To Appear in SPAA 2001, Thirteenth ACM Symposium on Parallel Algorithms and Architectures

Latency Effects on Reachability in Large-scale Peer-to-Peer Networks *

(Extended Abstract)

Fred S. Annexstein, Kenneth A. Berman, Mihajlo A. Jovanović
Department of ECECS, University of Cincinnati, Cincinnati, OH 45221
fred.annexstein@uc.edu, ken.berman@uc.edu, mjovanov@ececs.uc.edu

ABSTRACT

In this paper we study the latency effects introduced in large scale internet applications. In particular, we study the effects of heterogeneous latency on reachability in decentralized, distributed networks operating under flooding protocols. We show that the standard protocol mechanisms of time-to-live (TTL) and unique message identification (UID), used to govern flooding message transmissions, can combine to potentially devastating effect on the reachability of message broadcast. We call this combined effect *short-circuiting*, and we investigate the consequences of this phenomenon of network communications. We show that in the worst case, short-circuiting resulting from heterogeneous latencies can near-completely eliminate the reach of broadcast messages, even with an ability to place $k \geq 1$ broadcast servers optimally. This dramatic negative effect on reachability shows that application protocol designs need to be sensitive to latency models. In addition, we show that short-circuiting can have a negative effect on well known approximation algorithms for maximizing reachability. We show that with respect to distance measures that account for short-circuiting, the standard k-center problem can not be approximated within $n^{1-\epsilon}$, unless $P = NP$. Our theoretical results suggest that it may be quite difficult to find near optimal solutions to certain internet server selection problems. We consider the significance of our results on a large-scale peer-to-peer searching application, that relies on related flooding protocols. We report measurements through experimental studies with both simulated networks and a large network application known as Gnutella. Our empirical results, using statistics obtained from both the simulations and real applications, support the conclusion that, on average, the real effects of short-circuiting are significant, but not devastating to the performance of an overall large-scale system.

*Research supported in part by NSF Grant No. CCR-9877139, NSF Grant No. CISE-9871345, and an Ohio Board of Regents' Research Investment Grant.

1. INTRODUCTION

Recently, the decentralized, distributed computing model called peer-to-peer (P2P) has attracted significant interest among application-level protocol designers [3, 1]. The peer-to-peer framework posits that a large number of autonomous hosts are able to interconnect, and therefore cooperate as a group with little or no centralized control. Without centralized control, the issue of message reachability becomes critical to the design. A number of P2P applications have adopted a flooding mechanism to forward messages in an effort to maximize reachability. In this paper, we study the message flooding operation as it is affected by latencies in the network. We show that standard mechanisms used to control the transmission overheads associated with flooding are significantly affected by heterogeneous latencies, which reside in large-scale P2P applications.

Traditionally, latency has been studied to model network performance as it relates to throughput. Network reachability has traditionally been studied through the analysis of distance in graphs. In this work, we point towards a novel fact that heterogeneous latencies can significantly impact reachability, independent of distance. We demonstrate this phenomenon by considering the interaction of popular mechanisms used to govern message transmissions during network flooding operations.

Flooding enables communication without routing tables or knowledge of the specific topology of the network, and thus is often a design choice for broadcasting in a dynamically changing network. Of course, unrestricted flooding requires tremendous number of message transmissions and network resources, and does not scale well. To address this issue, protocols for flooding are designed with additional governing mechanisms used to limit redundant message sends. There are three standard governing mechanisms. Each mechanism operates by executing a conditional test to determine whether or not a node should continue the flood. A node that continues the flood will forward the message to each of its neighbors, except the one from which the message was received.

Mechanism 1. Time-to-Live Bounds Time-to-Live is a mechanism that prevents messages from traveling farther than a specified number of hops, defined by an initial TTL value. TTL bounds are implemented by including in each message header a TTL value field. When a node receives a message it first checks to see if its TTL value is greater than zero. If it is greater than zero, the node continues the flood with a decremented TTL. Otherwise the message is dropped.

Mechanism 2. Unique Message Identification Unique Message Identification is a mechanism that prevents unique messages from being transmitted more than once from each node. This mechanism is implemented by including in each message header a UID (a unique ID label, or unique sequence number). When a node receives a message it checks to see if it has previously seen that message. If it has, the message is dropped and not forwarded. Otherwise, the node stores the new UID in a local table, and then continues the flood.

Mechanism 3. Path Identification Path Identification is a mechanism that prevents message paths from looping. This mechanism is implemented by including in each message a header that records which nodes of the network have already encountered the message. When a node receives a message it checks the header to see if it has previously seen the message. If it has, the message is dropped and not forwarded. If it has not, the node adds its name to the header, and then continues the flood.

We now describe how latencies in the network can negatively impact the performance of flooding operations using these mechanisms. We define latency as the time it takes a message to traverse a link in the network. We interpret the latency of a message path as the sum of the latencies on the links involved in that path. When Mechanisms 1 and 2 are implemented together they can impact network reachability via a phenomenon we call the *short-circuiting effect*. This effect characterizes a situation where nodes that are within the TTL bound of the broadcast source still fail to receive the message, and is described as follows.

The Short-circuiting Effect. Consider a message broadcast from a source node a, and consider a path P joining nodes $a = u_0$ and $b = u_p$, $P = \{u_0, u_1, \ldots, u_p\}$. It is possible that there may be no throughput of the broadcast messages from a to b along P, even if the hop-length p of the path P is less than or equal to the initial TTL value t. This can result from heterogeneous latencies in the network, as the following scenario shows. Suppose there exists a message path Q of length $q > p$ from a to some intermediate node $x = u_i$ of P, having a strictly smaller latency. Then a broadcast message originating from a, and following path P will be killed (by Mechanism 2) when it reaches x, since it is the duplicate of the earlier arriving message originating from a, but following path Q. Notice that there may also be no throughput along path R consisting of the path Q together with the subpath of P from x to b. This effect results from the fact that R may have a hop-length strictly greater than t, and hence, by Mechanism 1 there is no throughput of the broadcast message originating at a along path R. And, indeed, there may be no throughput of the broadcast message along *any path* from a to b. We use the term *short-circuiting* to refer to this negative impact on reachability caused by heterogeneous latencies.

For the remainder of the paper we will consider broadcasts as operating under the combination of Mechanisms 1 and 2. Note that short-circuiting like effects can not be caused by the combination of Mechanisms 1 and 3, since, in that case, all loop-free paths within the TTL bound are valid message paths.

Somewhat surprisingly, short-circuiting effects can actually decrease the reachability of the broadcast messages when larger TTL values are used. And, in the worst case, short-circuiting can catastrophically reduce reachability, as the following example shows.

PROPOSITION 1. *For each $2 < t < n$, there is an $n+1$-node network G, with edge latencies, and a source vertex v_0 with the property that (a) a broadcast with an initial TTL value of $2 \le t' < t$, will reach at least $n - t + t'$ nodes of G, and (b) a broadcast with an initial TTL value of t, will reach only t nodes of G.*

Proof. We construct G as follows. Start with a length-$t+1$ cycle $a = v_0, v_1, \ldots, v_t, v_0$. Add $n - t$ new vertices $u_1, \ldots, u_{n-t}$, each joined by a single edge to v_t. Assign a high latency to the edge $v_0 v_t$, so that it dominates the latency of the t-path $v_0 v_1 \cdots v_t$. Note that with a TTL of $2 \le t' < t$, a broadcast from v_0 will reach all the $n - t$ nodes u_i, in addition to the t' nodes of the t-path. However, because of short-circuiting, a broadcast message from v_0 with TTL of t is able to reach only the t nodes of the cycle. □

The catastrophic potential illustrated by Proposition 1 leads us naturally to consider the possibility of relocating the flooding originator, or *server*, to a better position relative to the latencies. (In the last example, moving the server from v_0 to v_t eliminates the effects of short-circuiting.) Thus, a central issue which we focus on in this paper is the problem of locating a set of servers for implementing a broadcast operation which maximizes the reachable horizon, i.e., the set of reachable nodes in the flooding operation. We will show in this paper that, even provided with the ability of freely locating flooding servers, short-circuiting can still potentially have devastating consequences. In a worst case scenario, short-circuiting can reduce the effective reachable horizon of broadcast messages significantly. We prove the following theorem that gives nearly tight upper and lower bounds on the minimum size of the reachable horizon, caused by short-circuiting, of a broadcast operation for an optimal placement of servers.

THEOREM 1. *Given integers $n, r, k, t \ge 1$, let $\mathcal{H}(n, r, k, t)$ denote the minimum size of the reachable horizon, caused by short-circuiting reductions, for an optimal placement of k servers broadcasting a message in an n-node network of radius-r, with initial TTL value of t. Then,*

$$c_1 n^{\frac{1}{r}} k^{1-\frac{1}{r}} \le \mathcal{H}(n, r, k, t) \le c_2 n^{\frac{1}{r}} (tk)^{1-\frac{1}{r}},$$

where c_1, c_2 are constants independent of n, r, k, t.

The proof of Theorem 1 is presented in Section 3. In addition to the extremal result above, we show that short-circuiting can have a dramatic effect on classical network optimization algorithms associated with server location problems. In particular, we show that the standard network optimization k-center problem is more complicated when applied to short-circuit induced distances. Without heterogeneous latencies, i.e., when short-circuiting is not present, there exist well known efficient 2-approximations for the k-center problem. In contrast, we have the following theorem, the proof of which appears in Section 4.

THEOREM 2. *For an n-node network, the problem of obtaining an $n^{1-\epsilon}$ approximation for the short-circuit k-center problem, for any $\epsilon > 0$, is NP-hard.*

Our theoretical results suggest that it may be quite difficult to find near optimal solutions to certain internet server selection problems. We therefore consider empirical evidence to test the significance of our results. We report a series of experimental results on large-scale broadcasting applications, that rely on related flooding protocols. We provide measurements obtained through experimental studies with both simulated networks and a large network application for peer-to-peer file sharing known as Gnutella. Our empirical results, using the statistics obtained from both the simulations and real applications, support the conclusion that, on average, the real effects of short-circuiting are significant, but not devastating to the performance of an overall large-scale system.

2. THE FORMAL MODEL

In this section we present notation and a formal model for analyzing the effects of latencies on broadcast flooding operations. We model networks by cost-weighted graphs G, where weights denote latency values on communication links (edges). As mentioned above, we consider the latency of paths as the sum of the latencies of the edges on the path. The flooding operation governed by the combination of mechanisms 1 and 2 in a network G is defined by the following protocol regimen. Packets in the network we will denote $p(u, t, h)$, with unique message identifier $UID = u$, initial TTL value $TTL = t$, and current hop-value $HOP = h$. The hop-value denotes the number of hops from the packet's source node. The values of u and t remain constant over the life of a broadcast operation. The h-value is increased with each hop, and is compared with t at each node to test the TTL limit. We will denote a packet (ready for broadcast) originating at node σ, with initial $TTL = t$, by $p(u_\sigma, t, 0)$. The broadcast regimen operates as follows, and defines the valid message paths associated with the transmission of the broadcast packet.

1. Source σ sends $p(u_\sigma, t, 0)$ to all the neighbors of σ, injecting the packet on all links connected to σ at the same time.
2. Nodes process packets on first-come-first-served basis as follows: when a node v receives packet $p(u_\sigma, t, h)$ it checks whether the UID u_σ has been seen previously. If it has, then the packet is dropped with no further processing.
3. If not, then v records u_σ in its local table, and checks whether or not $t = h$. If $t > h$, then v replicates and forwards the message $p(u_\sigma, t, h+1)$ (with incremented hop count) to all its neighbors (except the node from which it received the packet). If $t = h$ then the packet is dropped and not forwarded.

When heterogeneous latencies are introduced into this model, the issue arises of which set of nodes are reachable by the flooding broadcast. To determine reachability it is not sufficient to consider only minimum-cost paths from the source σ. We now define needed terminology.

DEFINITION 1. *For a given latency-weighted network, the t-**horizon** $\mathcal{H}(\sigma, t)$ from a source node σ, is the set of all nodes v which receive a broadcast packet $(p(u_\sigma, t, -))$ from σ with $TTL = t$. The t-**neighborhood** $\mathcal{N}(\sigma, t)$ from a source node σ, is the is the set of all nodes within a hop-distance of t from σ. Likewise, for a set of source nodes S, we denote by $\mathcal{H}(S, t)$ and $\mathcal{N}(S, t)$, respectively, the t-horizon of S and the t-neighborhood of S, respectively, where we assume that the broadcast is initiated by each $\sigma \in S$ simultaneously.*

For a given set of latencies in a network, we can formally characterize the t-horizon $\mathcal{H}(S, t)$ of a set of source nodes S as follows. Let H_0 be the set of nodes reachable by shortest-latency paths from any node in S within the TTL bound. Now let H_1 be the set of nodes reachable by shortest-latency paths from S within the TTL bound that are not short-circuited by nodes in H_0. Continuing iteratively, let H_i be the set of nodes reachable by shortest-latency paths within the TTL bound not short-circuited by any node in the set $H_0 \cup H_1 \cup \cdots \cup H_{i-1}$. The union $\bigcup_{i \geq 0} H_i$ is the t-horizon $\mathcal{H}(S, t)$.

We show that there is an efficient algorithm which will precisely compute the set $\mathcal{H}(S, t)$. Furthermore, we show more generally that the hop-distances involved with broadcasting can be calculated efficiently.

DEFINITION 2. *For each $v \in \mathcal{H}(\sigma, t)$, the short-circuit hop-distance, **SCH-distance** $D_t(\sigma, v)$ is the hop-length of the path to v determined by a broadcast of $p(u_\sigma, t, 0)$. That is, $D_t(\sigma, v)$ is the number of hops h recorded in the header of the first message received by v with $ID = u_\sigma$. Likewise, $D_t(S, v)$ is defined for multi-source sets S.*

We show that with a modification of Dijkstra's algorithm the hop-distances $D_t(S, v)$ can be calculated efficiently. The following procedure computes hop-distances for each node by maintaining a priority queue of latency values $L(v)$, representing the time the message first reaches node v. These values are used to determine the delivery order of messages under the broadcast model.

Procedure **Short-circuit Hop-distance** (G, σ, t)

Input: An ℓ-weighted graph G, source vertex σ, and TTL bound t

Output: The short-circuited hop-distances $D_t(\sigma, v)$ from σ to each vertex v.

Step 1. Initialize $L(\sigma) = D_t(\sigma) = 0$, and for all $v \neq \sigma$ set $L(v) = D_t(v) = \infty$. Mark σ.

Step 2. Greedily choose an unmarked node v, so that $L(u) + \ell(u, v)$ is minimized over all possible pairs (u, v), where $D_t(u) + 1 \leq t$

Step 3. Set $L(v) = L(u) + \ell(u, v)$, set $D_t(v) = D_t(u) + 1$, and Mark v.

Step 4. If any unmarked nodes remain, go to Step 2.

Proof of Correctness. Omitted. □

3. THE EXTREMAL EFFECTS OF SHORT-CIRCUITING

In this section we prove Theorem 1, by proving the upper and lower bounds in two separate lemmas.

3.1 The Upper Bound

LEMMA 1. *Given integers $n, r, k, t \geq 1$,*

$$\mathcal{H}(n, r, k, t) \leq cn^{\frac{1}{r}}(tk)^{1-\frac{1}{r}},$$

for a constant $c \leq 8$.

Proof. We prove this lemma by showing that there exists a latency-weighted n-node network G of radius r, with the property that for any placement of k servers at a subset $S \subset V(G)$, the t-horizon of S is small.

To construct G, consider a latency-weighted graph F_t on $t+1$ vertices $u_0, \ldots, u_t$ defined as follows. F_t consists of a path $P = \{u_0, u_1, \ldots, u_t\}$ of length t joining $a = u_0$ and $b = u_t$, all of whose edges have latency weight zero. Further, we join a to all vertices u of P giving the edge au a latency weight equal to the length of the path joining a to u in the path P. Do the same for vertex b. We call the graph F_t the *t-flat-tire* graph, and we use it a a building block for the construction of G. The t-flat-tire graph F_t has the following strong short-circuiting property.

LEMMA 2. *For any $0 \leq h < t$, a message broadcast in F_t from node a with an initial hop count of h will reach node b with a hop count of precisely t.* □

From Lemma 2 it follows that by linking flat-tire graphs together we can produce large graphs with large short-circuiting effects, independent of broadcast server location. For example, consider a chain linking together n/t copies of the F_t graph. This n-node graph has the property that any choice of k server locations will have a horizon of size at most $3kt$. To prove the upper bound for a general radius r, we consider a specific construction as follows.

Consider a tree T having n_T vertices V. Let V_T denote the vertex set of T. For $v \in V_T$ let d_v denote the degree of vertex v. We define the *t flat tire extension of T*, denoted by $T^{(t)}$ to be the graph obtained from T by replacing each edge of T with a flat tire graph F_t, such that the end nodes of each flat tire are placed at the parent and child nodes of each tree edge. It is easily verified that the number $n = n_{T^{(t)}}$ of nodes of $T^{(t)}$ satisfies

$$n = t(n_T - 1) + 1. \tag{1}$$

The following Lemma characterizes the size of the t-horizon of the graph $T^{(t)}$, and follows as an immediate consequence of Lemma 2.

LEMMA 3. *Suppose $v \in V \subset V_T$ is an end node of a flat tire in $T^{(t)}$. The t-horizon $\mathcal{H}(v, t)$ of v is the union of the nodes in all the t flat tires for which v is an end node, i.e., $\mathcal{H}(v, t) = td_v + 1$. On the other hand, if $v \in V - V_T$, then v is contained in the interior of some flat tire having end nodes a and b, in which case the t-horizon of v is the union of the nodes in all the t flat tires for which either a or b is an end node, i.e., $\mathcal{H}(v, t) = t(d_a + d_b) - t + 1$.* □

Now consider a tree T of depth r, whose root vertex has y children and whose vertices at level i all have x children, $i = 1, \ldots, t-1$. Then, the number n_T of nodes of T satisfies:

$$n_T = 1 + y + yx + \cdots + yx^{r-1} > yx^{r-1} + 1. \tag{2}$$

Substituting equation (1) into inequality (2), we obtain $n = t(n_T - 1) + 1 \geq tyx^{r-1}$. Therefore,

$$1 \leq y \leq nt^{-1}x^{1-r}. \tag{3}$$

Now consider any placement of k servers at a subset S of vertices in $T^{(t)}$. It follows immediately from Lemma 3 that the size of the t-horizon $\mathcal{H}(S, t)$ of S is given by the following.

$$\mathcal{H}(S, t) \leq ty + (2tx + t + 1)k \leq ty + 4tkx. \tag{4}$$

Applying inequality (3) we obtain

$$\mathcal{H}(S, t) \leq ty + 4tkx \leq nx^{1-r} + 4tkx. \tag{5}$$

We now choose x to minimize $\mathcal{H}(S, t)$, i.e., $x = (n/4tk)^{\frac{1}{r}}$. Substituting this choice in (5) we obtain

$$\mathcal{H}(S, t) \leq 2(n^{\frac{1}{r}}(4tk)^{1-\frac{1}{r}}). \tag{6}$$

Note that so long as $n > 4kt$, then $x \geq 1$ and $y \geq 1$. However, for $n \leq 4kt$ then the upper bound follows trivially since the right side is at least n. So Lemma 1 follows. □

3.2 The Lower Bound

LEMMA 4. *Given integers* $n, r, k, t \geq 1$,

$$\mathcal{H}(n, r, k, t) \geq cn^{\frac{1}{r}} k^{1-\frac{1}{r}},$$

for a constant $c \geq 1/4$.

Proof.

We prove this result by showing that in any weighted n-node network G of radius r, there exists a set $S \subset V(G)$ of k servers that have t-horizon that matches the bound. To show this we need only apply the fact that each server can reach all the nodes in its immediate neighborhood which, of course, is independent of t. All nodes in the immediate neighborhood of a server must be reachable in a broadcast, since they cannot be short-circuited.

Now assume that G is an n-node graph of radius r, and let T_G be any radius-r spanning tree of G. We choose a set of k servers as follows. First select one node of minimum radius in T_G. Greedily select the remaining $k-1$ servers by choosing the nodes of highest degree in T_G. Denote by D_0 this set of k nodes. For each $1 \leq i \leq t$, let D_i denote those nodes of T_G at distance exactly i from D_0. Let $d_i = |D_i|$.

LEMMA 5. *The average degree in* T_G *of nodes in* D_0 *is at most* $2(d_1/d_0 + 1)$.

Proof. The sum of degrees of nodes in D_0 can be bounded as follows:

$$\sum_{v \in D_0} deg(v) \leq 2|E(T_G[D_0 \cup D_1])| \leq$$

$$2|V(G[D_0 \cup D_1])| = 2(d_0 + d_1).$$

The first inequality states that the sum of degrees is at most twice the number of edges in the induced subgraph on $D_0 \cup D_1$, which follows from Euler's formula. The second inequality follows from the fact that the induced graph $T_G[D_0 \cup D_1]$ is a forest; and, the last equality follows by definition. The lemma follows immediately by obtaining the average by dividing by d_0. □

Continuing now with the proof of Lemma 4; by the greedy selection of D_0, we have the immediate corollary.

COROLLARY 1. *For each* $v \in V - D_0$, $deg(v) \leq 2(d_1/d_0 + 1)$. □

Let $m = 2(d_1/d_0 + 1)$; from the corollary we have that for each $i \geq 1$,

$$d_i \leq md_{i-1} \leq m^2 d_{i-2} \leq \cdots \leq m^{i-1} d_1$$

Since the radius of D_0 in T is r, we have that

$$n \leq d_0 + d_1 + d_1 m + d_1 m^2 + \ldots + d_1 m^{r-1}$$

So,

$$n \leq d_0 + d_1(rm^{r-1}) \leq (d_0 + d_1)(r2^{r-1}(d_0 + d_1/d_0)^{r-1})$$

and so

$$n \leq 2^{r-1}(r+1)(d_0 + d_1)^r / d_0^{r-1}$$

and hence

$$(d_0 + d_1) \geq (n^{\frac{1}{r}} d_0^{1-\frac{1}{r}}) / (2^{1-\frac{1}{r}}(r+1)^{\frac{1}{r}}) \geq \frac{1}{4} n^{\frac{1}{r}} d_0^{1-\frac{1}{r}}.$$

The proof of the lower bound follows, since, for any $t \geq 1$, the t-horizon of D_0 is $\mathcal{H}(D_0, t) \geq d_0 + d_1 \geq \frac{1}{4} n^{\frac{1}{r}} d_0^{1-\frac{1}{r}}$. □

4. NON-APPROXIMABILITY OF MAXIMIZING T-HORIZON

We show in this section that short-circuiting can have a dramatic effect on classical network optimization algorithms. In addition, we show that short-circuiting can have a negative effect on well known approximation algorithms for maximizing reachability. Often server location problems are well approximated by greedy approaches [4, 6]. Consider the following greedy method.

Procedure **Greedy k-Maximum t-Horizon** (G, k, t)
Input: A graph G, and integers k, t
Output: A k subset of nodes $S \subset V(G)$
Objective: Maximize the set $\mathcal{H}(S, t)$ of nodes reachable from S within hop-distance of t

Initialize L to empty set
For $i = 1$ to k do
Find h_i so that $\mathcal{H}(L \cup h_i, t)$ is maximized
Set $L = L \cup h_i$

THEOREM 3. *Given a network with homogeneous latencies, i.e., all latencies are the same. The algorithm* **Greedy k-Max** t**-Horizon** *yields a polytime constant ratio* $(e/e-1)$*-approximation algorithm.*

Proof. If a network has homogeneous latencies, then no short-circuiting occurs. Therefore the following independence on the t-horizon holds: $\mathcal{H}(\{h_1, h_2\}, t) = \mathcal{H}(h_1, t) \cup \mathcal{H}(h_2, t)$. A well known, and more general, result states that if a set function R is a nondecreasing, submodular function, then the greedy k-covering algorithm yields a constant $(e/e-1)$-approximation algorithm [8]. A set function R is submodular if for all $S \subset T$ and $h \notin T$, then $R(T \cup \{h\}) - R(T) \leq R(S \cup \{h\}) - R(S)$. Since we have shown in **Procedure Short-circuit Hop-distance** that computing the t-horizon from a set S can be computed in polytime, we have that the Greedy k-Maximum t-Horizon is a polytime constant factor approximation algorithm, when applied to networks with homogeneous latencies. □

Although, the submodular property holds in some models of reachability, it does not hold, in general, for heterogeneous latency values where interference arises. We conjecture that no polytime constant factor approximation algorithm exists for the general k-Maximum t-Horizon problem. As evidence of the difficulty of this problem we show now that a closely related problem has a strong non-approximability characterization.

The classical k-center problem is defined as follows: given a weighted graph G and an integer k, select a set of k nodes S which minimizes the maximum (shortest-path) distance from any node in the network to S. We refer to the analogous problem where short-circuit hop distances are considered, as opposed to shortest-path distances, as the *short-circuit k-center* problem, i.e., given a latency-weighted graph G, an integer k, find a set of k nodes S, so that the maximum SCH-distance $D(S, v)$ is minimized. Although the classical k-center problem is NP-hard, it has well known 2-approximation algorithms [4]. In contrast, Theorem 2 shows that no good approximation algorithms are likely to exist for the short-circuit k-center problem.

Proof of Theorem 2. For any $\epsilon > 0$, we exhibit a polynomial time reduction from the dominating number problem, a classical NP-hard problem, to the problem of computing an $n^{1-\epsilon}$ approximation to the short-circuited k-center problem.

Consider an instance of the dominating number problem. We are given a graph $G = (V, E)$ of n vertices, and the problem is to find the size of the minimum size *dominating set*, i.e., a set $D \subset V(G)$ of vertices such that every other vertex is adjacent to a vertex in D.

For an integer m, where $m = n^k$ is polynomial in n (with k to be determined), we construct the graph $\mathcal{G}_m$. To begin, let $\mathcal{G}_m$ contain two copies of the graph G, $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$. For each edge $e = u_1v_1 \in E_1$, subdivide e into $m + 1$ edges by adding m bivalent vertices, i.e., replace e with a path P_e of length $m + 1$ joining u_1 and v_1. Assign each of these new edges on P_e a small latency, i.e., a latency of 0. Now join each of the corresponding vertices $u_2, v_2 \in V_2$ to each of the $m + 2$ verticies of P_e, and give each of these new edges a large latency, i.e., a latency of $M < \infty$. Give each edge in E_2 of G_2 a latency of 1.

Note that the number of vertices of $\hat{n} = |V(\mathcal{G}_m)|$ satisfies:

$$\hat{n} \leq C(n, 2)(2m + 1) + n \leq n^2 m,$$

for any $m \geq n$.

Now choose $m = n^k$, where $k \geq 2/\epsilon - 2$. Hence,

$$m \geq \hat{n}^{1-\epsilon}.$$

For a given subset of the vertices $S \subset V(\mathcal{G}_m)$, define $radius(S)$ to be the maximum SCH-distance $D_t(S, v)$ over all vertices v, i.e., the $radius(S)$ is the smallest integer r such that every vertex in V can be reached by a vertex in S along a shortest-latency path from S in at most r hops.

LEMMA 6. *Given a subset of the vertices $S \subset V(\mathcal{G}_m)$, transform S into a set S' as follows: for each $v \in S \bigcap V_1$, replace v in S with a vertex $v' \in V_2$, where edge $vv' \in E(\mathcal{G}_m)$. Then $radius(S') \leq radius(S)$.*

Proof. Consider the vertex x which realizes the $radius(S')$, i.e., x has maximum SCH-distance to S'. It follows from the construction of S' that the SCH-distance of x to S is at least as large. □

By Lemma 6 we can assume wlog that all selected sets S returned by an approximation algorithm for the short-circuited k-center problem are contained in V_2.

LEMMA 7. *Given a subset of the vertices $S \subset V(\mathcal{G}_m)$, if $radius(S) \leq m$ then S is a dominating set of $\mathcal{G}_m$, and hence a dominating set of G.*

Proof. If S is not a dominating set then there exists a vertex $v \in V_2$ that has no neighbor in S. Hence, the corresponding vertex $v' \in V - 1$ has SCH-distance $D(S, v') \geq m + 1$. □

LEMMA 8. *Given a subset of the vertices $S \subset V(\mathcal{G}_m)$, then $radius(S) = 1$ if, and only if, S is a dominating set of $\mathcal{G}_m$.*

Proof. The if part is immediate. Now, if S is a dominiating set then all the SCH-distances $D(v, S) = 1$, and thus $radius(S) = 1$, since $S \subset V_2$ is assumed and all edges leaving V_2 are uniformly large. □

Now the following polytime procedure will compute the dominating number of a graph G using an m-approximation for the short-circuit k-center problem. First, construct the graph $\mathcal{G}_m$ from G as described above. Then, iteratively, for each k, $0 \leq k \leq n$, call the m-approximation for short-circuit k-center problem on $\mathcal{G}_m$ to obtain set S_k. If $radius(S_k) = 1$ then return that k is the dominating number of G and exit; else return that G has no dominating set of size k, and continue with next value of k.

The correctness of the procedure follows from Lemmas above. The proof follows. □

We conjecture that maximum t-horizon problem is similarly non-approximable.

5. EMPIRICAL STUDIES

Our original interest in the effects of short-circuiting arose out of experimental evidence associated with the performance of a large-scale peer-to-peer file sharing application called Gnutella. Gnutella applies flooding in its search strategy, and this flooding is governed in its current implementation by Mechanisms 1 and 2 as described in Section 1. Gnutella client software generally enforces limits on TTL values to at most 7, and its routing protocol makes it potentially vulnerable to the short-circuiting effects we have described. During experiments that involved crawling and mapping the entire Gnutella network, we noted that the number of reachable hosts reported by a client was substantially less than on off-line analysis of the generated topology map. We consistently noted discrepancies of approximately one half. After conjecturing that short-circuiting may play a substantial role is such discrepancies, we attempted to try to prove this empirically.

Level-1 Short-circuiting. During the month of October 2000, at a time when the system was relatively stable, we devised an experimental method of discovering a subset of the effects of short-circuiting. We call these *level-1 short-circuiting* effects. These are the effects of short-circuiting caused by the paths interfering at the first level of a broadcast tree, that is, in our experiments we compare the 7-horizon of a message broadcast from v with the 6-horizon of distinct message broadcasts from the neighbors of v. The idea is that sending messages with distinct ID labels will prevent them from interfering with each other, and thereby allows us to measure a subset of the total short-circuiting effect. The actual number of hosts reached by the broadcast of the shared message is compared to a union of host sets reached by the set of distinct broadcast messages. More refined estimates of short-circuiting effects are obtained by comparing the hop counts of messages responding to a shared broadcast to the hop counts of messages responding to distinct broadcasts: if the former is larger than the minimum of the later, than we posit that short-circuiting has occurred. The following Figure 1 shows the results of 70 trials of an experiment of this nature. We note that the observed reductions average 55%.

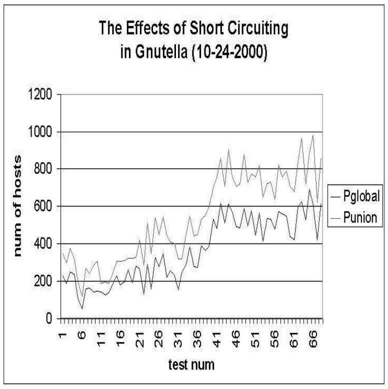

Figure 1: The results of level-1 short-circuiting effects on the broadcast horizon on the Gnutella network, October 2000. The y-axis represents the broadcast horizon size, and the x-axis labels each of 70 broadcast trials. The top line is the resulting horizon from multiple distinct broadcasts from the same source, and the lower line is the resulting horizon from a single broadcast message from a single source. The discrepancy represents "level-1 short-circuiting" effects.

Gnutella is an active project and has changed considerably since October of 2000. Recently we have seen the emergence of a new generation of clients with built-in application logic designed to promote overall network health by conserving bandwidth usage. While such clients have succeeded in allowing the Gnutella network to scale-up, they have also created a serious obstacle to conducting sophisticated experimental studies on the network. We have found it difficult to repeat experiments similar to the one reported in Figure 1 in the current system.

Today it is even more challenging to isolate the effects of short-circuiting on the Gnutella network application than it has in the past. Our difficulties are illustrated by the fact that during March 2001 several identical tests of horizon size, performed consecutively, can differ drastically in their results. Figure 2 shows the size of the broadcast horizon size over time using a client with four connections. Each data point represents the horizon size for a particular broadcast trial, with trials performed consecutively in six minute intervals.

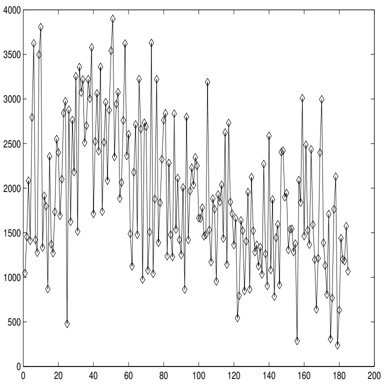

Figure 2: Horizon-size variation over time with broadcasting client using multiple connections on the Gnutella network, March 2001. The y-axis represents the horizon size, and the x-axis labels each of 180 broadcast trials, performed consecutively in six minute intervals.

We attribute these inconsistent results illustrated in Figure 2 to the highly dynamic nature of the network and constantly changing network conditions and topology. We therefore must leave it as open research question to analyze precisely the effects of short-circuiting on an active peer-to-peer network application such as Gnutella.

Because of the difficulties with measuring short-circuiting effects directly on the application, we turned our attention to a series of network simulation studies in which we were able to precisely isolate the effects of short-circuiting on theoretical network topologies.

Network Simulation Studies. Simulated network studies allow us to isolate the effects of short-circuiting on fixed topologies and fixed latency distributions. We report on a series of experiments ranging over a number of graph topologies, including random graphs and structured graphs.

Although we were unable to obtain conclusive evidence on short-circuiting effects on the Gnutella application, we have applied some of the topological data generated during our preliminary investigations. In our initial work we have generated topological "snapshots" of the Gnutella network obtained by a network crawler that performs topology discovery in parallel. Details of the crawler and pictures of the Gnutella topology can be found in [5]. Analysis of the obtained data reveals interesting structural properties of the network, including its strong "small-world" characteristics and a power law distribution of node degrees and latencies. We use the term "small-world" loosely to mean that the network possesses both small diameter and is also highly clustered. Graphs with "small world" properties have received recent interest relative to modeling the Internet and the Web [7, 9, 2]. In our network simulations we have focused attention on the Gnutella snapshots and the popularly used Watts-Strogatz model [9]. For the sake of comparison, we have also run experiments on some well-known regular topologies such as the mesh and the hypercube.

To model network latencies we used several classes of weights representing various commonly used Internet connection bandwidths. We conducted our experiments by using random distributions of these weights.

Observations and Conclusions. Although our theoretical results (Theorem 1) suggest that the effects of short-circuiting can be devastating, our empirical results suggest that in practice these effects are not so pronounced. We have observed the most significant impact on "small-world" topologies such as our Gnutella snapshots and Watts-Strogatz network models. For certain values of t, for these graphs, we have seen average reduction in t-horizon sizes of over 70%, and in the worst case this reduction is as large as 90%, where we are comparing the t-neighborhood size to the t-horizon size. Furthermore, in such graphs the difference between the worst observed t-horizon and the best observed t-horizon can differ by factors of 4 or more.

In our experimental studies we have observed that both random graphs and highly structured graphs such as the mesh and hypercube tend to have, on average, less pronounced short-circuiting effects, as compared with "small-world" graphs.

In general, for a fixed $TTL = t$, the distribution of t-horizon sizes tend to be normally distributed with small variance, independent of network topology. We have also observed that, independent of topology, mean reduction ratios are dependent on the TTL$= t$. Our results suggest that the reduction ratio increases as t increases, until certain thresholds are reached, usually at about the point t is equal to half the network radius or diameter, after which the reduction decreases.

Statistical Tables. We present the statistics of our simulation studies as tables, which report the amount of reduction in the reachability of a broadcast message caused by short-circuiting, given by randomly chosen latencies on a fixed topology. Each table is associated with a fixed topology and a fixed broadcast source node. Each row of the table represents results from 100 trials, each trial using randomly generated latencies. In each row we report for a fixed $TTL = t$, the worst, average, and best observed t-horizon sizes, and t-neighborhood size (which is equal to t-horizon when using uniform latencies). We then give the worst-case reduction ratios (**WRR**), which is obtained by by dividing the worst t-horizon size by the t-neighborhood size, and the mean reduction ratio (**MRR**), which is obtained by dividing the average t-horizon size by the t-neighborhood size. Our tables consider each value of t from 1 up to the radius of the network.

Table 1 shows the results for an instance of the Gnutella snapshot topology with 1125 nodes obtain via a crawl during October 2000. Table 2 represents the results for the Watts-Strogatz small-world topology. The histogram shown in Figure 3 represents the distribution of t-horizon reduction ratio values over the set of many trials using a TTL value of $t = 10$, which is the value of t for which the reduction ratio is observed to be the most severe. Note that the distribution is strongly clustered around the mean of approximately 28%. Table 3 presents the statistical results for a random graph. Similarly, Table 4 and Table 5 shows the results for the 2-D Mesh and the Hypercube topologies, respectively.

6. REFERENCES

[1] *The Gnutella Homepage.* http://gnutella.wego.com.

[2] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener. Graph structures in the web. *Computer Networks*, 33(1-6):309–20, June 2000.

[3] I. Clarke, O. Sandberg, B. Wiley, and T. W. Hong. Freenet: A distributed anonymous information storage and retrieval system. In *ICSI Workshop on Design Issues in Anonymity and Unobservability*, Berkeley, CA, 2000. International Computer Science Institute.

[4] D. Hochbaum, editor. *Approximation Algorithms for NP-Hard Problems.* PWS Publishing Co., 1997.

[5] M. Jovanovic, F. Annexstein, and K. Berman. Scalability issues in large peer-to-peer networks - a case study of gnutella. Technical report, University of Cincinnati, January 2001.

[6] S. Khuller, A. Moss, and J. Naor. The budgeted maximum coverage problem. *Information Processing Letters*, 70(1):39–45, April 1999.

[7] J. Kleinberg. The small-world phenomenon: An algorithmic perspective. Technical report, Cornell University Computer Science Department, October 1999.

[8] G. Nemhauser and L. Woolsey. Maximizing submodular set functions: formulations and analysis of algorithms. In *Studies of Graphs and Discrete Programming*, pages 279–301. North-Holland, Amsterdam, 1981.

[9] D. J. Watts and S. H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393:440–442, June 1998.

7. APPENDIX: STATISTICAL TABLES

TTL	Worst	Avg	Best	Nbhd	WRR	MRR
1	2	2	2	2	100%	100%
2	4	4	4	4	100%	100%
3	10	10	10	10	100%	100%
4	65	92	113	113	58%	81%
5	214	492	689	844	25%	58%
6	246	589	843	1107	22%	53%
7	419	806	1040	1124	37%	72%
8	566	915	1071	1125	50%	81%

Table 1: t-horizon sizes and short-circuiting effects on the Gnutella topology (nodes = 1125, edges = 4080)

TTL	Worst	Avg	Best	Nbhd	WRR	MRR
1	8	8	8	8	100%	100%
2	18	21	24	25	72%	84%
3	24	47	66	69	35%	68%
4	43	84	124	138	31%	61%
5	67	150	238	310	22%	48%
6	121	274	424	678	18%	40%
7	278	498	723	1399	20%	36%
8	434	819	1364	2771	16%	30%
9	765	1388	2307	5018	15%	28%
10	977	2148	3420	7729	13%	28%
11	2030	3153	4549	9449	21%	33%
12	2252	4290	5812	9928	23%	43%
13	3692	5519	6599	9994	37%	55%
14	4995	6392	7563	10000	50%	64%

Table 2: t-horizon sizes and short-circuiting effects on the Watts-Strogatz topology ($nodes = 10000, k = 3, p = 0.2$)

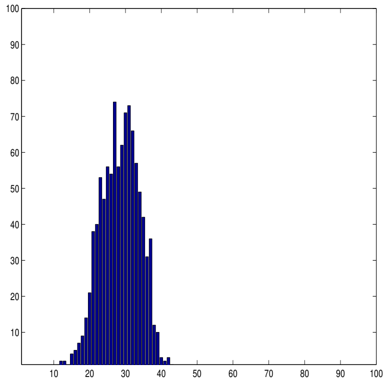

Figure 3: Histogram of trials with $TTL = 10$ from experiment on the Watts-Strogatz topology ($nodes = 10000, k = 3, p = 0.2$), same topology and TTL as used in row 10 of Figure 2. The y-axis represents the number of trials, and the x-axis represents the percentage reduction ratio. Note the narrow variance around the mean (MRR) of approximately 28%.

TTL	Worst	Avg	Best	Nbhd	WRR	MRR
1	6	6	6	6	100%	100%
2	54	54	54	54	100%	100%
3	405	410	419	419	97%	98%
4	1473	2216	2606	2851	52%	78%
5	4686	5986	6875	9021	52%	66%
6	6557	8143	8809	9998	66%	81%
7	8113	9060	9443	10000	81%	91%

Table 3: t-horizon sizes and short-circuiting effects on a random topology (nodes = 10000, edges = 40000)

TTL	Worst	Avg	Best	WRR	MRR
1	3	3	3	100%	100%
2	6	6	6	100%	100%
3	8	9	10	80%	90%
4	11	13	15	73%	87%
5	13	18	21	62%	86%
6	18	23	28	64%	82%
7	21	29	36	58%	81%
8	25	36	45	56%	80%
9	32	44	55	58%	80%
10	40	52	66	61%	79%

Table 4: t-horizon sizes and short-circuiting effects on the 2D-Mesh of size $N = 6X6$.

TTL	Worst	Avg	Best	Nbhd	WRR	MRR
1	14	14	14	14	100%	100%
2	92	92	92	92	100%	100%
3	258	315	368	378	68%	83%
4	685	858	1008	1093	63%	78%
5	1120	1750	2139	2380	47%	74%
6	2243	3079	3544	4096	55%	75%
7	2796	4422	5298	5812	48%	76%
8	3970	5813	6644	7099	56%	82%
9	6023	6844	7424	7814	77%	88%
10	6259	7558	7950	8100	77%	93%
11	6930	7907	8147	8178	85%	97%
12	7877	8108	8187	8191	96%	99%
13	8050	8174	8192	8192	98%	100%

Table 5: t-horizon sizes and short-circuiting effects on Hypercube of size $N = 2^{13}$.