GRAPHS AND THEIR APPLICATIONS (4)

K. M. KOH*, F. M. DONG AND E. G. TAY

11. Distance between Two Vertices

Consider the connected graph H of Figure 11.1.

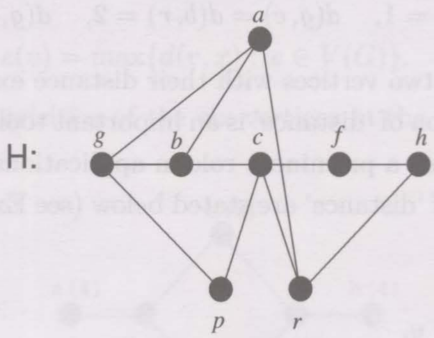

Figure 11.1

As H is connected, every two vertices in H are joined by at least one path. Take, for instance, the two vertices g and h, and observe that $gpcbarh$ is a $g-h$ path in H. The **length** of a path is defined as the number of edges contained in it. Thus the length of the above $g-h$ path is 6. Some other $g-h$ paths of different lengths are given below:

$g-h$ path	Length
$gabcrh$	5
$gbcfh$	4
$garh$	3

The smallest length of the above $g-h$ paths is 3. Is there any $g-h$ path of length less than 3 in H? The answer is 'NO'. Thus, 3 is the **minimum** among the lengths of all $g-h$ paths in H. In this situation, we say that the **distance** between g and h is 3, and we write $d(g,h) = 3$.

In general, let G be a connected multigraph and let $u, v \in V(G)$. The **length** of the $u-v$ walk: $u = v_0v_1v_2\cdots v_k = v$ is defined as k (note that the v_i's are not necessarily distinct; also, for the definition of a 'walk', see page 12 in [9]). In particular, the length of a path is the number of edges contained in it. The **distance** between u and v in G, denoted by $d_G(u,v)$, or simply $d(u,v)$ if the graph G under consideration is

*Corresponding author *Email address*: matkohkm@nus.edu.sg.

Graphs and Their Applications

clear from the context, is defined as the **minimum** of the lengths of all $u-v$ paths in G. Thus, for the graph H of Figure 11.1, we have:

$$d(c,f)=d(p,c)=1, \quad d(g,c)=d(b,r)=2, \quad d(g,h)=d(a,f)=3;$$

and that there exist no two vertices with their distance exceeding 3.

As we shall see, the notion of 'distance' is an important tool in studying the structure of a graph, and plays also a prominent role in applications of graphs.

Some basic properties of 'distance' are stated below (see Exercise 11.4). Let $x, y, z \in V(G)$. Then

(11.1) $d(x,x)=0$,
(11.2) $d(x,y)>0$ if $x \neq y$,
(11.3) (**the symmetric property**) $d(x,y)=d(y,x)$,
(11.4) (**the triangle inequality**) $d(x,y)+d(y,z) \geq d(x,z)$.

Exercise 11.1. *In the graph H of Figure 11.1, find $d(b,p)$, $d(b,h)$ and $d(a,p)$.*

Exercise 11.2. *In the graph H of Figure 11.1, find*
(i) three vertices x, y, z such that $d(x,y)+d(y,z)=d(x,z)$;
(ii) three vertices x, y, z such that $d(x,y)+d(y,z)>d(x,z)$.

Exercise 11.3. *Let G be a connected multigraph and let $u, v \in V(G)$. Show that every $u-v$ walk in G always contains a $u-v$ path.*

Exercise 11.4. *Prove the results (11.1) - (11.4).*

Exercise 11.5. *For each integer $n>1$, construct a graph of order n such that $d(x,y)=1$ for any two distinct vertices x, y in G.*

Exercise 11.6. *For each integer $n>1$, construct a graph G of order n such that for each integer k with $0<k<n$, there exist two vertices x, y in G with $d(x,y)=k$.*

Exercise 11.7. *Let G be a connected multigraph and let $u, v \in V(G)$. Show that for each integer k with $0<k<d(u,v)$, there exists $w \in V(G)$ such that $d(u,w)=k$.*

Exercise 11.8. *Let H be a connected subgraph of a connected graph G. Show that $d_G(u,v) \leq d_H(u,v)$ for any two vertices u, v in H.*

12. Eccentricity, Radius, Diameter and Centre

Consider the vertex a in the graph H of Figure 11.1. Which vertices in H are furthest from a? It can be checked that f is the only such vertex. What is $d(a,f)$? The answer is 3. This '3', which measures the distance between a and a vertex furthest from a, is called the **eccentricity** of the vertex a in H.

Graphs and Their Applications

In general, let G be a connected multigraph and $v \in V(G)$. The **eccentricity** of v, denoted by $e(v)$, is the distance between v and a vertex furthest from v in G. That is,

(12.1) $$e(v) = \max\{d(v,x) : x \in V(G)\}.$$

As an example, the eccentricities of the six vertices in the graph G of Figure 12.1 are shown in parentheses.

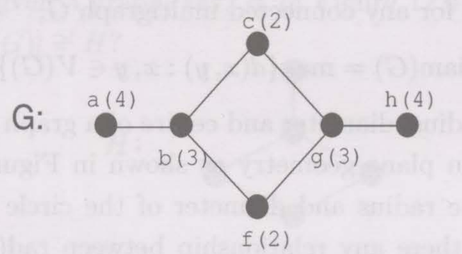

Figure 12.1

Among the six eccentricities shown in the figure, we notice that '2' is the smallest while '4' is the largest. In this situation, we say that the **radius** of G is 2 and the **diameter** of G is 4. Note also that there are two vertices in G, namely c and f, with least eccentricity (i.e., $e(c) = e(f) = 2$). Each of them is called a **central** vertex, and the set of these two central vertices is called the **centre** of G.

In general, given a connected multigraph G, the **radius** of G, denoted by rad(G), is defined by

(12.2) $$\mathrm{rad}(G) = \min\{e(v) : v \in V(G)\}$$

and the **diameter** of G, denoted by diam(G), is defined by

(12.3) $$\mathrm{diam}(G) = \max\{e(v) : v \in V(G)\}.$$

A vertex w in G is called a **central** vertex if $e(w) = \mathrm{rad}(G)$, and the **centre** of G, denoted by $C(G)$, is the set of all central vertices of G.

In the graph G of Figure 12.2, the eccentricities of the twelve vertices are shown in parentheses.

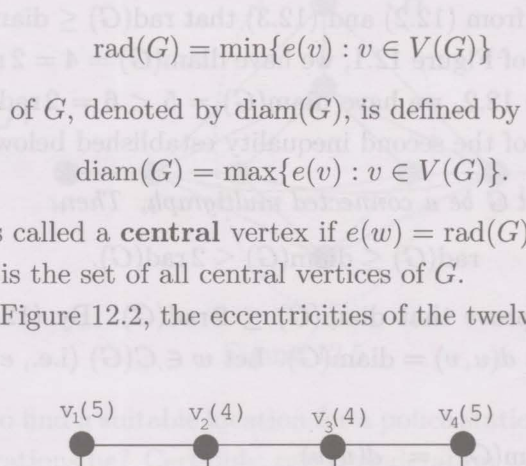

Figure 12.2

Graphs and Their Applications

By (12.2) and (12.3), $\mathrm{rad}(G) = 3$ and $\mathrm{diam}(G) = 5$. The graph G has two central vertices, namely, v_6 and v_7. Thus, $C(G) = \{v_6, v_7\}$. If G represents the street network of a small town, then, geographically, the junctions v_6 and v_7 are really situated at the 'town centre'.

While $\mathrm{diam}(G) = 5$ by (12.3) as pointed out above, we note also that there exist at least two vertices, for instance, v_1 and v_{12}, in G such that $d(v_1, v_{12}) = 5$, and $d(x, y) \leq 5$ for any other two vertices x, y in G. As a matter of fact, it can be shown (see Exercise 12.2) that for any connected multigraph G,

$$\mathrm{diam}(G) = \max\{d(x,y) : x, y \in V(G)\}. \tag{12.4}$$

The above notions of radius, diameter and centre of a graph are, actually, borrowed from those of a circle in plane geometry as shown in Figure 12.3, where r and d denote, respectively, the radius and diameter of the circle with centre O. While $d = 2r$ for a circle, is there any relationship between $\mathrm{rad}(G)$ and $\mathrm{diam}(G)$ for a graph G?

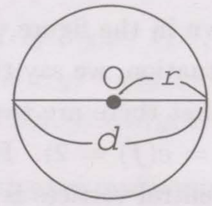

Figure 12.3

It follows immediately from (12.2) and (12.3) that $\mathrm{rad}(G) \leq \mathrm{diam}(G)$. On the other hand, for the graph G of Figure 12.1, we have $\mathrm{diam}(G) = 4 = 2\,\mathrm{rad}(G)$, whereas for the graph G of Figure 12.2, we have $\mathrm{diam}(G) = 5 < 6 = 2\,\mathrm{rad}(G)$. Indeed, these are just two instances of the second inequality established below.

Theorem 12.1. *Let G be a connected multigraph. Then*

$$\mathrm{rad}(G) \leq \mathrm{diam}(G) \leq 2\,\mathrm{rad}(G). \tag{12.5}$$

Proof. We need only prove that $\mathrm{diam}(G) \leq 2\,\mathrm{rad}(G)$. By (12.4), let u, v be two vertices in G such that $d(u, v) = \mathrm{diam}(G)$. Let $w \in C(G)$ (i.e., $e(w) = \mathrm{rad}(G)$). We then have

$$\begin{aligned}
\mathrm{diam}(G) &= d(u,v) \\
&\leq d(u,w) + d(w,v) \quad \text{(by (11.4))} \\
&= d(w,u) + d(w,v) \quad \text{(by (11.3))} \\
&\leq e(w) + e(w) \quad \text{(by (12.1))} \\
&= 2e(w) \\
&= 2\,\mathrm{rad}(G),
\end{aligned}$$

Graphs and Their Applications

as was to be shown. □

In the graph G of Figure 12.1, its centre $C(G)$ consists of two vertices, and the induced subgraph $[C(G)]$ (see page 112 in [10]) is isomorphic to O_2. In the graph G of Figure 12.2, its centre $C(G)$ also consists of two vertices, but the induced subgraph $[C(G)]$ is isomorphic to K_2. One can see also from Exercise 12.1 that the subgraphs induced by the centres of the three graphs are pairwise non-isomorphic. Suppose now we are given a graph, say, H of Figure 12.4, and we ask: Is there a graph G such that $[C(G)] \cong H$?

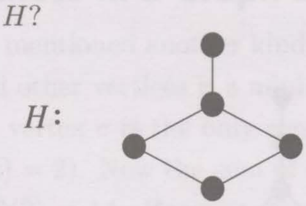

Figure 12.4

Construct a new graph G from H by adding four new vertices w, x, y, z such that x is adjacent only to w and all vertices in H, and, similarly, y is adjacent only to z and all vertices in H as shown in Figure 12.5. It can be checked that $[C(G)] = H$. In [11], Kopylov and Timofeev stated, without proof, that **for any graph H, there always exists a graph G such that $[C(G)] \cong H$.** The above method of constructing G from H was, however, due to Hedetniemi as pointed out in [2].

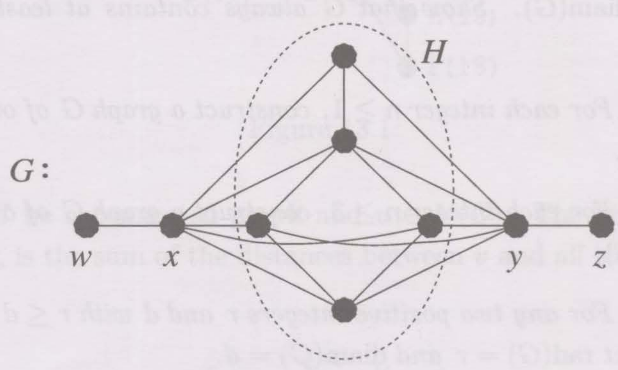

Figure 12.5

Suppose we want to find a suitable location for a police station in a new town. What would our considerations be? Certainly, one consideration would be to minimize the longest time taken to reach any other place in town. Using a graph to model the new town, this requirement can be viewed as minimizing the eccentricity of the vertex in which we will build the police station. Thus, we will build the police station on a centre of the graph. If, however, we wish to build a shopping mall, our considerations would be different. One consideration could be to be as near as possible to as many people as possible. Using a graph to model the new town, this

Graphs and Their Applications

requirement can be viewed in one way as minimizing the total distance of all vertices to the vertex in which we will build the mall. In some way, this vertex is a 'centre' of the graph, though not exactly in the way as defined above. These examples show that questions of centrality with regard to the location of facilities can be studied using graphs as models.

Exercise 12.1. *For each of the graphs G shown in Figure 12.6, find*
(i) the eccentricity of each vertex in G;
(ii) $\mathrm{rad}(G)$;
(iii) $\mathrm{diam}(G)$;
(iv) $C(G)$.

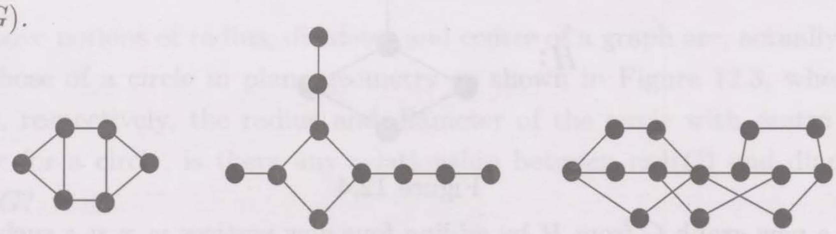

Figure 12.6

Exercise 12.2. *Show that the result (12.4) holds for any connected multigraph G.*

Exercise 12.3. *A vertex v in a connected multigraph G is called a* **peripheral** *vertex if $e(v) = \mathrm{diam}(G)$. Show that G always contains at least two peripheral vertices.*

Exercise 12.4. *For each integer $n \geq 1$, construct a graph G of order n such that* $\mathrm{rad}(G) = \mathrm{diam}(G)$.

Exercise 12.5. *For each integer $n \geq 3$, construct a graph G of order n such that* $\mathrm{diam}(G) = 2\,\mathrm{rad}(G)$.

Exercise 12.6. *For any two positive integers r and d with $r \leq d \leq 2r$, construct a graph G such that $\mathrm{rad}(G) = r$ and $\mathrm{diam}(G) = d$.*

Exercise 12.7. *Let xy be an edge in a connected multigraph. Show that*

$$-1 \leq e(x) - e(y) \leq 1.$$

Exercise 12.8. *Let G be a connected multigraph with $\mathrm{rad}(G) = r$ and $\mathrm{diam}(G) = d$. Show that for each integer k with $r \leq k \leq d$, there exists a vertex w in G such that $e(w) = k$.*

Exercise 12.9. *Let G be a graph of order $n \geq 1$ such that $d(x) \geq \frac{n-1}{2}$ for each vertex x in G, where $d(x)$ is the degree of x in G. Must G be connected? What can be said about $\mathrm{diam}(G)$?*

Graphs and Their Applications

Exercise 12.10. *Let G be a connected graph of order $n \geq 3$ and let $\Delta(G)$ denote the maximum of the vertex degrees in G, i.e., $\Delta(G) = \max\{d(x) : x \in V(G)\}$.*
(i) If $\Delta(G) = n-1$, find $\mathrm{diam}(G)$.
(ii) If $\Delta(G) = n-2$, what can be said about $\mathrm{diam}(G)$*?*
(iii) If $\Delta(G) = n-2$ and $\mathrm{diam}(G) = 2$, show that $|E(G)| \geq 2(n-2)$.

13. The Sum of Distances in a Graph and the Wiener Index

At the end of Section 12, we mentioned another kind of 'centre', which was a vertex whose sum of distances to all other vertices is a minimum. Consider the graph G of Figure 13.1. It is clear that vertex c is the only centre by the definition of having the smallest eccentricity ($e(c) = 2$). Now the sum of the distances between c and all the other vertices is $2(1) + 6(2) = 14$. However, the sum of the distances between vertex d and all the other vertices is $5(1) + 2(2) + 1(3) = 12$, which makes d more suitable as a 'centre' for other purposes (such as for a shopping mall).

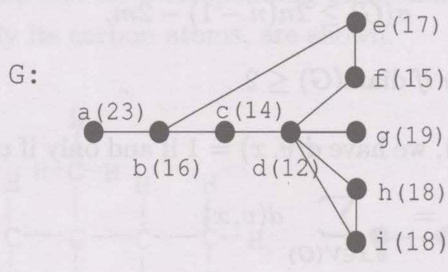

Figure 13.1

In general, let G be a connected graph and $v \in V(G)$. The **transmission** of v, denoted by $\sigma(v)$, is the sum of the distances between v and all the other vertices in G. That is,

$$\sigma(v) = \sum_{x \in V(G)} d(v, x). \tag{13.1}$$

As an example, the transmissions of the nine vertices in the graph G of Figure 13.1 are shown in parentheses.
Of greater interest to researchers is the notion of the transmission of the graph itself, which is defined to be the sum of transmissions of all the vertices in the graph. Thus, for a graph G, the **transmission** of G is the value

$$\sigma(G) = \sum_{v \in V(G)} \sigma(v). \tag{13.2}$$

Graphs and Their Applications

Observe that

$$\begin{aligned}
\sigma(G) &= \sum_{v\in V(G)} \sigma(v) && \text{(by (13.2))}\\
&= \sum_{v\in V(G)}\sum_{x\in V(G)} d(v,x) && \text{(by (13.1))}\\
\text{(13.3)} \qquad &= \sum_{v,x\in V(G)} d(v,x),
\end{aligned}$$

where the last sum is taken over all **ordered** pairs of vertices v and x in G. Thus, the transmission of G is the sum of the distances between all ordered pairs of vertices in G.

The following two theorems give a lower bound (Entringer, Jackson and Snyder [7]) and an upper bound (Entringer, Jackson and Snyder [7], Doyle and Graver [6], and Lovasz [[12], p.276]) for $\sigma(G)$.

Theorem 13.1. *If a connected graph G has n vertices and m edges, then*

$$\sigma(G) \geq 2n(n-1) - 2m.$$

Equality occurs if and only if $\mathrm{diam}(G) \leq 2$.

Proof. For any $v, x \in V(G)$, we have $d(v,x) = 1$ if and only if $vx \in E(G)$. By (13.3),

$$\begin{aligned}
\sigma(G) &= \sum_{v,x\in V(G)} d(v,x)\\
&= \sum_{\substack{v,x\in V(G)\\ vx\in E(G)}} d(v,x) + \sum_{\substack{v,x\in V(G)\\ vx\notin E(G)}} d(v,x)\\
&\geq \sum_{\substack{v,x\in V(G)\\ vx\in E(G)}} 1 + \sum_{\substack{v,x\in V(G)\\ vx\notin E(G)}} 2\\
&= 2m + 2\times 2\times\left(\binom{n}{2} - m\right)\\
&= 2n(n-1) - 2m,
\end{aligned}$$

where the equality holds if and only if $d(v,x) = 2$ for all $v, x \in V(G)$ with $vx \notin E(G)$, i.e., $\mathrm{diam}(G) \leq 2$. □

Theorem 13.2. *If G is a connected graph with n vertices, then*

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

Moreover, this bound is achieved if and only if G is a path.

Graphs and Their Applications

We omit the proof in this article but shall revisit this theorem when we discuss the notion of trees in a future article.

For more comprehensive information on the theoretical results pertaining to 'distances', the reader may find the book [1] useful.

Transmission numbers have been investigated by several authors under different names. The term 'transmission' is due to Christofides [3] while Harary [8], motivated by certain sociometric problems, used the term 'status'.

In 1947, Harold Wiener [13] introduced the quantity W, eventually named the Wiener index, in his paper entitled *Structural determination of paraffin boiling points.* To explain the different boiling points of various saturated paraffins, it was reasoned that compounds with a less 'compact' molecular structure would boil at higher temperatures since they were more likely to be entangled during motion. Thus, W was conceived as the sum of distances between all pairs of vertices in the molecular graph of an alkane, with the aim of providing a measure of the compactness of the respective hydrocarbon molecule. (The Wiener index is indeed the transmission number of a graph!) In Figure 13.2, the structural formula of 3, 4-dimethylhexane and its corresponding molecular graph, which is actually a representation of only its carbon atoms, are shown.

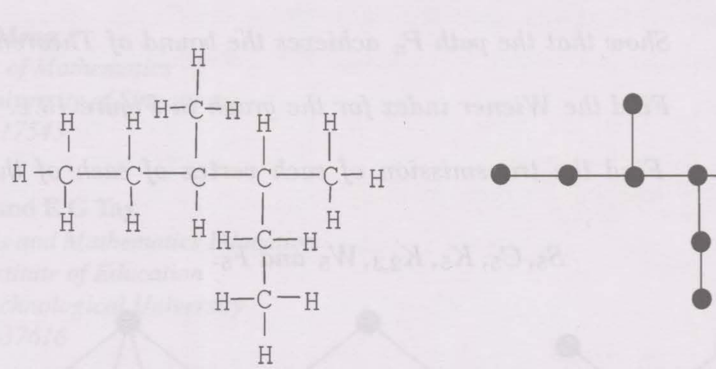

Figure 13.2

Wiener showed that there is an excellent correlation between W and the boiling points. He proposed a formula that closely approximates the boiling point (B) of an alkane as

$$B = \alpha W + \beta P + \gamma,$$

where α, β and γ are empirical constants and P, the polarity number, is the number of pairs of vertices whose distance is equal to 3.

Weiner subsequently showed strong correlations between W and other physico-chemical properties of organic compounds, such as molar volumes, refractive indices, heats of isomerisation and heats of vaporization of alkanes.

Graphs and Their Applications

More recent work on the Wiener index has demonstrated that it measures the area of the surface of the respective molecule and thus reflects its compactness. Physical and chemical properties of organic substances which are expected to depend on the area of the molecular surface are thus further expected to correlate well with W and so it has been reported for heats of formation, vaporization and atomization, density, boiling point, critical pressure, refractive index, surface tension and viscosity. Lest it be concluded that W correlates with everything, results with melting points have not been satisfactory.

Attempts have been made to use the Wiener index in designing new drugs. Correlations have been established between W and cytostatic and antihistaminic activities of certain pharmacologically interesting compounds. More recently, the Wiener index was employed in the study of the n-octanol/water partition coefficient, an indicator of transport characteristics and interaction between receptor and bioactive molecule. This coefficient is a physico-chemical parameter of significant importance for the forecasting of pharmacological activity of many compounds.

For a more detailed survey of the Wiener index and its applications in Chemistry, the reader is referred to the articles by Diudea and Gutman [4], and Dobrynin, Entringer and Gutman [5].

Exercise 13.1. *Show that the path P_n achieves the bound of Theorem 13.2.*

Exercise 13.2. *Find the Wiener index for the graph in Figure 13.2.*

Exercise 13.3. *Find the transmission of each vertex of each of the following graphs*

$$S_5, C_5, K_5, K_{2,3}, W_5 \text{ and } F_5.$$

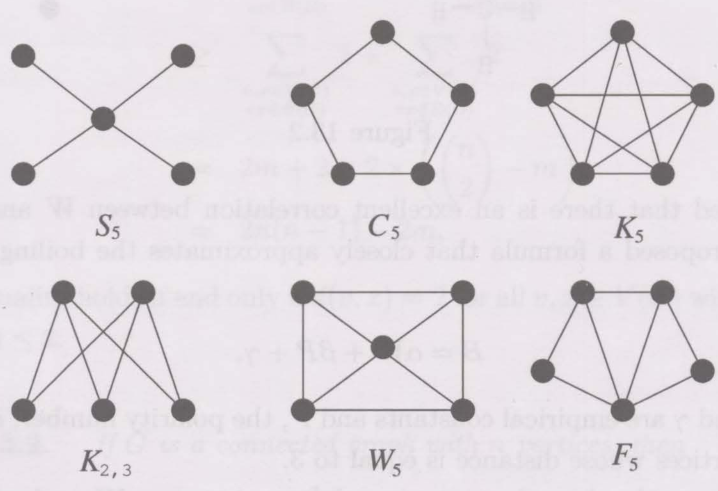

Figure 13.3

Graphs and Their Applications

References

[1] F. Buckley and F. Harary, *Distances in Graphs*, Addison-Wesley, Reading, MA (1990).
[2] F. Buckley, Z. Miller and P.J. Slater, On graphs containing a given graph as center, *J. Graph Theory* **5** (1981) 427-434.
[3] N. Christofides, *Graph Theory: An algorithmic approach*, Academic, London (1975).
[4] M.V. Diudea and I. Gutman, Wiener-type topological indices, *Croatica Chemica Acta* **71** (1) (1998) 21-51.
[5] A.A. Dobrynin, R. Entringer and I. Gutman, Wiener index of trees: Theory and applications, *Acta Appl. Math.* **66**(3) (2001) 211-224.
[6] J.K. Doyle and J.E. Graver, Mean distance in a graph, *Discrete Math.* **17** (1977) 147-154.
[7] R.C. Entringer, D.E. Jackson and D.A. Snyder, Distance in graphs, *Czech. Math. J.* **26** (101) (1976) 283-296.
[8] F. Harary, Status and contrastatus, *Sociometry* **22** (1959) 23-43.
[9] K.M. Koh, Graphs and their applications (2), *Mathematical Medley* **30** (1) (2003) 11-22.
[10] K.M. Koh, F.M. Dong and E.G. Tay, Graphs and their applications (3), *Mathematical Medley* **30** (2) (2003) 102-116.
[11] G.N. Kopylov and E.A. Timofeev, Centers and radii of graphs, *Usp. Mat. Nauk* **32** (1977) 226.
[12] L. Lovász, *Combinatorial Problems and Exercises*, Akademiai Kiado, Budapest (1979).
[13] H. Wiener, *J. Am. Chem. Soc.* **69** (1947), 17-20.

Koh Khee Meng
Department of Mathematics
National University of Singapore
Singapore 117543

F M Dong and E G Tay
Mathematics and Mathematics Education
National Institute of Education
Nanyang Technological University
Singapore 637616