From Hall's matching theorem to optimal routing on hypercubes

Shuhong Gao*, Beth Novick*† and Ke Qiu‡

April 30, 1997

Abstract

We introduce a concept of so-called disjoint ordering for any collection of finite sets. It can be viewed as a generalization of a system of distinctive representatives for the sets. It is shown that disjoint ordering is useful for network routing. More precisely, we show that Hall's 'marriage' condition for a collection of finite sets guarantees the existence of a disjoint ordering for the sets. We next use this result to solve a problem in optimal routing on hypercubes. We give a necessary and sufficient condition under which there are internally node-disjoint paths each shortest from a source node to any other s ($s \leq n$) target nodes on an n-dimensional hypercube. When this condition is not necessarily met, we show that there are always internally node-disjoint paths each being either shortest or near shortest, and the total length is minimum. An efficient algorithm is also given for constructing disjoint orderings and thus disjoint short paths. As a consequence, Rabin's information disposal algorithm may be improved.

1 Introduction

A permutation of the elements of a finite set is called an *ordering*. Suppose X and Y are two sets ordered as $O_1 = (x_1, x_2, \cdots, x_k)$ and $O_2 = (y_1, y_2, \cdots, y_\ell)$ where $k = |X|$ and $\ell = |Y|$. We say that O_1 and O_2 are *disjoint* if for every $1 \leq t \leq \min(k, \ell)$

$$\{x_1, x_2, \cdots, x_t\} \quad \neq \quad \{y_1, y_2, \cdots, y_t\} \tag{1}$$

as sets, unless $t = k = \ell$. Note that X and Y can be the same set and still have disjoint orderings. For instance, if $X = Y = \{1, 2\}$ then $(1, 2)$ and $(2, 1)$ are disjoint. If $X = Y = \{1\}$ then, by our definition, the trivial ordering (1) is disjoint to itself. We say that a collection of finite sets have a *disjoint ordering* if each set has an ordering and all

*Department of Mathematical Sciences, Clemson University, Clemson, South Carolina 29634-1907
†Support of NSF grant DMS-9409274 acknowledged.
‡Jodrey School of Computer Science, Acadia University, Nova Scotia, Canada

the orderings are pairwise disjoint. In particular, as long as we require all singletons in the collection to be distinct, then the first elements of a disjoint ordering form a system of distinct representatives. For example, the following are four sets and a disjoint ordering for them.

$$\begin{aligned} X_1 &= \{1,2,4\} & \qquad O_1 &= (1,2,4) \\ X_2 &= \{1,3,4\} & O_2 &= (3,4,1) \\ X_3 &= \{1,2,3\} & O_3 &= (2,3,1) \\ X_4 &= \{1,2,4\} & O_4 &= (4,1,2) \end{aligned}$$

Note that the initial elements of the ordering, i.e., $1,3,2,4$, form a system of distinct representatives for $\{X_1, X_2, X_3, X_4\}$.

A well-known theorem of P. Hall [3], often called Hall's matching theorem, says that a family of finite sets has a system of distinct representatives (SDR) if and only if the union of any k sets contains at least k distinct elements. The condition in Hall's theorem is known as the *marriage condition*. Obviously the marriage condition is necessary for the existence of a disjoint ordering for a collection of finite sets, since the latter implies the existence of a system of distinct representatives. Surprisingly, as we will show in Theorem 1 of Section 2, this condition is also sufficient. In Section 3, we give an efficient algorithm for finding disjoint orderings. In Section 4, we use disjoint orderings to construct disjoint short paths on hypercube graphs.

In [7], Rabin designs an information dispersal algorithm (IDA) for efficient and accurate transmission of large files in a parallel computer or a distributed network. To send a large file from one node to another node in a network, Rabin tactically divides the file into many pieces and these pieces are sent separately to a target node in two stages: first to randomly chosen intermediate nodes, and then to the target node. The paths traveled by the pieces in each stage are required to be node disjoint (except the source and target nodes). The delay time for each stage is measured by the length of the longest path in the corresponding stage. So it is desirable to construct disjoint paths from one node to many other nodes such that the longest path is shortest possible. Rabin showed that for an n-dimensional hypercube graph there are always disjoint paths from any node to any other n nodes with each of length at most $n+1$. So the total delay time for the two stages is $2(n+1)$. Our contribution is to show, in Section 4, that if the intermediate nodes each have distance at most m from a given node then there are disjoint paths each of length at most $m+2$ from the node to the intermediate nodes. Thus if the intermediate nodes are chosen so that their maximum distance to both the source and target nodes is at about $n/2$ then the total delay time of Rabin's IDA will be about n, which is just half of Rabin's delay time.

More precisely, in Section 4, we prove the following results. A disjoint ordering of subsets can be converted into a collection of disjoint paths each being individually shortest. As a consequence, the marriage condition gives a necessary and sufficient condition for the existence of internally node-disjoint shortest paths from a source node to any other s ($s \leq n$) target nodes on an n-hypercube. When this condition is not met, we show that there are always internally node-disjoint paths each being either shortest or near shortest and the total length being minimum. The algorithm from Section 3 is adapted to constructing these short paths.

2 Strengthening Hall's matching theorem

In this section we prove the following strengthening of Hall's theorem.

Theorem 1. *For any collection of nonempty finite sets $X_1, X_2, \cdots, X_s$, in which all singletons are distinct, there is a disjoint ordering if and only if*

$$\text{for any } 1 \le i_1 < i_2 < \cdots < i_k \le s, \qquad \left| \bigcup_{\ell=1}^{k} X_{i_\ell} \right| \ge k. \tag{2}$$

We need only to prove the sufficiency of the condition (2). In fact, we will prove the following sharper result.

Lemma 1. *If $X_1, X_2, \cdots, X_s$ is a collection of finite sets satisfying (2) with SDR $t_1, t_2, \cdots, t_s$, then there is a disjoint ordering of $X_1, X_2, \cdots, X_s$, using $\{t_1, t_2, \cdots, t_s\}$ as the set of its initial elements.*

Suppose that X_i is ordered as $\{o_{i1}, o_{i2}, \cdots, o_{i\,n_i}\}$, $1 \le i \le s$ where $n_i = |X_i|$. The above statement does *not* require $t_i = o_{i1}$, $1 \le i \le s$, only that $\{t_1, t_2, \cdots, t_s\}$ and $\{o_{11}, o_{21}, \cdots, o_{s1}\}$ be equal as *sets*. To see that the requirement $t_i = o_{i1}$, $1 \le i \le s$ would be too restrictive, consider the example

$$X_1 = \{1, 2, 3\}, \qquad X_2 = \{1, 2\}, \qquad X_3 = \{1, 3\}$$

with SDR $t_1 = 1$, $t_2 = 2$ and $t_3 = 3$.

We need some notations first. For any ordering $O = (o_1, \cdots, o_n)$, denote $\{O\}_k = \{o_1, \cdots, o_k\}$, i.e., the set of the first k elements in O, and $(O)_k = (o_1, \cdots, o_k)$, the k elements of O in the same order. If $k > n$, we understand that $\{O\}_k = \{O\}_n$ and $(O)_k = (O)_n$.

Proof (of Lemma 1). We prove the lemma by contradiction. Let $\mathcal{F} = \{X_1, X_2, \cdots, X_s\}$ be a collection of sets with SDR $t_1, t_2, \cdots, t_s$ for which the theorem fails. Of all such collections, assume that we have chosen $\mathcal{F}$ with $m = \sum_{i=1}^{s} |X_i|$ smallest possible. Clearly not all sets in $\mathcal{F}$ have cardinality 1, for in that case the ordering $O_i := (t_i)$, for $i = 1, 2, \cdots, s$, would be disjoint. Without loss of generality, we assume that $|X_s| = v \ge |X_i|$ for $1 \le i \le s-1$. Then $v > 1$. Proceed in two cases.

Case I *There exists $a \in X_s$, $a \ne t_s$, such that $X_s \setminus \{a\} \notin \{X_1, X_2, \cdots, X_{s-1}\}$.* Then the collection of sets $\{X_1, X_2, \cdots, X_{s-1}, X_s \setminus \{a\}\}$ still has $t_1, t_2, \cdots, t_s$ as their SDR and, by the minimality of m, has a disjoint ordering $\hat{O} = \{\hat{O}_1, \hat{O}_2, \cdots, \hat{O}_s\}$ using this SDR as initial elements. Construct an ordering $O = \{O_1, O_2, \cdots, O_s\}$ of $X_1, X_2, \cdots, X_s$ as follows: $O_i = \hat{O}_i$ for $1 \le i < s$ and $O_s = \hat{O}_s$ with the element a appended. Then O is a disjoint ordering. As $v \ge |X_i|$ for $1 \le i \le s-1$, we only need to verify that

$$\{O_s\}_{v-1} = \{\hat{O}_s\}_{v-1} = X_s \setminus \{a\} \neq \{O_i\}_{v-1} = \{\hat{O}_i\}_{v-1} \tag{3}$$

for $1 \le i \le s-1$ with $|X_i| \ge v-1$. If $|X_i| = v-1$, then $\{O_i\}_{v-1} = X_i$ and (3) holds by our assumption in this case. If $|X_i| = v$ then (3) holds because $\hat{O}_s$ and $\hat{O}_i$ are disjoint. Therefore this case is impossible.

Case II *For all $a \in X_s \setminus \{t_s\}$, the set $X_s \setminus \{a\} \in \{X_1, X_2, \cdots, X_{s-1}\}$.* Since the system $X_1, X_2, \cdots, X_s$ has an SDR, the condition (2) implies that for each $a \in X_s \setminus \{t_s\}$ the set $X_s \setminus \{a\}$ occurs among $X_1, X_2, \cdots, X_{s-1}$ exactly once. We may assume, without loss of generality, that

$$\{X_s \setminus \{a\} : a \in X_s \setminus \{t_s\}\} = \{X_{s-v+1}, \cdots, X_{s-2}, X_{s-1}\}.$$

As $t_i \in X_i \subseteq X_s$, $s-v+1 \le i \le s$, we have $X_s = \{t_{s-v+1}, \cdots, t_{s-2}, t_{s-1}, t_s\}$ and for $1 \le i \le s-v$, $t_i \notin X_s$. Now let $\{O_1, \cdots, O_{s-v}\}$ be a disjoint ordering for $\{X_1, X_2, \cdots, X_{s-v}\}$ with $\{t_1, \cdots, t_{s-v}\}$ as initial elements, guaranteed by the minimality of m, where O_i denotes the ordering of X_i for $1 \le i \le s-v$. Since the initial element of each such O_i belongs to $\{t_1, \cdots, t_{s-v}\}$, O_i is disjoint to any ordering of X_j for any $1 \le i \le s-v$ and $s-v+1 \le j \le v$. Hence any disjoint ordering of X_j, $s-v+1 \le j \le s$, together with O_i, $1 \le i \le s-v$, form a disjoint ordering for $X_1, \cdots, X_s$ with $\{t_1, \cdots, t_s\}$ as its set of initial elements. However, a disjoint ordering for X_j, $s-v+1 \le j \le s$, can be constructed as follows:

$$\begin{aligned} O_s &= (t_{s-v+1}, \cdots, t_{s-1}, t_s), \\ O_j &= (t_{k_j+1}, \cdots, t_s, t_{s-v+1}, \cdots, t_{k_j-1}), \quad s-v+1 \le j \le s-1 \end{aligned}$$

where $s-v+1 \le k_j \le s-1$ is the unique integer with $t_{k_j} \notin X_j$.

Putting together, $\{O_1, \cdots, O_s\}$ is a disjoint ordering for $\mathcal{F} = \{X_1, \cdots, X_s\}$. This contradicts our assumption that the theorem fails for $\mathcal{F}$, completing the proof. □

We remark that Qiu and Novick [6] show that condition (2) is sufficient for a somewhat less general problem.

3 Algorithm and complexity

The proof of Lemma 1 is constructive. In this section, we convert it into an efficient algorithm for finding a disjoint ordering for any collection of sets with an SDR. Note that writing out the ordering alone needs $O(n^2)$ time and our algorithm finds a disjoint ordering in time $O(n^4)$ in the worst case where n is the number of distinct elements among the sets. Note that the input size is $N = \sum_{i=1}^{s} |X_i| = O(n^2)$, where $X_1, \cdots, X_s \subseteq X = \{1, 2, \cdots, n\}$

are the input sets. Thus our algorithm has a running time of $O(N^2)$. For comparison, note that the best known algorithm for finding a maximum cardinality matching in a bipartite graph G has time $O(m\sqrt{n})$ [2], where m is the number of edges and n the number of nodes in G. Thus the best known time for finding a maximum SDR for $X_1, \cdots, X_s$ is $O(n^2\sqrt{n}) = O(N^{\frac{5}{4}})$.

Suppose that $X_1, X_2, \cdots, X_s$ are ordered such that $|X_1| \leq |X_2| \leq \cdots \leq |X_s|$ with SDR $t_j \in X_j$, $1 \leq j \leq s$. Let $|X_j| = n_j$, $1 \leq j \leq s$. The idea is to start with an arbitrary ordering O_1 of X_1 with t_1 as its initial element, and find a disjoint ordering for $\{X_1, X_2\}$. Then extend it to a disjoint ordering for $\{X_1, X_2, X_3\}$, etc. At a typical step, a disjoint ordering $\{O_1, O_2, \cdots, O_{\ell-1}\}$ of $\{X_1, X_2, \cdots, X_{\ell-1}\}$ has been found and one needs to find a disjoint ordering for $\{X_1, X_2, \cdots, X_\ell\}$.

Let $O_\ell = (o_{\ell 1}, \cdots, o_{\ell v})$ denote a desired ordering for X_ℓ, where $v = |X_\ell|$. We first choose $o_{\ell v}$ to be any element $a \in X_\ell \setminus \{t_\ell\}$ such that

$$X_\ell \setminus \{a\} \neq \{O_j\}_{v-1} \quad \text{for all } 1 \leq j \leq \ell - 1.$$

Then choose $o_{\ell\, v-1}$ to be any $a \in X_\ell \setminus \{o_{\ell v}, t_\ell\}$ such that

$$X_\ell \setminus \{a, o_{\ell v}\} \neq \{O_j\}_{v-2} \quad \text{for all } 1 \leq j \leq \ell - 1.$$

Continue until there is no such a, at which point the situation is as follows. For some $1 \leq i \leq v$, we have found distinct elements $o_{\ell\, i+1}, \cdots, o_{\ell\, v} \in X_\ell \setminus \{t_\ell\}$ such that

$$X_\ell \setminus \{o_{\ell\, k}, \cdots, o_{\ell\, v}\} \neq \{O_j\}_{k-1} \quad \text{for all } 1 \leq j \leq \ell - 1 \text{ and } i + 1 \leq k \leq v$$

and for each $a \in X_\ell \setminus \{o_{\ell\, i+1}, \cdots, o_{\ell\, v}, t_\ell\}$ there is $1 \leq j_a \leq \ell - 1$ such that

$$X_\ell \setminus \{a, o_{\ell\, i+1}, \cdots, o_{\ell\, v}\} = \{O_{j_a}\}_{i-1}.$$

If $i = 1$ then $O_\ell = (t_\ell, o_{\ell\, 2}, \cdots, o_{\ell\, v})$ is an ordering of X_ℓ and is disjoint to all O_j for $1 \leq j \leq \ell - 1$. So assume that $i > 1$. Let $Y = X_\ell \setminus \{o_{\ell\, i+1}, \cdots, o_{\ell\, v}\}$. By the technique used in Case II in the proof of Lemma 1, we construct a disjoint ordering for the collection of sets Y and $\{O_{j_a}\}_{i-1}$, $a \in Y \setminus \{t_\ell\}$. Actually, it is not necessary to find j_a first. We can put $(o_{\ell\, 1}, \cdots, o_{\ell\, i})$ to be any permutation of Y to get an ordering $O_\ell = (o_{\ell\, 1}, \cdots, o_{\ell\, i}, o_{\ell\, i+1}, \cdots, o_{\ell\, v})$ for X_ℓ. We then reorder $\{O_j\}_{i-1}$, $1 \leq j \leq \ell - 1$, as follows. for some $a \in Y \setminus \{t_\ell\}$ if and only if $o_{j1} \in Y$ by the proof of Case II of Lemma 1. Thus if $o_{j1} \notin Y$ then O_j is not touched. If $o_{j1} \in Y$ then replace $(O_j)_{i-1}$ by

$$(o_{\ell\, u+1}, \cdots, o_{\ell\, i-1}, o_{\ell\, 1}, \cdots, o_{\ell\, u-1}),$$

where $1 \leq u \leq i$ is the unique integer such that $o_{\ell\, u} \neq t_\ell$ and $o_{\ell\, u} \notin \{O_j\}_{i-1}$. Thus the ordering of X_j is

$$(o_{\ell\, u+1}, \cdots, o_{\ell\, i-1}, o_{\ell\, 1}, \cdots, o_{\ell\, u-1}, o_{j\, i}, \cdots, o_{j\, n_j}).$$

This will produce a disjoint ordering for $X_1, X_2, \cdots, X_\ell$. We implement the steps more explicitly in the following algorithm.

Algorithm 1: Ordering with SDR

Input: A collection of finite sets $X_1, X_2, \cdots, X_s$ with SDR.
Output: A disjoint ordering:
$O_j = (o_{j1}, o_{j2}, \cdots, o_{jn_j})$ being the order of X_j where $|X_j| = n_j$, $1 \le j \le s$.
Step 1. Reorder X_j's such that $|X_1| \le |X_2| \le \cdots \le |X_s|$.
Step 2. Find an SDR: $t_j \in X_j$, $1 \le j \le s$.
Step 3. Set $O_1 = (o_{11}, o_{12}, \cdots, o_{1n_1})$ to be any permutation of X_1 with $o_{11} = t_1$.
Step 4. For ℓ from 2 to s do
Step 4.1. Set $i \leftarrow n_\ell$, $Y \leftarrow X_\ell$.
While there is $a \in Y \setminus \{t_\ell\}$ such that $Y \setminus \{a\} \notin \Big\{\{O_1\}_{i-1}, \cdots, \{O_{\ell-1}\}_{i-1}\Big\}$,
set $o_{\ell i} \leftarrow a$, $Y \leftarrow Y \setminus \{a\}$ and $i \leftarrow i - 1$.
Step 4.2. If $i = 1$ then set $o_{\ell 1} \leftarrow t_\ell$.
Step 4.3. Otherwise $i > 1$. Then
Step 4.4. Set $(o_{\ell 1}, \cdots, o_{\ell i})$ to be any permutation of Y.
For j from 1 to $\ell - 1$ reorder $\{O_j\}_{i-1}$ as follows:
If $o_{j1} \in Y$ then
find $1 < u \le i$ such that $o_{\ell u} \ne t_\ell$ and $o_{\ell u} \notin \{O_j\}_{i-1}$ and
replace $(o_{j1}, \cdots, o_{j\,i-1})$ by $(o_{\ell\,u+1}, \cdots, o_{\ell i}, o_{\ell 1}, \cdots, o_{\ell\,u-1})$.
end if
end for loop
end outer for loop
Step 5. Return the list $O_j = (o_{j1}, o_{j2}, \cdots, o_{j\,n_j})$ for $1 \le j \le s$.

Theorem 2. *Algorithm 1 finds a disjoint ordering in time $O(sn^3)$ where n is the number of distinct elements among the sets X_i's.*

Proof . The correctness of the algorithm follows from the proof of Lemma 1 and the discussion above. The dominant costs are at Steps 2, 4.1 and 4.4. Step 2 can be done in time $O(n^2\sqrt{n})$, see [2]. At Step 4.1, for each $a \in Y \setminus \{t_\ell\}$, one needs to compare $Y \setminus \{t_\ell\}$ with $\ell - 1$ other sets. Since all the sets have size $i - 1 \le n$, each comparison of two sets can be done using $O(n)$ comparisons on numbers in $\{1, \ldots, n\}$. For each $2 \le \ell \le s$, Step 4.1 can be done in time $O(|X_\ell|(\ell-1)n) = O(n^3)$, as $|X_\ell| \le n$ and $\ell \le n$. For each $2 \le \ell \le s$, Step 4.4 needs at most $O(\ell^2) = O(n^2)$ operations (for updating at most $i \le \ell$ sequences of length at most $i \le \ell$). The total cost for Steps 4.1 and 4.4 is $O(sn^3 + sn^2) = O(sn^3)$, which dominates the total cost for other steps. Therefore Algorithm 1 needs at most $O(sn^3)$ operations on numbers in $\{1, \ldots, n\}$. □

4 Optimal routing on hypercubes

An n-dimensional hypercube, or n-cube, is an undirected graph C_n whose node set consists of all n-tuples of 0's and 1's of length n and two nodes are adjacent if and only if the two tuples differ at exactly one position, that is, one can be obtained from the other by flipping (1 to 0, or 0 to 1) one coordinate. It is well-known that C_n has connectivity n and diameter n. There are always internally node disjoint paths from any node to any other n nodes on an n-dimensional hypercube. In Rabin's application as we mentioned in the introduction, it is desirable to have disjoint paths with the longest one shortest possible. The reader is referred to the excellent survey paper [4], where D. F. Hsu discusses some related problems for general graphs. For more information on hypercube-based parallel computers, consult Leighton's book [5].

One can go further to ask whether it is possible to have disjoint paths with the longest one shortest possible and each one individually shortest as well. Of course, if each path is shortest, i.e., whose length is equal to the distance of the nodes connected by the path, then the longest length is automatically shortest possible. In some sense, such a collection of paths is best possible or optimal. In general, one would not expect them to exist. The question is to characterize exactly when shortest disjoint paths exist and how to construct them. In this section, we solve this question completely for hypercubes by Corollary 1 of Theorem 3. Here disjoint ordering plays an essential role in the proof of Theorem 3.

We need more notation and terminology. A *partial* SDR for any family $\mathcal{F}$ of sets is any SDR of a subfamily of $\mathcal{F}$. The *deficiency* of $\mathcal{F}$ is defined to be $s - m$, where m is the size of a largest partial SDR. A partial SDR of largest size is also called a maximum partial SDR. A path from a node u to another v is called *shortest* if its length is equal to the distance from u to v, and is called *near-shortest* if it is not shortest but with length at most 2 more than the distance.

Theorem 3. *Let v be a node and $v_1, \ldots, v_s$ any other $s \leq n$ nodes on C_n, not necessarily distinct. For $1 \leq i \leq s$, let X_i denote the set of coordinate positions $j \in \{1, \ldots, n\}$ where v and v_i differ. Suppose the system $\{X_1, \ldots, X_s\}$ of sets has deficiency d and no identical singletons. Then any collection of internally node-disjoint paths from v to $v_1, \ldots, v_s$ has total length at least $\sum_i^n |X_i| + 2d$. Furthermore, this lower bound is tight and obtainable with a collection of paths each being either shortest or near-shortest.*

Proof . Note that a path on C_n can be described by a node followed by a sequence coordinate positions where the path travels. More explicitly, for any node v and a sequence $a_1, a_2, \ldots, a_t$ of elements in $\{1, \ldots, n\}$, let $P(v; a_1, a_2, \ldots, a_t)$ denote the path, starting at v, on which the i-th node, i from 1 to t and v being the 0th node, is obtained from the previous one by flipping the a_i-th coordinate. Here the coordinate positions are counted from left to right. For example, if $v = (0011)$ then $P(v; 1, 2, 3, 4, 1)$ denotes the path

$$(0011) \to (1011) \to (1111) \to (1101) \to (1100) \to (0100).$$

Obviously, $P(v; a_1, a_2, \ldots, a_t)$ has length t.

First we will establish the lower bound. Let $P_1, \ldots, P_s$ be any collection of internally node-disjoint paths from v to the target nodes $v_1, \ldots, v_s$. Suppose the paths are represented by coordinate sequences. For each $1 \leq i \leq s$, since X_i is the set of coordinate positions where v and v_i differ, every element in X_i has to be in the coordinate sequence of P_i. Let j_i be the first coordinate on the path P_i. Since $P_1, \ldots, P_s$ are disjoint and no two identical nodes in $\{v_1, \dots, v_s\}$ have distance 1 to v (i.e., no identical singletons among $X_1, \dots, X_s$), $j_1, \dots, j_s$ must be distinct. But $X_1, \dots, X_s$ has deficiency d, there are at least d values of $i \in \{1, \dots, n\}$ such that $j_i \notin X_i$. For any $j_i \notin X_i$, the j_i-th coordinate has to be flipped back somewhere on the path P_i to reach v_i, i.e., j_i occurs at least twice in the coordinate sequence of P_i. So the length of such a path P_i is at least $|X_k| + 2$ and the total length is at least $\sum_i^s |X_i| + 2d$.

Next we will establish the tightness of the bound. Suppose $t_1 \in X_{k_1}$, $t_2 \in X_{k_2}$, $\cdots$, $t_m \in X_{k_m}$ is a maximum partial SDR for the system of subsets $\{X_1, \dots, X_s\}$. We can assume that there is no $j \in \{1, \dots, s\} \setminus \{k_1, \dots, k_m\}$ such that X_j is a proper subset of X_{k_i} for some i and the system

$$X_{k_1}, \dots, X_{k_{i-1}}, X_j, X_{k_{i+1}}, \dots, X_{k_m}$$

has an SDR of size m. If this condition is not satisfied, we can replace X_{k_i} by X_j and we still have a maximum partial SDR for $\{X_1, \dots, X_s\}$. Repeat replacing until there is no such j. This process has to stop as the total cardinality of the sets with representatives decreases by one at each replacement. Without loss of generality, we may assume that $k_1 = 1$, $\dots$, $k_m = m$. Then $X_j \subseteq \{t_1, \dots, t_m\}$ for $m < j \leq s$. Since $d = s - m \leq n - m$, there are d different elements $a_1, \dots, a_d$ in $\{1, \dots, n\} \setminus \{t_1, \dots, t_m\}$. Consider the subsets X_1, $\dots$, X_m, $X_{m+1} \cup \{a_1\}$, $\dots$, $X_s \cup \{a_d\}$. Note that $t_1 \in X_1$, $\dots$, $t_m \in X_m$, $a_1 \in X_{m+1} \cup \{a_1\}$, $\dots$, $a_d \in X_s \cup \{a_d\}$ form a complete SDR. By Lemma 1, there is a disjoint ordering $O_1, \dots, O_s$ for these subsets. It is straightforward to check that the paths $P(v; O_i)$, $1 \leq i \leq s$, are internally node-disjoint. Then the paths

$$P(v; O_1), \dots, P(v; O_m), P(v; O_{m+1}, a_1), \dots, P(v; O_s, a_d)$$

are from v to $v_1, \dots, v_s$, respectively. We need to show that they are internally node-disjoint. It is sufficient to show that no v_j, $1 \leq j \leq s$, becomes an internal node. First, for $1 \leq j \leq m$ and $1 \leq i \leq s - m$, we have $X_j \neq X_{m+i} \cup \{a_i\}$, since otherwise we would have

$$t_1 \in X_1, \cdots, t_{j-1} \in X_{j-1}, a_i \in X_j, t_{j+1} \in X_{j+1}, \cdots, t_m \in X_m, t_j \in X_{m+i}$$

which is a system of SDR of size $m+1$, contradicting the fact that our SDR is maximum. So v_j is not the end node of $P(v; O_{m+i})$ and thus not an internal node of $P(v; O_{m+i}, a_i)$. Secondly, for $m < j \leq s$, we claim that v_j is not an internal node of $P(v; O_i)$ for any $1 \leq i \leq m$. In fact, if v_j is an internal node on $P(v; O_i)$ for some $m < j \leq s$ and $1 \leq i \leq m$, then X_j is a proper subset of X_i and contains the initial element of the ordering O_i of X_i. But then the system $X_1, \dots, X_{i-1}, X_j, X_{i+1}, \dots, X_m$ has an SDR of size m, i.e., the initial elements of $O_1, \dots, O_m$. This is contradictory to our assumption on the chosen SDR.

Note that P_j has length $|X_j|$ for $1 \leq j \leq m$, and $|X_j| + 2$ for $m < j \leq n$. So the theorem follows. $\square$

Corollary 1. *Using the notation in Theorem 3, a necessary and sufficient condition for there to be internally node-disjoint shortest paths from the source node v to any $s \leq n$ other target nodes $v_1, \dots, v_s$ in C_n is that $X_1, \dots, X_s$ satisfy Hall's marriage condition.*

Proof . Note that the distance from v to v_i is equal to $|X_i|$, $1 \le i \le s$. Also $d = 0$ if and only if the system of sets $X_1, \dots, X_s$ has a complete set of distinct representatives. The latter happens if and only if Hall's marriage condition is satisfied. $\square$

The proof of Theorem 3 yields the following result [7, Lemma 3], which is attributed to M. Ben-Or.

Corollary 2. *For any node v and any collection of n other nodes $v_1, \dots, v_n$ on C_n, there exists n internally node disjoint paths from v to $v_1, \dots, v_n$ with each of length at most $n + 1$.*

Proof . Using the notation in the proof of Theorem 3. For $1 \le j \le m$, $P(v; O_j)$ has length $|X_j| \le n$. If $d > 0$ then $|X_{m+i}| \le m < s \le n$, as $X_{m+i} \subseteq \{t_1, \dots, t_m\}$ for $1 < i \le s - d$. So $P(v; O_{m+i}, a_i)$ has length $|X_{m+i}| + 2 \le m + 2 \le n + 1$. Therefore every path has length at most $n + 1$. $\square$

Corollary 3. *For any two distinct nodes u and v in C_n of distance d, there are n internally node disjoint paths from u to v each of length at most $\min\{d + 2, n + 1\}$.*

Proof . Let $X_1 = \cdots = X_n = X$, the set of coordinate positions where u and v differ. Then $d = |X| \le n$. If $d = n$ then $X = \{1, 2, \dots, n\}$. In this case, the system $X_1, X_2, \dots, X_n$ has a complete set of representatives, and Corollary 3 follows from Corollary 1. If $d < n$ then $|X_i| + 2 = d + 2 \le n + 1$ for $1 \le i \le n$, and Theorem 3 applies. $\square$

By the proof of Theorem 3, it should be clear that a collection of disjoint paths described in Theorem 3 can be found in time $O(n^4) = O(\log^4 N)$ where $N = 2^n$ is the number of nodes in C_n. The algorithm is straightforward and is omitted here.

As we mentioned in the introduction, Rabin's IDA may be improved by choosing (randomly) the intermediate nodes such that each has distance about $n/2$ to both the source and target nodes. Then the total delay time is reduced to about n. With this modification of choosing intermediate nodes, it seems that Rabin's results, Theorems 1 and 2 in [7], on reliability and fault tolerance of the network still hold. But this is yet to be analyzed rigorously.

Acknowledgments. We thank Doug Shier for his helpful comments on the paper. A preliminary version of the paper was presented at the 28-th Southeastern International Conference on Combinatorics, Graph Theory, and Computing, Florida Atlantic University, Boca Raton, March 3–7, 1997.

References

[1] J.A. Bondy, and U.S.R. Murty, *Graph Theory with Applications,* Macmillan, Lon-

don, 1976.

[2] Z. Galil, "Efficient algorithms for finding maximal matchings in graphs," *ACM Computing Surveys*, vol. 18, no. 1, 1986, pp 23–38.

[3] P. Hall, "On representatives of subsets," *J. London Math. Soc.*, vol. 10, 1935, pp 26–30.

[4] D.F. Hsu, "On Container Width and Length in Graphs, Groups, and Networks," *IEICE Transactions, on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E77-A, No.4, April 1994, pp 668–680.

[5] F.T. Leighton, *Introduction to Parallel Algorithms and Architectures: Arrays, Trees, Hypercubes*, Morgan Kaufmann Publishers, San Mateo, California, 1992.

[6] K. Qiu and B. Novick, "Disjoint Paths in Hypercubes," *Congressus Numeratium*, vol. 119, 1996, pp 105–112.

[7] M. Rabin, "Efficient dispersal of information for security, load balancing, and fault tolerance," *Journal of the Associattion of Computing Machinery*, vol. 36, no. 2, 1989, pp 335–348.