Discrete Mathematics 25 (1979) 65–74.

PATH COVERINGS OF THE VERTICES OF A TREE*

Peter J. SLATER
Applied Mathematics Division 5641, Sandia Laboratories, Albuquerque, NM 87185, USA

Received 7 January 1977
Revised 11 April 1978

Consider a collection of disjoint paths in graph G such that every vertex is on one of these paths. The size of the smallest such collection is denoted $i(G)$. A procedure for forming such collections is established. Restricting attention to trees, the range of values for the sizes of the collections obtained is examined, and a constructive characterization of trees T for which one always obtains a collection of size $i(T)$ is presented.

1. Introduction

Consider a finite, simple graph G (i.e., an undirected graph without loops and without multiple edges) whose vertex set and edge set will be denoted by $V(G)$ and $E(G)$, respectively. If $|V(G)|$, the cardinality of $V(G)$, is p, then one can consider G as a spanning subgraph of the complete graph on p vertices, denoted K_p. Consider any Hamiltonian cycle C in K_p. Either C is a Hamiltonian cycle of G or one has the following. If C contains n edges of $K_p - G$, then C contains exactly n disjoint paths in G (some of which may be trivial), and these paths contain every vertex of G.

Define the *Hamiltonian completion number* of G, denoted $\mathrm{hc}(G)$, to be the minimum number of edges which must be added to G in order to obtain a Hamiltonian cycle. Define the *path-covering number* of G, denoted $i(G)$, to be the minimum number of vertex-disjoint paths which contain $V(G)$. With the above comments in mind, it is straightforward to obtain the first result.

Proposition 1. *Either* $\mathrm{hc}(G)=0$ *and* $i(G)=1$, *or else* $\mathrm{hc}(G)=i(G)$.

In Ore's work [9] on degree conditions for Hamiltonian paths he introduced vertex disjoint path coverings of $V(G)$ such that the paths contain a maximum number of edges. While this is equivalent to the definition of $i(G)$, he used it to derive a theorem only for the case $i(G)=1$.

* This work was supported by the U.S. Energy Research and Development Administration (ERDA) under Contract No. AT (29–1)–789.

Boesch, Chen and McHugh [1] studied the path-covering number, and at the same time Goodman and Hedetniemi [5] introduced the problem in the form of a Hamiltonian completion. Work on the path-covering number includes that of Skupień [10] who (independent of Boesch, Chen and McHugh) examined sufficient conditions for a graph G to have $i(G) \leq n$, and Noorvash [8] who examined $g(p, n)$ which is the minimum integer so that every graph G with p vertices and at least $g(p, n)$ edges has $i(G) \leq n$. The Hamiltonian completion problem has been studied by Goodman, Hedetniemi and Slater in [6, 7, 11], where each path in a collection of disjoint paths covering $V(G)$ is called an "island" (and hence we use the notation $i(G)$).

A graph is *randomly path-Hamiltonian from vertex* v iff a Hamiltonian path always results upon starting at the vertex v and successively proceeding to any adjacent vertex not yet encountered. A graph G is *randomly circuit-Hamiltonian from vertex* v iff G is randomly path-Hamiltonian from v and every Hamilton path starting at v is contained in a Hamiltonian circuit. Chartrand and Kronk [2] characterized the graphs which are randomly path-Hamiltonian (randomly circuit-Hamiltonian) from every vertex, and, along with Lick in [4] obtained corresponding results for directed graphs. The graphs which are randomly circuit Hamiltonian from some vertex are characterized in [3], and Thomassen [13] characterizes the graphs which are randomly path-Hamiltonian from some vertex.

In [14], Thomassen characterizes the graphs G in which the following procedure always results in a Hamiltonian path. As he notes, this is possible iff every path of G is contained in some Hamiltonian path of G.

Procedure P. Select first any vertex $v_0 \in V(G)$, then select any vertex v_1 adjacent to v_0 (if any exist), then (if possible) a vertex $v_2 \in V(G) - \{v_0, v_1\}$ which is adjacent to v_1, etc. If this stops with v_k, then select (if possible) a vertex $v_{-1} \in V(G) - \{v_0, v_1, \ldots, v_k\}$ which is adjacent to v_0, then (if possible) a vertex $v_{-2} \in V(G) - \{v_{-1}, v_0, \ldots, v_k\}$ which is adjacent to v_{-1}, and so on.

If $v \in V(G)$, then the neighborhood of v, denoted $N(v)$, is the set of vertices adjacent to v. Suppose $P = v_1, v_2, \ldots, v_k$ is a path in graph G, and hence $V(P) = \{v_1, v_2, \ldots, v_k\}$. Call P a *blocked path* in G iff $N(v_1) \subseteq V(P)$ and $N(v_k) \subseteq V(P)$. Thus Procedure P is simply to construct a blocked path in G. Note that in determining when a graph is randomly path-Hamiltonian from v, a path $P = v, v_1, \ldots, v_k$ is considered to be "blocked from v" iff $N(v_k) \subseteq V(P)$, and $N(v)$ may, or may not, be contained in $V(P)$.

Consider the following procedure for obtaining a disjoint path cover of $V(G)$.

Procedure Q. Select P_1 to be any blocked path in G. Having chosen $P_1, \ldots, P_{k-1}$, if $V(G) - \bigcup_{j=1}^{k-1} V(P_j) \neq \emptyset$, then let P_k be any blocked path in $G - \bigcup_{j=1}^{k-1} P_j = G_k$.

It should be emphasized that Procedure Q produces a finite *sequence* of disjoint paths, say $S = (P_1, P_2, \ldots, P_k)$. If $\{P_1, P_2, \ldots, P_k\}$ is a disjoint path cover of $V(G)$, then $S=(P_1, P_2, \ldots, P_k)$ is called *sequentially blocked*, or a *blocked sequence*, iff sequence S can be obtained from G via Procedure Q (that is, iff path P_i is blocked in G_i). Calling k the *length* of S, one sees that Thomassen has characterized those graphs G for which every blocked sequence of G has length one.

Let $i'(G)$ denote the minimum possible value for the length of a blocked sequence of G, and let $I(G)$ denote the corresponding maximum possible value. This paper is an introductory study of the range of values that can be obtained for the lengths of blocked sequences of a graph G. Restricting my attention to trees, an interpolation theorem is presented (the result for arbitrary graphs remaining as a conjecture at this point), and a constructive characterization of trees T for which $i'(T)=I(T)$ is made.

First, $i'(G)$ and $i(G)$ will be related.

If $S=\{P_1, P_2, \ldots, P_n\}$ is a collection of disjoint islands (paths) which cover $V(T)$, call S a *blocked set* iff for each pair v_i and v_j of vertices such that v_i is an endpoint of P_i and v_j is an endpoint of P_j with $i \neq j$ one has $v_iv_j \notin E(G)$. Now $i(G)$ is clearly the minimum number of islands in a blocked set of G. Since any blocked sequence of length k clearly gives us a blocked set with k islands, one has $i(G) \leq i'(G)$. Note, however, that not every blocked set leads directly to a blocked sequence.

Theorem 2. *For any graph G, $i'(G)=i(G)$.*

Proof. Suppose $S=\{P_1, P_2, \ldots, P_n\}$ is a blocked set with $n=i(G)$ islands. Let $v(j, 1)$ and $v(j, 2)$ be the endpoints of P_j, where $v(j, 1)=v(j, 2)$ iff P_j is a singleton. Let $P'_j=P_j$ for $1 \leq j \leq n$, and let $G'_k=G-\bigcup_{j=1}^{k-1} P'_j$. Let t be the smallest value for which P'_t is not blocked in G'_t.

Assuming $v(t, 1)$ is adjacent to some vertex w in G'_{t+1}, one can assume the islands are labelled so that $w \in P'_{t+1}$. Since $n=i(G)$, w is not $v(t+1, 1)$ or $v(t+1, 2)$. Let x be the vertex adjacent to w on the subpath of P'_{t+1} from w to $v(t+1, 2)$. Change P'_t to be the path from $v(t+1, 1)$ to w, and the edge $w\, v(t, 1)$, and the old path P'_t from $v(t, 1)$ to $v(t, 2)$. Change P'_{t+1} to be its subpath from x to $v(t+1, 2)$. The new collection $\{P'_1, \ldots, P'_n\}$ is also a disjoint path cover of G with $i(G)$ elements, and hence it must also be a blocked set.

Since a finite number of iterations of this procedure will make P'_t blocked in G'_t, one eventually has P'_j blocked in G'_j for $1 \leq j \leq n$. Now $(P'_1, \ldots, P'_n)$ is sequentially blocked, and so $i'(G)=n=i(G)$.

2. An interpolation theorem for trees

First, some terminology (introduced in [12]) will be developed for trees. If v is a vertex of tree T, then a *branch of T at v* is defined to be a maximal subtree containing v as an endpoint. That is, a branch of T at v is the subgraph induced by v and one of the components of $T-v$. If v *has degree* d, written $\deg(v)=d$, then v has d different branches. A branch B of T at v which is a path will be called a *branch path at* v iff $\deg(v)\geq 3$. Vertex v will be called the *stem* of the branch path at v. If v is the stem of at least one branch path, then the subgraph of T consisting of v and all its branch paths will be called a *leaf with stem* v. Thus a leaf with k branch paths is homeomorphic to $K_{1,k}$. Note that distinct leaves of T must be disjoint. A leaf whose stem v has $\deg(v)$ or $\deg(v)-1$ branch paths will be called an *end leaf.*

Letting D denote the number of vertices in T of degree at least three, the straightforward proof of the next lemma is omitted.

Lemma 3. *Suppose tree* T *is not a path. If* $D=1$, *then there is exactly one (end) leaf. If* $D\geq 2$, *then there are at least two end leaves whose stems,* v_1 *and* v_2, *have* $\deg(v_1)-1$ *and* $\deg(v_2)-1$ *branch paths, respectively.*

It is fairly easy to see that if v is the stem of a leaf in tree T and has branch paths B_1 and B_2, then there is a path covering of $V(T)$ with $i(T)$ elements in which one of the islands is the path $B_1\cup B_2$. Using this one proves the following lemma.

Lemma 4. *Let* T *be a tree. If* $D=0$, *then* $i(T)=1$. *If* $D=1$, *let* v *be the vertex of degree at least three, then* $i(T)=\deg(v)-1$, *and one island consists of two (arbitrary) branches from* v. *If* $D\geq 2$, *let* L *be an end leaf with stem* v. *Suppose* $\deg(v)=k+1$ *(and so* v *has* k *branch paths). Then* $i(T)=i(T-L)+k-1$, *and one island can be taken to be two of the branch paths from* v.

This lemma leads to an algorithm for determining $i(T)$ for any tree T. This algorithm is a slight modification of the algorithm presented in [1] and in [5]. The difference is that here we select v to be the stem of an end leaf L rather than the stem of an arbitrary leaf containing two or more branch paths. In this way, after we cover L (with $k-1$ disjoint paths) $T-L$ is also a tree. That is, $T-L$ is also connected.

If L is any end leaf of T and the stem of L has degree $k+1$ in T, then it is easy to see that $I(T)\geq I(T\text{-}L)+k-1$. Since every tree has an end leaf we get by Lemma 4:

Lemma 5. *Let* T *be a tree with at least two vertices of degree* 3. *Then there is a sequence of trees* $T_0, T_1, \ldots, T_s$ *such that* $T_s=T$, T_0 *is a homeomorph of* $K_{1,k}$ *and* T_{i+1} *is obtained from* T_i *by adding a leaf and letting its stem be adjacent to some vertex of* T_i. *For any such sequence* $T_0,\ldots,T_s$ *we have*

$$0=I(T_0)-i(T_0)\leq I(T_1)-i(T_1)\leq\cdots\leq I(T_s)-i(T_s).$$

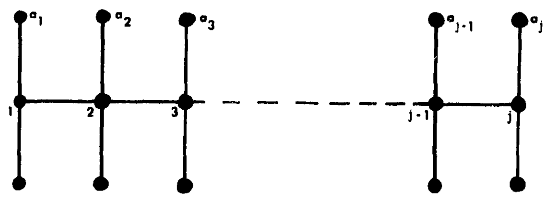

Fig. 1. Tree T_j with $i(T)=j$ and $I(T_j)=2j-1$.

For the tree T_j of Fig. 1, let P_k be the path containing vertex k and the two endpoints adjacent to this vertex. Now $S_1=\{P_1,\ldots,P_j\}$ is a path covering of T_j with $i(T_j)=j$ islands. For $n\geq 2$, let S_n contain the paths $P'_n=a_1, 1, \ldots, n, a_n$, and $P_{n+1},\ldots,P_j$, and the remaining singleton vertices. Now S_n is a blocked sequence with $1+(j-n)+(2n-2)=n+j-1$ islands for $1\leq n\leq j$. In particular, S_j has $I(T_j)=2j-1$ islands.

The next result shows that the above example is a demonstration of an interpolation theorem for trees (an ITT result). That is, if we present two blocked sequences of tree T with i_1 and i_2 members, then if $i_1\leq h\leq i_2$ then we are guaranteed the existence of a blocked sequence with h members.

Theorem 6 (ITT). *If $i(T)\leq h\leq I(T)$, then there exists a blocked sequence S_h for T of length h.*

Proof. To prove the theorem, induction on $p=|V(T)|$ will be used. If $D=0$ or 1, then $i(T)=I(T)$, and we are done. Thus we can assume that the number of vertices of degree at least three is two or more. Clearly it suffices to show that if $S=(P_1,\ldots,P_{j+1})$ is a blocked sequence for T of length $j+1\geq i(T)+2$, then we can construct a blocked sequence $S'=(P'_1,\ldots,P'_j)$ of length j.

Let L be an end leaf with stem vertex v, and let $B_1,\ldots,B_k$ be all the branch paths from v. Thus $\deg(v)=k+1$, and we let the vertex in $T-L$ which is adjacent in T to v be labelled w. Let P_t be the path in S which contains v.

First, assume that $w\notin P_t$, and one can assume P_t contains B_1 and B_2. Reorder S to form $S^*=(P^*_1,\ldots,P^*_{j+1})$ where

$$P^*_1=P_t,\ P^*_2=B_3-v,\ldots,P^*_{k-1}=B_k-v,$$

and the remaining paths for S^* are the remaining paths in S arranged in the same order. Clearly S^* is a blocked sequence since S is. Now $i(T-L)=i(T)-k+1$, and $(P^*_k,\ldots,P^*_{j+1})$ is a blocked sequence for $T-L$ of length $j+1-k+1\geq i(T)+2-k+1=i(T-L)+2$. By induction we have a blocked sequence S'' for $T-L$ of length $j+1-k$. Juxtaposing $(P^*_1,\ldots,P^*_{k-1})$ and S'', we have a blocked sequence of length $(k-1)+(j+1-k)=j$ for T.

Second, assume $w\in P_t$ and there is a vertex $x\in N(w)$ such that x is an endpoint of P_s with $s\neq t$. (Necessarily, $s>t$.) Let A and B be the components of $T-L-wx$ containing w and x, respectively. Create S' as follows. First list $k-1$ paths which cover L. Within the next groups described, paths appear in the same order as in S.

List the paths of S in A which appeared before P_t; list the paths of S in B which appeared before P_s; let the next path be P_s and edge xw and the part of P_t in $T-L$; list the remaining paths of S. Now S' is a blocked sequence since S is, and the length of S is j.

Third, assume $w \in P_t$ and $x \in N(w) \cap T-L$ implies that x is not an endpoint of a path in S except possibly P_t. As in the first case, S^* is formed with

$$P_1^* = B_1 \cup B_2,\ P_2^* = B_3 - v, \ldots, P_{k-1}^* = B_k - v,$$

and the remaining paths for S^* are the remaining paths in S after replacing P_t by $P_t \cap (T-L) = P_t'$. In the order in which they appear in S, one lists all paths in the components of $T-L-w$ which do not intersect P_t. Then list all the remaining paths (with P_t replaced by P_t') in the order in which they appear in S. Now S^* is a blocked sequence of length $j+1$. Following the same argument as the first case, we obtain a blocked sequence of length j for T.

From the proof of the previous theorem one can derive the inequality $I(T) \leq I(T-L)+k$ where k is the number of branch paths at stem v of end leaf L. Since $i(T)=i(T-L)+k-1$, and since $i(T)=1$ implies $I(T)=1$, one obtains by induction:

$$I(T) \leq 2i(T)-1$$

for any tree T. Fig. 1 shows that this is the best possible.

Conjecture. *If $i(G) \leq h \leq I(G)$, then there exists a blocked sequence S_h for graph G of length h.*

3. Randomly island decomposable trees

As a consequence of the ITT result one has that an application of Procedure Q can result in a blocked sequence whose length differs from $i(T)$ by any value from 0 to $I(T)-i(T)$. In this section a constructive characterization will be presented for trees T with the property that any application of Procedure Q produces an $i(T)$ island decomposition.

Call graph G a *randomly island decomposable* graph iff $i(G)=I(G)$. For short, G will be called an RID graph. Let $\mathcal{R}$ denote the class of RID trees. As indicated in the proof of Theorem 6, we will have to consider vertices which are not in an end leaf, but which are adjacent to the stem (or stems) of an end leaf (end leaves).

First, two methods of extending a tree $T' \in \mathcal{R}$ to a larger tree $T \in \mathcal{R}$ will be described.

Operation 1. Given a tree T', select a vertex $v \in T'$ such that $\deg(v)=1$ or 0, and let T be the tree obtained from T' by adding a leaf L_1 whose stem s_1 is made adjacent to v. (See Fig. 2, (a) and (b).)

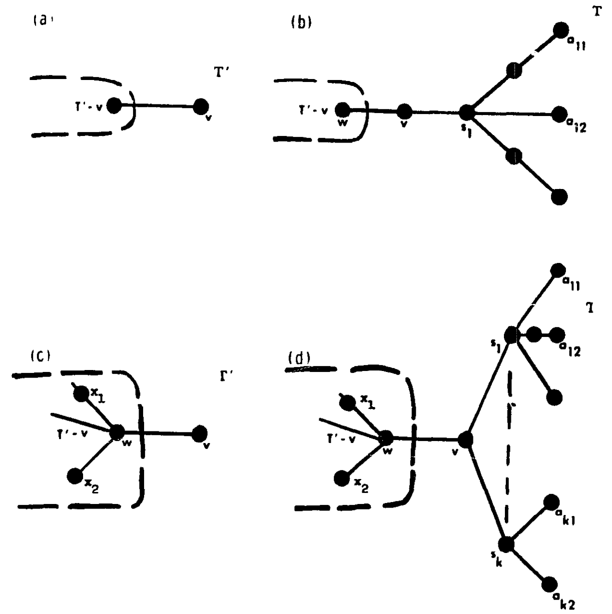

Fig. 2. Examples of leaf addition.

Operation 2. Given a tree T', select a vertex $v \in T'$ such that $\deg(v) = 1$. Let w be adjacent to v in T', and assume that w is also adjacent to two other vertices, say x_1 and x_2, of degree at most two. Let $L_1, \ldots, L_k$ $(k \geq 2)$ be leaves with stems $s_1, \ldots, s_k$. Let T be the tree obtained from T' by adding leaves $L_1, \ldots, L_k$ and edges $vs_1, \ldots, vs_k$. (See Fig. 2, (c) and (d).)

It is required only that s_i have degree at least two in L_i, and then its degree in T will be at least three. For leaves $L_1, L_2, \ldots, L_k$ it is assumed that L_i has $t_i + 1$ endpoints, two of which are a_{i1} and a_{i2}. Let $\sigma = \sum_{i=1}^{k} t_i$.

Proposition 7. *If tree T is obtainable from subtree T' by Operation* 1, *then $T \in \mathfrak{R}$ iff $T' \in \mathfrak{R}$.*

Proof. Assume T is RID. Since one can let P_1 be the path from a_{11} to a_{12}, and $P_2, \ldots, P_{t_1}$ be the remaining paths in L_1, and any extension of $(P_1, \ldots, P_{t_1})$ to a blocked sequence for T must have length $i(T)$, then any blocked sequence for $T - L_1$ must have $i(T) - t_1$ paths. Hence T' is also RID.

Assume T' is RID, and let $(P_1, \ldots, P_n)$ be a blocked sequence for T. If s_1 and v are on different paths, then one can assume P_1 is the a_{11} to a_{12} path and $P_1, \ldots, P_{t_1}$ cover L_1. Now T' is RID implies that $n = t_1 + i(T') = i(T)$. If s_1 and v are on the same path, then one must have, for example, a_{11} on this path and t_1 other paths in L_1. Remove these t_1 paths and reduce the path containing v and s_1 to its vertices in T'. Since v is an endpoint of T', one has a blocked sequence for T' of length $n - t_1 = i(T')$, and hence $n = t_1 + i(T') = i(T)$. Thus T is also RID.

Next assume $\deg(x_2) \geq 3$ and let $P_2, \ldots, P_d$ and $P_1', \ldots, P_k'$ be as described. Let $\bar{P}$ be a path from v through w and x_1 to an endpoint e of T. Let P_1'' be as described, and let $\bar{P}_2''$ be the path from w to e. Now $(P_2, \ldots, P_d, P_1', \ldots, P_k', \bar{P})$ and $(P_2, \ldots, P_d, P_1'', \bar{P}_2'')$ have extensions to blocked sequences for T of different lengths. This contradiction implies that $\deg(x_2) \leq 2$, and the proof is complete.

References

[1] F.T. Boesch, S. Chen, and J.A.M. McHugh, On covering the points of a graph with point disjoint paths, in: R.A. Bari and F. Harary, eds., Graphs and Combinatorics (Springer-Verlag, Berlin, 1974) 201–212.

[2] G. Chartrand and H.V. Kronk, Randomly traceable graphs, Siam. J. Appl. Math. 16 (1968) 696–700.

[3] G. Chartrand and H.V. Kronk, On a special class of Hamiltonian graphs, Comment. Math. Helv. 44 (1969) 84–88.

[4] G. Chartrand, H.V. Kronk and D.R. Lick, Randomly Hamiltonian digraphs, Fund. Math. 65 (1969) 223–226.

[5] S.E. Goodman and S.T. Hedetniemi, On the Hamiltonian completion problem, in: R.A. Bari and F. Harary, eds., Graphs and Combinatorics (Springer-Verlag, Berlin, 1974) 262–272.

[6] S.E. Goodman, S.T. Hedetnieme, and P.J. Slater, Advances on the Hamiltonian completion problem, J. Assoc. Comput. Mach. 22 (3) (1975) 352–360.

(7) S.E. Goodman, S.T. Hedetniemi, and P.J. Slater, Some results on the optional and non-optional Hamiltonian completion problem, in: F. Hoffman et al; eds., Proceedings of the Fifth Southeastern Conference on Combinatorics, Graph Theory and Computing (Utilitas Mathematica, Winnipeg, 1974) 423–448.

[8] S. Noorvash, Covering the vertices of a graph by vertex-disjoint paths, Pac. J. Math. 58 (1) (1975) 159–168.

[9] O. Ore, Arc coverings of graphs, Ann. Mat. Pura Appl. 55 (1961) 315–332.

[10] Z. Skupień, Hamiltonian circuits and path coverings of vertices in graphs, Colloquium Mathematicum 30 (1974) 295–316.

[11] P.J. Slater, S.E. Goodman, and S.T. Hedetniemi, On the optional Hamiltonian completion problem, Networks 6 (1976) 35–51.

[12] P.J. Slater, Leaves of trees, in: F. Hoffman et al., eds., Proceedings of the Sixth Southeastern Conference on Combinatorics, Graph Theory and Computing, (Utilitas Mathematica, Winnipeg, 1975) 549–559.

[13] C. Thomassen, On randomly Hamiltonian graphs, Math. Ann. 200 (1973) 195–208.

[14] C. Thomassen, Graphs in which every path is contained in a Hamilton path, J. für Reine Angew Math. 268/269 (1974) 271–282.