August 25, 1986 LIDS-P-1594

A Systematic Approach to Reliable Multistage Interconnection Network Design*

C.-C. Jay Kuo†

Abstract

A systematic approach to the topological design of reliable multistage interconnection networks (MIN's) is proposed. First, we describe a procedure to derive topologies and routing schemes for unique-path MIN's. This procedure includes the specification of a switching model by a state transition process, the realization of the switching model by a buddy-type network, and the layout of the buddy-type network in the plane. By this approach, the topological equivalence of different MIN's can be easily understood. Then, we discuss the design of multiple-path (reliable) MIN's by generalizing the above procedure. Two reliable MIN's, i.e. the extra stage cube and the multipath omega, are shown to be special examples of our general designs.

Index terms - Fault tolerance, multistage interconnection networks, redundant path networks, buddy-type networks, parallel processing, MIMD Machines, topological equivalence, Extra Stage Cube, Multipath Omega.

* This work was supported in part by the Army Research Office under Grant No. DAAG29-84-K-0005 and in part by the Advanced Research Projects Agency monitored by ONR under contract N00014-81-K-0742.

† Laboratory for Information and Decision Systems and Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139.

1. Introduction

An MIMD machine with common shared memory basically contains three parts: processor elements (PE's), memory modules (MM's), and the interconnection network (IN) between PE's and MM's. The interconnection network is crucial for the performance of such a system. One important class of interconnection networks is known as *multistage interconnection networks* (MIN's) [1]. Since MIN's have a good balance between low cost, short communication time, blocking probability, and require a simple routing strategy, they have received a large amount of attention and have been studied for more than two decades [2]. A current active research problem in this area is the analysis and design of *fault-tolerant,* or *reliable,* MIN's.

Many different types of fault-tolerant MIN's such as the *extra stage cube* [3] and the *multipath omega* [4] have been proposed and analyzed. These fault-tolerant MIN's are very different from each other by appearance, and are derived from different methods. One objective of this paper is to find a common principle behind these different designs so that we can obtain a more general design methodology and have a better understanding of the structure of these MIN's.

To study a MIN, the conventional approach begins by specifying its configuration in the plane, and then analyzes its topological properties and describes some control scheme so that the given MIN can achieve the desired function [2]. In this paper, we adopt a completely different point of view. Our approach first specifies the desired function. In order to achieve this function, an appropriate network topology is constructed, from which various MIN configurations in the plane can be derived. Hence, our approach is from a *synthetic* point of view while the conventional approach is primarily from an *analytic* point of view.

Our design procedure can be briefly stated as follows. First, the function of a MIN is described by a switching model specified by a state transition process. Then, a buddy-type network is constructed to support the switching model. All topological information of a MIN is included in its buddy-type network. However, if we want to obtain a MIN in conventional form, i.e. the planar configuration, the layout of its buddy-type network in the plane has to be considered finally. The above procedure not only provides a systematic way to MIN design but also helps us understand the topological properties of MIN's better.

In Section 2, we describe a design procedure for MIN's which have a unique-path between each PE and each MM. By this procedure, we can derive the banyan network [5], the omega network [6], the flip network [7], and the indirect binary n-cube network [8] in a unified framework. Although the topological equivalence of the above MIN's is well known [9] [10] [11], we show the same result from a *synthetic* point of view.

The design procedure is generalized to multiple-path, or fault-tolerant, MIN's in Section 3. The generalization is primarily achieved by introducing extra intermediate states into the state transition process. The corresponding buddy-type network and its layout can be obtained from the modified state transition process in a straightforward way. We use two important fault-tolerant MIN's, i.e. the extra stage cube and the multipath omega, as examples to show how to apply our systematic design procedure. These two MIN's are chosen since they have been examined more thoroughly than others and since they will be used in real machines. The extra stage cube is the interconnection network of the PASM prototype machine [12] while the multipath omega is proposed for the Cedar supercomputer [13].

Conclusions and extensions are given in Section 4.

2. A Systematic Approach to Unique-Path MIN Design

The conventional approach used to study MIN's includes two steps: First, to characterize the network configuration in the plane by graphic or algebraic tools; second, to analyze its topological properties and to design routing schemes based on the given network configuration. One graphic tool is the *switching diagram* formed by switching boxes and links. Although the graphic tool is easy to understand, some important information is hidden behind it. For example, it is well known that some MIN's whose graphic representations look very different are in fact topologically equivalent. The algebraic tool is some kind of permutation function describing the *interconnection pattern* between two consecutive switching stages. For example, "shuffle", "bit reversal", "butterfly", "exchange", "cube", and "PM2I" can all be precisely defined [10] [14]. Hence, one MIN can be completely characterized by its number of stages, the sizes of the switching boxes at each stage, and the interconnection patterns between consecutive stages described by algebraic tools. There is a one-to-one correspondence between graphic and algebraic tools and both are commonly used to describe MIN's. That is, the graphic tool is used for illustration while the algebraic tool is used for manipulation.

There are some disadvantages with the conventional approach. First, this approach does not suggest a way to synthesize a network topology. Hence, it is only used as an analytical tool. Second, the tools used to describe one MIN primarily concentrate on its interconnection patterns between two consecutive stages since they appear to be the most complicated part of the entire MIN. However, it will be shown later that the interconnection pattern is only a direct consequence of different kind of arrangement of switching boxes so that the complication can be avoided.

To overcome the above two disadvantages, we develop a systematic approach to MIN design, which includes three steps: the specification of a switching model by a state transition process, the realization of the switching model by a buddy-type network, and the layout of the buddy-type network in the plane. In this section, we discuss the design procedure for unique-path MIN's.

2.1 Specification of a Switching Model By a State Transition Process

We use $\overline{N} = \{ 0 , 1 , 2 , . . . , N-1\}$ and $\overline{M} = \{ 0 , 1 , 2 , . . . , M-1\}$ to denote the sets of inputs and outputs. Each member of $\overline{N}$ is called an *input state* and each member of $\overline{M}$ is called an *output state.* Then, we can use a state transition process to model the *switching* between $\overline{N}$ and $\overline{M}$. The *state transition process* is defined to be the set formed by all state transition paths from any input state to any output state. If every input state can transit to any output state in one step, there are NM different state transition paths. Therefore, the state transition process is

$$S_{CB} = \{ (n , m) \mid n \in \overline{N} , m \in \overline{M} \} .$$

The input set $\overline{N}$, the output set $\overline{M}$, and the state transition process S_{CB} form a *switching model.* For convenience, we use a simplified notation

$$X_{CB} = (n \rightarrow m)$$

to denote this model. X_{CB} is in fact the switching model for an $N \times M$ crossbar.

Assume N and M can be represented as products of K positive integers, i.e.

$$N = N_1 \times N_2 \times \cdots \times N_K \quad \text{and} \quad M = M_1 \times M_2 \times \cdots \times M_K.$$

Let

$$\overline{N}_k = \{ 0 , 1 , 2 ,\dots , N_k - 1 \} \quad k = 1 , 2 , \dots , K ,$$

$$\overline{M}_k = \{ 0 , 1 , 2 ,\dots , M_k - 1 \} \quad k = 1 , 2 , \dots , K .$$

Then, their Cartesian products, i.e.

$$I = \overline{N}_1 \times \overline{N}_2 \times \cdots \times \overline{N}_K \quad \text{and} \quad O = \overline{M}_1 \times \overline{M}_2 \times \cdots \times \overline{M}_K , \tag{2.1}$$

give a K-dimensional input space I and a K-dimensional output space O. Since $|I| = |\overline{N}|$ and $|O| = |\overline{M}|$, there exist one-to-one mappings from $\overline{N}$ to I and from $\overline{M}$ to O respectively. In other words, states in the one-dimensional spaces $\overline{N}$ and $\overline{M}$ are mapped into states in the K-dimensional spaces I and O.

We define the k-th step input space I^k and output space O^k as

$$I^k = \overline{M}_1 \times \overline{M}_2 \times \cdots \times \overline{M}_{k-1} \times \overline{N}_k \times \overline{N}_{k+1} \times \dots \times \overline{N}_K , \tag{2.2a}$$

$$O^k = \overline{M}_1 \times \overline{M}_2 \times \cdots \times \overline{M}_{k-1} \times \overline{M}_k \times \overline{N}_{k+1} \times \dots \times \overline{N}_K . \tag{2.2b}$$

It is easy to check that

$$I^1 = I , \quad O^k = I^{k+1} \text{ for } k = 1 , 2 , \dots , K-1 , \quad O^K = O .$$

Given an arbitrary input state $q^0 = (n_1, n_2, \cdots, n_K) \in I$ and any output state $q^K = (m_1, m_2, \cdots, m_K) \in O$, consider the following state transition process,

$$S_{MIN} = \{ (q^0, \cdots, q^k, \cdots, q^K) \mid q^0 \in I , \; q^k \in O^k \;\; 1 \le k \le K \}. \tag{2.3}$$

The state transition process S_{MIN} is obtained by dividing the state transition from q^0 to q^K into K steps whereby only one index can change at one step. Hence, the input state q^0 has to go through $K-1$ *intermediate states,* i.e.

$$q^k = (m_1 , \cdots , m_k , n_{k+1} , \cdots , n_K) , \ k = 1, 2, ..., K-1$$

before it accesses the output state q^K. Equations (2.1), (2.2), and (2.3) form a switching model denoted by

$$X_{MIN} = (n_1 \rightarrow m_1 ; \cdots ; n_k \rightarrow m_k ; \cdots ; n_K \rightarrow m_K) , \qquad (2.4)$$

and the *k-th state transition step* is denoted by

$$q^{k-1} \rightarrow q^k \quad or \quad n_k \rightarrow m_k \ .$$

Another convenient tool to describe the above switching model is the *state transition graph,* in which we use a circle to denote a state and a directed link to denote a state transition step. The state transition graph corresponding to (2.4) is given in Figure 1. X_{MIN} is a switching model of a K-stage MIN.

To summarize, the switching model X of a network consists of its input space I, output space O, and a state transition process S which is a set formed by specified state transition paths. Design parameters for specifying the switching model include the dimension of input and output spaces (K), the size of each dimension (N_k and M_k), and the state transition sequence in the state transition process S.

2.2 Buddy-type Network Realization

The switching model is only a mathematical description of how the input and output spaces are divided and how the sequence of state transition is performed. However, a network is made of hardware which supports all required state transition paths. In this section, a *constructive* procedure is given so that we can associate the switching model with a certain topology. This procedure is called the *realization* of a switching model.

The realization of the switching model X_{CB} of a crossbar is straightforward. Elements of input set $\overline{N}$ and output set $\overline{M}$ are realized as *input* and *output terminals*. Each state transition path is realized by a *switching link* which connects the corresponding input and output terminals.

The realization of the switching model X_{MIN} of a MIN given by (2.4) can be divided into two parts.

(1) Switching Links and Switching boxes

Consider the k-th switching step $q^{k-1} \rightarrow q^k$, where $1 \leq k \leq K$. Elements of input set I^k and output set O^k are realized as input and output terminals at stage k. However, unlike the crossbar network, not every input state q^{k-1} can transit to every output state q^k. To build appropriate switching links between input and output terminals, the key observation is that the elements of I^k and O^k can be grouped in such a way that the state transition only occurs within each individual small group.

Consider the following set

$$\overline{U}^k \equiv \overline{M}_1 \times \overline{M}_2 \times \cdots \times \overline{M}_{k-1} \times \overline{N}_{k+1} \times ... \times \overline{N}_K ,$$

which is a $(K-1)$-dimensional subspace of I^k and O^k with cardinality

$$U^k = |\overline{U}^k| = M_1 \times M_2 \times \cdots M_{k-1} \times N_{k+1} \times ... \times N_K .$$

The K-dimensional space I^k can be viewed as the Cartesian product of the $(K-1)$-dimensional space $\overline{U}^k$ and the 1-dimensional space $\overline{N}_k$,

$$I^k = \bar{U}^k \times \bar{N}_k \ .$$

Similarly, O^k can be viewed as

$$O^k = \bar{U}^k \times \bar{M}_k \ .$$

Since the state transition at stage k only changes the k-th coordinate, elements of the space $\bar{U}^k$ remain the same at this stage. This means that there is no state transition from the elements of I^k to the elements of O^k, if they have different values in the space $\bar{U}^k$. Therefore, we can divide elements of I^k and O^k into U^k disjoint groups, each of which is characterized by an index $\mathbf{u}^k \in \bar{U}^k$. Within each group, there are N_k inputs and M_k outputs which have to be fully connected. Hence, an $(N_k \times M_k)$-crossbar, formed by N_k M_k switching links, can be used to realize each individual group. Totally, U^k such crossbars are needed for stage k. A switching box with fan-in N_k and fan-out M_k is used to represent such a crossbar in switching diagrams.

In addition to the switching link (or intra-box link), we need another kind of link called the *interconnection link* (or inter-box link). The switching link is used to realize a state transition step whereby its front and rear ends represent two different states. In contrast, the interconnection link is used to connect the output of the switching boxes at stage k and the input of the switching boxes at stage $k+1$, which are of the same state. Hence, an interconnection link and its both ends all represent the same intermediate state in the switching model. Conventionally, an appropriate group of switching links is symbolized by a switching box, and only interconnection links are drawn in switching diagrams. We described above how to construct switching links and switching boxes. Next, we examine how to construct interconnection links between switching boxes of two consecutive stages.

(2) Buddy-type Interconnection

Although there exist different interconnection schemes, we are primarily interested in the most natural and popular one, i.e. the *buddy-type* interconnection [11]. The resulting network is called the *buddy-type network*. Most MIN's discussed in the literature are buddy-type networks.

Consider two consecutive stages k and $k+1$. The output set of stage k is O^k and the input set of stage $k+1$ is I^{k+1}, which can be represented as

$$O^k = \bar{U}^k \times \bar{M}_k = (\bar{V}^{k,k+1} \times \bar{N}_{k+1}) \times \bar{M}_k ,$$

$$I^{k+1} = \bar{U}^{k+1} \times \bar{N}_{k+1} = (\bar{V}^{k,k+1} \times \bar{M}_k) \times \bar{N}_{k+1} ,$$

where

$$\bar{V}^{k,k+1} \equiv \bar{M}_1 \times \bar{M}_2 \times \cdots \times \bar{M}_{k-1} \times \bar{N}_{k+2} \times ... \times \bar{N}_K ,$$

with cardinality

$$V^{k,k+1} = | \bar{V}^{k,k+1} | = M_1 \times M_2 \times \cdots \times M_{k-1} \times N_{k+2} \times ... \times N_K .$$

An intermediate state $q^k = (m_1, \cdots, m_k, n_{k+1}, \cdots, n_K)$ is realized as both an output terminal m_k of the switching box with index ($\mathbf{v}^{k,k+1}$; n_{k+1}) at stage k and an input terminal n_{k+1} of the switching box with index ($\mathbf{v}^{k,k+1}$; m_k) at stage $k+1$, where $\mathbf{v}^{k,k+1} \in \bar{V}^{k,k+1}$. Since these two terminals represent the same state, an interconnection link is used to connect these two terminals.

In order to build appropriate interconnection links, we have to classify switching boxes at stages k and $k+1$ into groups. That is, the switching boxes at stage k with the same index $\mathbf{v}^{k,k+1} \in \bar{V}^{k,k+1}$ should be grouped together whereby each group has N_{k+1} switching boxes. Similarly, the switching boxes at stage $k+1$ with the same index $\mathbf{v}^{k,k+1} \in \bar{V}^{k,k+1}$ should be grouped together whereby each group has M_k switching boxes. Then, the interconnection rule is that the N_{k+1} switching boxes at stage k with group index $\mathbf{v}^{k,k+1}$ should be fully connected to the M_k switching boxes at stage $k+1$ with the same group index $\mathbf{v}^{k,k+1}$. It is easy to see that no links are needed between groups with different index $\mathbf{v}^{k,k+1}$, since the output of stage k cannot be the same state as the input of stage $k+1$ among these groups. For groups with the same index $\mathbf{v}^{k,k+1}$, the output terminal m_k of the box with index n_{k+1} at stage k corresponds to the same state as the input terminal n_{k+1} of the box with index m_k at stage $k+1$. Therefore, for groups with the same index, the switching boxes at stage k should be fully connected to the switching boxes at stage $k+1$, and there are $M_k \; N_{k+1}$ interconnection links. By repeating the above constructive procedure for $k = 1, 2, ..., K-1$, we can obtain a buddy-type network.

If the switching model (2.4) is realized by a buddy-type network, the uniqueness of its physical path between a given input-output pair, say from input terminal $(n_1, \ldots, n_K)$ to output terminal $(m_1, \ldots, m_K)$, can be easily checked. Its switching link at stage k is uniquely determined by input terminal $(m_1, \cdots, m_{k-1}, n_k, \cdots, n_K)$ and output terminal $(m_1, \cdots, m_k, n_{k+1}, \cdots, n_K)$, and its interconnection link between stages k and $k+1$ is uniquely determined by switching box $(m_1, \cdots, m_{k-1}, n_{k+1}, \cdots, n_K)$ at stage k and switching box $(m_1, \cdots, m_k, n_{k+2}, \cdots, n_K)$ at stage $k+1$. The *topology* of one MIN means the incidence relationship among its switching boxes and interconnection links. Then, the buddy-type network obtained from the above procedure *completely* characterizes its topology.

The difference between a MIN and a crossbar can be compared as follows. For an $(N \times M)$-crossbar, we need NM switching links, each of which is a dedicated path for an input-output pair. In contrast, for a buddy-type MIN corresponding to (2.4), $\sum_{k=1}^{K} U^k N_k M_k$ switching links and $\sum_{k=1}^{K-1} V^{k,k+1} M_k N_{k+1}$ interconnection links are required. Consider a special case where $N = M = 2^K$ and $N_i = M_i = 2$, $1 \le i \le K$. Then, there are N^2 switching links in a crossbar while there are $2N\log N$ switching links and $N\log N - N$ interconnection links in a MIN. Hence, hardware complexity for a crossbar and a MIN is $O(N^2)$ and $O(N\log N)$ respectively.

Though fewer links are required in a MIN, each link has to be shared by several input-output paths. A switching link at stage k is determined by two intermediate states, $q^{k-1}=(m_{1,} \cdots, m_{k-1}, n_k, \cdots, n_K)$ and $q^k=(m_{1,} \cdots, m_k, n_{k+1}, \cdots, n_K)$, so it is shared by all paths from input states $(n'_{1,} \cdots, n'_{k-1}, n_k, \cdots, n_K)$ to output states $(m_{1,} \cdots, m_k, m'_{k+1}, \cdots, m'_K)$, where $n'_i \in N_i$, $1 \le i \le k-1$ and $m'_j \in M_j$, $k+1 \le j \le K$. There are totally $N_1 \cdots N_{k-1} M_{k+1} \cdots M_K$ paths sharing this switching link. An interconnection link between stages k and $k+1$

represents an intermediate state $q^k = (m_{1,} \cdots, m_k, n_{k+1}, \cdots, n_K)$ which is an intermediate state for all state transition paths from input states $(n'_{1,} \cdots, n'_k, n_{k+1}, \cdots, n_K)$ to output states $(m_{1,} \cdots, m_k, m'_{k+1}, \cdots, m'_K)$, where $n'_i \in N_i$, $1 \leq i \leq k$ and $m'_j \in M_j$, $k+1 \leq j \leq K$. Therefore, all these paths share this common interconnection link and there are totally $N_1 \cdots N_k\ M_{k+1} \cdots M_K$ such paths. Since different paths share the same link in MIN's, hardware complexity is reduced. Nevertheless, it results in a nonzero block probability.

2.3 Two-dimensional Layout and Topological Equivalence

Strictly speaking, the topological design of a MIN only requires the above two steps, the specification of a switching model and the realization of the model by a buddy-type network. The third step, i.e. the two-dimensional layout, is introduced primarily for two reasons. First, we want to relate the buddy-type network to conventional MIN's. Second, the layout of a MIN in the plane becomes necessary, if it is implemented by the printing circuit board (PC board) or the integrated circuit (IC) technology.

Conventionally, switching diagrams of MIN's are drawn in the plane, where the horizontal dimension is used to arrange the switching sequences, or stages, and the vertical dimension is used to order input terminals, output terminals, and switching boxes at each stage. For such switching diagrams, there essentially exist one-dimensional orderings between input terminals, between output terminals, and between switching boxes at every stage. In other words, instead of applying indices with *multidimensional* coordinates, the conventional approach uses indices with *one-dimensional* coordinate for inputs, outputs, and switching boxes.

For a given buddy-type network, we can layout many different configurations in the plane. Consider a K-stage buddy-type network, input and output terminals of which are characterized by K-dimensional indices $\mathbf{n}$ and $\mathbf{m}$, and switching boxes of which are characterized by a $(K-1)$-dimensional index $\mathbf{u}^k$. To layout a MIN in the plane is equivalent to finding a one-to-one mapping which maps K-dimensional or $(K-1)$-dimensional indices into 1-dimensional indices. There are totally $N!M!\prod_{k=1}^{K} U^k!$ mappings, since there are N, M, and U^k elements in the input set I, the output set O, and the space $\bar{U}^k$ respectively. As to the terminals ($\mathbf{u}^k$; n_k) (the terminals ($\mathbf{u}^k$; m_k)) which are associated with switching box $\mathbf{u}^k$ and do not belong to the set of input terminals (output terminals), they are conventionally ordered by the remaining index n_k (m_k) along the input (output) side of box $\mathbf{u}^k$.

Some particular mappings give familiar types of MIN's such as the omega network and the SW banyan network. Nevertheless, these MIN's are all topologically equivalent since the mappings do not change the incidence relationship among switching boxes and interconnection links.

2.4 An Example

We use an example to demonstrate the design of various unique-path MIN's by the above approach. Suppose we want to design a 4-stage interconnection network with 16 inputs and 16 outputs.

2.4.1 Specification of the Switching Model

Consider the following decompositions:

$$N = 16 = 2 \times 2 \times 2 \times 2 \ , \quad \text{and} \quad M = 16 = 2 \times 2 \times 2 \times 2 \ .$$

Then, we have

$$\overline{N_i} = \overline{M_i} = \{ 0 \ , 1 \} \ , \quad i = 1, 2, 3, 4 \quad ,$$

$$I = \{ (n_1 \ , n_2 \ , n_3 \ , n_4) \mid n_i \in \overline{N_i} \ , \ \textit{where } i = 1, 2, 3, 4 \} \ ,$$

$$O = \{ (m_1 \ , m_2 \ , m_3 \ , m_4) \mid m_i \in \overline{M_i} \ , \ \textit{where } i = 1, 2, 3, 4 \} \ ,$$

and a typical state transition process is given by

$$S = \{ (q^0, \cdots ,q^k, \cdots ,q^4) \mid q^0 \in I \ , q^k \in O^k \quad 1 \le k \le 4 \} \ ,$$

where

$$O^k = \overline{M}_1 \times \overline{M}_2 \times \ \cdots \ \times \overline{M}_{k-1} \times \overline{M}_k \times \overline{N}_{k+1} \times ... \times \overline{N}_K \ .$$

So, the switching model is

$$X = (n_1 \rightarrow m_1 \ ; n_2 \rightarrow m_2 \ ; n_3 \rightarrow m_3 \ ; n_4 \rightarrow m_4) \ . \tag{2.5}$$

2.4.2 Buddy-type Network Realization

Based on the switching model (2.5), we can construct a buddy-type network which has 16 input terminals, 16 output terminals, and 4 stages. The input and output terminals are with indices **n** and **m**,

$$\mathbf{n} = (n_1 , n_2 , n_3 , n_4) ,$$

$$\mathbf{m} = (m_1 , m_2 , m_3 , m_4) .$$

At stage k, there are eight switching boxes with indices $\mathbf{u}^k$:

$$\mathbf{u}^1 = (n_2 , n_3 , n_4) ,$$

$$\mathbf{u}^2 = (m_1 , n_3 , n_4) ,$$

$$\mathbf{u}^3 = (m_1 , m_2 , n_4) ,$$

$$\mathbf{u}^4 = (m_1 , m_2 , m_3) .$$

Each switching box is a 2 × 2 crossbar; that is, there are 4 switching links within a box. As to the interconnection links, the interconnection rule is:

(1) $\mathbf{u}^1$ and $\mathbf{u}^2$ are connected if they have the same values for both n_3 and n_4;

(2) $\mathbf{u}^2$ and $\mathbf{u}^3$ are connected if they have the same values for both m_1 and n_4;

(3) $\mathbf{u}^3$ and $\mathbf{u}^4$ are connected if they have the same values for both m_1 and m_2.

2.4.3 Two-dimensional Layout

The one-to-one mapping from $\mathbf{n}$, $\mathbf{m}$, and $\mathbf{u}^k$, to one-dimensional indices can be denoted by

$$(a_1 , a_2 , a_3 , a_4) \rightarrow a_4 a_3 a_2 a_1 ,$$

where a_i is either 0 or 1 and $a_4 a_3 a_2 a_1 = 8a_4 + 4a_3 + 2a_2 + a_1$ is a binary representation. We show how six well known MIN's can be obtained by different mappings in Table 1. Their corresponding switching diagrams are given in Figure 2. They are all topologically equivalent since they are derived from the same buddy-type network.

The complicated interconnection patterns of interconnection links shown in Figure 2 have been the main focus in the literature on the topology of MIN's and these patterns are also used to classify different types of MIN's. However, from our point of view, the complication is simply due to the effect of layout in the two-dimensional plane. The distinction between them disappears when we go to the original buddy-type network.

	Baseline	Reverse Baseline	Data Manipulator	Regular SW Banyan	Omega	Flip
input	$n_4n_3n_2n_1$	$n_4n_3n_2n_1$	$n_2n_3n_4n_1$	$n_4n_3n_2n_1$	$n_1n_2n_3n_4$	$n_4n_3n_2n_1$
stage 1	$n_4n_3n_2$	$n_4n_3n_2$	$n_2n_3n_4$	$n_4n_3n_2$	$n_2n_3n_4$	$n_4n_3n_2$
stage 2	$m_1n_4n_3$	$n_4n_3m_1$	$m_1n_3n_4$	$n_4n_3m_1$	$n_3n_4m_1$	$m_1n_4n_3$
stage 3	$m_1m_2n_4$	$n_4m_1m_2$	$m_1m_2n_4$	$n_4m_2m_1$	$n_4m_1m_2$	$m_2m_1n_4$
stage 4	$m_1m_2m_3$	$m_1m_2m_3$	$m_1m_2m_3$	$m_3m_2m_1$	$m_1m_2m_3$	$m_3m_2m_1$
output	$m_1m_2m_3m_4$	$m_1m_2m_3m_4$	$m_1m_2m_3m_4$	$m_3m_2m_1m_4$	$m_1m_2m_3m_4$	$m_4m_3m_2m_1$

Table 1. Several isomorphic MINs

There exist nonbuddy-type networks. One example is the CC banyan network [5]. The 4-stage CC banyan, which is shown in Figure 3, is not isomorphic to the six MIN's given in Figure 2. Although characterized by the same switching model (2.5), it cannot be realized by the buddy-type interconnection scheme described in Section 2.2.

2.4.4 The Self-Routing Scheme

The routing scheme for a buddy-type network is very simple, which is usually called the *self-routing* scheme, and can be done distributedly. Each request from an input terminal only has to contain the address of the output terminal coded as a *routing tag,* say ($m_1 m_2 \cdots m_K$). At stage k, the k-th index m_k can be used to make the switching decision. Used indices can be thrown away, since they are not used any more. It is easy to see that the routing scheme depends on the switching model X only. As a consequence, when we specify the switching model, the routing scheme for the network is determined at the same time.

3. A Systematic Approach to Multiple-Path MIN Design

In Section 2, we developed a systematic approach to unique-path MIN design. In this section, we generalize this approach to multiple-path, or fault-tolerant, MIN design.

3.1 The Extra Intermediate State Method

The uniqueness of the path between every input-output pair of a buddy-type network comes from the fact that there exists a unique state transition path from every input state to every output state in the switching model. By the *extra intermediate state* method, we provide some extra intermediate states in the state transition process so that there exist some distinct state transition paths from an input state to an output state. These extra intermediate states can be characterized by the *extra index*.

Suppose we treat all input-output pairs in the same fashion so that we can focus on a typical pair, say from input $(n_1, \cdots, n_k, \cdots, n_K)$ to output $(m_1, \cdots, m_k, \cdots, m_K)$. A unique-path switching model is given by

$$X = (n_1 \rightarrow m_1 ; \cdots ; n_k \rightarrow m_k ; \cdots ; n_K \rightarrow m_K) .$$

Now, we add an extra index $r_k \in \{ 0 , 1 \}$ so that transition in the k-th coordinate is composed of two steps: first, from a state with input index n_k to a state with extra index r_k; then, from a state with extra index r_k to a state with output index m_k. These two transition steps are denoted by $n_k \rightarrow r_k$ and $r_k \rightarrow m_k$ respectively. Unlike n_k and m_k, which are uniquely specified by the input and output states, the extra index r_k can take two different values. Hence, we obtain two distinct state transition paths.

There exists a natural constraint for the two transition steps in the k-th coordinate. That is, $n_k \rightarrow r_k$ should proceed before $r_k \rightarrow m_k$. Although it is allowable that the transition $r_k \rightarrow m_k$ follows directly the transition $n_k \rightarrow r_k$, it is not a good fault-tolerant design. It is usually preferable to put these two transition steps apart as far as possible, which will be discussed in detail in Section 3.2.

Several examples are given below to illustrate the extra intermediate state method. Let

$$I = \overline{N}_1 \times \overline{N}_2 \times \overline{N}_3 \ , \qquad O = \overline{M}_1 \times \overline{M}_2 \times \overline{M}_3 \ , \tag{3.1}$$

where $\overline{N}_i = \overline{M}_i = \{ 0 , 1 \}$, $i = 1 , 2 , 3$. Three extra indices r_1, r_2 and t_1 are introduced, where $r_1 , r_2 \in \{ 0 , 1 \}$ and $t_1 \in \{ 0 , 1 , 2 \}$. Several multiple-path switching models are given by

$$Model\ A : \ X_A = (n_1 \rightarrow r_1 \ ; r_1 \rightarrow m_1 \ ; n_2 \rightarrow r_2 \ ; n_3 \rightarrow m_3 \ ; r_2 \rightarrow m_2) \ , \tag{3.2a}$$

$$Model\ B : \ X_B = (n_1 \rightarrow r_1 \ ; n_2 \rightarrow r_2 \ ; n_3 \rightarrow m_3 \ ; r_2 \rightarrow m_2 \ ; r_1 \rightarrow m_1) \ , \tag{3.2b}$$

$$Model\ C : \ X_C = (n_1 \rightarrow r_1 \ ; n_2 \rightarrow r_2 \ ; n_3 \rightarrow m_2 \ ; r_1 \rightarrow m_1 \ ; r_2 \rightarrow m_2) \ , \tag{3.2c}$$

$$Model\ D : \ X_D = (n_1 \rightarrow r_1 \ ; r_1 \rightarrow t_1 \ ; n_2 \rightarrow m_2 \ ; n_3 \rightarrow m_3 \ ; t_1 \rightarrow m_1) \ , \tag{3.2d}$$

$$Model\ E : \ X_E = (n_1 \rightarrow (r_1,t_1) \ ; n_2 \rightarrow m_2 \ ; n_3 \rightarrow m_3 \ ; (r_1,t_1) \rightarrow m_1) \ . \tag{3.2e}$$

Their state transition graphs are depicted in Figure 5 (a)-(e), where typical directed links are labeled by their corresponding state transition steps and nodes are labeled by their corresponding extra indices. From these switching models, we can write down their k-th step input space I^k and output space O^k easily. We do not allow the extra indices r_1, r_2, and t_1 to appear in the input and output states, since we are primarily concerned with the redundancy of communication resources (networks) and not the redundancy of input and output resources (processor elements and memory modules).

There are some easy ways to manipulate switching models. Two state transition steps $a_1 \rightarrow b_1$ and $a_2 \rightarrow b_2$ are called *independent,* if a_1, a_2, b_1, and b_2 are different from each other. The combination of two or more consecutive independent state transition steps, say

$$(\cdots ; a_1 \rightarrow b_1 ; a_2 \rightarrow b_2 ; \cdots) \tag{3.3}$$

into one state transition step, say

$$(\cdots ; (a_1,a_2) \rightarrow (b_1,b_2) ; \cdots) \tag{3.4}$$

is called *compression.* On the other hand, the decomposition of a state transition step into two or more state transition steps is called *expansion.* For example, the first and last two state transition steps of model X_B are independent of each other. By compression, we can combine them respectively and obtain a new switching model

$$\textit{Model } F : X_F = ((n_1,n_2)\rightarrow(r_1,r_2) ; n_3\rightarrow m_3 ; (r_2,r_1)\rightarrow(m_2,m_1)) . \tag{3.5}$$

We can also view that model X_B is obtained from model X_F by expansion. Notice that the switching model X_{MIN} of a unique-path MIN is in fact an expansion of the switching model X_{CB} of a crossbar.

3.2 Optimal Design of Multiple-Path Switching Models

The fault-tolerant property of a multiple-path switching model can be easily understood by examining its state transition graph. Since a state corresponds to a node and a state transition step corresponds to a directed link, a fault in a state and a state transition step is equivalent to the removal of a node and a directed link in the state transition graph respectively.

The state transition graph is a special type of digraph [15]. Recall that the *distance* from node a to node b is the number of directed links of any shortest path from a and b. Then, the following lemma characterizes the special property of the state transition graph.

Lemma 1: Nodes at the same distance from the input node are characterized by the same set of indices. If this set contains extra indices, the number of these nodes is equal to the number of different possible values assumed by the extra indices. Otherwise, it is 1.

Proof: Since there is a one-to-one correspondence between a switching model and a state transition graph, if two nodes are at the same distance from the input node, then their corresponding states are obtained from the input state through the same number of state transition steps. Hence, they are characterized by the same set of indices. The number of these nodes is the same as that of all possible intermediate states determined by the set of indices. If the set does not contain any extra index, there exists a unique intermediate state specified by input and output indices. If the set contains some extra indices, the number of intermediate states is that of different possible values assumed by these extra indices.

Q.E.D.

The *node-connectivity* (*link-connectivity*) of a state transition graph is the minimum number of nodes (links) whose removal disconnects the output state node from the input state node. Since the effect of removing a node is the same as that of removing all its incident links, the node connectivity of a state transition graph can never exceed its link connectivity. Therefore, we say a state transition graph is *C-connected* if its node connectivity is C. It is easy to see that the state transition graph A is 1-connected, the state transition graphs B, C, D are 2-connected, and the state transition graph E is 6-connected. If the number of nodes at distance d from the the input node is denoted by $N(d)$, we have

$$C = \min_{0 < d < D} N(d) \tag{3.6}$$

where C is the node connectivity and D is the distance between the input and output nodes.

State transition paths are called *disjoint,* if they have no links and no nodes other than the input and output nodes in common. The relationship between the node connectivity and the number of disjoint state transition paths is stated in Lemma 2.

Lemma 2: The node connectivity of a state transition graph is the same as the number of its disjoint state transition paths from the input node to the output node.

Proof: Suppose the node connectivity of a given state transition graph is C and the number of its disjoint state transition paths is C'. By definition, the removal of C appropriate nodes from this state transition graph disconnects the input node from the output node, so it cannot have more than C disjoint state transition paths, i.e. $C' \leq C$. On the other hand, from the structure of the switching model and (3.6), we can construct at least C disjoint state transition paths from the input to the output. Hence, $C' \geq C$. Combining the above arguments, we have $C' = C$

Q.E.D.

The main result is that, in order to increase the connectivity of a state transition graph by introducing a set of extra indices, the best design is to bring them to the state vector as early as possible and remove them as late as possible. We summarize it below.

Theorem 3: For all state transition graphs given by switching models which contain a set of extra indices $\{ r_1 , \cdots , r_p \}$, the state transition graph corresponding to the switching model

$$X_{opt} = (n_1 \rightarrow (r_1, \cdots ,r_p) \, ; n_2 \rightarrow m_2 \, ; \; \cdots \; ; n_K \rightarrow m_K \, ; (r_1, \cdots ,r_p) \rightarrow m_1)$$

has the maximum number of node connectivity.

Proof: It is easy to see that X_{opt} gives the largest number of disjoint state transition paths in its state transition graph. Hence, by Lemma 2, we know the graph has the maximum number of node connectivity.

Q.E.D.

From Lemmas 1 and 2, we know that the compression of a switching model does not change the number of disjoint paths in its state transition graph. Therefore, any switching model obtained by compressing X_{opt} also has the maximum number of node connectivity.

A *R-path switching model* is a model whose state transition graph has R disjoint state transition paths. A R-path switching model can at least tolerate $R-1$ faults since its state transition graph is R-connected by Lemma 2. To design a R-path switching model, the simplest scheme is to use one extra index $r_1 \in \{ 0 , 1 , \cdots R-1 \}$ and to use the switching model:

$$X = (n_1 \rightarrow r_1 ; n_2 \rightarrow m_2 ; \cdots ; n_k \rightarrow m_k ; \cdots ; n_K \rightarrow m_K ; r_1 \rightarrow m_1) \quad (3.7)$$

By compressing (3.7), we can obtain many other R-path switching models.

3.3 Realization, Layout, and Routing for Multiple-path Buddy-type MIN's

According to the constructive procedure described in Section 2.2, the realization of the R-path switching model (3.7) by a buddy-type network is straightforward. The realization procedure can be summarized as:

(1) the unchanged indices at stage k can be used as indices for switching boxes, and the input and output terminals are fully connected within each switching box;

(2) the unchanged indices at stages $k-1$ and k can be used as indices for groups formed by switching boxes at these two stages, and the switching boxes at stage $k-1$ are fully connected to the switching boxes at stage k by interconnection links if they have the same group indices.

By mapping the multidimensional indices to the one-dimensional indices, we can layout the buddy-type network in the two-dimensional plane. The details can be found in Section 2.3.

A simple routing scheme for the above R-path network is similar to that for the unique-path network, except that we have to generate a random number from the set $\{ 0 , 1 , \cdots R-1 \}$ for the extra index r_1 so that we can choose one of R disjoint paths between every input-output pair under the fault-free condition. Suppose a single fault is detected and it is known which path is disconnected, we still can randomly choose one of the $R-1$ remaining fault-free paths. It seems feasible to use more complicated routing schemes to determine a "good" path among all redundant paths to avoid congestion occurring in some spots of the network. However, we will not go to the issue in this paper.

3.4 Two Design Examples

Let us consider the design of a R-path MIN with N inputs and M outputs. We can use the above systematic approach to derive the extra stage cube and the multipath omega networks.

3.4.1 Extra Stage Cube

Step 1: Suppose we can factorize N and M as follows:

$$N = N_1 \times N_2 \times \cdots \times N_K \ , \quad M = M_1 \times M_2 \times \cdots \times M_K \ .$$

According to the discussion in Section 3.2, we choose the switching model

$$X_{ESC} = (\, n_1 \rightarrow r_1 \, ; n_2 \rightarrow m_2 \, ; \ \cdots \ ; n_k \rightarrow m_k \, ; \ \cdots \ ; n_K \rightarrow m_K \, ; r_1 \rightarrow m_1 \,), \tag{3.8}$$

where $n_k \in \{ 0, \cdots ,N_k-1 \}$, $m_k \in \{ 0, \cdots ,M_k-1 \}$, and $r_1 \in \{ 0, \cdots ,R-1 \}$.

Step 2: We use a $(K+1)$-stage buddy-type network to realize X_{ESC}. Switching boxes at stages k are characterized by indices $\mathbf{u}^k$,

$$\mathbf{u}^1 = (\, n_2 \, , \ldots , \, n_K \,) \, , \tag{3.9a}$$

$$\mathbf{u}^k = (\, r_1 \, , m_2 \, , \ldots , \, m_{k-1} \, , n_{k+1} \, , \ldots , \, n_K \,) \, , \quad 2 \le k \le K \ , \tag{3.9b}$$

$$\mathbf{u}^{K+1} = (\, m_2 \, , \ldots , \, m_K \,) \, . \tag{3.9c}$$

where switching box $\mathbf{u}^1$ is a $(N_1 \times R)$-crossbar, switching box $\mathbf{u}^k$, $k = 2, \cdots , K$ is a $(N_k \times M_k)$-crossbar, and switching box $\mathbf{u}^{K+1}$ is a $(R \times M_1)$-crossbar. Switching boxes at two consecutive stages have $K-2$ indices in common, and they should be connected if they have the same values for all these common indices.

Step 3: To obtain the cube-type configuration in the plane, we use the following mapping,

$$\mathbf{n} = (\, n_1 \, , \ldots , \, n_K \,) \rightarrow n_2 n_3 \cdots n_{K-1} n_K n_1 \quad ,$$

$$\mathbf{u}^1 = (\, n_2 \, , \ldots , \, n_K \,) \rightarrow n_2 n_3 \cdots n_{K-1} n_K \quad ,$$

$$\mathbf{u}^k = (\, r_1 \, , m_2 \, , \ldots , \, m_{k-1} \, , n_{k+1} \, , \ldots , \, n_K \,) \rightarrow r_1 m_2 \cdots m_{k-1} n_{k+1} \cdots n_K \ ,$$

$$2 \le k \le K+1 \ ,$$

$$\mathbf{u}^{K+1} = (m_2 , \ldots , m_K) \rightarrow m_2 m_3 \cdots m_{K-1} m_K \ ,$$

$$\mathbf{m} = (m_1 , \ldots , m_K) \rightarrow m_2 m_3 \cdots m_{K-1} m_K m_1 \ .$$

Nevertheless, the third step is not crucial in the above design procedure.

Since we can eliminate the extra index r_1 in (3.8) and realize it with a K-stage unique-path network, and since the cube-type configuration is chosen in step 3, the resulting network is called the *extra stage cube.* It is easy to see that from any input to any output there are R disjoint state transition paths characterized by $r_1 = 0 , \cdots , R-1$. From (3.9), we know that two disjoint state transition paths do not share common switching boxes except at stages 1 and $K+1$. They neither share a common interconnection link since each interconnection link is uniquely specified by switching boxes at its both ends. So, they are also physically disjoint. The extra stage cube can tolerate at least $R-1$ faults, provided that the first and last stages are fault-free.

Since all redundant paths for a given input-output pair are contained in the same switching boxes at the first and last stages, one single fault in the switching boxes of these two stages can disconnect the path for this input-output pair. As a consequence, we have to increase the reliability of the components of these two stages by some schemes. For example, the extra stage cube uses some multiplexing scheme so that the message can bypass faulty switching boxes at these two stages [3].

3.4.2 Multipath Omega

Step 1: Suppose we can factorize N and M as follows:

$$N = (N_1 \times \cdots \times N_b) \times (N_{b+1} \times \cdots \times N_{2b}) \times \cdots$$

$$\cdots \times (N_{(p-1)b+1} \times \cdots \times N_{pb}) \times (N_{pb+1} \times \cdots \times N_{pb+q}) ,$$

$$M = (M_1 \times \cdots \times M_b) \times (M_{b+1} \times \cdots \times M_{2b}) \times \cdots$$

$$\cdots \times (M_{(p-1)b+1} \times \cdots \times M_{pb}) \times (M_{pb+1} \times \cdots \times M_{pb+q}) ,$$

$$R = R_1 \times \cdots \times R_{b-q} .$$

Then, we choose the switching model

$$X_{MPO} = ((n_1, n_2, \cdots, n_b) \rightarrow (r_1, r_2, \cdots, r_{b-q}, m_{b-q+1}, \cdots, m_b) ;$$

$$(n_{b+1}, \cdots, n_{2b}) \rightarrow (m_b, \cdots, m_{2b}) ; \cdots ;$$

$$(n_{(p-1)b+1}, \cdots, n_{pb}) \rightarrow (m_{(p-1)b}, \cdots, m_{pb}) ;$$

$$(n_{pb+1}, \cdots, n_{pb+q}, r_1, \cdots, r_{b-q}) \rightarrow (m_{pb+1}, \cdots, m_{pb+q}, m_1, \cdots, m_{b-q})) . \quad (3.10)$$

where $n_k \in \{ 0, \cdots, N_k - 1 \}$, $m_k \in \{ 0, \cdots, M_k - 1 \}$, and $r_k \in \{ 0, \cdots, R_k - 1 \}$.

Step 2: X_{MPO} is realized by a $(p+1)$-stage buddy-type network.

Step 3: By some appropriate mapping, we obtain the modified multipath omega network described in [4].

The motivation for selecting the switching model (3.10) can be intuitively given as follows. Suppose $N = M = 2^K$ and only one standard type of switching box with size $2^b \times 2^b$, where $1 < b < K$, is used to implement the network. In order to fully utilize this type of switching box, we have to merge b input or output binary indices into a single vector index. In general, the relationship between K and b can be related by

$$K = p\ b + q \quad ,$$

where p and q are integer, $p > 0$, and $b > q \geq 0$. If K happens to be multiples of b, then q is zero and we can divide the input and output coordinates into p groups, each of which contains b binary indices. Then, a unique-path network made of p-stage $(2^b \times 2^b)$-switching boxes can be obtained. If q is nonzero, we can divide the original K binary indices into $p+1$ groups, where p groups contain b binary indices and one group contains q binary indices. For those groups containing b binary indices, they can make full use of the given $(2^b \times 2^b)$-switching boxes. For the group which has q binary indices, it only uses part of the given switching box with 2^{b-q} input terminals and 2^{b-q} output terminals left. Therefore, we can introduce $b-q$ extra binary indices r_1, r_2, $\cdots$,r_{b-q} to fully utilize the unused portion.

The modified omega network is a *uniform network* which contains only one type of switching boxes. However, it is also feasible to design nonuniform multipath omega networks which have several types of switching boxes [4]. We can derive these various types of multipath omega by the same procedure described above.

3.4.3 Comparison

We can use a simple example, i.e. a 2-path MIN with 8 inputs and 8 outputs to illustrate the above two designs. The switching model of the extra stage cube is

$$X_{ESC} = (\ n_1 \rightarrow r_1\ ;\ n_2 \rightarrow m_2\ ;\ n_3 \rightarrow m_3\ ;\ r_1 \rightarrow m_1\ )\ . \qquad (3.11a)$$

Its corresponding state transition graph and switching diagram is given in Figure 5.

The switching model of the multipath omega is

$$X_{MPO} = ((n_1,n_2) \rightarrow (r_1,m_2) \; ; \; (n_3,r_1) \rightarrow (m_3,m_1)) \; . \qquad (3.11b)$$

Its corresponding state transition graph and switching diagram is given in Figure 6.

If we examine their buddy-type networks or their two-dimensional layouts, the extra stage cube and the multipath appear quite differently. However, comparing their switching models (3.11a) and (3.11b), we find that the difference is basically in different implementation of the first and the last two switching steps. The sequential implementation gives the extra stage cube while the parallel implementation gives the multipath omega. They are an *expansion/compression* pair.

Detailed cost and performance analysis of these two designs is beyond the scope of this paper. Roughly speaking, a MIN using *larger* switching boxes has lower block probability and requires less access time from the input to the output. This can be easily understood by examining the extreme cases, i.e. the $(2^K \times 2^K)$-crossbar and the K-stage MIN. The tradeoff for better performance is higher hardware complexity.

4. Conclusions and Extensions

We proposed a systematic approach to the topological design of reliable MIN's in this paper. This approach decomposes the design procedure into three steps: the specification of a switching model, the buddy-type network realization, and the two-dimensional layout.

Both cost and performance have to be considered in selecting an appropriate switching model. Since hardware cost has been reduced tremendously due to the VLSI technology, there is a new tradeoff between cost and performance. In order to get better performance, MIN's with larger switching boxes and fewer number of stages may be preferred. However, more quantitative analysis is still required. The performance analysis, comparison, and simulation of MIN's are all current research problems.

In terms of network realization, it seems reasonable to stick to the buddy-type network unless there is a reason in favor of some particular nonbuddy-type network. A systematic procedure for constructing buddy-type networks is clearly given here. However, how to realize a switching model by a nonbuddy-type network systematically is still an open question.

If switching boxes are connected by wires, there is no layout problem. However, since MIN's are often implemented either on PC board or with IC technology, the layout is an important issue in practice. On one hand, we want to minimize the layout area. On the other hand, the interconnection pattern should be as regular as possible to simplify the layout. For example, in terms of the regularity between stages, both the omega and the flip networks have the same type of interconnection pattern between any two consecutive stages (See Figure 2). Nevertheless, they consume more area compared to the other four types of MIN's given in Figure 2. A more detailed study in the layout problem is required.

Acknowledgements

The author wishes to thank helpful comments from Professor John N. Tsitsiklis, and Doctors Jean Regnier and Kevin W. K. Tsai. The author is also grateful to Professor Bernard C. Levy for his support and encouragement.

References

[1] T.-Y. Feng, "A Survey of Interconnection Networks," *Computer*, vol. 14, no. 12, pp. 12-27, Dec. 1981.

[2] C.-L. Wu and T.-Y. Feng, *Tutorial: Interconnection Networks for Parallel and Distributed Processing.* Silver Spring, MD: IEEE Computer Society Press, 1984.

[3] G. B. Adams, III and H. J. Siegel, "The Extra Stage Cube: The Fault-Tolerant Interconnection Network for Supersystems," *IEEE Trans. on Computers*, vol. C-31, no. 5, pp. 443-454, May 1982.

[4] K. Padmanabhan and D. H. Lawrie, "A Class of Redundant Path Multistage Interconnection Networks," *IEEE Trans. on Computers*, vol. C-32, no. 12, pp. 1099-1108, Dec. 1983.

[5] L. R. Goke and G. J. Lipovski, "Banyan Networks for Partitioning Multiprocessor Systems," in *Proc. of 1st Ann. Symp. on Computer Architecture*, 1973, pp. 21-28.

[6] D. H. Lawrie, "Access and Alignment of Data in an Array Processor," *IEEE Trans. on Computers*, vol. C-24, no. 12, pp. 1145-1155, Dec. 1975.

[7] K. E. Batcher, "The Flip Network in STARAN," *Proc. 1976 Int. Conf. on Parallel Processing*, pp. 65-71.

[8] M. C. Pease, III, "The Indirect Binary n-Cube Microprocessor Array," *IEEE Trans. on Computers*, vol. C-26, no. 5, pp. 458-473, May 1975.

[9] C.-L. Wu and T.-Y. Feng, "On a Class of Multistage Interconnection Networks," *IEEE Trans. on Computers*, vol. C-29, no. 8, pp. 694-702, Aug. 1980.

[10] D. S. Parker, "Notes on Schuffle/Exchange-Type Switching Netwroks," *IEEE Trans. on Computers*, vol. C-29, no. 3, pp. 213-222, Mar. 1980.

[11] D. P. Agrawal, "Graph Theoretical Analysis and Design of Multistage Interconnection Networks," *IEEE Trans. on Computers*, vol. C-32, no. 7, pp. 637-648, Jul. 1983.

[12] N. J. Davis IV and H. J. Siegel, "The PASM Prototype Interconnection Network Design," in *1985 AFIPS Conf. Proc.*, pp. 183-190.

[13] K. Padmanabhan, "Fault Tolerant Design and Reliability Considerations for the Interconnection Network in Cedar," Cedar Document No. 37, Laboratory for Advanced Supercomputers, University of Illinois at Urbana-Champaign, Apr. 1984.

[14] H. J. Siegel, "A Model of SIMD Machines and a Comparison of Various Interconnecton Networks," *IEEE Trans. on Computers*, vol. C-28, no. 12, pp. 907-917, Dec. 1979.

[15] F. Harary, *Graph Theory.* Readings, MA: Addison-Wesley, 1969.

Figure Captions

Figure 1: The state transition graph for X_{MIN}.

Figure 2: A class of isomorphic MINs: (a) baseline, (b) reverse baseline, (c) data manipulator, (d) regular SW banyan, (e) omega, and (f) flip MINs.

Figure 3: The switching diagram of a CC banyan network.

Figure 4: State transition graphs for Equation (3.2)

Figure 5: The extra stage cube: (a) switching diagram, and (b) state transition graph.

Figure 6: The multipath omega: (a) switching diagram, and (b) state transition graph.

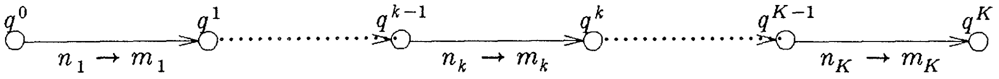


Figure 1: The state transition graph of X_{MIN}

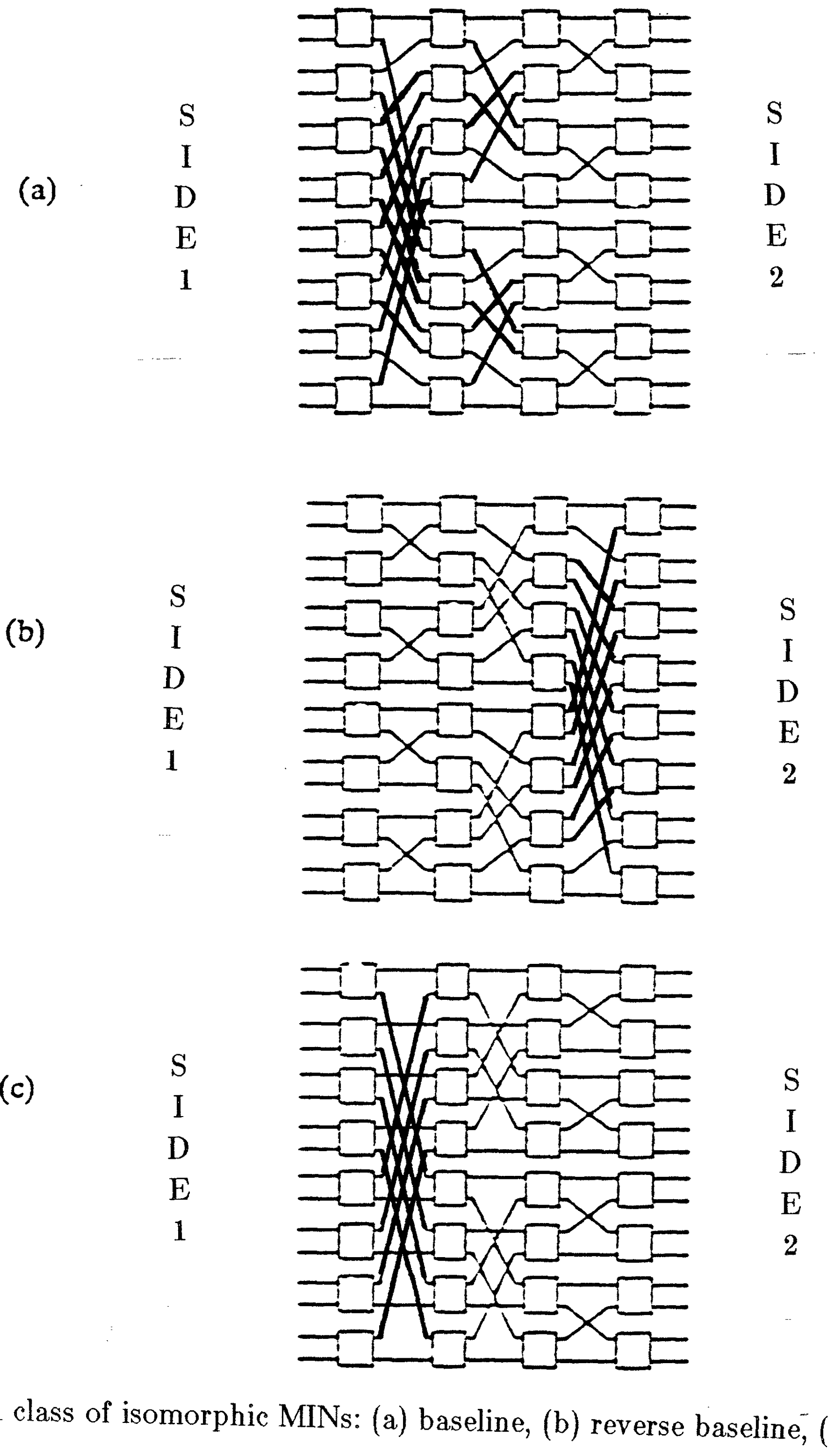


Figure 2: A class of isomorphic MINs: (a) baseline, (b) reverse baseline, (c) data manipulator,

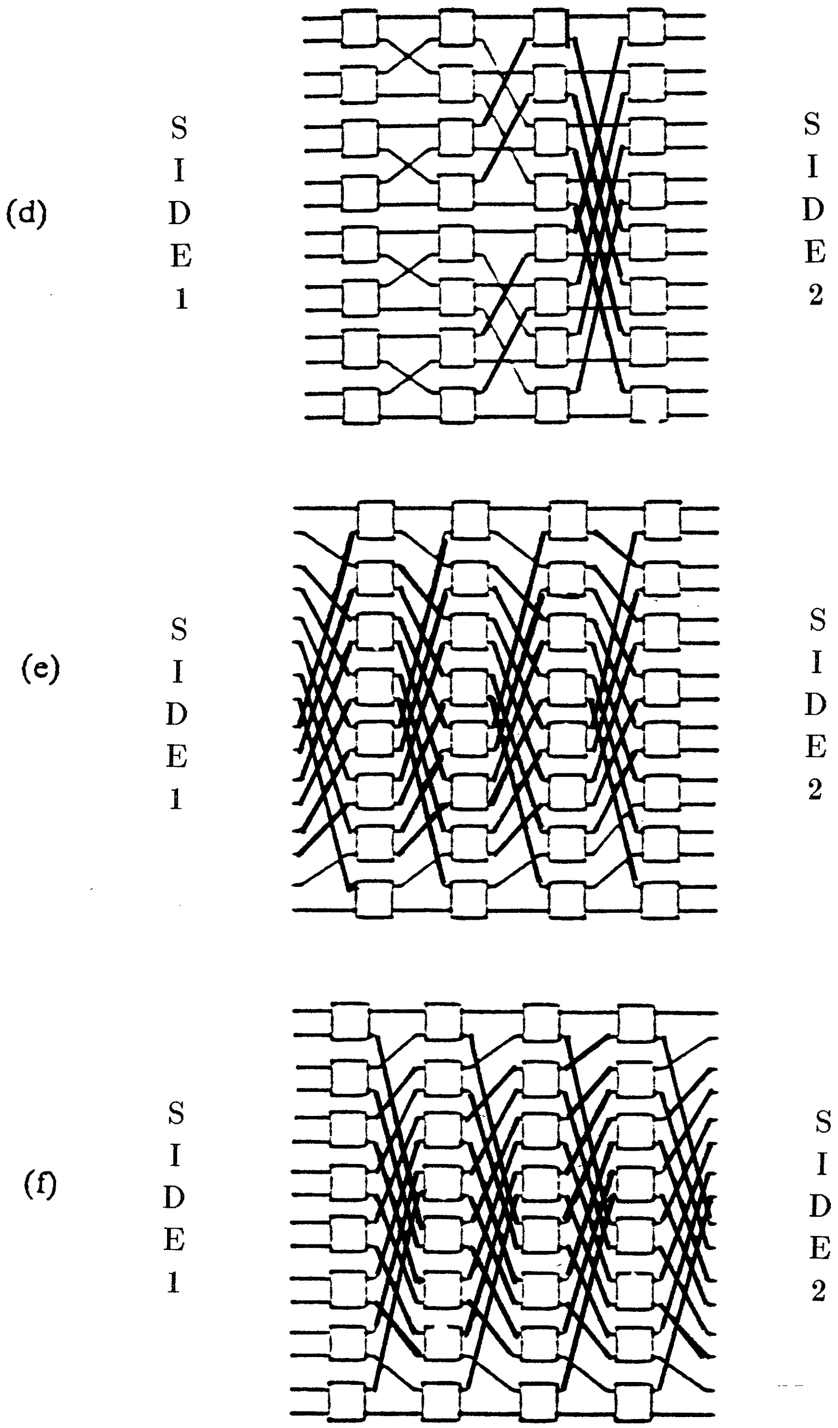


Figure 2 (Cont.): (d) regular SW banyan, (e) omega, and (f) flip MINs.

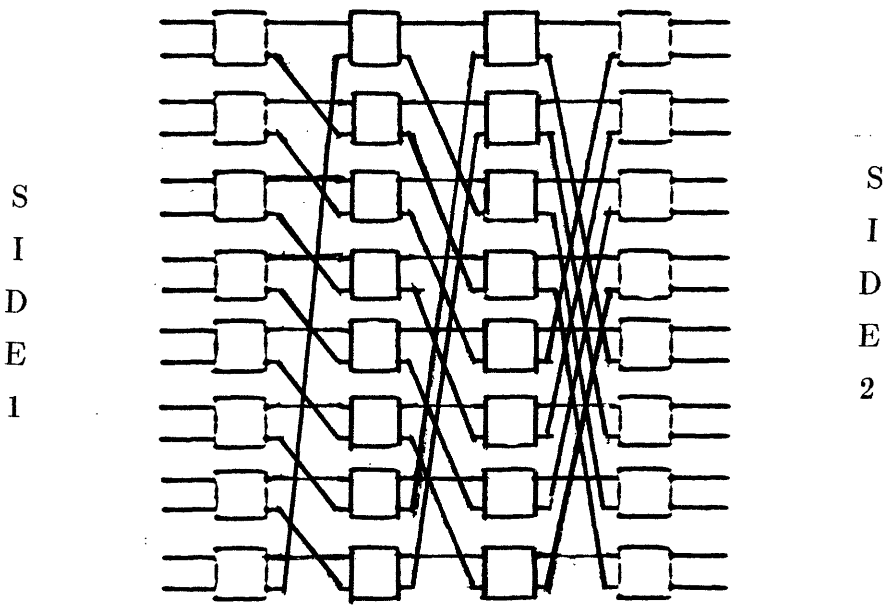


Figure 3: The switching diagram of a CC banyan network.

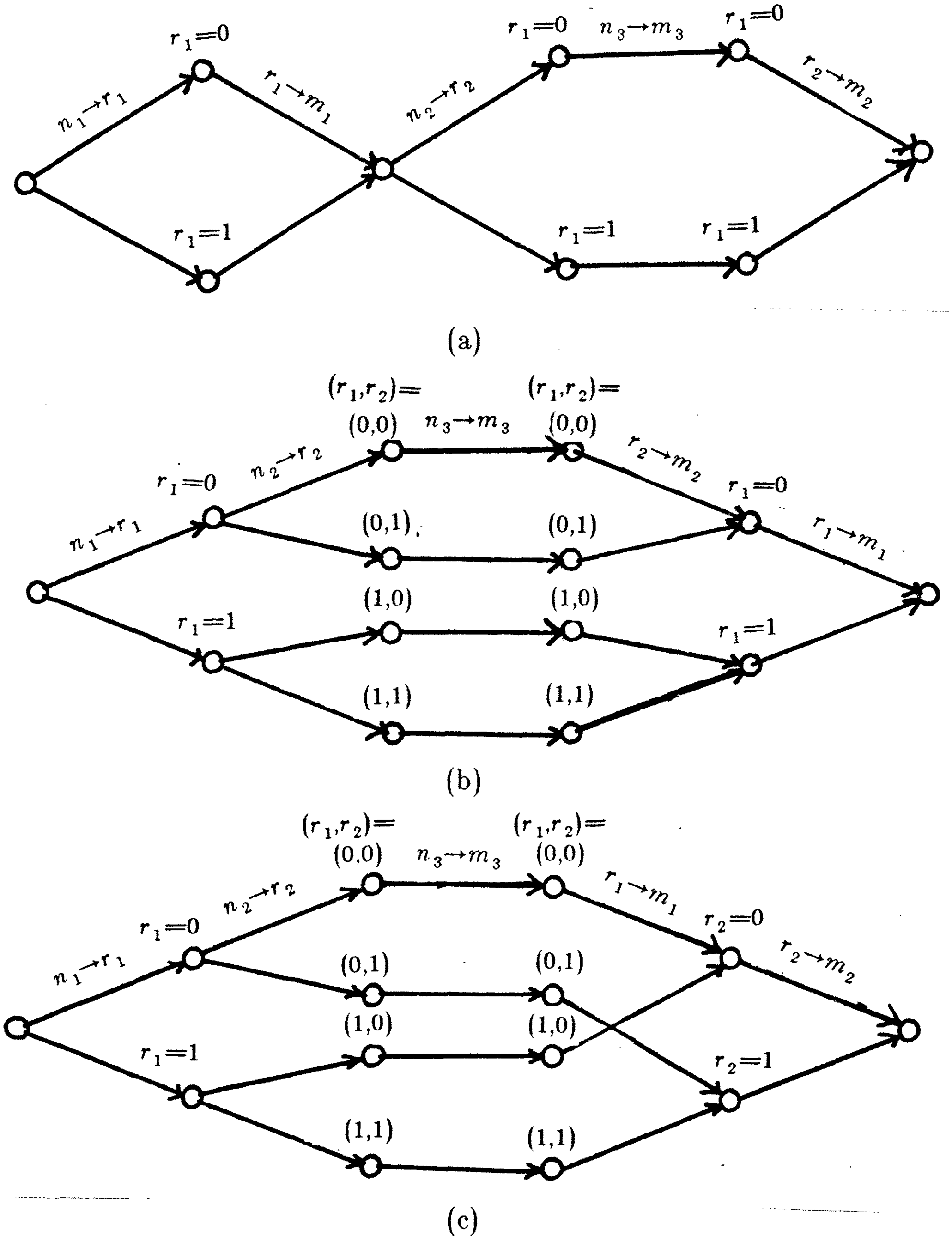
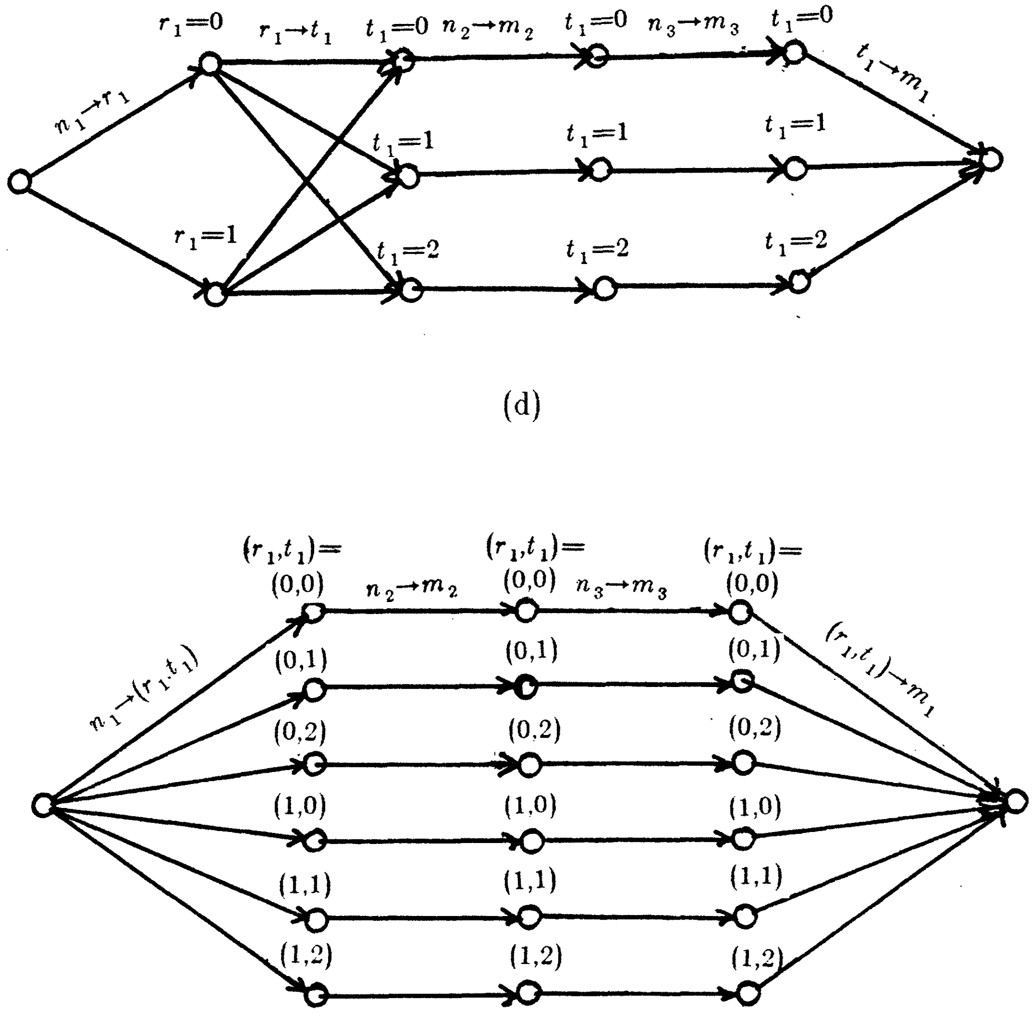


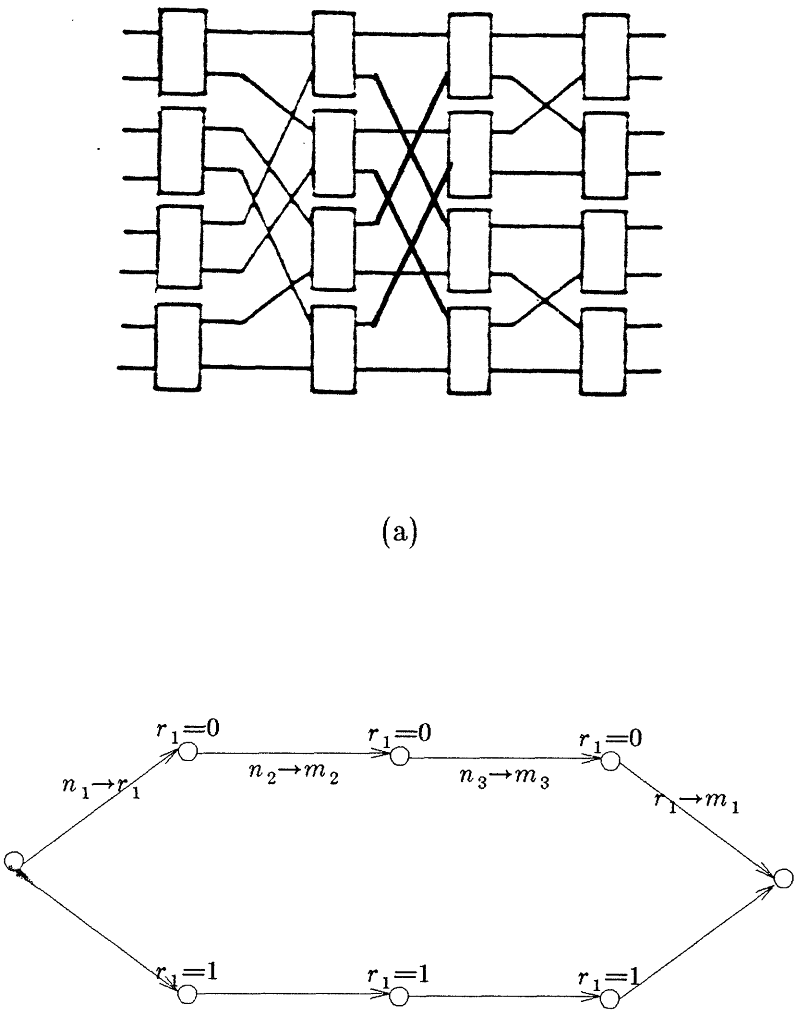
Figure 4: State transition graphs of Equation (3.2)



(d)

(e)

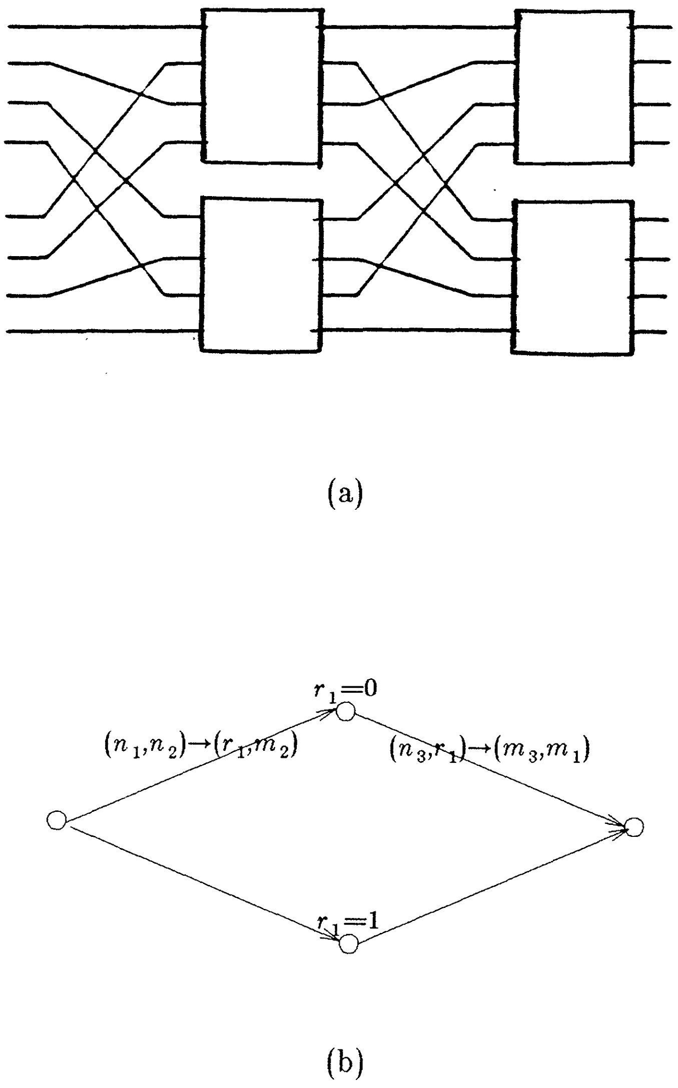
Figure 4 (Cont.): State transition graphs of Equation (3.2)



(a)

(b)

Figure 5: The extra stage cube (a) switching diagram (b) state transition graph.


$r_1=0$
$(n_1,n_2)\rightarrow(r_1,m_2)$
$(n_3,r_1)\rightarrow(m_3,m_1)$
$r_1=1$

(a)

(b)

Figure 6: The multipath omega (a) switching diagram (b) state transition graph.