PACIFIC JOURNAL OF MATHEMATICS
Vol. 58, No. 1, 1975

COVERING THE VERTICES OF A GRAPH BY VERTEX-DISJOINT PATHS

SHAHBAZ NOORVASH

Define the path-covering number $\mu(G)$ of a finite graph G to be the minimum number of vertex-disjoint paths required to cover the vertices of G. Let $g(n,k)$ be the minimum integer so that every graph, G, with n vertices and at least $g(n,k)$ edges has $\mu(G) \leqq k$. A relationship between $\mu(G)$ and the degree sequence for a graph G is found; this is used to show that

$$\tfrac{1}{2}(n-k)(n-k-1)+1 \leqq g(n,k) \leqq \tfrac{1}{2}(n-1)(n-k-1)+1$$

A further extremal problem is solved.

1. Introduction. A *graph* G is a finite collection $\mathscr{V}(G)$ of vertices (or points) some pairs of which are joined by a single edge; the collection of edges is denoted by $\mathscr{E}(G)$. H is a *subgraph* of G if $\mathscr{V}(H) \subseteq \mathscr{V}(G)$ and $\mathscr{E}(H) \subseteq \mathscr{E}(G)$. If H and K are two vertex-disjoint graphs, $H \cup K$ is the graph with $\mathscr{V}(H \cup K) = \mathscr{V}(H) \cup \mathscr{V}(K)$ and $\mathscr{E}(H \cup K) = \mathscr{E}(H) \cup \mathscr{E}(K)$; $H + K$ is $H \cup K$ together with all $|\mathscr{V}(H)|\ |\mathscr{V}(K)|$ possible choices of edges joining a vertex of H to a vertex of K. $\bar{G}$ denotes the complement of G; Γ_n denotes the complete graph with n vertices and $\Gamma_{m,n}$ denotes the complete bipartite graph, $\bar{\Gamma}_m + \bar{\Gamma}_n$.

Let G be a graph. A path of length n in G is an ordered sequence $P = \langle a_1, a_2, \cdots, a_n \rangle$ of distinct points, where if $n \geqq 2$, a_i is adjacent to a_{i+1} for $1 \leqq i \leqq n-1$. $\langle a_1, a_2, \cdots, a_n \rangle$ is the same path as $\langle a_n, a_{n-1}, \cdots, a_1 \rangle$. If P and Q are paths, by $P * Q$ we shall mean that one end-point, a of P, is adjacent to one end-point, b of Q, and that $P * Q$ is formed by joining a to b. More specifically we may write $Pa * bQ$ or $P * bQ$ or $Pa * Q$ to specify, in varying degrees, which end-point of P is joined to which end-point of Q. Also, $\langle a_1, a_2, \cdots, a_n \rangle * \langle b_1, b_2, \cdots, b_m \rangle = \langle a_1, a_2, \cdots, a_n, b_1, b_2, \cdots, b_m \rangle$ where a_n must be adjacent to b_1. A *Hamilton-path* is a path of length $|\mathscr{V}(G)|$. A *path-cover* of G is a collection, $\mathscr{S}$, of vertex-disjoint paths such that every vertex of G lies on some path in $\mathscr{S}$. The *path-covering number*, denoted by $\mu(G)$, of G is defined by:

$$\mu(G) = \text{Min}\{|\mathscr{S}|: \mathscr{S} \text{ is a path-cover of } G\}.$$

A *minimal path-cover* (M.P.C.) of G is a path-cover, $\mathcal{S}$ of G, with $|\mathcal{S}| = \mu(G)$.

We note that $\mu(G)$ is an invariant of G and remark that a graph, G, has a Hamilton-path if and only if $\mu(G) = 1$. It has been shown by Nash-Williams [1] and others that the problem of classifying all Hamiltonian graphs is equivalent to that of classifying all graphs which have a Hamilton-path. Thus a classification of all graphs with $\mu(G) = k$ ($k = 1, 2, 3, \cdots$) would also solve the Hamiltonian problem as a special case.

Historically, O, Ore [3] first introduced the graphical invariant μ. In [2] some elementary properties of μ are derived. In §2 we generalize a result of O. Ore (Theorem 2.1 in [3]) and in §3 we consider two extremal problems involving μ.

2. Valency considerations. In this section we derive a connection between the path-covering number and the degree sequence of a graph. We begin with some definitions:

DEFINITION 2.1. Let A be a finite set and f a real-valued function defined on the collection of subsets of A. For $B \subseteq A$ and for any integer i with $1 \leqq i \leqq |B|$, define the function S_i by:

$$S_i(f, B) = \sum_{\substack{C \subseteq B \\ |C| = i}} f(C).$$

DEFINITION 2.2. If G is a graph, $B \subseteq \mathcal{V}(G)$, and either $H \subseteq \mathcal{V}(G)$ or H is a subgraph of G, then define the generalized valence function, ρ, by

$\rho_H(B) =$ the number of vertices of H which are adjacent to every member of B.

If x is a vertex of G, then we write $\rho(x)$ for $\rho_G(\{x\})$.

DEFINITION 2.3. Let G be a graph and $X \subseteq \mathcal{V}(G)$ with $|X| = k \geqq 2$. Define:

$$D(G, X) = \frac{1}{k} S_1(\rho_G, X) - \sum_{i=1}^{k} (-1)^i \left(\frac{k-i}{k}\right) S_i(\rho_G, X).$$

The following lemma is easily verified:

LEMMA 2.4. *If* $X = \{x_1, x_2, \cdots, x_k\}$, *and* $1 \leqq m \leqq k-1$, *then*

$$\sum_{i=1}^{k} S_m(f, X - \{x_i\}) = (k-m) S_m(f, X).$$

We now state the main result of this section:

THEOREM 2.5. *Let* G *be a graph with* $\mu = \mu(G) \geqq 2$, $|\mathscr{V}(G)| = n$ *and* k *an integer with* $2 \leqq k \leqq \mu$, *then there exists a set* X *consisting of* k *mutually non-adjacent vertices of* G, *satisfying*:

$$\mu \leqq n - D(G, X). \tag{2.6}$$

Note that the case $k = 2$ reduces to the result of Ore (Theorem 2.1 in [3]):

$$\mu \leqq n - \rho(x_1) - \rho(x_2).$$

Proof. Let $\mathscr{S} = \{P_1, P_2, \cdots, P_\mu\}$ be a M.P.C. for G. For each $1 \leqq i \leqq k$, let x_i be an end-vertex of P_i. Since $\mathscr{S}$ is a M.P.C., x_i is not adjacent to x_j for $i \neq j$.

Let $X = \{x_1, x_2, \cdots, x_k\}$. We first show that for $1 \leqq i \leqq k$ and $1 \leqq j \leqq \mu$, the inequality:

$$\rho_{P_j}(\{x_i\}) \leqq |P_j| - \left(1 - \sum_{l=1}^{k-1} (-1)^l S_l(\rho_{P_j}, X - \{x_i\})\right) \tag{2.7}$$

holds. Let P_j be the path $\langle a_1, a_2, \cdots, a_t \rangle$, let $1 \leqq m \leqq k$, $m \neq i$, and consider the following cases:

(i) $i = j$. In this case assume that $x_i = a_1$.

(ii) $m = j$. In this case assume that $x_m = a_t$.

(iii) $m \neq j$ and $i \neq j$.

Let

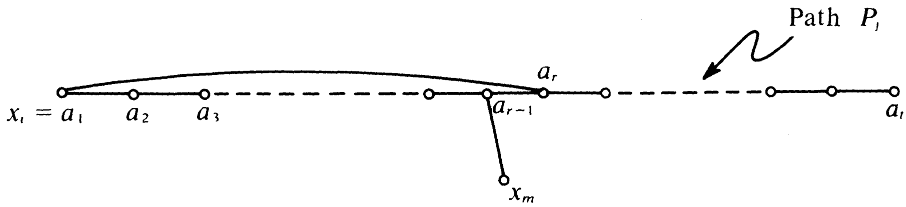
$$A = \{r\colon a_r \text{ is adjacent to } x_i\},$$

$$B_m = \{r\colon a_{r-1} \text{ is adjacent to } x_m\}$$

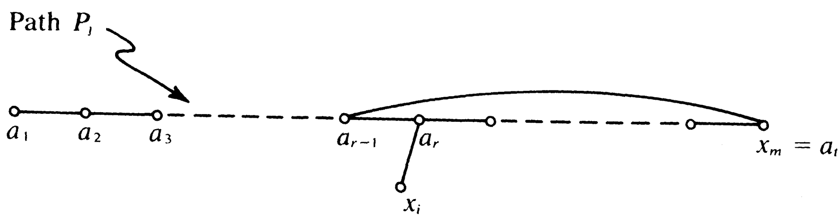
and

$$B = \bigcup_{\substack{1 \leqq m \leqq k \\ m \neq i}} B_m.$$

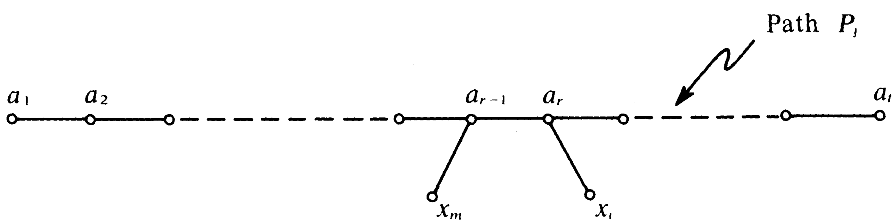
We claim that $A \cap B_m = \phi$, for if $r \in A \cap B_m$, then in each case we can construct a path-cover, $\mathscr{T}$ for G, as follows (see Figure 2.8):



Case (i)



Case (ii)



Case (iii)

FIGURE 2.8

In case (i), let:

$$\mathscr{T} = \mathscr{S} \cup \{\langle a_t, a_{t-1}, \cdots, a_r, x_i, a_2, a_3, \cdots, a_{r-1}\rangle * x_m P_m\} - \{P_i, P_m\}.$$

In case (ii), let:

$$\mathcal{T} = \mathcal{S} \cup \{\langle a_1, a_2, \cdots, a_{r-1}, x_m, a_{t-1}, a_{t-2}, \cdots, a_r\rangle * x_iP_i\} - \{P_i, P_m\}.$$

In case (iii), let:

$$\mathcal{T} = \mathcal{S} \cup \{\langle a_1, \cdots, a_{r-1}\rangle * x_mP_m, \langle a_t, a_{t-1}, \cdots, a_r\rangle * x_iP_i\} - \{P_i, P_j, P_m\}.$$

In either case, $|\mathcal{T}| = |\mathcal{S}| - 1 < |\mathcal{S}|$, contradicting the minimality of $\mathcal{S}$. Hence $A \cap B_m = \phi$. Also, in each case $a_1 \notin A$; so $A \subseteq P_j - B \cup \{a_1\}$. This gives $|A| \leqq |P_j| - |B \cup \{a_1\}|$, since $B \cup \{a_1\} \subseteq P_j$. But then, since $a_1 \notin B$, we get:

$$|A| \leqq |P_j| - (1 + |B|). \tag{2.9}$$

For $1 \leqq m \leqq k$, let:

$$C_m = \{r\colon a_r \text{ is adjacent to } x_m\}.$$

Then since x_m is not adjacent to a_1, $|C_m| = |B_m|$ and:

$$|B| = \left| \bigcup_{\substack{1 \leqq m \leqq k \\ m \neq i}} B_m \right| = \left| \bigcup_{\substack{1 \leqq m \leqq k \\ m \neq i}} C_m \right|$$

$$= \sum_{l=1}^{k-1} (-1)^{l+1} \sum_{\substack{1 \leqq m_1 < m_2 < \cdots < m_l \leqq k \\ m_1, m_2, \cdots, m_l \neq i}} |C_{m_1} \cap C_{m_2} \cap \cdots \cap C_{m_l}|$$

$$= -\sum_{l=1}^{k-1} (-1)^l S_l(\rho_{P_j}, X - \{x_i\}). \tag{2.10}$$

So since $|A| = \rho_{P_j}(\{x_i\})$, (2.7) follows from (2.9) and (2.10). Summing (2.7) for $1 \leqq i \leqq k$ and applying Lemma 2.4, we get:

$$S_1(\rho_{P_j}, X) \leqq k|P_j| - \left(k - \sum_{l=1}^{k-1} (-1)^l (k - l) S_l(\rho_{P_j}, X)\right). \tag{2.11}$$

Summing (2.11) for $1 \leqq j \leqq \mu$, we get:

$$S_1(\rho_G, X) \leqq kn - \left(k\mu - \sum_{l=1}^{k-1} (-1)^l (k-l) S_l(\rho_{G_l}, X)\right).$$

from which (2.6) follows.

3. Extremal problems.

DEFINITION 3.1. Let k and n be integers with $1 \leqq k \leqq n$. Define:

$$g(n,k) = \text{Min}\{m\colon \text{ every graph, } G, \text{ with } |\mathscr{V}(G)| = n \text{ and } |\mathscr{E}(G)| \geqq m \text{ has } \mu(G) \leqq k\}.$$

In this section we determine bounds for $g(n,k)$. See **[4]** for techniques in proving the following:

LEMMA 3.2.

$$\sum_{i=1}^{k=1} (-1)^i \left(\frac{k-i}{k}\right) \binom{k}{i} = -1 \quad \textit{if} \quad k \geqq 2, \tag{3.3}$$

$$\sum_{i=2}^{k} (-1)^i (k-i+1) \binom{k}{i-1} = k \quad \textit{if} \quad k \geqq 2, \tag{3.4}$$

$$\sum_{i=2}^{j} (-1)^i (k-i+1) \binom{j-1}{i-1} = k \quad \textit{if} \quad 3 \leqq \mathrm{j} \leqq \mathrm{k}. \tag{3.5}$$

LEMMA 3.6. *Let K be a graph with $|\mathscr{V}(K)| = s \geqq 1$, and let k be an integer with $k \geqq 2$, and suppose $H = \bar{\Gamma}_k + K$, then*:

$$D(H, \mathscr{V}(\bar{\Gamma}_k)) = 2s.$$

Proof. For $1 \leqq i \leqq k-1$ and $B \subseteq \mathscr{V}(\bar{\Gamma}_k)$ with $|B| = i$, each member of B is adjacent to every member of $\mathscr{V}(K)$. There are $\binom{k}{i}$ choices for B and $|\mathscr{V}(K)| = s$; thus:

$$S_i(\rho_H, \mathscr{V}(\bar{\Gamma}_k)) = s\binom{k}{i}.$$

This gives:

$$D(H, \mathscr{V}(\bar{\Gamma}_k)) = \frac{s}{k}\binom{k}{1} - \sum_{i=1}^{k-1}(-1)^i s\left(\frac{k-i}{k}\right)\binom{k}{i}$$

$$= s\left[1 - \sum_{i=1}^{k-1}(-1)^i\left(\frac{k-i}{k}\right)\binom{k}{i}\right]$$

$$= 2s, \quad \text{using} \quad (3.3).$$

THEOREM 3.7. *For* $1 \leqq k \leqq n$,

$$g(n,k) \leqq \tfrac{1}{2}(n-1)(n-k-1)+1. \tag{3.8}$$

Proof. Let G be a graph with $|\mathscr{V}(G)| = n$, and $|\mathscr{E}(G)| \geqq \frac{1}{2}(n-1)(n-k-1)+1$. Suppose $\mu(G) > k$ and $X = \{x_1, x_2, \cdots, x_k, x_{k+1}\}$ is a set of mutually nonadjacent vertices of G.

G may be considered to have been obtained from the complete graph Γ_n through the elimination of at most:

$$\tfrac{1}{2}n(n-1) - \tfrac{1}{2}(n-1)(n-k-1) - 1 = \tfrac{1}{2}(n-1)(k+1) - 1$$

edges. $\frac{1}{2}k(k+1)$ are removed in obtaining, from Γ_n, the graph H in which only members of X are nonadjacent. Thus, to obtain G from H, at most:

$$\tfrac{1}{2}(n-1)(k+1) - 1 - \tfrac{1}{2}k(k+1) = \tfrac{1}{2}(n-k-1)(k+1) - 1 \tag{3.9}$$

edges are removed from H.

We wish to compute $D(G,X)$. By Lemma 3.6,

$$D(H,X) = 2(n-k-1). \tag{3.10}$$

Now suppose that at some stage in the transformation from H to G, we have obtained a graph K with $\mathscr{E}(H) \supseteq \mathscr{E}(K) \supseteq \mathscr{E}(G)$ and $\mathscr{V}(K) = \mathscr{V}(H) = \mathscr{V}(G)$. Let $L = K - e$ where $e \in \mathscr{E}(K) - \mathscr{E}(G)$. We wish to know the effect, $f(e) = D(L,X) - D(K,X)$, on D, of removing the edge e. Since e is an edge of H, it cannot join two points of X. If neither end-point of e is in X, then $f(e) = 0$ since $S_i(\rho_K, X) = S_i(\rho_L, X)$ for $1 \leqq i \leqq k$. Now suppose that one end-point, y_1, of e is in X and that the other end-point, v, is not in X. Let $y_1, y_2, \cdots, y_j$ be the points of X which are adjacent to v in the graph K. Note that $1 \leqq j \leqq k+1$.

If $1 \leqq i \leqq j$ and $B \subseteq \{y_2, y_3, \cdots, y_j\}$ with $|B| = i-1$, and $C = B \cup \{y_1\}$, then $|C| = i$ and v is adjacent to every member of C in the graph K but not in the graph L. There are $\binom{j-1}{i-1}$ choices for such a set C. Furthermore, for any other combination of a vertex, t, and a set $A \subseteq X$ with $|A| = i$, t is adjacent to every member of A in the graph L. Thus:

$$S_i(\rho_L, X) - S_i(\rho_K, X) = \begin{cases} -\binom{j-1}{i-1} & \text{for} \quad i \leqq i \leqq j \\ 0 & \text{for} \quad i > j. \end{cases}$$

This gives:

$$f_j = f(e) = D(L, X) - D(K, X)$$

$$= \begin{cases} -\left[\frac{1}{k+1} - \sum_{i=1}^{k} (-1)^i \left(\frac{k-i+1}{k+1}\right) \binom{k}{i-1}\right] & \text{if} \quad j = k+1 \\ -\left[\frac{1}{k+1} - \sum_{i=1}^{j} (-1)^i \left(\frac{k-i+1}{k+1}\right) \binom{j-1}{i=1}\right] & \text{if} \quad 1 \leqq j \leqq k \end{cases}$$

$$= \begin{cases} -\frac{1}{k+1}\left[k+1 - \sum_{i=2}^{k} (-1)^i (k-i+1) \binom{k}{i-1}\right] & \text{if} \quad j = k+1 \\ -\frac{1}{k+1}\left[k+1 - \sum_{i=2}^{k} (-1)^i (k-i+1) \binom{j-1}{i-1}\right] & \text{if} \quad 2 \leqq j \leqq k \\ -1 & \text{if} \quad j = 1 \end{cases}$$

$$= \begin{cases} -\frac{1}{k+1} & \text{if} \quad 3 \leqq j \leqq k+1 \\ -\frac{2}{k+1} & \text{if} \quad j = 2 \\ -1 & \text{if} \quad j = 1 \end{cases}$$

using (3.4) and (3.5).

Notice that $f_1 \leqq f_2 \leqq \cdots \leqq f_k \leqq f_{k+1} < 0$ and that in order to realize the effect f_j, edges with effects $f_{k+1}, f_k, \cdots, f_{j+1}$ must first be removed. Hence when $(k+1)$ edges are removed, the combined effect is at least:

$$\sum_{i=1}^{k+1} f_i = -2.$$

So if r edges are removed in obtaining G from H,

$$D(G,X)-D(H,X) \geqq -\frac{2r}{k+1}. \tag{3.11}$$

Using (3.9) and (3.10) in (3.11) now gives:

$$D(G,X) \geqq [2(n-k-1)-(n-k-1)+2/(k+1)] > n-k-1. \tag{3.12}$$

But Theorem 2.5 guarantees the existence of a set X as constructed above, and satisfying:

$$D(G,X) \leqq n-\mu(G) \leqq n-k-1.$$

This contradicts (3.12) and completes the proof of the theorem.

COROLLARY 3.13. *For* $n \geqq 4$, $g(n,n-3)=n$.

Proof. The bipartite graph $\Gamma_{1,n-1}$ is a graph with n vertices, $(n-1)$ edges and path-covering number $(n-2)$. Thus $g(n,n-3) \geqq n$. The reverse inequality is given by Theorem 3.7.

To obtain a lower bound for $g(n,k)$, consider the graph $G = \Gamma_{n-k} \cup \bar{\Gamma}_k$; then $\mu(G)=k+1$, while $|\mathscr{V}(G)|=n$ and $|\mathscr{E}(G)| = \frac{1}{2}(n-1)(n-k-1)$. This gives:

PROPOSITION 3.14. *For* $n > k \geqq 1$

$$g(n,k) \geqq \tfrac{1}{2}(n-k)(n-k-1)+1. \tag{3.15}$$

The following proposition gives some results that are easily verified:

PROPOSITION 3.15.

(i) $g(n,n)=0$, $g(n+1,n)=1$, $g(n+2,n)=2$ *for* $n \geqq 1$
(ii) $g(6,2)=7$
(iii) $g(n+1,k+1) \geqq g(n,k)$ *for* $n \geqq k \geqq 1$.

Part (iii) can be seen by letting $G = H \cup \{x\}$ where H is a graph with n vertices, $g(n,k)-1$ edges, and $\mu(H)=k+1$, and x is an isolated vertex with $x \notin |$ ith $x \notin \mathscr{V}(H)$. Then G has $(n+1)$ vertices, $g(n,k)-1$ edges, and $(G)=k+2$.

In the case $k = 1$, the upper bound in (3.8) is seen to be the same as the lower bound in (3.15) and hence equality holds for $g(n, k)$ in both inequalities. However, Corollary 3.13 shows that the upper bound in (3.8) and not the lower bound in (3.15) is achieved in the case $k = n - 3$. ' Part (ii) of Proposition 3.15 shows a case where the lower bound and not the upper bound is achieved. It is conjectured that for small values of k, $g(n, k)$ is close to the lower bound in (3.15), while for large values of $k, g(n, k)$ is closer to the upper bound in (3.8).

We now turn to another extremal problem. Let v and n be integers with $0 \leqq v \leqq n$. Define:

$$h(n, v) = \operatorname{Min}\{k\colon \text{every graph, } G, \text{ with } |\mathcal{V}(G)| = n \text{ and } \rho(x) \geqq v \text{ for every } x \in \mathcal{V}(G), \text{ has } \mu(G) \leqq k\}.$$

THEOREM 3.16.

$$(n, v) = \begin{cases} 1 & \text{if } v \geqq \frac{n}{2} \\ n - 2v & \text{if } v < \frac{n}{2}. \end{cases}$$

Proof. The case $v \geqq \frac{n}{2}$ and the upper bound $h(n, v) \leqq n - 2v$ if $v < \frac{n}{2}$ follows from 0. Ore's result (the note to Theorem 2.5). If $v < \frac{n}{2}$, let $K = \Gamma_{v,n-v}$. Then clearly $|\mathcal{V}(K)| = n$ and $\rho(x) \geqq v$ for every $x \in \mathcal{V}(G)$; and in **[2]** (Theorem 2.2.10) we show that $\mu(K) = n - 2v$. Hence

$$h(n, v) \geqq n - 2v$$

completing the proof of the theorem.

REFERENCES

1. C. St. J. A. Nash-Williams, *Hamilton arcs and circuits, Recent Trends in Graph Theory*, Berlin-Heidelberg-New York, Springer-Verlag, (1971), 197–210.
2. S. Noorvash, *Covering the vertices of a graph by paths, Hamilton cycles and 1–factors, with related tree enumeration problems*, (Ph.D. Thesis, U.C.L.A.) 1972, 8–38.
3. O. Ore, *Arc coverings of graphs*, Ann. Mat. Ser. IV **55** (1961), 315–321.
4. J. Riordan, *An Introduction to Combinatorial Analysis*, New York, Wiley, 1958.

Received November 30, 1973.

UNIVERSITY OF ARIZONA