



ACADEMIC
PRESS

Available at
www.ComputerScienceWeb.com
POWERED BY SCIENCE @ DIRECT®

Information and Computation 184 (2003) 109–128

Information
and
Computation

www.elsevier.com/locate/ic

Analysis of a randomized rendezvous algorithm

Yves Métivier,* Nasser Saheb and Akka Zemmari

LaBRI Université Bordeaux I - ENSEIRB, 351 cours de la Libération, Talence 33405, France

Received 19 April 2002

Abstract

In this paper we propose and analyze a randomized algorithm to get rendezvous between neighbours in an anonymous graph. We examine in particular the probability to obtain at least one rendezvous and the expected number of rendezvous. We study the rendezvous number distribution in the cases of chain graphs, rings, and complete graphs. The last part is devoted to the efficiency of the proposed algorithm.

© 2003 Elsevier Science (USA). All rights reserved.

Keywords: Randomized algorithm; Analysis of algorithm; Efficiency; Graph; Rendezvous

1. Introduction

1.1. The distributed network model

Our model follows standard models for distributed systems given in [15,18]. The communication model is a point-to-point communication network which is represented as a simple connected undirected graph where vertices represent processors and two vertices are linked by an edge if the corresponding processors have a direct communication link. The network is anonymous: unique identities are not available to distinguish processes. Processors communicate by message passing, and each processor knows from which channel it receives a message. We consider the asynchronous model: processors cannot access a global clock and a message sent from a processor to a neighbour arrives within some finite but unpredictable time (asynchronous message passing).

* Corresponding author. Fax: +33-5-56-84-66-99.

E-mail addresses: metivier@labri.fr (Y. Métivier), saheb@labri.fr (N. Saheb), zemmari@labri.fr (A. Zemmari).

1.2. The problem

In synchronous message passing, the sender and the receiver must both be ready to communicate. A communication takes place only if the participant processors are waiting for the communication: this is termed a rendezvous.

Angluin [1] proved that there is no deterministic algorithm to implement synchronous message passing in an anonymous network that passes messages asynchronously (see [18] p. 316). In this paper, we consider the following distributed randomized procedure where every message will be a single bit.

Each vertex v repeats forever the following actions:

the vertex v selects one of its neighbours $c(v)$ chosen at random;

the vertex v sends 1 to $c(v)$;

the vertex v sends 0 to its neighbours different from $c(v)$;

the vertex v receives messages from all its neighbours.

(* There is a rendezvous between v and $c(v)$ if v receives 1 from $c(v)$ *)

A randomized algorithm to obtain rendezvous

1.3. Main results

The aim of this paper is the analysis of this procedure. It is based on the consideration of *rounds*: in order to measure the performance of the algorithm in terms of the number of rendezvous taking place, we assume that at some instant each node sends and receives messages. Thus this parameter of interest, which is the (random) number of rendezvous, is the maximal number (i.e., under the assumption that all nodes are active) authorized by the algorithm.

The first investigations, related to the number of rendezvous, are carried out on the properties of the expected number of rendezvous. We get the asymptotic lower bound $1 - e^{-1/2}$ for the probability of a success in a round. Sharper lower bounds are obtained for the classes of graphs with bounded degrees. As a direct consequence of the definitions, we compute easily the probability of a rendezvous for vertices, from which we derive the expected waiting time for a vertex to get a rendezvous. Elementary computations provide the expected waiting time between two rendezvous for edges. We also study the rather surprising effect of adding a new edge on the number of rendezvous. It is shown that the impact is not monotone. In some cases it is positive in others negative, both on the expected number of rendezvous and on the probability of a success. The asymptotic distribution of the rendezvous number is fully characterized for the class of complete graphs and the class of ring graphs. We use the expected number of rendezvous in a round to define the efficiency of the algorithm. It can be interpreted as the degree of parallelism authorized by the algorithm. We get a lower bound in the case of the trees.

Many problems have no solution in distributed computing [11]. The introduction of randomization makes it possible to address tasks (problems) that admit no deterministic solutions; for instance, the election problem in an anonymous network. The impossibility result on the election problem comes from the fact that the symmetry between the processors cannot be broken in an anonymous network that passes messages asynchronously.

Many papers and results are based on the same model. During a basic computation step, two adjacent vertices exchange their labels and then compute new ones. For example, in [1] an election algorithm is given for complete graphs or in [5,12] election algorithms are given for prime rings (rings having a prime size). In these cases, our randomized algorithm may be considered as a basic step for the

implementation of these algorithms in an anonymous asynchronous system where processors communicate with asynchronous message passing.

General considerations about randomized distributed algorithms may be found in [18] and some techniques used in the design and for the analysis of randomized algorithms are presented in [13,9].

Our paper is organized as follows. Section 2 contains basic notions. Section 3 gives general results. Section 4 studies the probability to get at least one rendezvous in particular cases. Section 5 gives a uniform lower bound for the success probability. Section 6 is devoted to the rendezvous number distribution. Section 7 gives a lower bound for the efficiency of our algorithm for trees.

A part of these results has been presented in [14].

2. Basic notions and notation on graphs

We use the standard terminology of graph theory [3]. A *simple graph* $G = (V, E)$ is defined as a finite set V of *vertices* together with a set E of *edges* which is a set of pairs of different vertices, $E \subseteq \{\{v, v'\} | v, v' \in V, v \neq v'\}$. The cardinality of V is called the *size* of the graph. A *tree* is a connected graph containing no cycle. A subgraph of G is a graph obtainable by the removal of a number of edges and/or vertices of G . The removal of a vertex necessarily implies the removal of every edge incident to it. A connected component of a graph is a maximal connected subgraph. A connected spanning subgraph of a connected graph G is a subgraph of G obtained by removing edges only and such that any pair of vertices remains connected. A spanning tree is a spanning subgraph which is a tree. A *forest* is a graph whose connected components are trees. A *spanning forest* of a graph G is a forest which contains the set of vertices of G . In this paper we consider only simple and connected graphs.

3. First results and remarks

Let $G = (V, E)$ be a connected simple graph of size $n > 1$. The purpose of this section is to provide a formal basis for the randomized procedure described in the introduction and to give simple general results on its analysis.

3.1. Definition and characterization of a call

Definition 1. Let $G = (V, E)$ be a graph. A call over $G = (V, E)$ is a function c from V into itself which maps each $v \in V$ to one of its neighbours.

Let c be a call, according to the definition, there is a rendezvous if and only if there exist two vertices v and w such that $c(v) = w$ and $c(w) = v$.

A call c over $G = (V, E)$ will be a *success*, if there is at least one rendezvous otherwise it will be a *failure*. It is convenient to represent a call c over G by a directed graph $G_c = (V, A)$, where A contains an arc from v to w if and only if $w = c(v)$.

Example 2. Let $V = \{a, b, \dots, q\}$, and let G be the complete graph over V . Fig. 1 represents the graph G_c for the call c over G with $c(a) = b, c(b) = f, c(f) = a, c(g) = f, c(h) = i, c(i) = h, \dots$

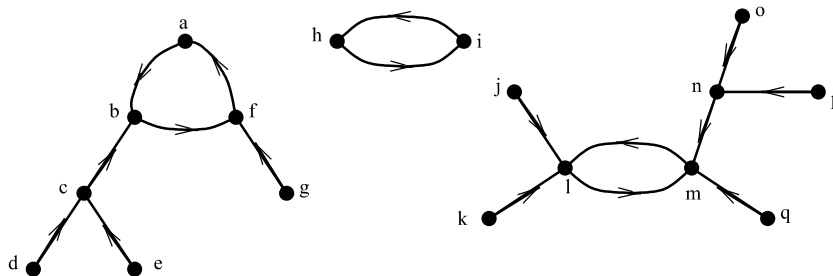


Fig. 1.

For this call there are two rendezvous, one between vertices h and i and one between vertices l and m .

Clearly G_c is a simple graph (i.e., without loop) whose vertices all have outdegree 1. It has, therefore, $n = |V|$ arcs. Moreover it is easy to see that:

Remark 3. Let c be a call over the graph G . Then c is a failure if and only if G_c has no cycle of length 2.

Lemma 4. If $G = (V, E)$ is a tree then any call over G is a success.

Proof. For any c over G , the directed graph G_c has $n = |V|$ arcs. Any arc corresponds to an edge from E , on which an orientation has been chosen. The graph G is a tree having n vertices thus $n - 1$ edges. Since each vertex chooses exactly one edge, at least one edge is chosen by two vertices and the lemma follows. \square

There is a one to one correspondence between calls over the graph G and objects obtained by the following construction:

- (1) let \mathcal{F} be a spanning forest of G such that each tree of \mathcal{F} contains at least two vertices,
- (2) for each tree T of \mathcal{F} , either we add an edge of G which does not belong to T , or we mark an edge of T ,
- (3) if we add an edge e to T then we choose an orientation of the cycle induced by the introduction of e .

A marked edge corresponds to a rendezvous between end-points of the edge. Thus the only vertices of a tree with a marked edge and participating to a rendezvous are the end-points of the marked edge; vertices of trees to which we add an edge have not obtained a rendezvous.

A quasi-tree τ is a connected graph containing exactly one cycle. A quasi-forest is a graph whose connected components are quasi-trees. Let G_c be the directed graph corresponding to a call over G without rendezvous. Let G'_c be the undirected graph associated with G_c . Then G'_c is a spanning quasi-forest of G . Conversely, let ϕ be a spanning quasi-forest of G , let τ be a quasi-tree of ϕ , we choose an orientation of the unique cycle which induces an orientation of each edge of the cycle; the other edges of τ are oriented in such a way that from any vertex of τ there is a path towards the cycle. We obtain a directed graph corresponding to a failure call and conversely any graph G_c of a failure call may be obtained by this way.

If we consider the case of K_n the complete graph of size n , a call corresponds to the combinatorial notion of endofunctions [4].

Let U be a set, an endofunction ψ on U is defined by a subset γ of $U \times U$ verifying:

$$\forall x \in U \quad \exists! y \in U \quad \text{such that} \quad (x, y) \in \gamma.$$

Obviously there is a one-to-one correspondence between a call over K_n and the set of endofunctions on integers of $[1, \dots, n]$ without fixpoint. A call is a success (resp. a failure) if and only if it corresponds to an endofunction without a fix-point containing at least one cycle of length 2 (resp. without cycle of length 2).

Example 5. If we consider the undirected graph corresponding to G_c of Fig. 1, it contains one quasi-tree defined by vertices a, b, c, \dots, g .

3.2. Probability of at least one rendezvous on the graph

We assume that all the adjacent vertices to v have the same chance equal to $1/d(v)$ to be chosen, where $d(v)$ is the degree of the vertex v . Thus any edge $e = \{v, w\} \in E$ has probability $1/d(v)$ to be the bearer of the unique message of v to w . The adjacent vertices v and w are said to *meet* each other, if v and w contact one other: there is a rendezvous. Throughout this study it is assumed that each vertex behaves independently in a memoryless manner.

Each vertex v has $d(v)$ possible choices, consider now the probability measure, which assigns to each call over G the probability $\alpha(G)$ equal to:

$$\alpha(G) = \prod_{v \in V} \frac{1}{d(v)}.$$

Let $s(G)$ be the probability of a success and $f(G)$ that of a failure. From Remark 3, we deduce:

Lemma 6. *We have:*

$$f(G) = \alpha(G)N(G),$$

and

$$s(G) = 1 - \alpha(G)N(G),$$

where $N(G)$ is the number of calls c over G for which G_c has no cycle of length 2.

The probability $f(G)$ may be obtained using quasi-forests. Let $F(G)$ be the set of spanning quasi-forests of G , if ϕ is a spanning quasi-forest of G then $|\phi|$ denotes the number of quasi-trees of ϕ . With these notations and using the characterization of failures by means of quasi-forests we obtain:

$$f(G) = \prod_{v \in V} \frac{1}{d(v)} \left(\sum_{\phi \in F(G)} 2^{|\phi|} \right).$$

In order to get an exact expression for the probability distribution of rendezvous number for a random call, we consider matchings. A matching over $G = (V, E)$ is a subset M of E such that for any pair e and e' in M , $e \cap e' = \emptyset$. To a matching M , we associate the rendezvous corresponding to meetings between end-points of edges of the matching, and this set of rendezvous is by definition the rendezvous over the matching M . Let $e = \{v, w\}$ be an edge, $e^{(1)}$ denotes the event of a rendezvous over e and $e^{(0)}$ the complementary. The probability of a rendezvous over e is

$$Pr(e^{(1)}) = Pr(\{v, w\}^{(1)}) = \frac{1}{d(v)d(w)}.$$

Let $M = \{e_1, \dots, e_k\}$ be a matching, in the same manner the probability $Pr(M)$ of the rendezvous over M is

$$Pr(M) = Pr(e_1^{(1)} \wedge e_2^{(1)} \wedge \dots \wedge e_k^{(1)}) = \prod_{\{v,w\} \in M} \frac{1}{d(v)d(w)} = \prod_{e \in M} Pr(e^{(1)}).$$

For the integer k , a k -matching over G is a matching of cardinality k . Let \mathcal{M}_k denote the set of all k -matchings. Let:

$$q_k = \sum_{M \in \mathcal{M}_k} Pr(M), \quad k = 0, 1, \dots, \lfloor n/2 \rfloor.$$

According to this definition, it should be noted that $q_0 = 1$. By a straightforward application of the Sieve principle ([17] p. 433), we have:

Proposition 7. *Let the sequence q_k , $k = 0, 1, \dots, \lfloor n/2 \rfloor$ be defined as above for the connected graph G of size n . Then, for the integer l in the above stated range, the probability of having exactly l rendezvous over G is:*

$$P_l = \sum_{l \leq i \leq \lfloor n/2 \rfloor} (-1)^{i+l} \binom{i}{l} q_i.$$

And the probability of a success is:

$$s(G) = P_1 = \sum_{0 \leq i \leq \lfloor n/2 \rfloor - 1} (-1)^i q_{i+1}.$$

It is also possible to derive rather simple expressions for the probability $s(G)$ of success and subsequently that of the expected number of necessary calls in the case of special classes of graphs. For instance:

Example 8. Let G be a ring graph (cycle) of size $n \geq 2$. The number $N(G)$, used in Lemma 6, is equal to 2. Hence

$$f(G) = \frac{1}{2^{n-1}},$$

and

$$s(G) = 1 - \frac{1}{2^{n-1}}.$$

The expected number of necessary calls to get a success is thus:

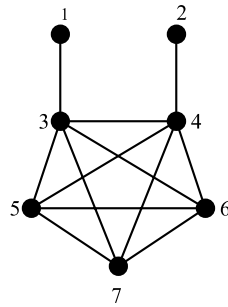


Fig. 2. An example where the addition of an edge may increase the probability of at least one rendezvous.

$$\left(1 - \frac{1}{2^{n-1}}\right) + 2\frac{1}{2^{n-1}}\left(1 - \frac{1}{2^{n-1}}\right) + 3\left(\frac{1}{2^{n-1}}\right)^2\left(1 - \frac{1}{2^{n-1}}\right) + \dots$$

i.e.:

$$\frac{2^{n-1}}{2^{n-1} - 1}.$$

The impact of the addition of an edge on the probability to have at least one rendezvous is not monotone.

- If we add an edge to a tree the probability of at least one rendezvous decreases.
- The graph G of Fig. 2, due to Austinat and Diekert [2], shows that the addition of an edge may increase this probability. In fact, for this graph, we have $s(G) = 1156/1600 = 0.7225$. Let G' denote the graph obtained from G by adding the edge $\{1, 2\}$. Then we have $s(G') = 4742/6400 = 0.7409\dots$

3.3. The expected time between two successive rendezvous for a vertex or over an edge

For a vertex v the probability $p(v)$ of a rendezvous involving this vertex can be computed easily thanks to the independence of the choice for vertices and thanks to the fact that events associated to rendezvous over incident edges are disjoint:

$$pr(v) = \sum_{e \text{ incident with } v} Pr(e^{(1)}),$$

i.e.,

$$p(v) = \frac{1}{d(v)} \sum_{w \text{ adjacent to } v} \frac{1}{d(w)}.$$

From this formula it is clear that $p(v) = 1$ if and only if all w adjacent to v are leaves.

Now, if we consider the successive calls we can define the expected time between two successive rendezvous for a vertex or over an edge.

Given the above relations we can see that the expected time between two successive rendezvous for a vertex v is

$$d(v) \frac{1}{\sum_{w \text{ adjacent to } v} \frac{1}{d(w)}}.$$

Thus, for a graph of degree at most d , the expected time between two successive rendezvous for a vertex is bounded by d .

The expected time between two successive rendezvous on an edge $e = \{v, w\}$ is $d(v)d(w)$. Once more, if we consider a graph of bounded degree this value is bounded by d^2 .

3.4. Expected number of rendezvous

Let X be the number of rendezvous of a call over G , the expected number of rendezvous over G , denoted $\overline{M}(G)$, is $E(X)$: the expected value of X . This parameter may be considered as a measure of the degree of parallelism of the rendezvous algorithm.

For each edge $e \in E$, we define χ_e as follows: if there is a rendezvous between end-points of e then $\chi_e = 1$, else $\chi_e = 0$. We have

$$X = \sum_{e \in E} \chi_e.$$

Now:

$$E(X) = \sum_{e \in E} E(\chi_e),$$

as

$$E(\chi_{\{v,w\}}) = \frac{1}{d(v)d(w)},$$

we get:

$$E(X) = \sum_{\{v,w\} \in E} \frac{1}{d(v)d(w)}.$$

Finally:

Proposition 9. *The expected number of rendezvous over the graph G is:*

$$\overline{M}(G) = \sum_{\{v,w\} \in E} \frac{1}{d(v)d(w)}.$$

Consider the following particular cases.

Example 10. If G is a complete graph of size $n \geq 2$, we have:

$$\overline{M}(G) = \binom{n}{2} \frac{1}{(n-1)^2} = \frac{n}{2(n-1)}.$$

Example 11. If G is a cycle of size $n \geq 2$, we have:

$$\overline{M}(G) = \frac{n}{4}.$$

Example 12. If $G = (V, E)$ has a degree bounded by d then

$$\overline{M}(G) \geq \frac{|E|}{d^2}.$$

If we consider the case of a tree T of size n with a degree bounded by d , we get:

$$\overline{M}(T) \geq \frac{n - 1}{d^2}.$$

In the case of regular graphs of degree d , we have:

$$\overline{M}(T) = \frac{n}{2d},$$

where n is the size of the graph.

We are interested in the impact of the addition of an edge on $\overline{M}(G)$. The above examples illustrate the fact that the number of edges does not necessarily favour the events of rendezvous. Nevertheless, Figs. 3 and 4 shows that the expected number of rendezvous is not monotone with respect to the addition of new edges.

Proposition 13 gives a lower bound for the number of rendezvous expectation.

Proposition 13. For a given fixed positive integer n , the complete graph K_n minimizes the expected number of rendezvous over graphs of size n . The minimal expected value realized by K_n is $n/2(n - 1)$.

Proof. Given a graph $G = (V, E)$, if we denote by $\overline{M}(G)$ the expected number of rendezvous in G , we have

$$\overline{M}(G) = \frac{1}{2} \sum_{v \in V} p(v).$$

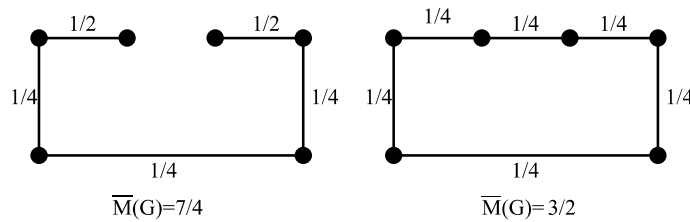


Fig. 3. An example where the expectation number of rendezvous decreases if we add an edge.

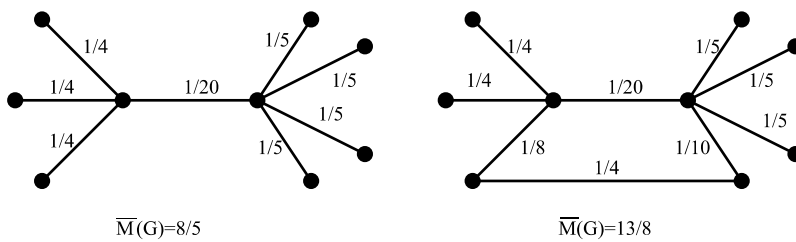


Fig. 4. An example where the expectation number of rendezvous increases if we add an edge.

Where $p(v)$ is as in Section 3.3. Since

$$p(v) = \frac{1}{d(v)} \sum_{\{w,v\} \in E} \frac{1}{d(w)},$$

and $d(w) \leq n - 1$, we have $p(v) \geq 1/(n - 1)$. Summing on all vertices, we get

$$\overline{M}(G) \geq \frac{n}{2(n - 1)}.$$

By Example 10, if G is the complete graph of size n , we have

$$\overline{M}(G) = \frac{n}{2(n - 1)}.$$

The proposition follows. \square

3.5. Expected number of rendezvous in trees

In the previous sections, we have seen that the tree maximizes the success probability, since it is 1. In this section, we are interested in the expected number of rendezvous if the graph $G = (V, E)$ is a tree.

Proposition 14. *Let $T = (V, E)$ be a tree, and $T' = (V', E')$ the tree obtained from T by the addition of a new leaf. If $\overline{M}(T')$ is the expected number of rendezvous over T' and $\overline{M}(T)$ those on T then $\overline{M}(T') \geq \overline{M}(T)$.*

Proof. Let a be the leaf added and v the vertex of T to which a is attached. Let d be the degree of v in T . Then we have

$$\overline{M}(T') = \overline{M}(T) + \frac{1}{d + 1} \left[1 - \frac{1}{d} \sum_{w \text{ adjacent to } v} \frac{1}{d(w)} \right].$$

Since $\sum_{w \text{ adjacent to } v} \frac{1}{d(w)} \leq d$, it follows that $\overline{M}(T') \geq \overline{M}(T)$. \square

By successive applications of Proposition 14, we obtain:

Corollary 15. *Let T be a tree whose maximum degree is k and whose diameter is D . Then*

$$\overline{M}(T) \leq \overline{M} \left(T \left(k, \frac{D}{2} \right) \right),$$

where $T(k, h)$ denotes the k -ary tree whose height is h .

4. Probability of success in particular cases

In this section we study the probability of getting at least one rendezvous over the graph. The considered graphs are special classes of graphs. Let e be an edge, we recall that $e^{(1)}$ denotes the event of a rendezvous over e and $e^{(0)}$ the complementary event.

4.1. Graphs with bounded degrees

We start by the case where $G = (V, E)$ is a graph of degree at most d . The following proposition gives a lower bound on the failure probability.

Proposition 16. *Let $G = (V, E)$ be a d -bounded degree graph, and $s(G)$ denotes its success probability. Then we have*

$$s(G) \geq 1 - \left(1 - \frac{1}{d^2}\right)^{|E|}.$$

Proof. If $E = \{e_1, e_2, \dots, e_m\}$, we have

$$\begin{aligned} f(G) &= Pr(e_1^{(0)} \wedge e_2^{(0)} \wedge \dots \wedge e_m^{(0)}) \\ &= Pr(e_1^{(0)}) Pr(e_2^{(0)} | e_1^{(0)}) Pr(e_3^{(0)} | e_1^{(0)} \wedge e_2^{(0)}) \dots Pr(e_m^{(0)} | e_1^{(0)} \wedge e_2^{(0)} \wedge \dots \wedge e_{m-1}^{(0)}). \end{aligned}$$

Nevertheless, if we did not get a rendezvous on an edge e , this favours the rendezvous on the incident edges and does not change the probability of rendezvous on other edges. To see this, consider an edge $e = \{v, w\}$. Let e_1, \dots, e_k be some edges incident to v and w different from e and f_1, \dots, f_l be some edges not incident to v and w . Clearly we have:

$$Pr(e^{(1)} | e_1^{(0)} \wedge \dots \wedge e_k^{(0)} \wedge f_1^{(0)} \wedge \dots \wedge f_l^{(0)}) = \frac{Pr(e^{(1)} \wedge e_1^{(0)} \wedge \dots \wedge e_k^{(0)} | f_1^{(0)} \wedge \dots \wedge f_l^{(0)})}{Pr(e_1^{(0)} \wedge \dots \wedge e_k^{(0)} | f_1^{(0)} \wedge \dots \wedge f_l^{(0)})}.$$

Since $Pr(e_1^{(0)} \wedge \dots \wedge e_k^{(0)} | f_1^{(0)} \wedge \dots \wedge f_l^{(0)}) \leq 1$, we have:

$$Pr(e^{(1)} | e_1^{(0)} \wedge \dots \wedge e_k^{(0)} \wedge f_1^{(0)} \wedge \dots \wedge f_l^{(0)}) \geq Pr(e^{(1)} \wedge e_1^{(0)} \wedge \dots \wedge e_k^{(0)} | f_1^{(0)} \wedge \dots \wedge f_l^{(0)}).$$

Once there is a rendezvous on the edge e there will be no rendezvous on the edges e_1, \dots, e_k . So we have:

$$Pr(e^{(1)} \wedge e_1^{(0)} \wedge \dots \wedge e_k^{(0)} | f_1^{(0)} \wedge \dots \wedge f_l^{(0)}) = Pr(e^{(1)} | f_1^{(0)} \wedge \dots \wedge f_l^{(0)}),$$

yielding

$$Pr(e^{(1)} | e_1^{(0)} \wedge \dots \wedge e_k^{(0)} \wedge f_1^{(0)} \wedge \dots \wedge f_l^{(0)}) \geq Pr(e^{(1)} | f_1^{(0)} \wedge \dots \wedge f_l^{(0)}).$$

The edges f_1, \dots, f_l and e being not adjacent, the fact that there is no rendezvous on f_1, \dots, f_l does not affect the probability of a rendezvous on the edge e . Therefore:

$$Pr(e^{(1)} | e_1^{(0)} \wedge \dots \wedge e_k^{(0)} \wedge f_1^{(0)} \wedge \dots \wedge f_l^{(0)}) \geq Pr(e^{(1)}).$$

Thus for all $e_j \in E$ $j = 2, \dots, m$:

$$Pr(e_j^{(0)} | e_1^{(0)} \wedge e_2^{(0)} \wedge \dots \wedge e_{j-1}^{(0)}) \leq Pr(e_j^{(0)}).$$

Hence

$$f(G) \leq Pr(e^{(0)})^{|E|}.$$

Since $Pr(e^{(0)}) \leq (1 - (1/d^2))$, we have

$$f(G) \leq \left(1 - \frac{1}{d^2}\right)^{|E|}.$$

The proposition follows. \square

The above bound becomes very interesting if the ratio $|E|$ over d is large. Indeed, the above formula shows that

$$f(G) \leq e^{-\frac{|E|}{d^2}}.$$

In particular in the case of d -regular graphs G , we have $|E| = nd/2$ and therefore:

Corollary 17. *Let G be a d -regular graph, the failure probability $f(G)$ satisfies:*

$$f(G) \leq e^{-\frac{n}{2d}}.$$

4.2. Complete graphs

For this class of graphs we have:

Proposition 18. *Let K_n be the complete graph of size n , then:*

- $s(K_n) = \sum_{k \geq 1} (-1)^{k+1} \frac{n!}{k!2^k(n-2k)!} \frac{1}{(n-1)^{2k}}$,
- $s(K_n)$ is asymptotically $1 - e^{-1/2}$,
- and the expected number of necessary calls to get a success is asymptotically $\sqrt{e}/(\sqrt{e} - 1)$.

Proof. For a fixed k , the k -matchings of K_n all have the same probability

$$q_k = \frac{1}{(n-1)^{2k}}.$$

On the other hand, an easy computation yields the k -matching number of K_n , which is $n!/(k!2^k(n-2k)!)$. We have thus:

$$s(K_n) = \sum_{k \geq 1} (-1)^{k+1} \frac{n!}{k!2^k(n-2k)!} \frac{1}{(n-1)^{2k}}.$$

The expected number of necessary calls to get a success is $1/s(K_n)$, where $s(K_n)$ is given by the above expression. The above expression is difficult to compute, nevertheless, if we use a combinatorial reasoning, we can estimate it asymptotically.

By the definition of a call, a call c over G is a failure, if G_c is without cycle of length 2. A translation of specifications of types shows that the exponential generating function $F(z)$ of the number of such graphs is $F(z) = c(t(z))$, where $t(z) = ze^{t(z)}$ is the EGF (exponential generating function) of the number of labeled trees and $c(z) = (1/(1-z))e^{-z-z^2/2}$ is the EGF of the number of cycles of length at least 3, see [8]. The unique singularity of $F(z) = c(t(z))$ is $z_0 = 1/e$, since $t(z_0) = 1$ if $z_0 = 1/e$.

In [7,8], the authors show that

$$t(z) \sim 1 - 2^{1/2} \sqrt{1 - ez}.$$

Hence,

$$F(z) \sim \frac{1}{\sqrt{2e^3}} (1 - ez)^{-1/2},$$

which yields the number of failure calls $N(K_n)$ on the complete graph K_n :

$$N(K_n) = n! [z^n] F(z) \sim \frac{1}{\sqrt{2e^3}} e^n \frac{1}{\sqrt{\pi n}}.$$

In order to get the failure probability over K_n , we have to divide $N(K_n)$ by $(n - 1)^n$ which is the total number of calls over K_n . Using Stirling formula, we derive

$$f(K_n) \sim e^{-1/2},$$

or

$$s(K_n) \sim 1 - e^{-1/2}.$$

And the expected number of necessary calls to get a success is asymptotically

$$\frac{\sqrt{e}}{\sqrt{e} - 1}. \quad \square$$

5. Uniform lower bound for the probability of a success

Proposition 16 gives a lower bound for the success probability if the graph is of maximum degree d . Corollary 17 shows how this bound is important if d is small enough in comparison with n . But this bound becomes uninteresting if d is too large and $|E|$ is not large enough. It is therefore interesting to find a uniform bound which does not depend on d or on $|E|$. The goal of this section is to give such a bound. Indeed, we have the following theorem:

Theorem 19. *The probability $s(G)$ of a success in a call over any graph $G = (V, E)$ is bounded from below by $1 - e^{-\overline{M}(G)}$, where $\overline{M}(G)$ denotes the expected number of rendezvous in G .*

Proof. By a similar proof for Proposition 16, we have

$$f(G) = Pr(e_1^{(0)} \wedge e_2^{(0)} \wedge \dots \wedge e_m^{(0)}) = Pr(e_1^{(0)}) Pr(e_2^{(0)} | e_1^{(0)}) Pr(e_3^{(0)} | e_1^{(0)} \wedge e_2^{(0)}) \dots Pr(e_m^{(0)} | e_1^{(0)} \wedge e_2^{(0)} \wedge \dots \wedge e_{m-1}^{(0)}).$$

With the same notations and using the same reasoning, we have that

$$f(G) \leq \prod_{i=1}^m (1 - Pr(e_i^{(1)})).$$

By virtue of Proposition 9, we have

$$\sum_{i=1}^m Pr(e_i^{(1)}) = \overline{M}(G).$$

Thus, the bound on $f(G)$ is maximal when

$$Pr(e_i^{(1)}) = \frac{\overline{M}(G)}{m}, \quad \forall i = 1, \dots, m.$$

Hence,

$$f(G) \leq \prod_{i=1}^m \left(1 - \frac{\overline{M}(G)}{m}\right) = \left(1 - \frac{\overline{M}(G)}{m}\right)^m \leq e^{-\overline{M}(G)}.$$

The theorem follows. \square

By Theorem 19 and Proposition 13, we obtain:

Corollary 20. (Robson¹) *The probability $s(G)$ of a success call over any graph $G = (V, E)$ is lower bounded by $1 - e^{-1/2}$.*

Corollary 21. *The expected number of necessary calls to get a success is upper bounded by $\sqrt{e}/(\sqrt{e} - 1)$.*

Remark 22. Since the acceptance of this paper, Martin Dietzfelbinger has proved that the complete graph K_n minimizes the probability of success on graphs of size n .

6. Rendezvous number distribution

The previous study can be refined by determining the distribution of the rendezvous number in a given graph. It seems also interesting to evaluate the asymptotic behaviour of this random variable for graphs of large size. Let us consider for instance star, complete and chain graphs. For the first class of graphs the number of rendezvous is always 1. For the second class it takes value in the integer interval $[0, n/2]$ and its mathematical expectation is $n/2(n-1)$. For the third class it takes value in $[1, n/2]$ with the mathematical expectation $(n+1)/4$. Although the computation of the distribution is feasible in principle, no simple method is available and a standard technique based on a direct numbering is quite complicated.

As a first attempt in this direction we calculate the asymptotic distribution of the rendezvous number for the two extreme cases of complete graphs and chain graphs (which is the same as for ring graphs). In the case of complete graphs (as n grows to infinity), this random number remains an integer finite valued random variable with a distribution which will be determined in the sequel; the same study for chain graphs reveals a quite different behaviour: the expected rendezvous number grows and the normalized distribution tends to a normal one.

¹ A first direct proof of the corollary was provided by J.M. Robson [16].

6.1. Rendezvous number in ring and chain graphs

We start with the study of chain graphs for which the rendezvous number distribution can easily be computed. We then prove that this number for the ring graphs is asymptotically the same as in the case of chain graphs. Let $G = (V, E)$ be a chain of size n and let X_n denote the random variable which counts the number of rendezvous in G . In this section we are interested in the asymptotic distribution of this r.v. It is supposed that all vertices are active. We first prove the following lemma which provides an exact expression for the probability of exactly k rendezvous on the chain graph G .

Lemma 23. *For any integer k , the probability of having exactly k rendezvous is*

$$Pr(X_n = k) = \frac{1}{2^{n-2}} \binom{n-1}{2k-1}.$$

Proof. Let $\phi_n(x)$ be the ordinary probability generating function for the r.v. X_n , i.e., $\phi_n(x) = \sum_{k=0}^{\infty} Pr(X_n = k)x^k$. It is technically convenient to consider also generating function $\psi_{n-1}(x)$ for another r.v. which counts the rendezvous number on the chain graph whenever one of the endpoints is passive. A combinatorial reasoning shows that we have the following recurrences:

$$\begin{cases} \phi_n(x) = \frac{1}{2}\phi_{n-1}(x) + \frac{1}{2}x\psi_{n-1}(x) & \forall n \geq 2 \\ \psi_n(x) = \frac{1}{2}\psi_{n-1}(x) + \frac{1}{2}\phi_{n-1}(x) & \forall n \geq 2, \end{cases}$$

with $\phi_1(x) = \psi_1(x) = 1$.

A straightforward technique provides the solution

$$\phi_n(x) = \frac{2\sqrt{x}}{1+\sqrt{x}} \left(\frac{1+\sqrt{x}}{2} \right)^n - \frac{2\sqrt{x}}{1-\sqrt{x}} \left(\frac{1-\sqrt{x}}{2} \right)^n,$$

from which we get easily $[x^k]\phi_n(x) = (1/2^{n-2})\binom{n-1}{2k-1}$. The lemma follows. \square

The above generating function can be used to compute interesting parameters of X_n . We have in particular:

Corollary 24. *The expectation of the random variable X_n defined above is $(n + 1)/4$ and its variance $(n - 1)/16$.*

More interesting, the generating function $\phi_n(x)$ can be used to show that the asymptotic behaviour of this r.v. is normal. Indeed, we have the following theorem. \square

Theorem 25. *The normalized variable defined by*

$$Y_n = \frac{4X_n - n}{\sqrt{n}},$$

has a distribution which tends to the normal distribution $\mathcal{N}(0, 1)$, i.e., for any real interval $[a, b]$

$$Pr(a < Y_n \leq b) \rightarrow \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx.$$

Proof. For any integer k let $j = 2k - 1$, and $l = n - 2k + 1$. Lemma 23 shows that

$$Pr(X_n = k) \sim \frac{1}{2} \frac{n!}{j!l!} \left(\frac{1}{2}\right)^j \left(\frac{1}{2}\right)^l.$$

Applying the Stirling formula we get

$$Pr(X_n = k) \sim \frac{1}{2} \sqrt{\frac{n}{2\pi jl}} \left(\frac{n}{2j}\right)^j \left(\frac{n}{2l}\right)^l.$$

And

$$\frac{jl}{n} \sim \frac{(\sqrt{nx} + n)(n - \sqrt{nx})}{4n} \sim \frac{n}{4}.$$

Therefore

$$Pr(X_n = k) \sim \frac{1}{\sqrt{2\pi n}} \left(\frac{n}{2j}\right)^j \left(\frac{n}{2l}\right)^l,$$

and

$$\ln \left(\frac{n}{2j}\right)^j \left(\frac{n}{2l}\right)^l \sim -\frac{x^2}{2}.$$

The theorem is now proved by the same reasoning as ([10] p. 22).

The theorem can be also proved by the transformation of the generating function $\phi_n(x)$ into the characteristic function by setting $x = e^{it}$. Indeed, it is easily shown that the characteristic function of the r.v. Y_n defined in the theorem tends to $e^{-(1/2)t^2}$, as n tends to infinity, and this last function is the characteristic function of the normal distribution. Therefore the normalized variable Y_n tends in law to a normal random variable. \square

Let now the r.v. Z_n denote the rendezvous number in a ring graph of size n . We have seen that it is of mathematical expectation $E(Z_n) = n/2$. Indeed we have

Theorem 26. Define the normalized r.v. by $V_n = (4Z_n - n)/\sqrt{n}$. Then, as $n \rightarrow \infty$:

$$Pr(a < V_n \leq b) \rightarrow \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx,$$

for any real interval $[a, b]$.

Proof. Consider a chain graph with vertices labelled $1, 2, \dots, n$ and edges $\{i, i + 1\}, 1 \leq i \leq n - 1$. The rendezvous number over the chain can be written as the sum $X_n = \sum_{i=1}^{n-1} R_i$, where the r.v. R_i is defined by

$$R_i = \begin{cases} 1 & \text{if there is a rendezvous over } \{i, i + 1\} \\ 0 & \text{otherwise.} \end{cases}$$

We have clearly $\sum_{i=2}^{n-2} R_i \leq X_n \leq \sum_{i=2}^{n-2} R_i + 2$. Consider now a ring graph with vertices labelled $1, 2, \dots, n$ and edges $\{i, i + 1\}, 1 \leq i \leq n - 1, \{n, 1\}$. The rendezvous number over the ring is equally written as the $Y_n = \sum_{i=1}^n S_i$, where the r.v. S_i is 1 if there is rendezvous over the edge $\{i, i + 1\}$ and 0

otherwise. We have $\sum_{i=2}^{n-2} S_i \leq Z_n \leq \sum_{i=2}^{n-2} S_i + 3$. Since $\sum_{i=2}^{n-2} R_i = \sum_{i=2}^{n-2} S_i$, we have $|Z_n - X_n| \leq 3$. Hence $|V_n - Y_n| \leq 12/\sqrt{n}$. Thus, as n grows, the difference between the two normalized r.v. tends to 0. The theorem follows. \square

6.2. Rendezvous number in complete graphs

In this section, we are interested in the asymptotic behaviour of the rendezvous number in complete graphs. Here $G = (V, E)$ is a complete graph, X_n the random variable which counts the rendezvous number on G . We prove that, asymptotically, X_n has a Poisson distribution of parameter $1/2$. Indeed, we have:

Theorem 27. *For every positive integer m , the probability for X_n to be equal to m is given by*

$$Pr(X_n = m) \rightarrow \frac{1}{m!2^m} e^{-1/2}, \quad \text{as } n \rightarrow \infty.$$

Proof. By the same argument as in the proof of the first point of Proposition 18 of Section 4.2, and using the Sieve principle [17], we get

$$Pr(X_n = m) = \sum_{k=m}^{\lfloor n/2 \rfloor} (-1)^{k+m} \binom{m}{k} \frac{n!}{k!2^k(n-2k)!} \frac{1}{(n-1)^{2k}}.$$

Using the Stirling formula, we get easily:

$$\begin{aligned} \lim_{n \rightarrow \infty} Pr(X_n = m) &= \sum_{k \geq m} (-1)^{k+m} \binom{m}{k} \frac{1}{k!2^k} \\ &= \frac{1}{m!2^m} \sum_{r \geq 0} (-1)^r \frac{1}{r!2^r} \\ &= e^{-1/2} \left(\frac{1}{2}\right)^m \frac{1}{m!}. \end{aligned}$$

Which yields the assertion. \square

7. Performance analysis

Let G be a graph, recall that a matching in G is a set of pairwise nonadjacent edges in the graph; the matching number of G , denoted $K(G)$, is the cardinality of a matching having the largest size.

Following definitions for approximation algorithms [6]: given any randomized algorithm \mathcal{A} for a rendezvous-type algorithm, its efficiency $\Delta_{\mathcal{A}}(G)$ over a graph G is the ratio

$$\Delta_{\mathcal{A}}(G) = \frac{\overline{M}_{\mathcal{A}}(G)}{K(G)},$$

where $\overline{M}_{\mathcal{A}}(G)$ is the expected number of rendezvous in a round of \mathcal{A} .

To illustrate the definition, reconsider the two extremal cases of complete graphs and star graphs. In both examples the expected rendezvous number is very low:

$$\overline{M}(G) = \binom{n}{2} \frac{1}{(n-1)^2} = \frac{n}{2(n-1)}$$

if G is a complete graph, and $\overline{M}(G) = (n-1)/(n-1) = 1$ if G is a star graph. In fact in the case of complete graphs the matching number is $\lfloor n/2 \rfloor$ which is much greater than the expected number for large values of n . In the second example the algorithm allows an expected number which is the same as the matching number.

The first consequence of the definition is that the proposed algorithm is not efficient for the class of complete graphs. A ‘reasonable’ extension of this fact leads to a similar conclusion for families of densely linked graphs. It seems however that it is not possible to do much better if we restrict the message size to one bit.

An important class of sparsely linked graph is the class of trees. The rest of the section is devoted to the search of a uniform lower bound for the efficiency of the introduced algorithm.

Given a graph $G = (V, E)$ with $|V| = n$ and $|E| = m$, let M be its incidence matrix. For two given enumerations v_1, \dots, v_n and e_1, \dots, e_m of vertices and edges, respectively, it is defined by

$$M[i, j] = \begin{cases} 0 & \text{if } v_i \text{ is not an endpoint of } e_j \\ 1 & \text{if } v_i \text{ is an endpoint of } e_j, \end{cases}$$

for $0 \leq i \leq n$ and $0 \leq j \leq m$. A subset F of E is a matching in G iff its indicator, the m -vector column x

$$x[j] = \begin{cases} 0 & \text{if } x_j \in F \\ 1 & \text{if } x_j \notin F, \end{cases}$$

for $0 \leq j \leq m$, satisfies $Mx \leq \mathbf{1}_n$, $x \in \{0, 1\}^m$, where $\mathbf{1}_n$ denotes the n -vector column $(1, \dots, 1)^T$. Therefore the matching number $K(G)$ is the optimal value of the objective function of the following integer programming problem [17].

$$\begin{aligned} \text{maximize:} & \quad \langle \mathbf{1}_m, x \rangle \\ \text{subject to:} & \quad Mx \leq \mathbf{1}_n \\ & \quad x \in \{0, 1\}^m. \end{aligned}$$

Thus the value of $K(G)$ cannot be greater than the value of the objective function of the same problem whenever one removes the entireness condition $x \in \{0, 1\}^m$. But this value is in turn bounded by the value of the objective function in the following dual problem for any feasible solution (the weak duality theorem, see [17]):

$$\begin{aligned} \text{minimize:} & \quad \langle \mathbf{1}_n, y \rangle \\ \text{subject to:} & \quad M^T y \geq \mathbf{1}_m \\ & \quad y \in \mathbb{R}^n, y \geq 0. \end{aligned}$$

Now, if we add the constraint $y \in \{0, 1\}$, the optimal solution will be the minimal vertex cover in G , see [17]. It follows that $K(G)$ is bounded by the cardinality of any vertex cover in G .

We may now state the main theorem of the section.

Theorem 28. *The efficiency $\Delta_{RV}(T)$ of the proposed randomized algorithm RV over any tree T is strictly greater than $1/3$.*

Proof. By comparison to the objective function in the dual problem, it suffices to prove that the tree T admits a vertex cover of cardinality less than $3\overline{M}(T)$. We assume that this is true for the trees of size less than n and prove it for a tree T of size n . On the other hand, since the theorem holds for the star graphs, we may assume that the diameter of T is greater than 2. We have to show that T admits a cover vertex of cardinality less than $3\overline{M}(T)$. Consider now the tree rooted by one of its centers. Since the diameter is greater than 2, there will be vertices a and a' such that a' is the father of a , all other neighbours of a are leaves and a' is not a leaf in T . Let d and d' be their degrees.

Let T' be the tree induced by removing a and its sons from the rooted tree T . A vertex cover for T can be obtained by adding the vertex a to a vertex cover for T' . Thus if the theorem holds for T' , we have to prove $\overline{M}(T) - \overline{M}(T') \geq 1/3$. A simple computation yields the difference. Indeed, we have

$$\overline{M}(T) - \overline{M}(T') = \frac{d-1}{d} + \frac{1}{dd'} - \sum_{i=1}^{d'-1} \left[\frac{1}{(d'-1)\delta_i} - \frac{1}{d'\delta_i} \right], \tag{1}$$

where δ_i 's are the degrees of the neighbour vertices of a' different from a . The sum

$$\sum_{i=1}^{d'-1} \left[\frac{1}{(d'-1)\delta_i} - \frac{1}{d'\delta_i} \right],$$

can be bounded by $1/d'$ and hence

$$\overline{M}(T) - \overline{M}(T') \geq \frac{d-1}{d} + \frac{1}{dd'} - \frac{1}{d'} = \left(1 - \frac{1}{d}\right) \left(1 - \frac{1}{d'}\right).$$

Now, if at least one of the degrees d or d' is greater than 2, the computed difference will be at least $1/3$. Otherwise $d = d' = 1$ and the difference (2) reduces to $(1/2) + (1/4) - (1/2\delta)$, where δ is the degree of the unique neighbour vertex of a' different from a . If $\delta \geq 2$ then the difference is greater than or equal to $1/2$ and we are done; otherwise the whole tree T is reduced to four vertices and a simple verification proves the theorem. \square

Remark 29. There are trees for which the efficiency is less than $1/2$. Up to now (and to our knowledge) no sharper lower bound (than $1/3$) is known.

Acknowledgments

The authors are grateful to Martin Dietzfelbinger and Anca Muscholl for helpful remarks and corrections. They are also grateful to the anonymous referees for their helpful comments.

References

[1] D. Angluin, Local and global properties in networks of processors, in: Proceedings of the 12th Symposium on Theory of Computing, 1980, pp. 82-93.
 [2] H. Austinat, V. Diekert, private communication.

- [3] C. Berge, *Graphes*, Gauthier-Villars, 1983.
- [4] F. Bergeron, G. Labelle, P. Leroux, *Combinatorial Species and Tree-like Structures*, Cambridge University Press, Cambridge, 1998.
- [5] J.E. Burns, J. Pachl, Uniform self-stabilizing rings, *ACM Trans. Program. Lang. Syst.* 11 (1989) 330–344.
- [6] T.H. Cormen, C.E. Leiserson, R.L. Rivest, *Introduction to Algorithms*, McGraw-Hill, New York, 1990.
- [7] P. Flajolet, A. Odlyzko, Random Mapping Statistics, *Advances in Cryptology, Lecture Notes in Computer Science*, vol. 434, 1990, pp. 329–354.
- [8] R. Sedgewick, P. Flajolet, *An Introduction to the Analysis of Algorithms*, Addison Wesley, Reading, MA, 1996.
- [9] R. Gupta, S.A. Smolka, S. Bhaskar, On randomization in sequential and distributed algorithms, *ACM Comput. Surv.* 26 (1) (1994) 7–86.
- [10] M. Loève, *Probability Theory* (1955).
- [11] N. Lynch, A hundred impossibility proofs for distributed computing, in: *Proceedings of the 8th Annual ACM Symposium on Principles of Distributed Computing*, 1989, pp. 1–28.
- [12] A. Mazurkiewicz, Solvability of the asynchronous ranking problem, *Inf. Process. Lett.* 28 (1988) 221–224.
- [13] R. Motwani, P. Raghavan, *Randomized Algorithms*, Cambridge University Press, Cambridge, 1995.
- [14] Y. Métivier, N. Saheb, A. Zemmari, Randomized rendezvous, in: *Proceedings of the 1st Colloquium on Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probabilities*, Trends in Mathematics, Birkhäuser, Basel, 2000, pp. 183–194.
- [15] D. Peleg, *Distributed Computing—A locality-sensitive approach*, SIAM, 2000.
- [16] J.M. Robson, private communication, 2000.
- [17] K.H. Rosen, *Handbook of Discrete and Combinatorial Mathematics*, CRC Press, 2000.
- [18] G. Tel, *Introduction to Distributed Algorithms*, Cambridge University Press, Cambridge, 2000.