Information and Computation 247 (2016) 141–151

Contents lists available at ScienceDirect

Information and Computation

www.elsevier.com/locate/yinco

A distributed enumeration algorithm and applications to all pairs shortest paths, diameter…

Y. Métivier *, J.M. Robson, A. Zemmari

Université de Bordeaux - Bordeaux INP, LaBRI, UMR CNRS 5800, 351 cours de la Libération, 33405 Talence, France

ARTICLE INFO

Article history:
Received 13 October 2015
Received in revised form 20 October 2015
Available online 31 December 2015

ABSTRACT

We consider the standard message passing model; we assume the system is fully synchronous: all processes start at the same time and time proceeds in synchronised rounds. In each round each vertex can transmit a different message of size $O(1)$ to each of its neighbours. This paper proposes and analyses a distributed enumeration algorithm of vertices of a graph having a distinguished vertex which satisfies that two vertices with consecutive numbers are at distance at most 3. We prove that its time complexity is $O(n)$ where n is the number of vertices of the graph. Furthermore, the size of each message is $O(1)$ thus its bit complexity is also $O(n)$. We provide some links between this enumeration and Hamiltonian graphs from which we deduce that this enumeration is optimal in the sense that there does not exist an enumeration which satisfies that two vertices with consecutive numbers are at distance at most 2.
We deduce from this enumeration, algorithms which compute all pairs shortest paths and the diameter with a time complexity and a bit complexity equal to $O(n)$.

1. Introduction

1.1. The problem

In this paper we consider the all pairs shortest paths problem in a distributed network. We assume that there exists a distinguished vertex (called *Leader*) so that there exist distributed algorithms for solving it. We are interested in optimal solutions in time and in number of bits for this problem. The solution presented in this paper is based on a non-trivial distributed enumeration algorithm which satisfies that two vertices with consecutive numbers are at distance at most 3.

1.2. The model

1.2.1. The network

We consider the standard message passing model for distributed computing. The communication model consists of a point-to-point communication network described by a connected graph $G = (V(G), E(G))$ $(= (V, E)$ for short) where the vertices V represent network processes and the edges E represent bidirectional communication channels. Processes communicate by message passing: a process sends a message to another by depositing the message in the corresponding

* Corresponding author.
E-mail addresses: metivier@labri.fr (Y. Métivier), robson@labri.fr (J.M. Robson), zemmari@labri.fr (A. Zemmari).

http://dx.doi.org/10.1016/j.ic.2015.12.004
0890-5401/

channel. In the sequel, we consider only connected graphs. We assume the system is fully synchronous, namely, all processes start at the same time and time proceeds in synchronised rounds.

1.2.2. Time complexity

A round (cycle) of each process is composed of the following three steps: 1. Send messages to (some of) the neighbours, 2. Receive messages from (some of) the neighbours, 3. Perform some local computation. As usual the time complexity is the number of rounds needed until every vertex has completed its computation.

1.2.3. Bit complexity

We follow the definition given in [20]. By definition, in a bit round each vertex can send/receive at most 1 bit from each of its neighbours. The bit complexity of algorithm $\mathcal{A}$ is the number of bit rounds to complete algorithm $\mathcal{A}$.

Remark 1. A round of an algorithm consists of 1 or more bit rounds. The bit complexity of a distributed algorithm is an upper bound on the total number of bits exchanged per channel during its execution. It is also an upper bound on its time complexity.

If we consider a distributed algorithm having messages of size $O(1)$ (and this is the case in this paper) then the time complexity and the bit complexity are equal modulo a multiplicative constant.

The bit complexity is considered as a finer measure of communication complexity and it has been studied for breaking symmetry or for colouring in [5,4] or in [20,11]. Dinitz et al. explain in [11] that it may be viewed as a natural extension of communication complexity (introduced by Yao [38]) to the analysis of tasks in a distributed setting. An introduction to this area can be found in Kushilevitz and Nisan [19].

1.2.4. Network and processes knowledge

The network $G = (V, E)$ is anonymous: unique identities are not available to distinguish the processes. We only assume that there is an elected (a distinguished) vertex denoted *Leader*. We do not assume any global knowledge of the network, not even its size or an upper bound on its size. The processes do not require any position or distance information. Each process knows from which channel it receives or to which channel it sends a message, thus one supposes that the network is represented by a connected graph with a port numbering function defined as follows (where $I_G(u)$ denotes the set of edges of G incident to u):

Definition 1. Given a graph $G = (V, E)$, a *port numbering* function δ is a set of local functions $\{\delta_u \mid u \in V\}$ such that for each vertex $u \in V$, δ_u is a bijection between $I_G(u)$ and $[1, \deg_G(u)]$.

1.2.5. All pairs shortest paths, diameter, girth, cut-edge and cut-vertex

We follow definitions given in [26]. A walk in a graph $G = (V, E)$ is a finite alternating sequence of vertices and edges, beginning and ending with a vertex and where each edge is incident with the vertices immediately preceding and following it. A trail is a walk in which no edge occurs more than once. A path is a trail in which all vertices are different, except that the initial and final vertices may be the same. A walk with at least 3 edges in which the first and last vertices are the same but all other vertices are distinct is called a cycle.

Definition 2. Let $G = (V, E)$ be a connected graph, let $u, v \in V$.

The distance between u and v in G, denoted $dist_G(u, v)$, is the length of a shortest path between u and v in G.

The eccentricity of v is the greatest distance from v to any other vertex.

The diameter of G, denoted $D(G)$, is the maximum distance between any two vertices of G, i.e., $D(G) = \max\{dist_G(u, v) \mid u, v \in V\}$.

The girth of G is the length of a shortest cycle of G.

A cut-vertex is a vertex whose removal increases the number of connected components.

A cut-edge is an edge whose removal increases the number of connected components.

The all pairs shortest paths (APSP for short) problem in G is to compute the length of shortest paths between any pair of vertices in G.

We use trees and we follow the presentation given in [9]. A tree is a connected acyclic graph. A rooted tree is a tree in which one of the vertices is distinguished from the others (called *Leader* in this work). A spanning-tree of a connected graph $G = (V, E)$ is a tree $T = (V, E')$ such that $E' \subseteq E$.

1.3. Our contribution

We present a distributed enumeration algorithm, denoted *DEA*, which assigns to each vertex of a graph G of size n having a distinguished vertex, denoted *Leader*, a unique integer of $\{1, 2, \ldots, n\}$ such that the distance between any two vertices