ELSEVIER

Discrete Mathematics 148 (1996) 9–14

DISCRETE MATHEMATICS

The partial gossiping problem☆

Gerard J. Chang *, Yuh-Jiuan Tsay

Department of Applied Mathematics, National Chiao Tung University, Hsinchu 30050, Taiwan, ROC

Received 6 December 1993

Abstract

This paper studies the following variation of the gossiping problem. Suppose there are n persons, each of whom knows a message. A pair of persons can pass all messages they have by making a telephone call. The partial gossiping problem is to determine the minimum number of calls needed for each person to know at least k messages. This paper gives a complete solution to this problem.

1. Introduction

Gossiping and broadcasting problems have been extensively studied for several decades; see [9] for a survey. In these problems, there are n persons, each of whom knows a unique message and is ignorant of the messages of the other people at the beginning. These messages are then spread by telephone calls. In each call, two persons exchange all information they have so far. The *gossiping problem* is to find the minimum number of calls that need to be made for all the people to know all the messages. It has been proved that the solution to the problem is $2n - 4$ for $n \geqslant 4$. For proofs and related topics, see [1–8, 10, 11, 13].

Many variations of the gossiping problem have been studied. Examples include restricting the calls to certain pairs of people, allowing conference calls, allowing only one-way calls, partial gossiping, and set-to-set broadcasting. This paper studies the partial gossiping problem introduced by Richards and Liestman [12]. The problem is to determine the minimum number $P(n, k)$ of calls required for each person to know at least k messages. For the case of $k = n$, the well-known result is

$$P(n,n) = 2n - 4 \quad \text{for } n \geqslant 4.$$

☆ Supported partly by the National Science Council of the Republic of China under grant NSC83-0208-M009-050.

* Corresponding author. E-mail: gjchang@math.nctu.edu.tw

SSDI 0012-365X(94)00257-6

Richards and Liestman [12] determined $P(n,k)$ for $k \leqslant 3$ and gave upper bounds for $k \geqslant 4$. This paper gives a complete solution to the problem.

2. Partial gossiping

We represent the n persons by the set $V = \{1, 2, \ldots, n\}$. To any sequence of calls

$$c(1), c(2), \ldots, c(t)$$

between these n persons, there corresponds a multigraph G_c whose vertex set is V and edge set contains these t calls. From now on, persons and vertices (resp. calls and edges) will be treated as interchangeable.

To establish the solution to $P(n,k)$, we first consider the following upper bounds.

Lemma 1. *$P(n,k) \leqslant \lceil \frac{2^{k-1}-1}{2^{k-1}} n \rceil$ for $2^{k-1} \leqslant n$.*

Proof. Write $n = 2^{k-1}n_1 + n_2$, where $1 \leqslant n_1$ and $0 \leqslant n_2 \leqslant 2^{k-1} - 1$. Partition V into $n_1 + 1$ disjoint sets $V_i = \{v_{i,1}, v_{i,2}, \ldots, v_{i,2^{k-1}}\}$, $1 \leqslant i \leqslant n_1$, and $V' = V - \bigcup_{1 \leqslant i \leqslant n_1} V_i$. For each i, make the following calls in $k-1$ iterations. In iteration r, $0 \leqslant r \leqslant k-2$, $v_{i,s}$ calls $v_{i,s+2^r}$ for $1 \leqslant s \leqslant 2^r$. 2^r calls are made in iteration r, and upon the completion of iteration r the first 2^{r+1} persons in V_i all know exactly $r+1$ messages. So, at the completion of these $k-1$ iterations, a total of $2^{k-1} - 1$ calls have been made and all persons in V_i know exactly k messages. Finally, each person in V' calls $v_{1,1}$ to learn k messages. Thus,

$$P(n,k) \leqslant (2^{k-1} - 1)n_1 + n_2 = \left\lceil \frac{2^{k-1} - 1}{2^{k-1}} n \right\rceil. \qquad \square$$

Lemma 2. *$P(n,k) \leqslant n + i$ if $0 \leqslant i \leqslant k-4$ and $i + 2^{k-i-2} \leqslant n$.*

Proof. Choose two disjoint subsets X and Y of V as follows:

$$X = \{x_1, x_2, \ldots, x_i\} \quad \text{and} \quad Y = \{y_1, y_2, \ldots, y_{2^{k-i-2}}\}.$$

Since $i \leqslant k-4$, $|Y| \geqslant 4$. Make the following calls in $k-2-i$ iterations. In the 0th iteration, each person in X calls y_1. In this iteration i calls are made and upon the completion of this iteration y_1 knows $i+1$ messages. In iteration 1, y_1 calls y_2, y_3 calls y_4, y_1 calls y_3, and y_2 calls y_4. Four calls are made in this iteration and at the completion of this iteration the first four persons of Y all know $i+4$ messages. In iteration r, $2 \leqslant r \leqslant k-i-3$, y_s calls y_{s+2^r} for $1 \leqslant s \leqslant 2^r$. In this iteration 2^r calls are made and at the completion of this iteration the first 2^{r+1} persons of Y all know exactly $i+r+3$ messages. So, at the completion of these $k-i-2$ iterations, a total of $i + 2^{k-i-2}$ calls have been made and all persons in Y know exactly k messages. Finally, each person in $V - Y$ calls y_1 to learn k messages. Thus,

$$P(n,k) \leqslant i + 2^{k-i-2} + (n - 2^{k-i-2}) = n + i. \qquad \square$$

Next, we shall establish lower bounds for $P(n,k)$.

Lemma 3. *Suppose c is a call sequence on V and T is a component of G_c that is a tree. If every vertex in T (respectively, except possibly one) knows at least k messages, then T has at least 2^{k-1} (respectively, $2^{k-1}-1$) vertices.*

Proof. The lemma is trivial for $k=1$. Suppose it is true for $k'=k-1$. Let (x,y) be the first call of c that is in T. Let c' be the call sequence that results from removing (x,y) from c. $T-(x,y)$ has exactly two components, T_x and T_y, such that x is in T_x and y is in T_y. T_x and T_y are the only two components of $G_{c'}$.

Suppose every vertex in T knows at least k messages. Since every vertex in T_x learns messages only from y or vertices in T_x through edges in T_x, every vertex in T_x under the call sequence c' knows at least $k-1$ messages. By the induction hypothesis, T_x has at least 2^{k-2} vertices. Similarly, T_y has at least 2^{k-2} vertices and so T has at least 2^{k-1} vertices.

Suppose every vertex of T, except possibly one, knows at least k messages. Similar to the above arguments, either T_x has at least 2^{k-2} vertices and T_y has at least $2^{k-2}-1$ vertices, or vice versa. So, T has at least $2^{k-1}-1$ vertices. □

Lemma 4. $P(n,k)\geqslant\lceil\frac{2^{k-1}-1}{2^{k-1}}n\rceil$ *for* $k\geqslant 1$.

Proof. Suppose c is an optimal call sequence for $P(n,k)$. If G_c has n_i components of i vertices for $i\geqslant 1$, then

$$\sum_i in_i = n.$$

Note that each component of i vertices has at least $i-1$ edges. By Lemma 3, each component of $i<2^{k-1}$ vertices is not a tree and so has at least i edges. Thus,

$$P(n,k)\geqslant \sum_{i<2^{k-1}} in_i + \sum_{i\geqslant 2^{k-1}} (i-1)n_i.$$

Since $i-1\geqslant((2^{k-1}-1)/2^{k-1})i$ for $i\geqslant 2^{k-1}$, we have

$$P(n,k)\geqslant\frac{2^{k-1}-1}{2^{k-1}}\sum_i in_i = \frac{2^{k-1}-1}{2^{k-1}}n.$$

Thus, the lemma holds. □

A vertex that knows k messages is called a *k-vertex*. The following *interchange rule* was introduced by Hajnal et al. [7] to prove $P(n,n)=2n-4$ when $n\geqslant 4$. It is also useful in the next lemma. Suppose c is a call sequence in which the a calls $c(i+1),\ldots,c(i+a)$ are vertex disjoint from the succeeding b calls $c(i+a+1),\ldots,$ $c(i+a+b)$. Then we can interchange the order of these two blocks of a and b calls;

i.e., if we make the same calls as in c but in the order

$$c(1), \ldots, c(i), c(i+a+1), \ldots, c(i+a+b),$$
$$c(i+1), \ldots, c(i+a), c(i+a+b+1), \ldots, c(t),$$

then the total information conveyed is exactly the same as for the sequence c. If c' is a sequence of calls obtained from c by a number of interchanges like that just described, we say c' is *equivalent* to c and write $c' \sim c$.

Lemma 5. *Suppose $0 \leqslant i \leqslant k-4$ and $n \leqslant i-2+2^{k-i-1}$. If $c(1), c(2), \ldots, c(i+j)$ is a sequence of $i+j$ calls, then there are at most j k-vertices after these $i+j$ calls. Further, if there are exactly j k-vertices, then there is an equivalent call sequence $c' \sim c$ in which the last j calls*

$$c'(i+1), c'(i+2), \ldots, c'(i+j)$$

are all between k-vertices. Consequently, $P(n,k) \geqslant n+i$ for $0 \leqslant i \leqslant k-4$ and $n \leqslant i-2+2^{k-i-1}$.

Proof. We shall prove the lemma by induction on j. The lemma is true for $1 \leqslant j \leqslant k-i-2$, since each component of G_c has at most $k-2$ edges and then no vertex can receive k messages. We now assume $j \geqslant k-i-1$ and the lemma holds for $j' = j-1$.

Suppose there are $j+1$ k-vertices after the $i+j$ calls. Since the last call $c(i+j)$ can produce at most two k-vertices, it follows from the induction hypothesis that there must be exactly $j-1$ k-vertices $x_1, x_2, \ldots, x_{j-1}$ after the first $i+j-1$ calls and the last call $c(i+j)$ is between two additional k-vertices x_j and x_{j+1}. By the second part of the lemma, we can assume that the $j-1$ calls $c(i+1), c(i+2), \ldots, c(i+j-1)$ are between the k-vertices $x_1, x_2, \ldots, x_{j-1}$. By the interchange rule, the last call $c(i+j)$ could be made before $c(i+1), c(i+2), \ldots, c(i+j-1)$. It follows that after the $i+1$ calls

$$c(1), c(2), \ldots, c(i), c(i+j)$$

there would be two k-vertices x_j and x_{j+1}, which contradicts the fact that $i+1 \leqslant k-3$. This proves that there are at most j k-vertices.

Next, suppose there are exactly j k-vertices but the last j calls of the given sequence c are not all between k-vertices. Choose the maximum index p, $1 \leqslant p \leqslant j$, such that $c(i+p)$ is adjacent to at most one k-vertex. Note that $c(i+p+1), \ldots, c(i+j)$ are all between k-vertices. If $c(i+p)$ is not adjacent to any k-vertex, then by the interchange rule, this call could be made last and there would be j k-vertices after only $i+j-1$ calls

$$c(1), \ldots, c(i+p-1), c(i+p+1), \ldots, c(i+j).$$

This contradicts the induction hypothesis and so we can assume that $c(i+p)$ is adjacent to exactly one k-vertex.

Consider the subgraph G' of G_c induced by the $j-p+1$ edges $c(i+p), \ldots, c(i+j)$. Let C be the component of G' containing the edge $c(i+p)$. Let $c'(1) \equiv c(i+p)$, $c'(2), \ldots, c'(r)$ be the edges of C in the order in which these calls are made and $c''(1)$, $c''(2), \ldots, c''(s)$ be the remaining edges of G' in order. There are two cases.

Case 1: $p = 1$ *and* G' *has only one component, which is* C. In this case $s = 0$ and $r = j$. Since the first i calls of c do not produce any k-vertices, all k-vertices must be in G', i.e., G' is a connected graph with exactly j edges and exactly $j+1$ vertices, all of which except one are k-vertices. Then G' is a tree. Consider the component C' of G_c that contains G'. If there are i' vertices of C' not in G', then there are at least i' edges of C' not in G'. However, G_c has only i edges not in G', so $i \geqslant i'$. Delete these edges from c to get a new call sequence c'. G' is a component of $G_{c'}$ in which each vertex except one knows at least $k-i'$ messages under c'. By Lemma 3, $j+1 \geqslant 2^{k-i'-1}-1$. Therefore,

$$i-2+2^{k-i-1} \geqslant n \geqslant i'+j+1 \geqslant i'-1+2^{k-i'-1} \geqslant i-1+2^{k-i-1},$$

a contradiction. Note that the last inequality follows from the fact that for $0 \leqslant i \leqslant k-4$, $i-1+2^{k-i-1}$ is decreasing in i.

Case 2: $p > 1$ *or* G' *has at least two components.* In this case $p > 1$ or $s \geqslant 1$, and then $r < j$. By the interchange rule, $c''(1)$ can be made before all calls in C and similarly for $c''(2), \ldots, c''(s)$. Thus, the original call sequence is equivalent to the call sequence

$$c(1), c(2), \ldots, c(i+p-1), c''(1), \ldots, c''(s), c'(1), \ldots, c'(r).$$

Since $c'(1)$ is adjacent to only one k-vertex, the component C contains at most r k-vertices (C has r edges and at most $r+1$ vertices). It follows that after the first $i+j-r$ calls in the above sequence, there are at least $j-r$ k-vertices. Therefore, by the induction hypothesis there must be exactly $j-r$ such k-vertices (and the component C contains exactly r k-vertices) and there is an equivalent re-ordering of these $i+j-r$ calls so that the last $j-r$ calls are between the $j-r$ k-vertices not in C. In this way we obtain an equivalent call sequence

$$c_1(1), \ldots, c_1(i+j-r), c'(1), \ldots, c'(r).$$

Since the $j-r$ calls $c_1(i+1), \ldots, c_1(i+j-r)$ are between k-vertices not in C, they are vertex disjoint from $c'(1), \ldots, c'(r)$. It follows, again by the interchange rule, that an equivalent sequence is

$$c_1(1), \ldots, c_1(i), c'(1), \ldots, c'(r), c_1(i+1), \ldots, c_1(i+j-r).$$

The first $i+r$ calls in the above sequence give rise to the r k-vertices in C. Therefore, by the induction hypothesis, these calls can be rearranged so that the last r calls are between k-vertices. After re-ordering the first $i+r$ calls in this way, we obtain an equivalent call sequence $c' \sim c$ in which the last j calls are between k-vertices. This completes the proof of the lemma. □

Now we are ready to conclude our solutions to $P(n,k)$.

Theorem 6. $P(n,k) = \lceil \frac{2^{k-1}-1}{2^{k-1}} n \rceil$ *for* $2^{k-1} - 1 \leqslant n$.

Proof. The theorem follows from Lemma 1, Lemma 2 with $i = 0$, and Lemma 4. □

Theorem 7. $P(n,k) = n + i$ *for* $0 \leqslant i \leqslant k-4$ *and* $i + 2^{k-i-2} \leqslant n \leqslant i - 2 + 2^{k-i-1}$.

Proof. The theorem follows from Lemmas 2 and 5. □

References

[1] B. Baker and R. Shostak, Gossips and telephones, Discrete Math. 2 (1972) 191–193.
[2] A. Bavelas, Communication patterns in task-oriented groups, J. Acoust. Soc. Amer. 22 (1950) 725–730.
[3] G. Berman, The gossip problem, Discrete Math. 4 (1973) 91.
[4] R. Bumby, A problem with telephones, SIAM J. Algebraic Discrete Math. 2 (1981) 13–19.
[5] N. Cot, Extensions of the telephone problem, Proc. 7th SE Conf. on Combinatorics, Graph Theory and Computing, Utilitas Mathematics, Winnipeg (1976) 239–256.
[6] M.C. Golumbic, The general gossip problem, Tech. Report IBM Research Report RC 4977, IBM, August 1974.
[7] A. Hajnal, E.C. Milner and E. Szemeredi, A cure for the telephone disease, Canad. Math. Bull. 15 (1972) 447–450.
[8] F. Harary and A.J. Schwenk, The communication problem on graphs and digraphs, J. Franklin Inst. 297 (1974) 491–495.
[9] S.M. Hedetniemi, S.T. Hedetniemi and A.L. Liestman, A survey of gossiping and broadcasting in communication networks, Networks 18 (1988) 319–349.
[10] D.J. Kleitman and J.B. Shearer, Further gossip problems, Discrete Math. 30 (1980) 151–156.
[11] W. Knodel, New gossips and telephones, Discrete Math. 13 (1975) 95.
[12] D. Richards and A.L. Liestman, Generalizations of broadcasting and gossiping, Networks 18 (1988) 125–138.
[13] R. Tijdeman, On a telephone problem, Nieuw Arch. Wisk. 19 (1971) 188–192.