Better bounds for perpetual gossiping

A.D. Scott

Department of Pure Mathematics
and Mathematical Statistics,
University of Cambridge,
16 Mill Lane, Cambridge, CB2 1SB, England.

Abstract. In the perpetual gossiping problem, introduced by Liestman and Richards, information may be generated at any time and at any vertex of a graph G; adjacent vertices can communicate by telephone calls. We define $W_k(G)$ to be the minimum w such that, placing at most k calls each time unit, we can ensure that every piece of information is known to every vertex within w time units of its generation. Improving upon results of Liestman and Richards, we give bounds on $W_k(G)$ for the cases when G is a path, cycle or hypercube.

§1. Introduction

In gossiping problems, each vertex of a graph knows a different piece of information which must be transmitted by telephone calls (along the edges of the graph) to every other vertex. Each telephone call involves exactly two vertices, each of which learns all the information known by the other vertex. A typical gossiping problem asks for the minimum number of calls required for every vertex to learn the information known to every other vertex; it has been shown by various authors (see [1], [3]) that $2n-4$ calls are required for the complete graph on n vertices (this is sometimes known as the 'gossiping dons' problem; see [2]). For a survey on gossiping and related problems see Hedetniemi, Hedetniemi and Liestman [3].

In the *perpetual* gossiping problem, information may be generated at any time and at any vertex of a graph and must be communicated to the rest of the graph as quickly as possible. More formally, information may be generated at any set of vertices at the beginning of each time unit, and calls are made during the time unit (we may assume that it is generated at every vertex at the beginning of each time unit). A *perpetual gossip scheme* for a graph G is a sequence $(E_i)_{i=1}^{\infty}$, where E_i is an independent set of edges in G (each vertex can be involved in at most one call per time unit); $(E_i)_{i=1}^{\infty}$ is a *k-call perpetual gossip scheme* if in addition $|E_i| \leq k$ for every i (at most k calls are made each time unit). A piece of information generated at vertex v at the beginning of time unit $i+1$ is *known to* vertex v' by time $i+w$ iff there is a sequence $\langle e_1, t_1\rangle, \ldots, \langle e_s, t_s\rangle$ such that $i+1 \leq t_i < \cdots < t_s \leq i+w$, e_j is an edge in E_{t_j} for $j = 1, \ldots, s$, and $e_1 \ldots e_s$ is a path from v to v'. If it is defined, we say that a perpetual gossip scheme P has *gossip window of size w* iff w is the smallest integer such that, for every i, every piece of information generated by time $i+1$ is known to every vertex by time $i+w$. It is easily seen that if, for a graph G, there is a k-call perpetual gossip scheme P with gossip window of size w, then there is a k call perpetual gossip scheme P' that has the same window size and is also periodic.

In this paper we consider the problem, introduced by Liestman and Richards [4], of determining the smallest window size of a k-call perpetual gossip scheme for a fixed graph G. Given a graph G, we define $W_k(G)$ to be the smallest integer w such that there is a k-call perpetual gossip scheme P for G with gossip window of size w. Liestman and Richards [4] gave bounds for $W_k(G)$ when G is a path, cycle, hypercube or complete graph. In this paper we give substantial improvements on some of these bounds. In particular, we determine $W_k(P_n)$ to within an additive constant, sharpen the lower bound on $W_k(C_n)$ and give asymptotically best possible bounds on $W_k(Q_n)$ for $k = o(2^n/n)$.

A lower bound on $W_k(G)$ is clearly given by $W_k(G) \geq \mathrm{diam}(G)$. As we shall remark below, for paths, cycles and hypercubes $W_k(G)$ is very close to $\mathrm{diam}(G)$; in fact, we have $W_k(G) \leq \mathrm{diam}(G) + 1$ for $k \geq n/2$.

We shall write $\langle e, t\rangle$ for a call made along edge e at time t; we say that $\langle e, t\rangle$ *carries* a piece of information a if one of the vertices of e knows a by time t.

We use standard notation [2]. We shall write P_n (C_n) for the path (cycle) on n vertices and Q_d for the cube on 2^d vertices.

§2. Paths

For $k \geq \lceil (n-1)/2 \rceil$, the path P_n satisfies $w_k(P_n) = 1$ (colour the edges of P_n alternately red and blue; the call scheme is obtained by alternating between all red and all blue edges). The range of interest is thus $k \leq \lceil (n-1)/2 \rceil$.

Liestman and Richards [4] prove that, for $n \geq 3$,

$$W_1(P_n) = 3n - 6$$

and, for $n \geq 3$ and $2 \leq k \leq \lceil \frac{n-1}{2} \rceil$,

$$n + \left\lceil \frac{n-1}{k} \right\rceil - 2 \leq W_k(P_n) \leq n + 2\left\lceil \frac{n-2}{k} \right\rceil - 2. \tag{1}$$

We prove that the upper bound is essentially best possible.

Theorem 1. *For $n \geq 3$ and any k,*

$$W_k(P_n) \geq n + \left\lceil \frac{2(n-4)}{k} \right\rceil - 4.$$

Proof. Let P be a path with n vertices, with endvertices A and B, and let the edges from A to B be labelled $1, \ldots, n-1$ in that order. Let C be an optimal

k-call perpetual gossiping scheme for P with gossip window of size $w = W_k(P)$. Let a_t and b_t denote the information generated at the beginning of the tth time unit at A and B respectively. We shall consider only information generated at A and B. Let us first consider information generated at A, and let C_A be a minimal subset of the call scheme C such that, for every t, a_t reaches B by time $t+w$. For every t, let $C_A(a_t)$ be the set of calls in C_A that first carries a_t along each edge. More precisely, $\langle i, s\rangle$ is in $C_A(a_t)$ iff

$$s = \inf\{u : \langle i, u\rangle \in C_A \text{ and } \langle i, u\rangle \text{ carries } a_t\}. \tag{2}$$

Clearly $C_A(a_t)$ is a path from A to B.

Now we claim that, for any s and t, either $C_A(a_s) = C_A(a_t)$ or $C_A(a_s)$ and $C_A(a_t)$ are disjoint. Indeed, suppose that

$$C_A(a_s) = \{\langle 1, s_1\rangle, \ldots, \langle n-1, s_{n-1}\rangle\}$$

and

$$C_A(a_t) = \{\langle 1, t_1\rangle, \ldots, \langle n-1, t_{n-1}\rangle\},$$

and that $s_j = t_j$, with j as small as possible. We shall show that $C_A(a_s) = C_A(a_t)$. Since $\langle j, s_j\rangle = \langle j, t_j\rangle$ carries both a_s and a_t, the call $\langle j+1, \min(s_{j+1}, t_{j+1})\rangle$ also carries both a_s and a_t. It follows from (2) that $s_{j+1} = t_{j+1}$, and so by an inductive argument we have $s_i = t_i$ for $i \geq j$. In particular, a_s and a_t reach B at the same time. Now suppose that $j > 1$. Without loss of generality, we may assume that $s_1 < t_1$. We claim that this contradicts the definition of C_A. Let ${C_A}' = C_A \backslash \langle 1, s_1\rangle$, and suppose that a_r does not reach B by time $r + w$ under the call scheme ${C_A}'$. Now if $\langle 1, s_1\rangle$ is not in $C_A(a_r)$ then $C_A(a_r) \subset {C_A}'$, and a_r reaches B by time $r+w$. Otherwise, $\langle 1, s_1\rangle \in C_A(a_r)$ and so $C_A(a_r)$ and $C_A(a_s)$ coincide in their first call. Thus, as we have shown, $C_A(a_r) = C_A(a_s)$, and so a_r reaches B at the same time as a_s and a_t. However, since $t_1 > s_1$ we have $C_A(a_t) \subset {C_A}'$, and so a_r reaches B in ${C_A}'$, by way of $C_A(a_t)$, at the same time as a_s and a_t, which is the same time that a_r reaches B in C_A Therefore we must have $j = 1$ and so $C_A(a_s) = C_A(a_t)$.

We have shown that the sets $C_A(a_t)$ partition C_A into a collection of paths from A to B. Now a given path from A to B takes time at least $n-1$, so the time between two paths leaving A is at most $w-n+1$. Let us define C_B analogously to C_A: we get a collection of paths from B to A, with at most $w-n+1$ time units between the beginning of two consecutive paths. Consider a path P from A to B in C_A, say starting at time $t+1$. Now P must finish, at the latest, at time $t+w$. Therefore any path in C_B that meets P must start no later than time $t-w+1$ and end no later than time $t+2w$. Suppose P meets p paths $Q_1,\ldots,Q_p$ from C_B. Since these paths are pairwise disjoint, $Q_1,\ldots,Q_p$ must together use $p(n-1)$ calls, all of which must occur between time $t-w+1$ and time $t+2w$. At most $3wk$ calls can occur in this period, so $3wk \geq p(n-1)$ and hence

$$p \leq \frac{3wk}{n-1}.$$

Let $p_0 = 3wk/(n-1)$. Since a path must leave each of A and B at least once every $w-n+1$ time units, and each path meets at most p_0 paths in the other direction, the average number of calls per time unit must be at least

$$\frac{2(n-1)-p_0}{w-n+1}.$$

This quantity must be at most k, and so

$$\begin{aligned} w &\geq n-1+\frac{2(n-1)}{k}-\frac{p_0}{k} \\ &\geq n-1+\frac{2(n-1)}{k}-\frac{3w}{n-1}. \end{aligned}$$

Thus

$$w\left(1+\frac{3}{n-1}\right) \geq n-1+\frac{2(n-1)}{k},$$

and so

$$\begin{aligned} w\left(1-\left(\frac{3}{n-1}\right)^2\right) &\geq \left(1-\frac{3}{n-1}\right)\left(n-1+\frac{2(n-1)}{k}\right) \\ &= n-4+\frac{2(n-4)}{k}. \end{aligned}$$

Hence

$$w \geq n-4+\left\lceil\frac{2(n-4)}{k}\right\rceil.$$

□

The upper bound in (1), which Liestman and Richards obtained by specifying a perpetual gossiping scheme, is probably best possible. This might follow from a more careful version of the argument above.

§3. Cycles

It is easily seen that cycles satisfy $W_k(C_n) \leq \lfloor n/2 \rfloor + 2$ for $k \geq n/2$, by taking a similar construction to that used for paths. Liestman and Richards [4] prove that, for $n \geq 3$,

$$W_1(C_n) = 2n - 3,$$

and for $n \geq 3$ and $2 \leq k \leq \lfloor \frac{n}{2} \rfloor$,

$$\left\lfloor \frac{n}{2} \right\rfloor + \left\lfloor \frac{n-1}{2k} \right\rfloor \leq W_k(C_n) \leq \left\lfloor \frac{n}{2} \right\rfloor + \left\lfloor \frac{n-1}{\lfloor k/2 \rfloor} \right\rfloor - 2 + f,$$

where $f = 0$ if n is even and $f = 2$ if n is odd. (Note that $\mathrm{diam}(C_n) = \lfloor n/2 \rfloor$). A careful examination of their construction for the upper bound shows that, in fact,

$$W_k(C_n) \leq \left\lfloor \frac{n}{2} \right\rfloor + \frac{n}{2\lfloor k/2 \rfloor} + \frac{n}{2\lceil k/2 \rceil} + c, \tag{3}$$

where c is a constant ($c = 3$ will do).

Our aim is to improve the lower bound. We begin with a result valid for all k.

Theorem 2. *For $n \geq 6$ and any k,*

$$W_k(C_n) \geq \left\lfloor \frac{n}{2} \right\rfloor + \frac{n}{k} + O(1).$$

Proof. Let A and B be points on C_n, distance $\lfloor \frac{n}{2} \rfloor$ apart, and let C be an optimal gossiping scheme for C_n with gossip window of size $w = W_k(C_n)$. We would like to be able to identify the two paths between A and B, to get a single path of length $n/2$, and then apply Theorem 1 to get the desired lower bound. However, this approach involves some technical problems: the paths may be different lengths, and (less trivially) a legitimate call scheme in the cycle may correspond to an illegitimate scheme in the path, since we could end up with simultaneous calls on adjacent edges.

This being the case, we instead mimic the method of proof of Theorem 1. Once again, let C_A be a minimal subset of the calls C such that, for every t, a_t reaches B by time $t+w$, and let C_B be defined analogously. A similar argument to that in the proof of Theorem 1 gives us a set of paths from A to B partitioning C_A and a set of paths from B to A partitioning C_B, where each path has length at least $\lfloor \frac{n}{2} \rfloor$. The same set of calculations as before, with $\lfloor n/2 \rfloor$ in place of n, yields

$$\begin{aligned} W_k(C_n) &\geq \left\lfloor \frac{n}{2} \right\rfloor + \left\lceil \frac{2(\lfloor n/2 \rfloor - 4)}{k} \right\rceil - 4 \\ &\geq \left\lfloor \frac{n}{2} \right\rfloor + \frac{n-9}{k} - 4. \end{aligned}$$

□

For $k \geq 5$ we can do rather better than this.

Theorem 3. *For $n \geq 3$ and $k \geq 5$ we have*

$$W_k(C_n) \geq \left\lfloor \frac{n}{2} \right\rfloor + \frac{3n}{2k} + O(1). \tag{4}$$

Proof. We may assume that $n \geq n_0$, for any fixed n_0, adjusting the $O(1)$ term if necessary. For $k \geq 5$ it follows from (3) that $W_k(C_n) \leq \frac{11n}{12} + O(1)$. Let n_0 be large enough such that $W_k(C_n) < \frac{12n}{13}$ for $n \geq n_0$; we shall assume $n \geq n_0$.

As before, let C be an optimal gossiping scheme for C_n with gossip window of size $w = W_k(C_n)$. Let $A_0, \ldots, A_{12}$ be thirteen points spread as evenly as possible around C_n, with A_i closest to A_{i+1} and A_{i-1} for each i (we take $A_0 \equiv A_{13}$). We shall consider only the information generated by A_i, for each i. Let a_t be the information generated at A_1 at time t, and let C be an optimal call scheme for C_n. Since $W_k(C_n) < \frac{12n}{13}$, a_t must first reach A_0 and A_2 along the shortest path to each of these vertices (going round the other way would take too many time units). Let P_1 be the path from A_0 to A_2 containing A_1 and let $\mathcal{C}_1$ be a minimal set of calls from C such that, for every t, a_t reaches each of A_0 and A_2 no later in $\mathcal{C}_1$ than in C. Let $\mathcal{C}_i$ be the analogous set of calls for A_i, for $i = 1, \ldots, 13$. As before, we see that $\mathcal{C}_i$ can be decomposed into a set of paths from A_i to A_{i-1} or from A_i to A_{i+1}.

Now let us consider a particular piece of information a_t. Suppose the first path in $\mathcal{C}_1$ from A_1 to A_0 starting at time t or later begins at time a_{t+t_1} and the first path to A_2 begins at time a_{t+t_2}. It is easily seen that a_t does not reach the whole of C_n before time $t+r$, where

$$r = \frac{n-1+t_1+t_2}{2}, \tag{5}$$

since at time r the paths through A_0 and A_2 have reached at most $r-t_1$ and $r-t_2$ vertices respectively. Now suppose that paths leave A_1 (to A_0 or A_2) on average every s time units (we may assume that this average exists, since we may assume that C is periodic). We claim that there is some t such that, if $t+t_1$ and $t+t_2$ are the starting times of the earliest paths from A_1 to A_0 and from A_1 to A_2 respectively, then $t_1+t_2 \geq 3s-2$. Indeed, suppose paths leave A_1 at times $s_1 \leq s_2 \leq \cdots,$. For $i \geq 1$, let $r_i = s_{i+1} - s_i$, so $r_i \geq 0$. The piece of information a_{i+1} leaves A_1 in one direction no earlier than s_{i+1}, and in the other direction no earlier than s_{i+2}. Thus the sum of the two waiting times is at most $(s_{i+1}-s_i-1)+(s_{i+2}-s_i-1) = 2r_i + r_{i+1} - 2$. Since the average value of r_i is s, the average of $2r_i + r_{i+1} - 2$ is $3s-2$, as claimed.

It follows from (5) that

$$w \geq \frac{n-1+3s-2}{2},$$

and so

$$s \leq \frac{2w-n+3}{3}. \tag{6}$$

Let P be any path from A_1 to A_0 in $\mathcal{C}_1$, and let Q be a path from A_0 to A_1 in $\mathcal{C}_0$ that meets P. Now if P starts at time t then Q must start no earlier than time $t-w+1$ and finish no later than time $t+2w$ (since each path takes no more than w time units). Since there are at most $3kw$ calls made in this time and each path requires at least $\lfloor \frac{n}{13} \rfloor$ calls, P can meet at most p paths, where

$$p \leq 3kw / \left\lfloor \frac{n}{13} \right\rfloor \tag{7}$$

(note that the paths met by P are pairwise disjoint). Now, summing the calls in $\bigcup_{i=1}^{13} \mathcal{C}_i$, it follows from (6) that the average number of calls per time unit is at least

$$13\left(\left\lfloor \frac{n}{13} \right\rfloor - \frac{p}{2}\right) / \left(\frac{2w-n+3}{3}\right).$$

Thus

$$k \geq \frac{39\lfloor \frac{n}{13} \rfloor - \frac{39p}{2}}{2w - n + 3}$$

and so

$$w \geq \frac{n}{2} + \frac{3n}{2k} + O(1),$$

since it follows from (3) and (7) that $p = O(k)$. The assertion of the theorem follows immediately. □

We conjecture that the upper bound given in (3), which follows from a perpetual gossiping scheme given by Liestman and Richards [4], is best possible. In order to prove this it seems necessary somehow to take account of the way that chains of calls running round C_n in opposite directions are 'staggered'.

§4. Hypercubes

Let Q_d denote the d-dimensional hypercube, which has diameter $\mathrm{diam}(Q_d) = d$. Liestman and Richards [4] prove that, for $d \geq 2$ and $1 \leq k \leq 2^{d-1}$,

$$W_k(Q_d) \leq \min \left\{ (d+1) \left\lceil \frac{2^{d-1}}{k} \right\rceil - 1, 2^{d-1} + \left\lfloor \frac{2^d}{\lfloor k/2 \rfloor} \right\rfloor - 2 \right\}.$$

and

$$W_k(Q_d) \geq \left\lfloor \frac{2^{d-1} - 1}{k} \right\rfloor + \lceil \log_2 k \rceil + \left\lceil \frac{2^d - 2^{\lceil \log_2 k \rceil}}{k} \right\rceil.$$

We determine the asymptotic value of $W_k(Q_d)$, provided that $k = k(d)$ does not grow too fast.

Theorem 4. *Let $k = k(d)$ satisfy $k(d) = o(2^d/d)$. Then*

$$W_k(Q_d) = (1 + o(1)) \frac{2^{d+1}}{k}.$$

Proof. We begin with the lower bound. Liestman and Richards [4] showed that $W_k(K_n) \geq \lceil \frac{2n-4}{k} \rceil$. Since Q_d can be identified with a subgraph of K_{2^d}, it is clear that

$$W_k(Q_d) \geq W_k(K_{2^d}) \geq \left\lfloor \frac{2^{d+1} - 4}{k} \right\rfloor = (1 + o(1)) \frac{2^{d+1}}{k}.$$

For the upper bound, we construct a perpetual gossiping scheme. One approach would be to split Q_d into many subcubes of equal size and conduct simultaneous Hamiltonian cycles in each of them. This would suffice for gossiping within each subcube, but not for gossiping between subcubes. Our idea is then to insert fairly frequent phases of gossiping between the subcubes such that different Hamiltonian cycles regularly exchange information.

More precisely, suppose $d \geq 3$ and let $i = \lceil \log_2 k \rceil$. Pick h such that $h = o(2^d/kd)$ and $h \to \infty$ as $n \to \infty$. Let $h_j = \lfloor 2^{d-i} j/h \rfloor$, for $j = 0, \ldots, h$. Fixing the first i coordinates, we obtain a $(d-i)$-dimensional subcube: let $R_1, \ldots, R_{2^i}$ be the subcubes obtained in this way. Similarly, let $S_1, \ldots, S_{2^{d-i}}$ be the i-dimensional subcubes obtained by fixing all but the first i coordinates. (Thus we have split $Q_d \cong Q_i \times Q_{d-i}$ into subcubes in two ways.) Let $v_1^{(j)}, \ldots, v_{2^{d-i}}^{(j)}$ be a Hamiltonian cycle in R_j, for $j = 1, \ldots, 2^i$; call this cycle C_j. Picking the Hamiltonian cycles appropriately and relabelling if necessary, we may assume that $V_i^{(j)}$ is the unique vertex in $R_j \cap S_i$. We split each Hamiltonian cycle into h paths by setting $P_s^{(j)}$ to be the portion of C_j from $v_{h_{s-1}}^{(j)}$ to $v_{h_s}^{(j)}$, where we take $h_{2^i} \equiv h_0$.

We construct a call scheme as follows. We begin with simultaneously tracing out the paths $P_1^{(j)}$, for $j = 1, \ldots, 2^i$. If $k = 2^i$ then we can do this easily. In the general case, we 'stagger' the cycles as follows. Let $e_1^{(j)} \ldots e_r^{(j)}$ be the edges of the path chosen in R_j, where $r = h_1 - h_0$. The call scheme is obtained by moving through the sequence $e_1^{(1)}, \ldots, e_1^{(2^i)}, e_2^{(1)}, \ldots, e_2^{(2^i)}, \ldots, e_r^{(1)}, \ldots, e_r^{(2^i)}$, taking k calls each time unit. This takes total time at most $\lceil (h_1 - h_0)2^i/k \rceil + 1$.

We now perform all-to-all gossiping on the i-dimensional cube S_{h_1} of endvertices of the paths $P_1^{(j)}$, in $2d$ time units (this can be done by making calls in S_{h_1} along all edges in a given direction, which takes at most 2 time units, then repeating for the other directions in S_{h_1}). We continue with the paths $P_2^{(j)}$ and then the cube S_{h_2}, and so on.

A vertex can remain free from calls for at most

$$\begin{aligned}\sum_{j=1}^{h} \left(\left\lceil \frac{(h_j - h_{j-1})2^i}{k} \right\rceil + 1 \right) + 2dh &\leq h \frac{(2^{d-i}/h)2^i}{k} + 2h + 2dh \\ &= \frac{2^d}{k} + 2h + 2dh\end{aligned}$$

time units. Once information has been transmitted from a vertex, it is transmitted to every other cycle within

$$\max_j \left(\left\lceil \frac{(h_j - h_{j-1})2^i}{k} \right\rceil + 1 \right) + 2d \leq (1 + o(1)) \frac{2^d}{hk} + 2d$$

time units, and is then transmitted to every other vertex in at most another

$$\frac{2^d}{k} + 2h + 2dh$$

time units. Thus the maximum time for a given piece of information to reach a given vertex is at most

$$\frac{2^{d+1}}{k} + 4h + 4dh + (1 + o(1)) \frac{2^d}{kh} + 2d = (1 + o(1)) \frac{2^{d+1}}{k}.$$

□

References

[1] R. Bumby, A Problem with Telephones, *SIAM J. Alg. Disc. Meth.* **2** (1981), 13-18.

[2] B. Bollobás, *Extremal Graph Theory*, Academic Press, 1978, *xx*+488pp.

[3] S.T. Hedetniemi, S.M. Hedetniemi and A.L. Liestman, A Survey of Broadcasting and Gossiping on Communication Networks, *Networks* **18** (1988), 314-349.

[4] A.L. Liestman and D. Richards, Perpetual Gossiping, *Parallel Processing, to appear.*