *For trees T, the diameter equals either $2\,\mathrm{rad}(T)$ or $2\,\mathrm{rad}(T) - 1$. In the first case the center is a single vertex, and in the second the centre is a pair of adjacent vertices.*

The *derivative* T' of a tree T is the tree obtained by deleting the leaves of T. Suppose the k^{th} derivative $T^{(k)}$ of T has been defined for some $k \geq 1$. Then the $(k+1)^{st}$ derivative of T is defined by $T^{(k+1)} = (T^{(k)})'$. It can be shown that the centre of a tree is $T^{(\lfloor diam(T)/2 \rfloor)}$.

For general graphs, Harary and Norman [40] showed the following:

Fact 3 *The centre of a graph forms a connected subgraph, and is contained inside a block of the graph.*

In general, there are no structural restrictions on the centre of a graph. Indeed, Hedetniemi (see [9]) showed that every graph is the centre of some graph.

2.2 Bounds for the Radius and Diameter

We look next at upper bounds involving the minimum or maximum degree.

Theorem 1. *For a graph G of order n:*

1. $\text{diam}(G) \leq n - \Delta(G) + 1$.
2. *[71]* $\text{rad}(G) \leq (n - \Delta(G))/2 + 1$.
3. *If* $\delta(G) \geq n/2$, *then* $\text{diam}(G) \leq 2$.

and these bounds are sharp.

One can also consider the problem of maximizing or minimizing these parameters given the number of vertices and edges. Both the radius and diameter are minimized by the star (and any supergraph thereof), so that case is not interesting.

For a given number n of vertices and $m \geq n - 1$ of edges, the diameter and radius are maximized by the 'path–complete' and the 'cycle–complete' graphs, respectively. These are defined as follows. A *cycle–complete* graph is obtained by taking disjoint copies of a cycle of even length and a complete graph, and joining three consecutive vertices of the cycle to all vertices in the complete graph. The radius is half the length of the cycle. This graph was introduced by Vizing [71]. An example is given in Figure 1.

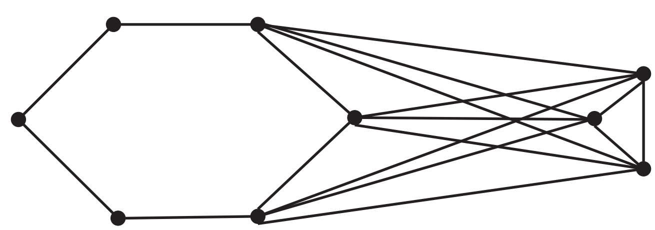

Fig. 1. A cycle–complete graph

A *path–complete* graph is obtained by taking disjoint copies of a path and complete graph, and joining an end vertex of the path to *one or more* vertices of the complete graph. It is not hard to show that there is a unique path–complete graph with n vertices and m edges; this is denoted as $PK_{n,m}$. This graph was introduced by Harary [38]. An example is given in Figure 2.

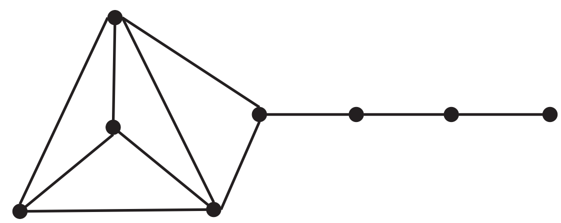

Fig. 2. The path–complete graph $PK_{8,11}$

Theorem 2. *1. [71] For a given number n of vertices and $m \geq n-1$ of edges, the radius is maximized by the cycle–complete graphs.*
2. [38] For a given number n of vertices and $m \geq n-1$ of edges, the diameter is maximized by the path–complete graph $PK_{n,m}$.

It is not hard to show that if a graph G has large diameter, then its complement $\bar{G}$ has small diameter:

Fact 4 *1. If* $\operatorname{diam}(G) > 3$*, then* $\operatorname{diam}(\bar{G}) \leq 2$.
2. If $\operatorname{diam}(G) = 3$*, then* $\operatorname{diam}(\bar{G}) \leq 3$.

2.3 Changes in Diameter and Radius with Edge and Vertex Removal

Removing an edge can never decrease the radius or the diameter of a graph. Indeed, removing a bridge disconnects the graph. So we consider cyclic edges.

Fact 5 *If e is a cyclic edge of graph G, then* $\operatorname{rad}(G) \leq \operatorname{rad}(G-e) \leq 2\operatorname{rad}(G)$ *and* $\operatorname{diam}(G) \leq \operatorname{diam}(G-e) \leq 2\operatorname{diam}(G)$.

Both upper bounds are attainable. Removing a cyclic edge can easily double the diameter; for example it does so in the odd cycles. Plesník [61] showed that if every edge's removal doubles the diameter, then the graph is a Moore graph, i.e., a graph of diameter d and girth $2d+1$ for some $d \geq 1$. For example, Moore graphs include the complete graphs, the odd cycles, the Petersen graph and the Hoffman–Singleton graph (see [7]). Removing a cyclic edge can also double the radius. Such a graph can be constructed by taking two equal-sized cycles, and sticking them together along one edge e.

There is a natural spanning tree with the same radius as the original graph, sometimes called a *breadth-first search* tree. This tree has diameter at most double its radius, and hence at most double the original radius. Buckley and Lewinter [8] determined which graphs have a diameter-preserving spanning tree.

Removing a vertex can both increase and decrease these parameters. Removing a cut-vertex from a disconnected graph results in a disconnected graph, so we do not consider such vertices..

Fact 6 *If v is a non-cut-vertex of graph G, then* $\mathrm{rad}(G-v) \geq \mathrm{rad}(G) - 1$ *and* $\mathrm{diam}(G-v) \geq \mathrm{diam}(G) - 1$.

The above bound is attained, for example, by an even-order path where v is an end-vertex of the path. There is no upper bound on $\mathrm{rad}(G-v)$ or $\mathrm{diam}(G-v)$ if G is 2-connected. To see this, let G be obtained from m cycles $C_i : v_{i_1} v_{i_2} \ldots v_{i_{2d+1}} v_{i_1}$ for $1 \leq i \leq m$ by identifying the vertices v_{i_1} for $1 \leq i \leq m$ into a vertex v and then adding the edges $v_{i_{2d+1}} v_{(i+1)_2}$ for $i = 1, 2, \ldots, m-1$. Then $\mathrm{diam}(G) = 2d$ and $\mathrm{diam}(G-v) = 2dm - 1$. Since m can be made as large as we wish, there is no upper bound on $\mathrm{diam}(G-v)$ in terms of $\mathrm{diam}(G)$. Figure 3 shows this construction with $d = 2$ and $m = 4$.

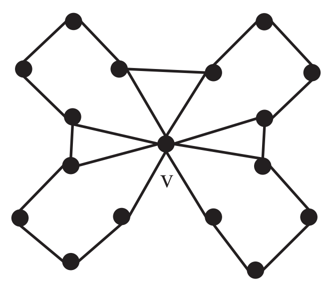

Fig. 3. Large increase in the diameter/radius of $G - v$

2.4 Matrices and Walks

A *walk* is a sequence of not necessarily distinct vertices such that each vertex in the sequence except the first one is adjacent to the previous one. Suppose the vertices of a graph G are labeled $v_1, \ldots, v_n$. Then the *adjacency matrix* A of G is defined as the $n \times n$ matrix whose (i,j)-th entry is 1 if there is an edge joining vertices v_i and v_j, and 0 otherwise. The *Laplacian matrix* is defined as $A - D$ where D is the diagonal matrix whose (i,i)-entry is the degree of the vertex v_i. The eigenvalues of A are referred to as the *eigenvalues* of G, and the eigenvalues of L as the *Laplacian eigenvalues* of G. The following fact is well known:

Fact 7 *The (i,j)-th entry of the power A^k gives the number of walks of length k from v_i to v_j.*

Bounds on the diameter of a graph in terms of the number of eigenvalues of these matrices follow from Fact 7 (see for example [17]):

Theorem 3. *Let G be a connected graph and b the number of distinct eigenvalues of G. Then*

$$\mathrm{diam}(G) \leq b - 1.$$

The same result holds if b is the number of distinct Laplacian eigenvalues of G.

3 Other Measures of Centrality

Apart from the centre of a graph, there are several other centrality measures. These have many applications. The centre is for emergency facility location: the response time must be minimized in the worst case. For biological graphs, the centre vertices might be the most important (see for example [42]).

Another measure of centrality is the 'median' of a graph. The *status* $\sigma(v)$ of a vertex v is the sum of the distances from v to all other vertices. The vertices having minimum status $\sigma(v)$ form the *median* of the graph, which is denoted by $M(G)$. For example, the median might be a good place to locate a mall: the average driving distance is minimized.

There is no intrinsic connection between the centre and the median. Indeed, they can be arbitrarily far apart. Slater [67] considered whether there are other measures of centrality that 'connect' the centre and the median of a graph. He defined for a graph G, integer $k \geq 2$ and vertex u of G,

$$r_k(u) = \max\Big\{ \sum_{s\in S} d(u,s) \mid S \subseteq V(G),\ |S| = k \Big\}.$$

Thus, $r_k(u)$ is the sum of the k largest vertex distances to u. In particular, $r_1(u)$ is its eccentricity, and $r_{n-1}(u)$ is its status.

The *k-centrum*, $C(G;k)$, of G is the subgraph induced by those vertices u for which $r_k(u)$ is a minimum. Thus, $C(G;2) = C(G)$ and $C(G;n-1) = M(G)$. It is shown in [67] that the k-centrum of every tree consists of a single vertex or a pair of adjacent vertices. Further, every vertex on the shortest path from the centre to the median of a tree belongs to the k-centrum for some k between 1 and n.

In the case of trees, another centrality measure is well-known. Suppose T is a tree and v a vertex of T. Then a *branch* at v is a maximal subtree containing v as a leaf. The *weight* at a vertex v of T is the maximum number of edges in any branch at v. A vertex of T is a *centroid vertex* of T if it has minimum weight among all vertices of T and the subgraph induced by the centroid vertices is called the *centroid* of T. That is, a centroid vertex is one whose removal leaves the smallest maximum component order. However, Zelinka [77] observed that the centroid and the median coincide.

Another centrality idea is a 'central path'. This concept can be defined in one of two ways that parallel the centre and median, respectively. If H is a subgraph of a graph G and v is any vertex of G, then the distance from v to H is the minimum distance from v to a vertex of H. The *eccentricity* of H is the distance of a vertex v furthest from H, and the *status* of H is the sum of the distances of every vertex of G to H. A path P is a *path centre* of G if

P has minimum eccentricity among all paths of G and has minimum length among all such paths. Similarly the *path median* of G is a path of minimum status in G. The path median of a network may indicate a good choice for a subway line, for example; in that case many individuals will use this line, so it is desirable that the average distance from the users' homes to the line is minimized. On the other hand the problem of finding the path centre of a network has application to the problem of building a highway between two major cities in such a way that the furthest distance from the highway to any town in a collection of other 'important' towns is minimized. The problem of finding the path centre of a tree was solved independently in [16] and [68]. Further types of centres are discussed in the book [7]. See [32] for more of the early history of centrality.

4 Special Graph Classes

4.1 Chordal Graphs

A *chordal graph* is one where every cycle of length greater than 3 has a chord. For example, trees and maximal outerplanar graphs are chordal. A *simplicial vertex* is one whose neighbourhood is complete. It was first shown by Dirac [21] and later by Lekkerkerker and Boland [49] that every chordal graph has a simplicial vertex (indeed at least two such vertices). Since induced subgraphs of chordal graphs are still chordal, chordal graphs have a *simplicial elimination ordering*; that is, an ordering $v_1, v_2, \ldots, v_n$ of the vertices such that the neighbourhood of v_i is complete in the induced graph $\langle\{v_i, v_{i+1}, \ldots, v_n\}\rangle$. Using induction one can show that every graph that has a simplicial elimination ordering is chordal.

The centre of chordal graphs was investigated by Laskar and Shier [48]. For example, they showed that the centre of a chordal graph is connected, and provided the following generalization of Fact 2:

Theorem 4. *For a chordal graph G,* $\operatorname{diam}(G) \geq 2\operatorname{rad}(G) - 1$.

4.2 Cartesian Products

Distance behaves nicely in Cartesian products of graphs. Suppose $G = F \times H$. Then, the distance between two vertices (u_1, v_1) and (u_2, v_2) in G is the sum of the distances $d_F(u_1, u_2) + d_H(v_1, v_2)$.

The n-cube Q_n ($n \geq 1$) is defined using Cartesian products of graphs as follows: $Q_1 \cong K_2$; and for $k \geq 1$, $Q_{k+1} \cong Q_k \times K_2$, i.e., Q_{k+1} is the Cartesian product of Q_k and K_2. Mulder [56] showed that the n-cube possesses interesting structural properties that can be described using distance notions. To this end, let u, v be vertices in a graph G. Then the *geodesic interval* between u and v, denoted by $I_g[u, v]$, is the collection of all vertices that belong to some $u - v$-geodesic (shortest path). A graph G is a *median graph* if for any three vertices u, v, w of G, $|I_g[u,v] \cap I_g[u,w] \cap I_g[v,w]| = 1$. For example, the 3-cube, shown in Figure 4, is a median graph.

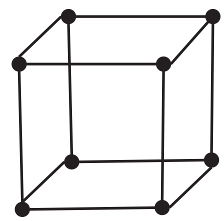

Fig. 4. The 3-cube: a median graph

The next result was established in [56]:

Theorem 5. *A graph G is isomorphic to Q_n if and only if G is a median graph with maximum degree n such that G contains two vertices at distance $diam(G)$ apart at least one of which has degree n.*

Hamming graphs are Cartesian products of complete graphs. Suppose $\{a_1, a_2, \ldots, a_n\}$ are positive integers and that $A_i = \{0, 1, \ldots, a_{i-1}\}$. Then the Hamming graph $H_{a_1,\ldots,a_n}$ is the graph with vertex set the Cartesian product $A_1 \times A_2 \times \ldots \times A_n$ in which two vertices are joined by an edge if and only if the corresponding vectors differ in exactly one coordinate. So $H_{a_1,a_2,\ldots,a_n} = K_{a_1} \times K_{a_2} \times \ldots \times K_{a_n}$ and the n-cube is the Hamming graph with $a_1 = a_2 = \ldots = a_n = 2$. The Hamming graph $H_{4,2}$ is shown in Figure 5. The distance between two vertices is equal to the *Hamming distance* between these vertices, namely, the number of coordinates in which the corresponding vectors differ. Hamming graphs are used in coding theory and have applications to other distance invariants as discussed in Section 8. Mulder [55] extended Theorem 5 to characterize Hamming graphs in terms of intervals and forbidden subgraphs.

4.3 Distance Hereditary Graphs

Howorka [41] defined a graph G to be *distance hereditary* if for each connected induced subgraph H of G and every pair u, v of vertices in H, $d_H(u,v) = d_G(u,v)$. Howorka [41] provided several conditions that characterize distance hereditary graphs:

Theorem 6. *For a graph G the following are equivalent:*

1. *G is distance hereditary;*
2. *every induced path of G is a geodesic;*

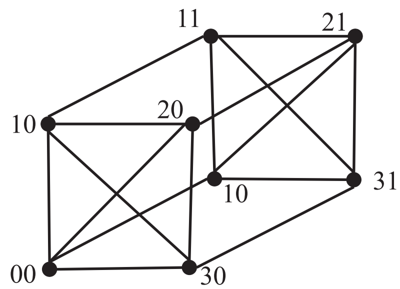

Fig. 5. The Hamming graph $H_{4,2}$

3. *every subpath of a cycle C of more than half C's length is induced;*
4. *every cycle C of length at least 5 has a pair of chords e_1, e_2 of C such that $C + \{e_1, e_2\}$ is homeomorphic with K_4.*

Distance hereditary graphs form a subclass of the perfect graphs (graphs where every induced subgraph has equal clique and chromatic numbers), and several NP-hard problems have efficient solutions for distance hereditary graphs. One such problem is discussed in Section 10. Another characterization of distance hereditary graphs that lends itself well to algorithmic applications was given independently by Bandelt and Mulder [2] and Hammer and Maffray [37]. A graph H is said to be obtained from a graph G by (i) *adding a leaf* v to some vertex v' of G, if v is added to G and joined to v' by an edge; and (ii) by *adding a twin* v to some vertex v' of G, if v is added to G and if v is joined to the vertices in the open or closed neighbourhood of v' in G, i.e., v is joined to the vertices in $N_G(v')$ or $N_G(v') \cup \{v'\}$, respectively.

Theorem 7. *A graph G is distance hereditary if and only if there is a sequence of graphs $G_1, G_2, \ldots, G_{n-1}$ such that $G_1 \cong K_2$, $G_n \cong G_{n-1}$ and for $2 \leq i \leq n-1$, G_i is obtained from G_{i-1} by adding a vertex as a leaf or a twin to some vertex of G_{i-1}.*

Another polynomial recognition algorithm for distance hereditary graphs that is useful for solving the 'Steiner problem' (see Section 10) for these graphs is described in [20]. Bandelt and Mulder [2] gave a characterization of these graphs in terms of forbidden subgraphs.

4.4 Random Graphs

The random graph G_p^n is obtained by starting with n vertices, and then for every pair of distinct vertices, making them adjacent with probability p (each decision independent).

Fact 8 *For any fixed p, G_p^n has diameter 2 with high probability (meaning the limit as $n \to \infty$ is 1).*

There is another version of random graphs, called power-law or scale-free graphs, which models the web better (see, for example, Bonato [6]).

5 Average Distance

The *average distance* of a graph $G = (V, E)$ of order n, denoted by $\mu(G)$, is the expected distance between a randomly chosen pair of distinct vertices; that is,

$$\mu(G) = \frac{1}{\binom{n}{2}} \sum_{u,v \subset V} d(u, v).$$

The study of the average distance began with the chemist Wiener [72], who noticed that the melting point of certain hydrocarbons is proportional to the sum of all distances between unordered pairs of vertices of the corresponding graph. This sum is now called the *Wiener number* or *Wiener index* of the graph and is denoted by $\sigma(G)$. (Note that $\sigma(G) = \binom{n}{2}\mu(G)$.) The average distance of a graph has been used for comparing the compactness of architectural plans [52]. Doyle and Graver [22] were the first to define $\mu(G)$ as a graph parameter.

Here are the values for some simple graphs:

Fact 9 *1. Complete graphs:* $\mu(K_n) = 1$ *(for* $n \geq 2$*).*
2. Complete bipartite graph: $\mu(K_{a,b}) = (ab + 2a^2 + 2b^2)/((a + b)(a + b - 1))$.
3. Path on n *vertices:* $\mu(P_n) = (n + 1)/3$.
4. Cycle on n *vertices:* $\mu(C_n) = (n + 1)/4$ *if* n *is odd, and* $n^2/(4(n - 1))$ *if* n *is even.*

The result on the path is a discrete version of the fact that a pair of randomly chosen points on a line of length 1 have expected distance 1/3.

One might expect some relationship between the radius or diameter of the graph and its average distance. However, Plesník [62] showed that, apart from the trivial inequality $\mu(G) \leq diam(G)$, no such relationship exists.

5.1 Bounds on the Average Distance

The following upper bound was established independently in [22, 24, 51].

Theorem 8. *If* G *is a connected graph of order* n*, then*

$$1 \leq \mu(G) \leq \frac{n+1}{3}.$$

Equality holds if and only if G *is a complete graph or a path.*

Plesník [62] improved this bound for 2-edge-connected graphs to approximately $n/4$. Mahéo and Thuillier (see [30]) showed that $\mu(n) \leq n/(2\kappa) + o(n)$ if G is κ-connected.

A straight-forward lower bound in terms of the order n and number of edges m follows from the fact that there are exactly m pairs of vertices distance 1 apart and the remaining pairs are distance at least 2 apart; see [24]:

Fact 10 *If G is a connected graph of n vertices and m edges, then*

$$\mu(G) \geq 2 - \frac{2m}{n(n-1)}.$$

Equality holds if and only if $\operatorname{diam}(G) = 2$.

Finding the upper bound for $\mu(G)$ in terms of n and m is much more difficult. Soltés [69] found a sharp upper bound:

Theorem 9. *The maximum average distance of a connected graph of order n and m edges, is achieved by the path–complete graph.*

There are a few other graphs achieving the maximum average distance: the extremal graphs were characterized in [35].

Plesník [62] found a sharp lower bound for the average distance of a graph in terms of the order and diameter. However, the problem of finding the exact maximum average distance among all graphs of a given order and diameter remains unsolved.

In Section 2.2 a bound on the diameter in terms of the distinct eigenvalues of the adjacency matrix and of the Laplacian was given. Rodriguez and Yebra [64] obtained a similar result for the average distance of G.

Theorem 10. *Let G be a connected graph of n vertices and m edges, and let b be the number of distinct eigenvalues of G. Then*

$$\mu(G) \leq b - \frac{2(b-1)m}{n(n-1)}.$$

The same result holds if b is the number of distinct Laplacian eigenvalues of G.

The spectrum (set of eigenvalues) does not necessarily determine a graph, not even if the graph is a tree. McKay (see [54]) and Merris [53] showed, however, that the average distance of a tree is determined by its spectrum.

Theorem 11. *Let T be a tree of order n and let $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ be the Laplacian eigenvalues of T. Then*

$$\mu(T) = \frac{2}{n-1} \sum_{i=2}^{n} \frac{1}{\lambda_i}.$$

Mohar [54] determined both upper and lower bounds on the average distance of a graph in terms of some of its Laplacian eigenvalues.

5.2 The Average Distance and Other Graph Invariants

Considerable interest in average distance was generated by conjectures made by the computer program GRAFFITI, developed by Fajtlowicz [27, 28]. One of these was that $\mu(G)$ is at most the independence number $\beta(G)$ of the graph. Chung [14] proved this conjecture:

Theorem 12. *For any connected graph G, $\mu(G) \leq \beta(G)$.*

This bound is attained if and only if G is a complete graph.

The computer program GRAFFITI also conjectured in [28] that every δ-regular connected graph of order n has average distance at most n/δ. GRAFFITI later restated this conjecture for graphs with minimum degree δ. This conjecture remained open for ten years until proven asymptotically by Kouider and Winkler [46]. Their result states:

Theorem 13. *Let G be a connected graph of order n and minimum degree δ. Then*

$$\mu(G) \leq \frac{n}{\delta + 1} + 2.$$

While this result is asymptotically stronger than the GRAFFITI conjecture, it does not imply it. The conjecture was finally proven by Beezer, Riegsecker and Smith [3].

5.3 Edge Removal and the Average Distance

In this section we look at the effect of edge removal on the average distance of a graph. Since the removal of a bridge disconnects the graph, we consider only cyclic edges e. We can measure the effect of e's removal from G by considering either the difference $\mu(G-e)-\mu(G)$ or the ratio $\mu(G-e)/\mu(G)$. In both cases the *best* edge is one whose removal minimizes the quantity in question, and the *worst* edge is one that maximizes the quantity. These questions are of importance in network design: how badly would an edge failure impact the network, or how much would the network suffer if we omitted a particular link to save costs.

If an edge $e = ab$ is removed, the distance between a and b increases and no other distances decrease. Thus, the difference $\mu(G-e)-\mu(G)$ for the best edge is at least $1/\binom{n}{2}$. Finding attainable upper bounds is more difficult. Favaron, Kouider and Mahéo [30] found the maximum value for the difference:

Theorem 14. *Let G be a graph of order n and e a cyclic edge of G. Then*

$$\mu(G - e) - \mu(G) \leq \frac{1}{3}(\sqrt{2} - 1)n + O(1).$$

We now consider the ratio $\mu(G-e)/\mu(G)$. If G is a cycle, then the removal of an edge increases the average distance by a factor of about 4/3. Winkler [73, 74] conjectured that for 2-edge-connected graphs this is the maximum possible ratio for the best edge. This became known as the 'four-thirds conjecture', and was eventually proven by Bienstock and Györi [5]. Soon thereafter, Györi [36] extended it and proved:

Theorem 15. *The four-thirds conjecture holds for all connected graphs that are not trees.*

The ratio $\mu(G-e)/\mu(G)$ can be arbitrarily large if a worst edge e is removed (see [17]). To see this, consider the graph G obtained from a cycle $C_r : v_0v_1v_2\ldots v_{r-1}v_0$, where $r \geq 6$ is a fixed integer, by replacing vertices v_0 and v_3 by complete graphs of order $\lfloor (n-r)/2 \rfloor$ and $\lceil (n-r)/2 \rceil$ whose vertices are adjacent to v_{r-1}, v_1 and v_2, v_4, respectively. If n is large, then two randomly chosen vertices are almost certainly in the union of the two complete graphs and the probability that they are in the same (different) complete graph(s) and thus have distance 1 (3) tends to $\frac{1}{2}$. Hence $\mu(G) = \frac{1}{2}3 + \frac{1}{2}1 + o(1)$. By a similar argument $\mu(G-e) = (r-3)/2 + o(1)$, where e is the edge v_1v_2. Choosing $r = \lfloor 2\sqrt{n} \rfloor$ gives $\mu(G-e)/\mu(G) = O(\sqrt{n})$. Favaron, Kouider and Mahéo [30] showed that this achieves the order of magnitude of the maximum possible value:

Theorem 16. *Let G be a connected graph of order n and e a cyclic edge of G. Then*

$$\frac{\mu(G-e)}{\mu(G)} \leq \frac{\sqrt{n}}{2\sqrt{3}} + O(1).$$

The *minimum average distance spanning tree* (or MAD tree) of a connected graph is a spanning tree having minimum average distance. Such a tree is also referred to as a minimum routing cost tree. It is surprising that the removal of a single best edge can increase the average distance by a factor of 4/3, but the removal of $m-n+1$ best edges (where m is the number of edges in the graph) increases the average distance by a factor less than 2. This fact was established by Entringer, Kleitman and Székely in [25] (and a related result is discussed in [76]).

Theorem 17. *Let G be a connected graph of order n. Then there exists a vertex v and spanning tree T_v that is distance preserving from v, such that*

$$\mu(T_v) \leq 2\frac{n-1}{n}\mu(G).$$

Johnson, Lenstra and Rinnooy-Kan [43] showed that the problem of finding a MAD tree in a graph is NP-hard.

5.4 Vertex Removal and Average Distance

We now turn our attention to the effect of vertex removal on the average distance. Unlike edge removal, vertex removal can both decrease or increase the average distance. For convenience, we express our results in terms of the Wiener index of the graph. Swart [70] showed that the maximum possible decrease occurs when an end-vertex is removed from a path.

Theorem 18. *Let G be a graph of order $n \geq 2$ and let v be a non-cut vertex of G. Then*

$$\sigma(G) - \sigma(G - v) \leq \frac{n(n-1)}{2},$$

with equality if and only if G is a path and v is an end-vertex of G.

Soltés [69] showed that the path–complete graphs are extremal for the ratio:

Theorem 19. *Let G be a graph of order n and $m \geq n-1$ edges, and let v be a non cut-vertex of G. Then*

$$\frac{\sigma(G-v)}{\sigma(G)} \geq \frac{\sigma(PK_{n-1,m-1})}{\sigma(PK_{n,m})}.$$

Soltés [69] also gave sharp upper bounds for $\sigma(G-v) - \sigma(G)$ in terms of the order and number of edges of G.

In some instances the removal of any vertex increases the average distance. For example, the cycle C_n leaves P_{n-1} after the removal of any vertex. Thus the average distance increases by a factor of nearly 4/3. Winkler [73, 74] conjectured that this is the worst increase. This vertex version of the 'four-thirds conjecture' was proven asymptotically by Bienstock and Györi [5]:

Theorem 20. *Every connected graph has a vertex whose removal increases the average distance by a factor of at most $\frac{4}{3} + O(n^{-5})$.*

Althöfer [1] proved the four-thirds conjecture for 4-connected graphs, and improved on it for graphs of higher connectivity.

6 Directed Graphs

Directed graphs behave somewhat differently to undirected graphs. In general we assume that the digraph D is strongly connected; that is, there is a directed path from each vertex to each other vertex. Note, however, that the radius of D might well exist even though the digraph is not strongly connected and the diameter does not exist. Further, it is possible that $\operatorname{diam}(D) \gg 2\operatorname{rad}(D)$.

An *oriented graph* is one obtained by assigning directions to each edge of an undirected graph. That is, it is a digraph without 2-cycles. Füredi et al. [33] provided a lower bound on the number of arcs in an oriented graph of diameter 2:

Theorem 21. *If an oriented graph has diameter* 2*, then* $m \geq (1-o(1))\, n \log n$*, where* n *is the order and* m *the number of arcs.*

Chvátal and Thomassen [15] studied the problem of taking an undirected graph and finding an orientation of minimum diameter. Earlier, Robbins [63] had shown that an undirected graph has a strong orientation if and only if it is bridgeless. Chvátal and Thomassen showed that a bridgeless graph of diameter d has an orientation of diameter at most $2d^2 + d$. They also showed that determining whether a graph has an orientation of diameter 2 is NP-complete.

6.1 The Average Distance of Digraphs

The average distance of digraphs has not received as much attention as the average distance of graphs. Ng and Teh [57] gave a lower bound for the average distance of a strong digraph D in terms of its order n and number of arcs m similar to the one given for undirected graphs; they showed that $\mu(D) \geq 2 - m/(n(n-1))$ with equality if and only if $\mathrm{diam}(D) = 2$. For $n \geq 3$ this bound is sharp if $m \geq 2n - 2$, the smallest number of arcs for which there exists a digraph of order n and diameter 2.

Plesník [62] proved the following upper bound on the average distance of a strong digraph.

Theorem 22. *Let* D *be a strong digraph of order* n*. Then* $\mu(D) \leq n/2$*. Equality holds if and only if* D *is a directed cycle.*

It is natural to ask whether there are upper bounds on the average distance of a digraph in terms of the order and minimum degree that are analogous to the ones for graphs, but it turns out that in general these bounds do not carry over for general digraphs.

In [19] the problem of finding an orientation of a 2-edge-connected graph that minimizes the average distance is studied.

6.2 Tournaments

A *tournament* is an oriented complete graph. Landau [47] showed that every tournament has radius at most 2. It is easy to construct a tournament on n vertices with diameter $n-1$. For $n \neq 4$, it is also easy to construct a tournament with diameter 2.

Plesník [62] gave bounds for average distance in strong tournaments.

Theorem 23. *Let* T_n *be a strong tournament of order* $n \geq 3$*. Then*

$$\frac{3}{2} \leq \mu(T_n) \leq \frac{n+4}{6}.$$

Equality holds if and only if T_n *has diameter* 2 *(this is possible only if* $n \neq 4$*) or if* T_n *is the unique strong tournament of diameter* $n-1$*, respectively.*

7 Convexity

Interval notions in graphs have led to the study of abstract convexity in graphs and structural characterizations of several interesting graph classes.

Suppose V is a collection of points and $\mathcal{M}$ a collection of subsets of V. Then $\mathcal{M}$ is a *convexity* if it contains both $\emptyset$ and V and it is closed under intersections. The elements of $\mathcal{M}$ are called *convex sets*. If $T \in \mathcal{M}$, then a point v of T is an *extreme point* of T if $T \setminus \{v\} \in \mathcal{M}$. If $S \subseteq V$, then the smallest convex set containing S is called the *convex hull* of S. A *convex geometry* is a convexity with the additional property that every convex set is the convex hull of its extreme points.

The most well-known graph convexity is defined in terms of geodesic intervals, which were introduced in Section 4.2. Suppose $G = (V, E)$ is a connected graph. Then a set $S \subseteq V$ is *g-convex* if $I_g[u, v] \subseteq S$ for all pairs $u, v \in S$. Let $\mathcal{M}_g(G)$ be the collection of all g-convex sets of G. Then $\mathcal{M}_g(G)$ is a convexity.

It is not difficult to see that a vertex v of a g-convex set S is an extreme vertex of S if and only if v is simplicial in $\langle S\rangle$, i.e., the neighbourhood of v in S induces a complete graph. Farber and Jamison [29] characterized the class of graph for which the g-convex sets form a convex geometry:

Theorem 24. *Let G be a connected graph. Then $\mathcal{M}_g(G)$ is a convex geometry if and only if G is chordal without an induced 3-fan (see Figure 6).*

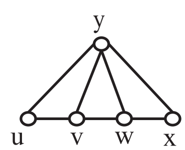

Fig. 6. The 3-Fan

Every shortest path is necessarily induced, but not conversely. This leads to another type of graph interval. The *monophonic interval* between a pair u, v of vertices in a graph G, denoted by $I_m[u, v]$, is the collection of all vertices that lie on an induced $u - v$ path in G. A set S of vertices in a graph is *m-convex* if $I_m[u, v] \subseteq S$ for all pairs $u, v \in S$. It is not difficult to see that the collection $\mathcal{M}_m(G)$ of all m-convex sets is a convexity, and that the extreme points of an m-convex set S are precisely the simplicial vertices of $\langle S\rangle$. Farber and Jamison [29] characterized those graphs for which the m-convex sets form a convex geometry:

Theorem 25. *Let G be a connected graph. Then $\mathcal{M}_m(G)$ is a convex geometry if and only if G is chordal.*

Dragan, Nicolai and Brandstädt [23] defined another type of graph interval. If u, v is a pair of vertices in a connected graph G, then the m^3-interval, denoted by $I_{m^3}[u, v]$, between u and v is the collection of all vertices that lie on some induced $u - v$ path of length at least 3. A set S of vertices in G is m^3-convex if $I_{m^3}[u, v] \subseteq S$ for all pairs u, v of vertices in S. For a graph G, let $\mathcal{M}_{m^3}(G)$ be the collection of all m^3-convex sets. This collection of sets is certainly a convexity. Further, it can be shown that a vertex v is an extreme point of an m^3-convex set S if and only if v is *semisimplicial*, i.e., not the centre of an induced P_4. The class of graphs for which the m^3-convex sets form a convex geometry was characterized in [23]:

Theorem 26. *Let G be a connected graph. Then $\mathcal{M}_{m^3}(G)$ is a convex geometry if and only if G is (house, hole, domino, A)-free (see Figure 7).*

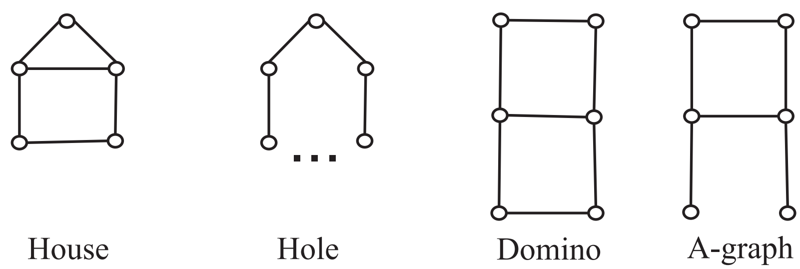

Fig. 7. Forbidden graphs

8 Metric Dimension

Distances in graphs have interesting applications. One such application is to uniquely locate the position of a vertex in a network using distances. A vertex v *resolves* a pair u, w of vertices in a connected graph G if $d(u, v) \neq d(w, v)$. A set of vertices S is a *resolving set* of G, if every pair of vertices in G is resolved by some vertex of S. A resolving set of minimum cardinality is called a *metric basis* of G and its cardinality the *metric dimension*, $dim(G)$, of G.

The metric dimension of a graph was introduced by Slater [66] and independently by Harary and Melter [39]. Slater referred to the metric dimension as the location number, and motivated its study by its application to the placement of a minimum number of sonar/loran detecting devices in a network so that the position of every vertex in the network can be uniquely described in terms of its distances to the detecting devices. A problem in pharmaceutical chemistry once again led to an independent discovery of the notion of a resolving set of a graph [12]. In [34] it was noted that the metric dimension of a graph is NP-hard.

The formula for the metric dimension of trees has been discovered independently by several authors (see [66, 39, 12]). The metric dimension of a nontrivial path is 1 since a leaf resolves the path. Suppose now that T is a tree that contains vertices of degree at least 3. A vertex v of degree at least 3 is an *exterior vertex* if there is some leaf u in T such that the $v - u$ path of T contains no vertices of degree exceeding 2 except for v. Let $ex(T)$ denote the number of exterior vertices of T and $\ell(T)$ the number of leaves of T. It turns out that a metric basis for a tree can be found by selecting, for each exterior vertex, all but one of its exterior leaves. That is, the metric dimension of T is given in the following:

Theorem 27. *For a tree T,* $\dim(T) = \ell(T) - ex(T)$.

Apart from trees, very few exact results for the metric dimension of graphs are known unless the graphs are highly structured (usually vertex transitive). It was claimed in [45] that the metric dimension of the cartesian product of k paths is k; but indeed it was only verified that k is an upper bound in this case. In [65] a connection between the metric dimension of the n-cube and the solution to a coin weighing problem was noted. This observation and results by Lindström [50] and Erdös and Rényi [26] show that $lim_{n\to\infty} \dim(Q_n) \log n/n = 2$, thereby disproving the claim about n-cubes made in [45].

Motivated by the connection between coin weighing problems/strategies for the Mastermind game and resolving sets in cartesian products of certain classes of graphs, the metric dimension of cartesian products of graphs was investigated in [10]. This paper introduces 'doubly resolving sets' as a useful tool for obtaining upper bounds on the metric dimension of graphs, particularly in Cartesian products of graphs.

9 Algorithms and Complexity

9.1 Shortest Paths

To compute the distance between two vertices in an unweighted graph, one can use a breadth-first search. To compute the distance between two vertices in a weighted graph, one can use Dijkstra's algorithm (which is in some sense a generalization of breadth-first search). Note that the algorithm actually finds the distance from a given start vertex to all other vertices.

```
ShortestPath (G:graph, a:vertex)
    for all vertices v do currDis(v) := infinity
    currDis(a) := 0
    remainder := [ all vertices ]
  while remainder nonempty do {
    let w be vertex with minimum value of currDis
    remainder −= [w]
    for all vertices v in remainder do
```

```
        currDis (v) := min ( currDis(v), currDis(w)+length(w,v) )
    }
```

The running time of the above implementation of Dijkstra's algorithm is $O(n^2)$. By using suitable data structures this can be brought down for sparse graphs to $O(m + n \log n)$, where m is the number of edges.

9.2 All Pairs Shortest Paths

Suppose we wanted instead to calculate the shortest path between every pair of vertices, for example, in order to compute the average distance. One idea would be to run Dijkstra with every vertex as a start vertex. This takes $O(n^3)$ time. There are two dynamic programming algorithms with similar running times. One is due to Bellman & Ford and the other Floyd & Warshall. A variant of the former is used in routing protocols in networks. We describe here the Floyd–Warshall algorithm [31].

Suppose the vertices are ordered 1 up to n. Then define

> $d_m(u, v)$ *as the length of the shortest path between* u *and* v *that uses only the vertices numbered* 1 *up to* m *as intermediates.*

The desired value is $d_n(u, v)$ for all u and v.

There is a formula for d_m in terms of d_{m-1}. For, the shortest u to v path that uses only vertices labeled up to m, either uses vertex m or it doesn't. Thus:

$$d_m(u, v) = \min \begin{cases} d_{m-1}(u, v) \\ d_{m-1}(u, m) + d_{m-1}(m, v) \end{cases}$$

The resultant program iterates m from $m = 0$ to $m = n - 1$.

10 Steiner Distances

Up to this point we have considered distance invariants that hinge on shortest paths between pairs of vertices. In this section we give a brief overview of related invariants that arise by considering the 'cheapest' subgraph that connects a given set of vertices.

10.1 Extending Distance Measures

Suppose G is a (weighted) graph and S a set of vertices in G. Then the *Steiner distance* for S, denoted by $d_G(S)$, is the smallest weight of a connected subgraph of G containing S. Such a subgraph is necessarily a tree, called a *Steiner tree* for S. The problem of finding a Steiner tree for a given set S of vertices is called the Steiner problem. In its two extremes, namely if $|S| = 2$ or $|S| = n$, the Steiner problem is solved efficiently by well-known algorithms, for example, Dijkstra's algorithm and Kruskal's minimum spanning tree algorithm, respectively. In general, however, this problem is NP-hard (see [34]), even for unweighted bipartite graphs. Winter's survey [75] provides a good overview of different heuristics that have been developed for the problem as well as exact solutions for various graph classes.

The radius, diameter and average distance have a natural extension. For a given vertex v in a connected (weighted) graph G and integer k ($2 \leq k \leq n$), the *k-eccentricity* of v, denoted by $e_k(v)$, is the maximum Steiner distance among all k-sets of vertices in G that contain v. The *k-radius*, $rad_k(G)$, of G is the minimum k-eccentricity of the vertices of G, and the *k-diameter*, $diam_k(G)$, of G is the maximum k-eccentricity. The *average Steiner k-distance*, $\mu_k(G)$ of G, is the average Steiner distance among all k-sets of vertices of G. The *k-distance* of a vertex v, denoted by $e_k(v)$, is the sum of the Steiner distances of k-sets of vertices containing v. The subgraph induced by vertices of minimum k-eccentricity is called the *k-centre* of G and is denoted by $C_k(G)$; the subgraph induced by the vertices of minimum k-distance is called the *k-median* and is denoted by $M_k(G)$.

The k-diameter is clearly an upper bound for the k-radius. No upper bound for the k-diameter as a function of k and the k-radius of an (unweighted) graph is known. For trees, the following generalization of Fact 2 was established in [13].

Theorem 28. *For a tree T of order n and integer $k \leq n$, $diam_k(T) \leq \frac{n}{n-1} rad_k(T)$.*

It was shown in [60], that the k-centre of a tree T can be found by successively pruning leaves. If T has at most $k-1$ leaves, then T is its own k-centre. If T has at least k leaves, then the k-centre is the i^{th} derivative of T where i is the smallest integer such that $T^{(i)}$ has at most $k-1$ leaves. This also shows that the k-centre of a tree is contained in the $(k+1)$-centre of a tree. (This containment does not hold in general graphs.) Moreover, it follows that k-centres of trees are connected.

The k-median of trees were shown in [4] to be connected. In the same paper it was shown that a tree H of order p is the k-median of a tree if and only if $p = 1, 2$ or k or if H has at most $k - p + 1$ leaves. An algorithm for finding the k-median of a tree T was also described. If a tree has order at least $2k-1$, then its k-median consists of a single vertex or a pair of adjacent vertices. The k-centre and k-median of a tree may be arbitrarily far apart (see [59]). Centrality structures that connect the k-centre and k-median of a tree are introduced and studied in [59].

The average Steiner k-distance of a graph was first defined in [18]. In the same paper it is shown that $\mu_k(G) \leq \mu_l(G) + \mu_{k+1-l}(G)$ for $2 \leq l \leq k-1$, and that the range of average Steiner k-distance of a graph is given by:

Theorem 29. *If G is a connected graph of order n and $2 \leq k \leq n$, then*

$$k-1 \leq \mu_k(G) \leq \frac{k-1}{k+1}(n+1),$$

with equality on the left if and only if G is $(n+1-k)$-connected or $k=n$, and equality on the right if and only if G is a path or $n=k$.

An efficient procedure that finds the average Steiner k-distance of a tree is described in [18]. This algorithm counts the number of k-sets such that a given edge belongs to a Steiner tree for the k-set. Moreover, it is shown that for a tree T, $\mu_k(T) \leq k\mu_l(T)/l$ for $2 \leq l \leq k-1$ with equality if and only if T is a star, and the lower bound given in Theorem 29 is improved to $k(1-1/n)$ for trees.

10.2 Steiner Intervals and Graph Convexity

The *Steiner interval* of a set X of vertices in a connected graph G, denoted by $I(X)$, is the collection of all vertices that belong to some Steiner tree for X. A set S of vertices is *k-Steiner convex*, denoted by g_k-convex, if $I(X) \subseteq S$ for all subsets X of S with $|X| = k$. Thus a g_2-convex set is a g-convex set. We call the extreme vertices of a g_k-convex set a *k-Steiner simplicial* vertex and abbreviate this by kSS. The $3SS$ vertices are characterized in [11] as the vertices that are **not** the centre of an induced claw, paw or P_4, see Figure 8.

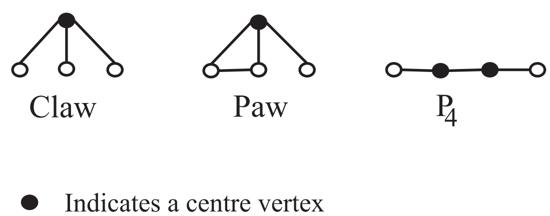

Fig. 8. Characterizing $3SS$ vertices

Thus the $3SS$ vertices are semisimplicial. Several graph convexities related to the g_3-convexity are introduced in [58] and those graphs for which these graph convexities form convex geometries are characterized in the same paper. We state here a characterization of those graphs for which the g_3-convex sets form a convex geometry. A replicated-twin C_4 is any one of the four graphs shown in Figure 9 where any subset of the dotted edges belongs to the graph. The collection of replicated-twin C_4's is denoted by $\mathcal{R}_{C_4}$.

Theorem 30. *Let G be a connected graph and $\mathcal{M}_{g_3}(G)$ the collection of all g_3-convex sets of G. Then $\mathcal{M}_{g_3}(G)$ is a convex geometry if and only if $diam(G) \leq 2$ and if G is (house, hole, 3-fan, $\mathcal{R}_{C_4}$)-free.*

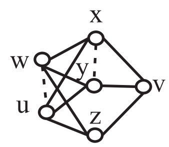

Fig. 9. The replicated twin C_4's

Acknowledgments

We would like to thank Peter Dankelmann for sharing his thoughts on average distance with us.

References

1. I. Althöfer. Average distances in undirected graphs and the removal of vertices. *J. Combin. Theory. Ser. B*, 48(1):140–142, 1990.
2. H.-J. Bandelt and H.M. Mulder. Distance-hereditary graphs. *J. Combin. Theory Ser. B*, 41:182–208, 1986.
3. R.A. Beezer, J.E. Riegsecker, and B.A. Smith. Using minimum degree to bound average distance. *Discrete Math.*, 226:365–371, 2001.
4. L.W. Beineke, O.R. Oellermann, and R.E. Pippert. On the Steiner median of a tree. *Discrete Appl. Math.*, 68:249–258, 1996.
5. D. Bienstock and E. Győri. Average distance in graphs with removed elements. *J. Graph Theory*, 12(3):375–390, 1988.
6. A. Bonato. A survey of models of the web graph. In *Combinatorial and algorithmic aspects of networking*, volume 3405 of *Lecture Notes in Comput. Sci.*, pages 159–172. Springer, Berlin, 2005.
7. F. Buckley and F. Harary. *Distance in Graphs*. Addison-Wesley Publishing Company Advanced Book Program, Redwood City, CA, 1990.
8. F. Buckley and M. Lewinter. A note on graphs with diameter-preserving spanning trees. *J. Graph Theory*, 12(4):525–528, 1988.
9. F. Buckley, Z. Miller, and P.J. Slater. On graphs containing a given graph as center. *J. Graph Theory*, 5:427–432, 1981.
10. J. Cáceres, C. Hernando, M. Mora, I.M. Pelayo, M.L. Puertas, C. Seara, and D.R. Wood. On the metric dimension of the cartesian product of graphs. *SIAM J. Discrete Math.*, 21(2):423–441, 2007.
11. J. Cáceres and O.R. Oellermann. On 3-Steiner simplicial orderings. To appear in *Discrete Math.*
12. G. Chartrand, L. Eroh, M.A. Johnson, and O.R. Oellermann. Resolvability in graphs and the metric dimension of a graph. *Discrete Appl. Math.*, 105:99–113, 2000.
13. G. Chartrand, O.R. Oellermann, S. Tian, and H.B. Zou. Steiner distance in graphs. *Časopis Pro Pěstování Matematiky*, 114(4):399–410, 1989.
14. F.R.K. Chung. The average distance and the independence number of a graph. *J. Graph Theory*, 12:229–235, 1988.
15. V. Chvátal and C. Thomassen. Distances in orientations of graphs. *J. Combin. Theory Ser B*, 24:61–75, 1978.

16. E.J. Cockayne, S.M. Hedetniemi, and S.T. Hedetniemi. Linear algorithms for finding the Jordan center and path center of a tree. *Transportation Sci.*, 15:98–114, 1981.
17. P. Dankelmann. Average distance in graphs – a survey. Private communication.
18. P. Dankelmann, O.R. Oellermann, and H.C. Swart. The average Steiner distance of a graph. *J. Graph Theory*, 22(1):15–22, 1996.
19. P. Dankelmann, O.R. Oellermann, and J-L. Wu. Minimum average distance of strong orientations of graphs. *Discrete Appl. Math.*, 143:204–212, 2004.
20. A. D'Atri and M. Moscarini. Steiner trees and connected domination. *SIAM J. Discrete Math.*, 17:521–538, 1988.
21. G.A. Dirak. On rigid circuit graphs. *Abh. Math. Sem. Univ. Hamburg*, 25:71–76, 1961.
22. J.K. Doyle and J.E. Graver. Mean distance in a graph. *Discrete Math.*, 7(2):147–154, 1977.
23. F.F. Dragan, F. Nicolai, and A. Brandstädt. Convexity and HHD-free graphs. *SIAM J. Discrete Math.*, 12:119–135, 1999.
24. R.C. Entringer, D.E. Jackson, and D.A. Snyder. Distance in graphs. *Czech. Math. J.*, 26:283–296, 1976.
25. R.C. Entringer, D.J. Kleitman, and L.A. Székely. A note on spanning trees with minimum average distance. *Bull. Inst. Combin. Appl.*, 17:71–78, 1996.
26. P. Erdös and A. Rényi. On two problems of information theory. *Magyar Tud. Akad. Mat. Kutató Int. Közl*, 8:229–243, 1963.
27. S. Fajtlowicz and W.A. Waller. On two conjectures of GRAFFITI. *Congr. Numer.*, 55:51–56, 1986.
28. S. Fajtlowicz and W.A. Waller. On conjectures of GRAFFITI II. *Congr. Numer.*, 60:187–197, 1987.
29. M. Farber and R.E. Jamison. Convexity in graphs and hypergraphs. *SIAM J. Alg. Discrete Meth.*, 7:433–444, 1986.
30. O. Favaron, M. Kouider, and M. Mahéo. Edge-vulnerability and mean distance. *Networks*, 19(5):493–504, 1989.
31. R.W. Floyd. Algorithm 97: Shortest path. *Comm. ACM*, 5:345, 1962.
32. L.C. Freeman. Centrality in social networks: Conceptual clarification. *Social Networks*, 1:215–239, 1978/1979.
33. Z. Füredi, P. Horak, C.M. Pareek, and X. Zhu. Minimal oriented graphs of diameter 2. *Graphs Combin.*, 14:345–350, 1998.
34. M.R. Garey and D.S. Johnson. *Computers and Intractibility: A Guide to the Theory of NP-Completeness.* W.H. Freeman and Company, New York, 1979.
35. W. Goddard, C.S. Swart, and H.C. Swart. On the graphs with maximum distance or k-diameter. *Math. Slovaca*, 55(2):131–139, 2005.
36. E. Györi. On Winkler's four thirds conjecture on mean distance in graphs. *Congr. Numer.*, 61:259–262, 1988.
37. P.L. Hammer and F. Maffray. Completely separable graphs. *J. Discrete Appl. Math.*, 27:85–99, 1990.
38. F. Harary. The maximum connectivity of a graph. *Proc. Nat. Acad. Sci. U.S.A.*, 48:1142–1146, 1962.
39. F. Harary and R.A. Melter. On the metric dimension of a graph. *Ars Combinatoria*, 2:191–195, 1976.
40. F. Harary and R.Z. Norman. The dissimilarity characteristic of Husimi trees. *Ann. of Math. (2)*, 58:134–141, 1953.

41. E. Howorka. A characterization of distance-hereditary graphs. *Quart. J. Math., Oxford*, 28:417–420, 1977.
42. H. Jeong, S.P. Mason, A.-L. Barabási, and Z.N. Oltvai. Lethality and centrality in protein networks. *Nature*, 411:41–42, 2001.
43. D.S. Johnson, J.K. Lenstra, and A.H.G. Rinnooy-Kan. The complexity of the network design problem. *Networks*, 8:279–285, 1978.
44. C. Jordan. Sur les assembalges des lignes. *J. Reine Angew. Math.*, 70:185–190, 1869.
45. S. Khuller, B. Raghavachari, and A. Rosenfeld. Landmarks in graphs. *Discrete Appl. Math.*, 70(3):217–229, 1996.
46. M. Kouider and P. Winkler. Mean distance and minimum degree. *J. Graph Theory*, 25:95–99, 1997.
47. H.G. Landau. On dominance relations and the structure of animal societies. III. The condition for a score structure. *Bull. Math. Biophys.*, 15:143–148, 1953.
48. R. Laskar and D. Shier. On powers and centers of chordal graphs. *Discrete Appl. Math.*, 6(2):139–147, 1983.
49. C.G. Lekkerkerker and J. Boland. Representation of a finite graph by a set of intervals on the real line. *Fund. Math.*, 51:45–64, 1962.
50. B. Lindström. On a combinatory detection problem. I. *Magyar Tud. Akad. Mat. Kutató Int. Közl*, 9:195–207, 1964.
51. L. Lovász. *Combinatorial Problems and Exercises*. Akadémiai Kiadó, Budapset, 1979.
52. L. March and P. Steadman. *The Geometry of Environment: An introduction to spatial organization in design*. M.I.T. Press, Cambridge, Mass., 1974.
53. R. Merris. An edge version of the matrix-tree theorem and the Wiener index. *Linear Multilinear Algebra*, 25(4):291–296, 1989.
54. B. Mohar. Eigenvalues, diameter, and mean distance in graphs. *Graphs Combin.*, 7(1):53–64, 1991.
55. H.M. Mulder. *The Interval Function of a Graph*. Mathematical Centre Tracts, Amsterdam, 1980.
56. H.M. Mulder. n-cubes and median graphs. *J. Graph Theory*, 4:107–110, 1980.
57. C.P. Ng and H.H. Teh. On finite graphs of diameter 2. *Nanta Math.*, 1:72–75, 1966/67.
58. M. Nielsen and O.R. Oellermann. Steiner trees and convex geometries. Submitted.
59. O.R. Oellermann. From Steiner centers to Steiner medians. *J. Graph Theory*, 20(2):113–122, 1995.
60. O.R. Oellermann and S. Tian. Steiner centers in graphs. *J. Graph Theory*, 14(5):585–597, 1990.
61. J. Plesník. Note on diametrically critical graphs. In *Recent advances in graph theory (Proc. Second Czechoslovak Sympos., Prague, 1974)*, pages 455–465. Academia, Prague, 1975.
62. J. Plesník. On the sum of all distance in a graph or digraph. *J. Graph Theory*, 8:1–24, 1984.
63. H.E. Robbins. Questions, discussions, and notes: A theorem on graphs, with an application to a problem of traffic control. *Amer. Math. Monthly*, 46(5):281–283, 1939.
64. J.A. Rodriguez and J.L.A. Yebra. Bounding the diameter and the mean distance of a graph from its eigenvalues: Laplacian versus adjacency matrix methods. *Discrete Math.*, 196(1-3):267–275, 1999.

65. A. Sebö and E. Tannier. On metric generators of graphs. *Math. Oper. Res.*, 29(2):383–393, 2004.
66. P.J. Slater. Leaves of trees. *Congr. Numer.*, 14:549–559, 1975.
67. P.J. Slater. Centers to centroids in graphs. *J. Graph Theory*, 2:209–222, 1978.
68. P.J. Slater. Centrality of paths and vertices in a graph: cores and pits. *The Theory and Applications of Graphs, Wiley, NY*, pages 529–542, 1981.
69. L. Šoltés. Transmission in graphs: a bound and vertex removing. *Math. Slovaca*, 41(1):11–16, 1991.
70. C.S. Swart. Distance measures in graphs and subgraphs. Master's thesis, University of Natal, Durban, 1996.
71. V.G. Vizing. On the number of edges in a graph with a given radius. *Dokl. Akad. Nauk SSSR*, 173:1245–1246, 1967.
72. H. Wiener. Structural determination of paraffin boiling points. *J. Amer. Chem. Soc.*, 69(1):17–20, 1947.
73. P. Winkler. Mean distance and the four thirds conjecture. *Congr. Numer.*, 54:53–61, 1986.
74. P. Winkler. Graph theory in memory of G.A. Dirac. In *Proceedings of Meeting in Sandbjerg, Denmark, 1985, Annals of Discrete Mathematics*. North-Holland, Amsterdam, 1989.
75. P. Winter. Steiner problem in networks: a survey. *Networks*, 17:129–167, 1987.
76. R. Wong. Worst-case analysis of network design problem heuristics. *SIAM J. Alg. Discrete Meth.*, 1:51–63, 1980.
77. B. Zelinka. Medians and peripherians of trees. *Arch. Math. (Brno)*, 4:87–95, 1968.

Index