Pacific Journal of Mathematics

RELATIONS BETWEEN PACKING AND COVERING NUMBERS OF A TREE

A. MEIR AND JOHN W. MOON

Vol. 61, No. 1 November 1975

PACIFIC JOURNAL OF MATHEMATICS
Vol. 61, No. 1, 1975

RELATIONS BETWEEN PACKING AND COVERING NUMBERS OF A TREE

A. MEIR AND J. W. MOON

Let P_k denote the size of the largest subset of nodes of a tree T with n nodes such that the distance between any two nodes in the subset is at least $k+1$; let C_k denote the size of the smallest subset of nodes of T such that every node of T is at distance at most k from some node in the subset. We determine various relations involving P_k and C_k; in particular, we show that $P_k + kC_k \leqq n$ if $n \geqq k+1$ and that $P_{2k} = C_k$.

1. Introduction. The *distance* between nodes x and y in a graph G is the number $d(x, y)$ of edges in any shortest path in G that joins x and y. (For definitions not given here see [1] or [5].) A subset $\mathscr{P}$ of nodes of G is a *k-packing* if $d(x, y) > k$ for all pairs of distinct nodes x and y of $\mathscr{P}$; the *k-packing number* of G is the number $P_k = P_k(G)$ of nodes in any largest k-packing in G. A subset $\mathscr{C}$ of nodes of G is a *k-covering* if for every node x in G there is at least one node y in $\mathscr{C}$ such that $d(x, y) \leqq k$; the *k-covering number* of G is the number $C_k = C_k(G)$ of nodes in any smallest k-covering of G.

Our object here is to establish various relations between $P_k(T)$ and $C_k(T)$ when T is a tree with n nodes. We consider the case $k = 1$ in §2 and determine those values of α and β for which there exists a tree T such that $P_1(T) = \alpha$ and $C_1(T) = \beta$. We derive upper bounds for $P_k(T)$ and $C_k(T)$ in §3. In §4 we show that $P_k(T) + kC_k(T) \leqq n$ for any tree T with n nodes when $n \geqq k+1$ and we show that this inequality is, in a sense, best possible. Finally, in §5 we show that $P_{2k} = C_k$.

The quantities $P_1(G)$ and $C_1(G)$ have been considered before under different names. For example, $P_1(G)$ and $C_1(G)$ are called the independence number and the domination number of G in [5; Chap. 13]; and they are called the coefficients of internal and external stability in [1; Chap. 4]. Some inequalities for $P_1(G)$ and $C_1(G)$ are given in [2; Chaps. 13 and 14] but some of these are unnecessarily weak when G is a tree.

2. Relations between P_1 and C_1. In what follows T will always denote an arbitrary tree with n nodes. For convenience, we shall frequently write P and C for $P_1(T)$ and $C_1(T)$.

THEOREM 1. *If $n \geqq 2$, then $P + C \leqq n$.*

Proof. If $\mathscr{P}$ denotes a 1-packing of P nodes in T then each node of $\mathscr{P}$ must be joined to at least one node not in $\mathscr{P}$ if $n \geqq 2$. Thus the $n - P$ nodes not in $\mathscr{P}$ constitute a 1-covering of T. Hence, $C \leqq n - P$, as required.

COROLLARY 1. *If $n \geqq 2$, then $1 \leqq C \leqq (1/2)n \leqq P \leqq n - 1$.*

Proof. It is obvious that $C \geqq 1$ and $P \leqq n-1$ when $n \geqq 2$. The remaining inequalities follow from Theorem 1 and the easily established fact that $C \leqq P$ (see [5; p. 211]); they may also be proved directly by observing that the sets of nodes of T whose distances from a given node x are odd or even, respectively, are both 1-packings and 1-coverings. We remark that the inequalities $C_1(G) \leqq (1/2)n \leqq P_1(G)$ hold for any nontrivial connected bipartite graph G with n nodes.

THEOREM 2. *If $n \geqq 1$, then $P + 2C \geqq n + 1$.*

Proof. Let $\mathscr{C}$ denote a 1-covering of C nodes of T and let R denote the subgraph determined by the $n - C$ nodes not in $\mathscr{C}$. If R has j components and e edges then $e = n - C - j$ (see [5; p. 68]) and it is easy to see that $P \geqq j$. Since each node of R is joined to at least one node of $\mathscr{C}$ and since T has $n - 1$ edges altogether it follows that

$$e \leqq (n - 1) - (n - C) = C - 1 .$$

Hence,

$$P \geqq j = n - C - e \geqq (n - C) - (C - 1) = n - 2C + 1 ,$$

as required. (It will follow from Theorem 3 that the inequalities $1/2(n + 1 - P) \leqq C \leqq n - P$, implied by Theorems 1 and 2 are, in a sense, best possible.)

The next result is obtained by combining the inequalities $P \geqq (1/2)n$ and $P + 2C \geqq n + 1$.

COROLLARY 2. *If $n \geqq 1$ and $0 \leqq \lambda \leqq 2$, then*

$$P + \lambda C \geqq \frac{1}{2}\Big(1 + \frac{1}{2}\lambda\Big)n + \frac{1}{2}\lambda \, ;$$

in particular,

$$P + C \geqq \left\{\frac{3}{4}n + \frac{1}{2}\right\} ,$$

where $\{x\}$ *denotes the least integer not less than* x.

It is not difficult to construct trees for which equality holds in the last inequality. We remark that it follows from results in [3] and [4] that the average value of $P + C$ over the n^{n-2} trees with n labelled nodes is approximately $.927n$ for large values of n.

THEOREM 3. *If* α *and* β *are positive integers such that*

$$\alpha \geqq \frac{1}{2}n \,, \tag{1}$$

$$\alpha + \beta \leqq n \,, \tag{2}$$

and

$$\alpha + 2\beta \geqq n + 1 \,, \tag{3}$$

then there exists a tree T *with* n *nodes such that* $P(T) = \alpha$ *and* $C(T) = \beta$.

Proof. Let $\nu = n - \alpha - \beta$. It follows from (1) that $\beta + \nu \leqq (1/2)n$ and this implies that $n + 1 - 2\beta - 2\nu \geqq 1$; furthermore, it follows from (3) that $\nu \leqq \beta - 1$ or $\beta - 1 - \nu \geqq 0$. Let T denote the tree constructed as follows: $n - 1$ nodes are split into ν sets of four nodes, $\beta - 1 - \nu$ sets of two nodes, and $n + 1 - 2\nu - 2\beta$ sets consisting of a single node; a path is formed on the nodes in each set and the node at one end of each of these paths is joined to an nth node. (The tree arising when $n = 13$, $\alpha = 7$, and $\beta = 4$ is illustrated in Figure 1.) It is not difficult to verify that this construction

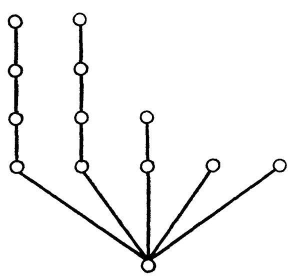

FIGURE 1

is indeed possible and that

$$P(T) = 2\nu + (\beta - 1 - \nu) + (n + 1 - 2\nu - 2\beta) = n - \beta - \nu = \alpha$$

and

$$C(T) = 1 + \nu + (\beta - 1 - \nu) = \beta \,,$$

as required.

3. Upper bounds for P_k and C_k.

In what follows k and n will denote arbitrary positive integers.

THEOREM 4. *If* $n \geqq [1/2(k+3)]$ *then*

$$P_k \leqq [2n/(k+2)] \tag{4}$$

if k is even, and

$$P_k \leqq [(2n-2)/(k+1)] \tag{5}$$

if k is odd.

Proof. If x is any node in any k-packing $\mathscr{P}$ with P_k nodes of T, let $N(x) = \{u: u \in T \text{ and } d(x, u) \leqq j\}$ where $j = [(1/2)k]$. Since the tree T is connected and has at least $[1/2(k+3)] \geqq j+1$ nodes, it follows that $|N(x)| \geqq j+1$ for all $x \in \mathscr{P}$. Furthermore, the sets $\{N(x): x \in \mathscr{P}\}$ are disjoint; for if $u \in N(x) \cap N(y)$ where $x \neq y$, then $d(x, y) = d(x, u) + d(u, y) \leqq 2j \leqq k$ and this would contradict the definition of $\mathscr{P}$. Hence,

$$n \geqq \sum_{x \in \mathscr{P}} |N(x)| \geqq P_k \cdot (j+1)$$

and this implies inequality (4).

If $k = 2j+1$ we may further assert that no edge joins a node u of any set $N(x)$ to a node v of any other set $N(y)$ where $x \neq y$; for if there were such an edge, then $d(x, y) = d(x, u) + d(u, v) + d(v, y) \leqq 2j + 1 = k$ and this would again contradict the definition of $\mathscr{P}$. If $P_k = 1$ inequality (5) certainly holds. If $P_k \geqq 2$ there must exist at least one node of T that is not in any set $N(x)$, where $x \in \mathscr{P}$, for T would not be connected otherwise. Hence,

$$n \geqq 1 + \sum_{x \in \mathscr{P}} |N(x)| \geqq 1 + P_k \cdot (j+1)$$

when $k = 2j+1$, and this implies inequality (5).

If $\mathscr{P}$ is any maximal k-packing of P_k nodes in a tree T, then $\mathscr{P}$ is also a k-covering of T; for if there were a node x in T such that $d(x, y) > k$ for each node y in $\mathscr{P}$ then $\mathscr{P} \cup \{x\}$ would be a larger k-packing in T which is impossible. This implies that $C_k \leqq P_k$ for any tree T (this result is given in [5; p. 211] when $k = 1$, as was mentioned earlier). Hence, Theorem 4 provides an upper bound for C_k also; a better bound is given in the following result.

THEOREM 5. *If* $n \geqq k+1$, *then*

$$C_k \leqq [n/(k+1)] \, . \tag{6}$$

Proof. Suppose one of the longest paths in T joins nodes x and y. Let

$$D_i = \{u\colon u \in T \quad \text{and} \quad d(x, u) \equiv i \,(\mathrm{mod}\,(k+1))\}$$

for $0 \leqq i \leqq k$. We may assume $D_i \neq \varnothing$ for each i, for otherwise the node x itself would constitute a k-covering and inequality (6) would certainly hold. We now show that each set D_i is a k-covering of T.

Let z denote any node of T and suppose $d(x, z) = l$. If $l \geqq i$ then $i + m(k+1) \leqq l < i + (m+1)(k+1)$ for some nonnegative integer m. Let u denote the unique node on the path joining x and z such that $d(x, u) = i + m(k+1)$; then $u \in D_i$ and $d(u, z) \leqq k$ as required. If $l < i$ let v denote the unique node on the path joining x and y such that $d(x, v) = i$; then $v \in D_i$ and

$$d(z, v) = d(z, y) - d(v, y) \leqq d(x, y) - d(v, y) = d(x, v) = i \leqq k \, ,$$

as required.

The k-coverings $\{D_i\colon 0 \leqq i \leqq k\}$ are disjoint and together they exhaust the nodes of T; hence, at least one of them has at most $[n/(k+1)]$ nodes. This suffices to complete the proof of the theorem.

It is not difficult to construct trees for which equality holds in (4), (5), and (6) for all admissible values of k and n.

4. A relation between P_k and C_k.

THEOREM 6. *If* $n \geqq k + 1$, *then*

$$P_k + kC_k \leqq n \, . \tag{7}$$

Proof. If $k = 1$ this is the same as Theorem 1, so we shall assume henceforth that $k \geqq 2$.

Let $\mathscr{P}$ denote a k-packing of P_k nodes in T. If $x \in \mathscr{P}$ let $E(x) = \{u\colon u \in T \text{ and } d(u, x) = 1\}$; these sets are nonempty and disjoint when $k \geqq 2$ and no edge joins two nodes of the same set $E(x)$. Select one node u_x from each set $E(x)$ and let R denote the graph obtained from T as follows: remove each node x of $\mathscr{P}$ and all edges incident with x, and insert new edges joining each node u_x to each of the other nodes of $E(x)$. It is not difficult to see that R is a tree with $n - P_k$ nodes.

If r and s are nodes in $E(x)$ and $E(y)$, respectively, where $x \neq$

y, then $d(r, s) \geqq k - 1$; for, if $d(r, s) \leqq k - 2$ then $d(x, y) \leqq k$ and this would contradict the definition of $\mathscr{P}$. This implies the following observation:

(*) If a path in R of length at most $k - 1$ contains a new edge of the type ru_x where $r \in E(x)$, then the path does not contain any nodes of any other set $E(y)$ where $y \neq x$.

Let $\mathscr{C}$ denote any smallest $(k - 1)$-covering of R. We shall show that the nodes of $\mathscr{C}$ constitute a k-covering of T. Let z denote any node of T. If $z \notin \mathscr{P}$ then $z \in R$ and there exists a node $v \in \mathscr{C}$ such that $d(v, z) \leqq k - 1$ in R. If there are no new edges in the path $p(v, z)$ from v to z in R then all the edges of $p(v, z)$ are in T and $d(v, z) \leqq k - 1$ in T also. If there is just one new edge in the path $p(v, z)$ of the type ru_x where $r \in E(x)$, then $d(v, z) \leqq k$ in T since the edge ru_x can be replaced by the two edges rx and xu_x in T. If there is more than one new edge in the path $p(v, z)$ then these new edges must all join pairs of nodes from the same set $E(x)$, in view of observation (*). But all new edges of this type are incident with the node u_x. Hence, there can be only two such edges in $p(v, z)$, they must occur consecutively, and they must be of the form ru_x and u_xs. But then $d(v, z) \leqq k - 1$ in T also since the edges ru_x and u_xs in $p(v, z)$ can be replaced by the edges rx and xs in T.

If $z \in \mathscr{P}$ then then there exist nodes $r \in E(z)$ and $v \in \mathscr{C}$ such that $d(v, r) \leqq k - 1$ in R and the path $p(v, r)$ from v to r does not pass through any other nodes of $E(z)$. This path cannot contain any new edges by observation (*). Hence, $d(v, z) = d(v, r) + 1 \leqq k$ in T, as required.

If $n = k + 1$ then $C_k = P_k = 1$ and inequality (7) certainly holds. If $n \geqq k + 2$, it follows from Theorem 4 that $n - P_k \geqq k$. Hence, when $n \geqq k + 2$, we may apply Theorem 5 to the tree R and conclude that $|\mathscr{C}| \leqq (n - P_k)/k$. Since $C_k \leqq |\mathscr{C}|$, this implies that $P_k + kC_k \leqq n$, as required.

We now show that inequality (7) is best possible when $n = m(k + 1)$ for $m = 1, 2, \cdots$. Let H denote the tree with n nodes constructed as follows: the n nodes are split into m sets of $k + 1$ nodes each; a path of length k is formed on the nodes in each set; and, finally, the nodes at one end of these paths are joined so as to form a path of length $m - 1$. (The tree H arising when $n = 20$ and $k = 3$ is illustrated in Figure 2.) It is not difficult to verify that $P_k + kC_k = m + km = n$ for the tree H. We leave it as an exercise for the reader to show that there exists a tree with n nodes for which $C_k =$

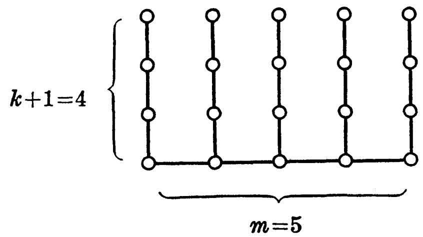

FIGURE 2

$[(n - P_k)/k]$ for arbitrary values of n and k such that $n \geqq k + 1$.

No inequality of the type

$$(1 + \varepsilon)P_k + (k - \varepsilon)C_k \leqq n ,$$

where ε is any positive constant, can be valid for all trees with sufficiently many nodes. To show this let J denote the tree with $n = m(k + 1) + 1$ nodes formed by joining a new node to one of the nodes of H in the way illustrated in Figure 3 when $n = 21$ and $k = 3$. It is easy to see that

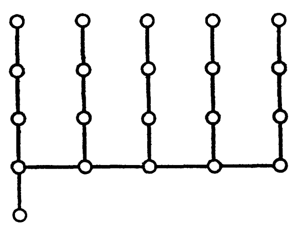

FIGURE 3

$$(1 + \varepsilon)P_k + (k - \varepsilon)C_k = (1 + \varepsilon)(m + 1) + (k - \varepsilon)m = n + \varepsilon$$

for the tree T. It might be of some interest to determine best possible upper bounds in terms of n for $lP_k + (k + 1 - l)C_k$ when $l > 1$.

It might also be of some interest to determine best possible upper bounds in term of n for $P_k + C_k$. It follows from Theorems 4 and 6 that $P_k + C_k \leqq 3n/(k + 2)$ when k is even, but this is probably not best possible in general.

There does not seem to be any natural nontrivial analogue of Theorem 2 when $k \geqq 2$, at least one that does not involve additional parameters or assumptions, since it is easy to construct trees for which $P_k = C_k = 1$ when $k \geqq 2$.

5. The equality of P_{2k} and C_k.

THEOREM 7. *If $k \geqq 1$, then $P_{2k} = C_k$.*

Proof. Let $\mathscr{P}$ denote a $2k$-packing consisting of P_{2k} nodes of the tree T and let $\mathscr{C}$ denote a k-covering consisting of C_k nodes of T. It is easy to see that for each node x in $\mathscr{C}$ the set $N(x) = \{u: u \in T \text{ and } d(x, u) \leqq k\}$ contains at most one node y in $\mathscr{P}$. Since every node y in $\mathscr{P}$ belongs to at least one set $N(x)$ it follows that $P_{2k} \leqq C_k$. It remains to show that $P_{2k} \geqq C_k$.

Let $\{x_0, x_1, \cdots, x_m\}$ denote the nodes of any longest path in the tree T. If $m \leqq 2k$, then $P_{2k} = C_k = 1$; so we may suppose that $m \geqq 2k + 1$. Let T' denote the smallest subtree of T containing all nodes z of T such that $d(x_k, z) > k$; that is, T' is the subtree determined by all nodes v of T such that either $d(x_k, v) > k$ or there exists some node, say z_v, such that $d(x_k, z_v) > k$ and the unique path joing z_v and x_m in T contains v. The subtree T' is nonempty since $d(x_k, x_m) > k$.

Let $\mathscr{P}'$ denote a largest $2k$-packing consisting of P'_{2k} nodes of T' and let $\mathscr{C}'$ denote a smallest k-covering consisting of C'_k nodes of T'. It is easy to see that $\mathscr{C} = \mathscr{C}' \cup \{x_k\}$ is a k-covering of T; consequently,

$$C_k \leqq C'_k + 1 . \tag{8}$$

Suppose there exists a node y in $\mathscr{P}'$ such that $d(x_k, y) \leqq k$. Let B_y denote the subtree of T' determined by all nodes s of T' such that the unique path from s to x_m contains y; in particular, the node z_y, defined earlier, is in B_y. We assert that y is the only node of $\mathscr{P}'$ in B_y. For, if there were a second such node, say w, then $d(w, y) \geqq 2k + 1$; this would imply that

$$\begin{aligned} d(w, x_m) &= d(w, y) + d(y, x_m) \geqq 2k + 1 + d(x_k, x_m) - d(y, x_k) \\ &\geqq 2k + 1 + m - k - k = m + 1 , \end{aligned}$$

contradicting the assumption that $\{x_0, x_1, \cdots, x_m\}$ was a longest path in T.

The foregoing observations imply that we may replace each node y in $\mathscr{P}'$ for which $d(x_k, y) \leqq k$ by a node z_y in T' for which $d(x_k, z_y) > k$ and still have a $2k$-packing. We may thus suppose, without loss of generality, that $d(x_k, y) > k$ for every node y in $\mathscr{P}'$; this implies that $d(x_0, y) > 2k$ for every node y in $\mathscr{P}'$. Thus the set $\mathscr{P}' \cup \{x_0\}$ is a $2k$-packing of T and, consequently,

$$P_{2k} \geqq P'_{2k} + 1 . \tag{9}$$

The tree T' has fewer nodes than T so we may assume, as our induction hypothesis, that

$$P'_{2k} \geqq C'_k . \tag{10}$$

It now follows, by inequalities (8), (9), and (10) that $P_{2k} \geqq C_k$, as required, and this completes the proof of the theorem.

Theorems 6 and 7 imply the following result.

COROLLARY 3. *If* $n \geqq k + 1$, *then*

$$P_k + kP_{2k} \leqq n ;$$

if $n > 2k + 1$, *then*

$$C_k + 2kC_{2k} \leqq n .$$

We remark that in general these packing and covering sets are not identical; in particular, for arbitrary k it is easy to construct a tree none of whose largest $2k$-packings are smallest k-coverings. Furthermore, trees are not the only graphs G with the property that $P_{2k}(G) = C_k(G)$ for all k. For example, any graph with a node joined to all the remaining nodes has this property. It seems difficult to characterize such graphs in general.

REFERENCES

1. C. Berge, *The Theory of Graphs and its Applications*, Methuen, London, 1962.
2. ———, *Graphs and Hypergraphs*, North-Holland, Amsterdam, 1973.
3. A. Meir and J. W. Moon, *The expected node-independence number of a random tree*, Nederl. Akad. Wetensch. Proc. Ser. A, **76** (1973), 335-341.
4. ———, *Packing and covering constants for certain families of trees*, (submitted).
5. O. Ore, *Theory of Graphs*, Amer. Math. Soc., Colloq. Publications 38, Providence, 1962.

Received November 11, 1974 and in revised form February 22, 1975. The preparation of this paper was facilitated by grants from the National Research Council of Canada.

UNIVERSITY OF ALBERTA

Printed in Japan by International Academic Printing Co., Ltd., Tokyo, Japan

Pacific Journal of Mathematics

Vol. 61, No. 1 November, 1975