A Note on the Irregularity of Graphs

F. K. Bell
Department of Mathematics
University of Stirling
Stirling FK9 4LA, Scotland

Submitted by Richard A. Brualdi

ABSTRACT

Denote by $\lambda_1(G)$ the largest eigenvalue of a real $(0,1)$-adjacency matrix of a graph G, and by $\bar{d}(G)$ the mean degree of G. Collatz and Sinogowitz proposed $\lambda_1(G)-\bar{d}(G)$ as a measure of irregularity of G. A second such measure is the variance of the vertex degrees of G. The most irregular graphs according to these measures are determined for certain classes of graphs, and the two measures are shown to be incompatible for some pairs of graphs.

We consider only finite, undirected graphs without loops or multiple edges. The *index* $\lambda_1(G)$ of a graph G is the largest eigenvalue of a real $(0,1)$-adjacency matrix of G. If $\bar{d}(G)$ denotes the mean of the vertex degrees of G, then $\lambda_1(G)\geqslant \bar{d}(G)$, with equality if and only if G is regular. (See e.g. [8, Theorem 3.8].) This result was proved in the fundamental paper [6] by Collatz and Sinogowitz, who proposed $\epsilon(G)=\lambda_1(G)-\bar{d}(G)$ as a measure of the "irregularity" of G, and asked which n-vertex connected graph (for given n) is most irregular according to this measure. An obvious alternative measure of irregularity is provided by $v(G)$, the variance of the vertex degrees of G. [If the degrees are $d_1,\dots,d_n$, then $v(G)=(1/n)\Sigma_{i=1}^{n}d_i^2-(1/n^2)(\Sigma_{i=1}^{n}d_i)^2$: note that G has the same variance as its complement $\bar{G}$.] In this note we compare ϵ and v as measures of irregularity, and we consider the problem of maximizing ϵ and v over the set $\mathscr{G}(n)$ of all n-vertex graphs and over its subset $\mathscr{H}(n)$ of connected n-vertex graphs.

In [6], $\max\{\epsilon(G):G\in\mathscr{H}(n)\}$ is denoted by $S(n)$. The authors observe that $S(n)=\sqrt{n-1}-2+2/n$ when $n\leqslant 5$, this bound being attained uniquely by the star $K_{1,n-1}$, and they ask whether this holds also when $n>5$. (This

655 Avenue of the Americas, New York, NY 10010 0024-3795/92/$3.50

problem appears also in [8, p. 266].) Cvetković and Rowlinson [10] show that the maximal value of ϵ is not always attained by a star. To demonstrate this they consider an example with $n=25$, but in fact a similar counterexample exists with n as small as 8. Let G be the star $K_{1,7}$, and G' the graph obtained from the complete graph K_5 by adding three pendant edges at a single vertex. Then $G, G' \in \mathscr{H}(8)$, and $\epsilon(G) \approx 0.896$, $\epsilon(G') \approx 0.912$. (We note that G and G' have the same variance 3.9375.) Rowlinson [11] considers certain maximal outerplanar graphs, and notes an example of two such graphs G_1, G_2 for which $\epsilon(G_1) > \epsilon(G_2)$, $v(G_1) = v(G_2)$. By deleting the edges of the unique Hamiltonian cycles of G_1, G_2 respectively, one clearly obtains graphs G_1^*, G_2^* with the same variance as G_1 and G_2, but it turns out that $\epsilon(G_1^*) < \epsilon(G_2^*)$. Moreover, if isolated vertices are removed from G_1^*, G_2^*, one obtains connected graphs G_1^{**}, G_2^{**} such that $\epsilon(G_1^{**}) < \epsilon(G_2^{**})$ and $v(G_1^{**}) = v(G_2^{**})$. These examples might suggest that as measures of irregularity v is coarser than ϵ; however, for some pairs of graphs the two measures are in fact incompatible. To illustrate this phenomenon, let $H = K_2 \triangledown \overline{K}_6$. (Here $\triangledown$ denotes the join, or complete product, of two graphs.) We find that $\epsilon(H) = 0.75$, $v(H) = 4.6875$, and therefore, comparing H with the graph G' defined above, we have $\epsilon(G') > \epsilon(H)$, $v(G') < v(H)$. This incompatibility is particularly striking because G' and H have the same number of vertices (8) and the same number of edges (13), and are both connected.

For given n and e, let $\mathscr{G}(n,e)$ be the set of all graphs with n vertices and e edges, and let $\mathscr{H}(n,e)$ be the set of all connected graphs in $\mathscr{G}(n,e)$. We use the terminology and notation of Ahlswede and Katona [2] in defining the *quasicomplete* graph C_n^e and the *quasistar* S_n^e, both of which belong to $\mathscr{G}(n,e)$. Let the n vertices be labelled $1, 2, \ldots, n$. For C_n^e, define integers d, t by

$$e = \binom{d}{2} + t, \qquad 0 \leqslant t < d;$$

connect together each pair of the vertices $1, 2, \ldots, d$, and connect also vertex $d+1$ with each of $1, 2, \ldots, t$. The quasistar S_n^e can be defined as the complement of $C_n^{\binom{n}{2}-e}$. The graph C_n^e therefore has a stepwise adjacency matrix (see [12]) in which the e ones above the principal diagonal occupy the minimum number of columns, while S_n^e has one in which they occupy the minimum number of rows. (Thus, for example, $S_8^{13} = K_2 \triangledown \overline{K}_6$.)

Proposition 1. *Let n and e be given, with*

$$e \leqslant \binom{n}{2}.$$

Then:

(i) $\max\{\epsilon(G): G \in \mathscr{G}(n,e)\}$ *is attained uniquely by* C_n^e;
(ii) $\max\{v(G): G \in \mathscr{G}(n,e)\}$ *is attained by one of* C_n^e *and* S_n^e.

In (ii), *the maximum is attained by* C_n^e *if*

$$e > \frac{1}{2}\binom{n}{2} + \frac{n}{2},$$

and by S_n^e *if*

$$e < \frac{1}{2}\binom{n}{2} - \frac{n}{2}.$$

Part (i) of this proposition follows directly from [12], because when n and e are fixed, the mean degree is fixed, so maximizing $\epsilon(G)$ is equivalent to maximizing $\lambda_1(G)$. For part (ii), we note that since n and e are fixed, maximizing $v(G)$ is equivalent to maximizing $\sum_{i=1}^n d_i^2$ and hence to maximizing

$$\sum_{i=1}^{n} \binom{d_i}{2},$$

the number of walks of length 2 in G. Accordingly (ii) follows from a result of Ahlswede and Katona [2, Theorem 2]. The same result was obtained independently by Brualdi and Solheid [5, Theorem 3.3]. Their paper is concerned with maximizing the sum of the entries of A^2, which is just $\sum_{i=1}^n d_i^2$. The identification of which of C_n^e and S_n^e has the larger variance, for arbitrary n and e, seems to be difficult; the partial result stated above is proved at some length in [2, Theorem 3].

We remark that in (ii) the maximum variance may be attained also by graphs other than C_n^e and S_n^e. As an illustration of this, suppose that the maximum is attained by C_n^e, and determine d and t by

$$e = \binom{d}{2} + t, \qquad 0 \leqslant t < d.$$

Suppose further that $1 < t \leqslant n - d$. If we move the 1's in column (and row) $d+1$ of the adjacency matrix of C_n^e to row (and column) 1 (positions $d+1, \dots, d+t$), then we obtain the adjacency matrix of a graph with the same variance as C_n^e. (Compare the concluding remarks of [1].)

In the next result we keep n fixed but allow e to vary.

PROPOSITION 2. *Given n, write*

$$m=\left[\tfrac{1}{2}(n+1)\right],\qquad N=\binom{m}{2},\qquad N'=\binom{m+1}{2},$$

$$r=\left[\tfrac{1}{4}(3n+2)\right],\qquad f=\binom{r}{2}.$$

Then:

(i) *One has*

$$\max\{\epsilon(G):G\in\mathscr{G}(n)\}=\begin{cases}\tfrac{1}{4}n-\tfrac{1}{2} & (n\ even),\\ \tfrac{1}{4}n-\tfrac{1}{2}+\tfrac{1}{4n} & (n\ odd).\end{cases}$$

This maximum is attained uniquely by C_n^N *if n is odd, and by* C_n^N *and* $C_n^{N'}$ *(only) if n is even.*

(ii) *One has*

$$\max\{v(G):G\in\mathscr{G}(n)\}=\frac{r}{n^2}(r-1)^2(n-r),$$

and this maximum is attained by C_n^f.

Before giving the proof of Proposition 2 we derive an upper bound for $\lambda_1(C_n^e)$. We write

$$e=\binom{d}{2}+t,$$

and it is convenient here to allow the value $t=d$; this gives the same value for e as replacing d by $d+1$ and setting $t=0$.

LEMMA 3. *If*

$$e=\binom{d}{2}+t\qquad(0\leqslant t\leqslant d)$$

then

$$\lambda_1(C_n^e)\leqslant d-1+\frac{t}{d},$$

with equality if and only if $t=0$ *or* $t=d$.

Proof. Write $\beta = t/d$, so that $0 \leqslant \beta \leqslant 1$, and define a polynomial F by

$$F(x) = x^3 + (2d-1)x^2 + (d^2 - d - \beta d)x - \beta^2 d^2.$$

From [12], $\lambda_1(C_n^e) = d - 1 + v$, where $v \in [0,1]$ is the largest root of F. Now

$$\begin{aligned} F(\beta) &= \beta^3 + (d - d^2 - 1)\beta^2 + (d^2 - d)\beta \\ &= \beta(1-\beta)[d(d-1) - \beta] \geqslant 0, \end{aligned}$$

and since $F(x)$ is an increasing function of x for $x > 0$, we deduce that $v \leqslant \beta$, as required. For equality we require $\beta = 0$ or 1, i.e. $t = 0$ or d. ∎

Proof of Proposition 2. By Proposition 1(i), $\max\{\epsilon(G) : G \in \mathscr{G}(n)\}$ is attained solely by C_n^e, for some value or values of e. Write

$$e = \binom{d}{2} + t \qquad (0 \leqslant t \leqslant d):$$

then, using Lemma 3,

$$\epsilon(C_n^e) = \lambda_1(C_n^e) - \frac{2e}{n} \leqslant d - 1 - \frac{d(d-1)}{n} + t\left(\frac{1}{d} - \frac{2}{n}\right).$$

If $d \geqslant \frac{1}{2}n$, then

$$\epsilon(C_n^e) \leqslant d - 1 - \frac{d(d-1)}{n} = \epsilon\left(C_n^{\binom{d}{2}}\right),$$

with equality if and only if $t = 0$ or $t = d = \frac{1}{2}n$. Similarly, if $d \leqslant \frac{1}{2}n$, then

$$\epsilon(C_n^e) \leqslant d - 1 - \frac{d(d-1)}{n} + d\left(\frac{1}{d} - \frac{2}{n}\right) = d - \frac{d(d+1)}{n} = \epsilon\left(C_n^{\binom{d+1}{2}}\right),$$

with equality if and only if $t = d$ or $t = 0$, $d = \frac{1}{2}n$. We therefore have to maximize

$$\epsilon\left(C_n^{\binom{d}{2}}\right) = \frac{(d-1)(n-d)}{n} \qquad \text{for} \quad 0 < d < n,$$

and (i) follows.

For (ii), we note first that since S_n^e is the complement of $C_n^{\binom{n}{2}-e}$, it follows from Proposition 1(ii) that $\max\{v(G): G \in \mathscr{G}(n)\}$ is attained by a graph C_n^e for some e. A simple calculation shows that with

$$e = \binom{d}{2} + t \qquad (0 \leqslant t \leqslant d)$$

we have

$$v(C_n^e) = \frac{1}{n^2}\left\{(n-4)t^2 + [n(2d-1) - 4d(d-1)]t + d(d-1)^2(n-d)\right\}.$$

For fixed d and fixed $n > 4$ this quadratic function of t takes its maximum value either when $t = 0$ or when $t = d$. We therefore have to maximize

$$v\left(C_n^{\binom{d}{2}}\right) = \frac{d}{n^2}(d-1)^2(n-d) \qquad (1 \leqslant d \leqslant n),$$

and it is easily shown that the maximum is given by $d = [\frac{1}{4}(3n+2)]$. The cases in which $n \leqslant 4$ are also simple to deal with. ■

Since a quasistar S_n^e is connected whenever $e \geqslant n-1$, an immediate consequence of part (ii) of Proposition 2 is the following:

Corollary. $\max\{v(G): G \in \mathscr{H}(n)\} = (r/n^2)(r-1)^2(n-r)$, *attained by* $S_n^{\binom{n}{2}-\binom{r}{2}}$.

The analogue of Proposition 1(ii) for the variance of connected graphs (when n and e are both fixed) follows from [5, Theorem 4.2] and Proposition 1(ii). In stating the result it is convenient to write $k = e - n$, as in [10, 3], and to denote the graph $K_1 \triangledown C_{n-1}^{k+1}$ by $G_{n,k}$.

PROPOSITION 4. *If n and e* ($=n-k$) *are such that*

$$n-1 \leqslant e \leqslant \binom{n}{2},$$

then $\max\{v(G): G \in \mathscr{H}(n,e)\}$ *is attained either by* $G_{n,k}$ *or by* S_n^e. *The maximum is attained by* $G_{n,k}$ *if*

$$e > \frac{1}{2}\binom{n}{2} + n - 1,$$

and by S_n^e *if*

$$e < \frac{1}{2}\binom{n}{2}.$$

To identity the graph(s) G in $\mathscr{H}(n,e)$ for which $\epsilon(G)$ is a maximum, we need to know which graphs in this set have maximal index. This is known when

$$\binom{n-1}{2} < e \leqslant \binom{n}{2},$$

because C_n^e, the graph of maximal index in $\mathscr{G}(n,e)$, is then connected. Aside from this, it is known only for certain values of k ($=e-n$): namely

$$k \in \{-1,0,1,2\}, \quad \text{and} \quad k = \binom{d-1}{2} - 1 \quad \text{for each } d \geqslant 5.$$

For $k \in \{-1,0,1,2\}$, the unique graph of maximal index is found in [4] to be $G_{n,k}$. [We remark that $G_{n,k}$ is also the graph of maximal index in the case $\binom{n-1}{2} < e \leqslant \binom{n}{2}$, because for such e it coincides with C_n^e.] For $k \leqslant n-3$, let $H_{n,k}$ be the graph obtained from the star $K_{1,n-1}$ by joining a vertex of degree 1 to $k+1$ other vertices of degree 1. It is shown in [3] that when

$$k = \binom{d-1}{2} - 1 \qquad (d \geqslant 5),$$

a graph of maximal index in $\mathscr{H}(n,e)$ is $G_{n,k}$ or $H_{n,k}$ according as $n \leqslant g(d)$ or $n \geqslant g(d)$, where $g(d) = \frac{1}{2}d(d+5) + 7 + 32/(d-4) + 16/(d-4)^2$. We therefore have:

PROPOSITION 5. *Let*

$$5 \leqslant d \leqslant n, \qquad k = \binom{d-1}{2} - 1, \qquad e = n + k.$$

Then $\max\{\epsilon(G) : G \in \mathscr{H}(n,e)\}$ *is attained uniquely by* $G_{n,k}$ *if* $n < g(d)$ *and uniquely by* $H_{n,k}$ *if* $n > g(d)$. *When* $n = g(d)$, *we have* $\epsilon(G_{n,k}) = \epsilon(H_{n,k})$, *and these are the only graphs* G *for which* $\epsilon(G)$ *is maximal.*

NOTE. For $n = g(d)$ we require $g(d)$ to be an integer, so that $d = 5$, 6, or 8.

It is interesting to compare Propositions 4 and 5. The maximal graphs obtained in both cases are $G_{n,k}$ and S_n^e, though in Proposition 5 the nonzero entries of the adjacency matrix of S_n^e above the diagonal are restricted to the first two rows. The conditions determining which of the two graphs is maximal are quite different. For example, $v(S_n^e)$ is maximal whenever $e < \frac{1}{4}n(n-1)$, whereas for $\epsilon(S_n^e)$ to be maximal in Proposition 5 it is necessary that $e < 2n$.

A straightforward calculation shows that when

$$k = \binom{d-1}{2} - 1 \qquad (d \geqslant 5),$$

we have

$$v(H_{n,k}) - v(G_{n,k}) = \frac{1}{4n}(d-1)(d-2)(d-3)(d-4).$$

Thus $H_{n,k}$ has greater variance then $G_{n,k}$ for each k. We can deduce from Proposition 5 an infinite family of pairs of connected graphs for which ϵ and v are incompatible as measures of irregularity: if

$$d \geqslant 5 \quad \text{and} \quad 2 + \binom{d-1}{2} \leqslant n < g(d)$$

then

$$\epsilon(H_{n,k}) < \epsilon(G_{n,k}), \qquad \text{while} \quad v(H_{n,k}) > v(G_{n,k}).$$

It seems likely that a result similar to Proposition 5 holds for arbitrary $k \geqslant 3$, but all that has so far been proved in this direction is the following:

PROPOSITION 6. *For fixed $k \geqslant 3$, and n sufficiently large (depending on k), $\max\{\epsilon(G): G \in \mathscr{H}(n, n+k)\}$ is attained uniquely by $H_{n,k}$.*

This is an immediate consequence of the main result of [10].

We now return to the function

$$S(n) = \max\{\epsilon(G): G \in \mathscr{H}(n)\}$$

defined by Collatz and Sinogowitz [6], and establish the following inequalities.

PROPOSITION 7. *For any $n \geqslant 3$,*

$$\frac{1}{4}n - \frac{3}{2} + \frac{2}{n} < S(n) < \frac{1}{4}n - 1 + \frac{1}{n}.$$

Proof. For the lower bound, consider the graphs $G_{n,k}$ where

$$k = \binom{d-1}{2} - 1, \qquad 3 \leqslant d < n.$$

If $e = n + k$ then

$$\epsilon(G_{n,k}) = \lambda_1(G_{n,k}) - 2 - \frac{d(d-3)}{n}.$$

Since $G_{n,k}$ has K_d as a proper subgraph, we have $\lambda_1(G_{n,k}) > d - 1$, so that

$$\epsilon(G_{n,k}) > \frac{(d-3)(n-d)}{n}.$$

If n is odd, we can take $d = (n+3)/2$, giving $\epsilon(G_{n,k}) > \frac{1}{4}n - \frac{3}{2} + 9/4n$; while if n is even, we can take $d = (n+2)/2$, giving $\epsilon(G_{n,k}) > \frac{1}{4}n - \frac{3}{2} + 2/n$.

It was proved by Yuan [13] that the index of a graph $G \in \mathscr{H}(n, e)$ satisfies $\lambda_1(G) \leqslant \sqrt{2e - n + 1}$, with equality if and only if G is complete or a star. Regarding e as a continuous variable, we find that the function

$\sqrt{2e-n+1}-2e/n$ takes its maximum value when $e=\frac{1}{8}(n^2+4n-4)$, and this gives the stated upper bound. ∎

We conclude with the remark that the maximal value $S(n)$ is attained by the star $K_{1,n-1}$ if and only if $n \leqslant 7$. That the star is maximal when $n \leqslant 5$ has already been mentioned, and it may be checked when $n = 6, 7$ by using the tables of graph spectra in [9, 7]. For $n = 8$, we have $\epsilon(G_{8,5}) > \epsilon(K_{1,7})$; and for $n = 9$, $\epsilon(G_{9,9}) > \epsilon(K_{1,8})$. Finally, if $n \geqslant 10$ the result follows from Proposition 7 above, because we then have $\sqrt{n-1}-2+2/n \leqslant \frac{1}{4}n-\frac{3}{2}+2/n$.

The author is indebted to Dr. P. Rowlinson for helpful comments and suggestions.

REFERENCES

1 R. Aharoni, A problem in rearrangements of (0,1) matrices, *Discrete Math.* 30:191–201 (1980).
2 R. Ahlswede and G. O. H. Katona, Graphs with maximal number of adjacent pairs of edges, *Acta Math. Acad. Sci. Hungar.* 32:97–120 (1978).
3 F. K. Bell, On the maximal index of connected graphs, *Linear Algebra Appl.*, 144:135–151 (1991).
4 R. A. Brualdi and E. S. Solheid, On the spectral radius of connected graphs, *Publ. Inst. Math. (Beograd)* 39(53):45–54 (1986).
5 R. A. Brualdi and E. S. Solheid, Some extremal problems concerning the square of a (0,1)-matrix, *Linear and Multilinear Algebra* 22:57–73 (1987).
6 L. Collatz and U. Sinogowitz, Spektren endlicher Grafen, *Abh. Math. Sem. Univ. Hamburg* 21:63–77 (1957).
7 D. Cvetković, M. Doob, I. Gutman, and A. Torgašev, *Recent Results in the Theory of Graph Spectra*, North-Holland, Amsterdam, 1988.
8 D. Cvetković, M. Doob, and H. Sachs, *Spectra of Graphs*, Academic, New York, 1980.
9 D. Cvetković and M. Petrić, A table of connected graphs on six vertices, *Discrete Math.* 50:37–49 (1984).
10 D. Cvetković and P. Rowlinson, On connected graphs with maximal index, *Publ. Inst. Math. (Beograd)* 44(58):29–34 (1988).
11 P. Rowlinson, On the index of certain outerplanar graphs, to appear.
12 P. Rowlinson, On the maximal index of graphs with a prescribed number of edges, *Linear Algebra Appl.* 110:43–53(1988).
13 H. Yuan, A bound on the spectral radius of graphs, *Linear Algebra Appl.* 108:135–139 (1988).

Received 12 August 1990; final manuscript accepted 14 November 1990