ARTICLE IN PRESS

Discrete Applied Mathematics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

Odd gossiping

Guillaume Fertin [a], Joseph G. Peters [b,*], Lynette Raabe [b], Charlie Xu [b,1]

[a] LINA - Université de Nantes, 2 rue de la Houssinière, B.P. 92208, 44322 Nantes Cedex 3, France

[b] School of Computing Science, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6

ARTICLE INFO

Article history:
Received 3 March 2015
Received in revised form 19 August 2015
Accepted 20 January 2016
Available online xxxx

Dedicated to the Memory of Levon Khachatrian

Keywords:
Gossiping
Information dissemination
Communication networks

ABSTRACT

Gossiping is an information dissemination process in which each node of a communication network has a piece of information that must be received by all other nodes. Most previous work on this problem has concentrated on networks for which the number of nodes n is even. We investigate networks for which n is odd. We use a *linear cost model* of communication in which the cost of communication is proportional to the amount of information transmitted. We study two variants of the problem. In *synchronous* gossiping, the pairwise communications are organized into *rounds* and all communications in a round start at the same time. In *asynchronous* gossiping, a pair of nodes can start communicating while communications between other pairs are in progress. We prove lower bounds on the total time to gossip for both synchronous and asynchronous gossiping. The asynchronous lower bound is achievable for some odd values of n, but we prove that no gossip algorithm for $n = 2^k - 1$ nodes, $k \geq 3$, can achieve the bound. For synchronous gossiping, we present an optimal algorithm for $n = 2^k - 1$. We conjecture that our synchronous lower bound is exact for all odd n.

1. Introduction

Information dissemination in communication networks emerged as a research topic more than 40 years ago. If one fixes the information dissemination process and the communication model, then essentially two main questions arise:

1. Given a network architecture, what is the optimal time to complete the information dissemination process in it?
2. Given the optimal time t_O to complete the information dissemination process in a complete network (i.e., a network having all possible connections), find the network(s) having the least connections, and still achieving a communication time t_O.

Broadcasting and *gossiping* are the most widely studied information dissemination processes in this setting. In the *broadcasting* problem, one node of the network has a piece of information, and the goal is to distribute this information to every other node. This problem has been extensively studied. Levon Khachatrian published four influential papers on this topic, two using the classical "telephone" model [14,15], and two using more specialized models for fault-tolerant broadcasting [1] and messy broadcasting [2]. The two papers using the classical model [14,15] are central to a rich theory of construction methods for *sparse broadcast graphs* which achieve communication time t_O with a small but not necessarily minimum number of connections (see [10] for the history of this research topic).

* Corresponding author.
E-mail addresses: guillaume.fertin@univ-nantes.fr (G. Fertin), peters@cs.sfu.ca (J.G. Peters), charliex@amazon.com (C. Xu).

1 Current address: Department of External Payment Services and Compliance, http://www.Amazon.com, Seattle.

http://dx.doi.org/10.1016/j.dam.2016.01.034
0166-218X/

In the *gossiping* problem, each node of a communication network has a piece of information that must be distributed to all of the other nodes. Gossiping is the generalization of broadcasting in which every node of a network broadcasts its information to every other node in the network. Many methods that have been developed for the broadcasting problem have direct analogs for the gossiping problem and vice versa. In this paper, we concentrate on the gossiping problem.

Like broadcasting, gossiping is a well-studied problem. There are many papers describing algorithms that minimize the gossip time on various interconnection networks such as hypercubes and meshes. See [8,10–13] for surveys of these results. There has been less study of the minimum time needed to gossip when the topology of the interconnection network does not restrict the communication patterns. Knödel [16] proved that the number of *rounds* of communication necessary to gossip is $\lceil \log_2 n \rceil$ when n is even, and $\lceil \log_2 n \rceil + 1$ when n is odd. He also proved sufficiency by describing gossip algorithms that meet the lower bounds on numbers of rounds. The *half-duplex* version of this problem, in which communication links can only be used in one direction at any given time, has also been studied [5,17]. All of these papers assume a *unit cost* model in which a communication takes one time unit independent of the amount of information being transmitted. In most applications, a unit cost model is not appropriate because the amount of information transmitted grows exponentially in most gossip algorithms.

Gossiping has recently been studied in the context of rumour spreading using randomized communication in constrained networks [3]. The potential applications of this approach include multicast and live streaming. However, the models are quite different from the models that we use in this paper. Gossiping has also been used in bioinformatics to study genome rearrangements, where two genomes of distant species are compared and a distance between them is computed [4,18].

In this paper, we use a classical *store-and-forward*, 1*-port, full-duplex* model in which each communication (or *call*) involves two nodes and the single communication link that connects them, each node communicates with at most one other node at any given time, and information can flow simultaneously in both directions along a link. Each node starts with a piece of information of length 1. Information can be concatenated into longer messages and sent in a single communication. We assume a *linear cost* model in which the time to send a message of length ℓ is $\beta + \ell\tau$ where β is the time for the leading edge of a message to propagate along a link from sender to receiver, and $\frac{1}{\tau}$ is the data rate of the link. A call involving messages of length ℓ can be thought of as a start-up period that takes time β followed by a sequence of ℓ *steps* each of which takes time τ. If the two nodes involved in a call send messages of different lengths, then the time for both nodes to complete the call is determined by the length of the longer message.

A linear cost model can be either *synchronous* or *asynchronous*. In the synchronous linear cost model, a gossip algorithm consists of a sequence of *rounds* of simultaneous pairwise communications. All calls in a round start at the same time. Calls in a round may end at different times, depending on the lengths of the messages, but no node can start a new call until all nodes are ready to start new calls. In the asynchronous linear cost model, a call can start as soon as both nodes are ready to communicate. Thus, a pair of nodes can start communicating while calls between other pairs are in progress. The unit cost model is always synchronous because each call takes one time unit.

Fraigniaud and Peters [9] investigated the structure of minimum-time gossip algorithms using a linear cost model. They established lower and upper bounds on the time to gossip when the number of nodes n is even and showed that there is a synchronous minimum-time gossip algorithm for every even n. They also gave examples to show that minimum-time gossip algorithms for some odd values of n must be asynchronous; any synchronous algorithm requires strictly more than minimum time.

In this paper, we study gossiping with n odd and a linear cost model. In the next section, we prove a general lower bound of $(\lceil \log_2 n \rceil + 1)\beta + n\tau$ on the time to gossip for any $\beta \geq 0$ and $\tau \geq 0$. This lower bound holds for all odd n for both the synchronous and asynchronous models. It is achievable in the asynchronous model for some odd values of n, but in Section 4 we prove that every gossip algorithm for $n = 2^k - 1$, $k \geq 3$, requires time strictly greater than $(\lceil \log_2 n \rceil + 1)\beta + n\tau$. For the synchronous model, we prove stronger lower bounds for all odd n, and we give a synchronous algorithm that achieves the lower bound for $n = 2^k - 1$. We conjecture that our synchronous lower bounds are achievable for all odd n.

2. Lower bounds

We say that a node is *idle* at any time that it is not involved in a communication. At any given time during a gossip algorithm for n odd, at least one node will be idle because each call involves a pair of nodes. Based on this fact, Knödel [16] showed that gossiping in the unit cost model requires $\lceil \log_2 n \rceil + 1$ rounds when n is odd. We begin this section with three theorems that establish lower bounds for all odd n in both the synchronous and asynchronous linear cost models. Knödel's result generalizes immediately to the following theorem for a linear cost model with $\beta > 0$ and $\tau = 0$. The only difference in this case is that every round of a gossip algorithm takes time β instead of time 1.

Theorem 1 *([16]). If $\tau = 0$, then $(\lceil \log_2 n \rceil + 1)\beta$ rounds are necessary and sufficient to gossip for any odd $n \geq 3$ and any $\beta > 0$. The bound can be achieved by a synchronous algorithm.*

For a linear cost model with $\beta = 0$ and $\tau > 0$, each node needs $n - 1$ steps to receive the information of all of the other nodes. When n is odd, at least one node is not communicating at any given time, and an argument similar to Knödel's proof establishes the following theorem.

Please cite this article in press as: G. Fertin, et al., Odd gossiping, Discrete Applied Mathematics (2016), http://dx.doi.org/10.1016/j.dam.2016.01.034