Polynomial Algorithms for Evaluation of Reliability of Parallel Computer Interconnection Systems

R. K. Dash and C. R. Tripathy

Abstract—Two new and efficient algorithms for evaluating the terminal reliability of parallel computer interconnection networks have been proposed. Both the algorithms are based on multiple decomposition approach. The former is meant for reliability evaluation of multi computers while the later is meant for multi stage interconnection networks. Using the first algorithm the reliability of some important multi computer networks based on cubic architecture have been evaluated. The second algorithm is a modified version of the first one, which takes into account the topological advantages of multi stage interconnection networks during the decomposition process. The terminal reliability of some important fault-tolerant multi-stage interconnection networks has been evaluated using the second algorithm for the purpose of comparison. The complexities of the proposed algorithms are found to be highly polynomial in nature. The proposed method of multiple decomposition is compared with some similar decomposition methods. The simulated results confirm that the proposed decomposition method is much better than its counter parts. The generality of the proposed method is assured by applying the same to evaluate the terminal reliability of Alternating group graph which is also an important interconnection network apart from cube like interconnection networks.

Index Terms—Probabilistic graph, minimal cut, network decomposition, parallel computer network, reliability.

I. INTRODUCTION

The design of reliable parallel computing systems is one of the most important issues facing system engineers today. Almost all of the parallel interconnection schemes may be broadly classified into two groups: static networks and dynamic networks. The dynamic networks are built out by crossed bar switches and are used in shared memory multiprocessors. Some of the important fault tolerant candidates of dynamic networks are Extra stage cube [1], Multi path omega network [1], Extra stage shuffle exchange network [2]. The main advantages associated with these networks are high bandwidth, low diameter and constant degree switches for which they have been used for various commercial machines.

A static network on the other hand has no crossbar type of switching elements and represents a fix scheme of interconnection connecting a collection of stand alone processors. This type of architecture is used as distributed memory multicomputers. Among various multicomputers, the cube based topologies [3] have proved themselves to be very powerful with the largest popularity due to many of their attractive properties including regularity, symmetry, small diameter, strong connectivity, recursive construction and partitionability. The important members of cube family include hypercube [3], crossed cube [4], varietal hypercube [5] and exchanged hypercube [6].

With the increase of size and complexity of these systems, their reliability becomes extremely important. There are many reliability measures out of which node-pair (two-terminal) reliability is an important performance measure in parallel computer interconnection systems. The node-pair (two-reliability) addresses the probability that a given source-destination pair has at least one fault free path between them.

Evaluation of reliability of parallel interconnection systems has been attempted by many researchers in the past [2], [7]-[17]. However, as the system size increases the reliability prediction using the above approaches becomes extremely difficult and sometimes becomes intractable. Different models which are used for reliability evaluation of networks are graph theoretic techniques [18], fault trees [19] and Markov models [20]. The exact reliability computation of fault tolerant parallel computing networks of larges sizes becomes difficult with existing tools. For example, in using graph theoretic techniques for a fault tolerant parallel computer of size larger than (32×32), the path enumeration and disjointing process takes enormous amount of computer time and memory. Similarly, in using Markov approaches the exponential growth of state-space for network of larger size makes the construction and solution of Markov chain computationally prohibitive. For larger sized networks the node/link complexity being quite high the evaluation of reliability using fault tree process becomes extremely difficult.

Kim et al. [21] have proposed an analytical model for computing the reliability of hypercube multi computers. However, the same method is not applicable for other parallel computer interconnection systems. In the literature, no other general method is available which can be utilized for computing the reliability of large parallel computer networks.

A common approach for solving large networks could be the multiple decomposition. In [22], a decomposition method has been proposed for finding the reliability of
small size networks. However, the method proposed in [22] suffers from certain limitations. The method has not been computerized. Its complexity which is an important measure for an efficient algorithm has not been evaluated. The said method as such can not be applied for parallel computer interconnection networks as the choice of cut is extremely difficult. Therefore, there is a need to go for some new and efficient decomposition methods for evaluating the reliability of parallel computer interconnection networks. Study literature further reports methods [23], [24] for decomposing small width graph. However, they don’t provide any technique reliability computation.

This motivates our study to look for some new methods to evaluate the reliability of parallel computer interconnection systems. The proposed methods yield unusually simpler symbolic reliability expression, which means less computer time for numerical evaluation and small round off errors in computed values. The complexities of the algorithms have been evaluated and it is polynomial in nature.

The two terminal reliability of a number of parallel interconnection systems have been evaluated using the proposed technique. The paper is organized as follows: in Section II, we describe the principle of multiple decomposition and propose two separate algorithms for reliability evaluation of parallel computer interconnection systems. Section III illustrates the proposed algorithms through simple examples. Section IV discusses the results for the purpose of comparison of various parallel computer networks. Finally, the conclusions are given in Section V.

II. PROPOSED METHOD FOR RELIABILITY EVALUATION OF PARALLEL COMPUTER INTERCONNECTION NETWORKS

A. Notations and Assumptions

Notations:

- $TR$: Terminal reliability of the network
- $N$: number of nodes in the network
- $C$: common node of decomposed sub graphs
- $S$: source node
- $t$: destination node
- $P_1$, $P_2$, $P_3$: paths of sub graph I, II, III
- $P$: cardinality of $P$, i.e., $|P|
- N_C(u)$: neighboring nodes of node of $u$
- $A$: adjacency matrix of graph $G$
- $λ$: link failure rate
- $Z$, $Z'$ indicator variable for successful and unsuccessful operation of component $Z$; $Z = 1$ and $Z' = 0$ if $x$ is good, and $Z = 0$ and $Z' = 1$ if $x$ is failed.
- $S$: indicator variable for success of the system in connecting its source node $x_1$ and sink node $x_t$.
- $S_{dis}$: disjoint sum-of-product
- $K$: minimal cut through which the graph is decomposed.
- $X_i$: indicator variable for the union of success events corresponding to all the paths in Sub graph I from the source node $S$ to node $x_i$, where $x_i \in C$.
- $Y_i$: indicator variable for the union of success events corresponding to all the paths in Sub graph I from node $x_i$ to sink node $t$, where $x_i \in C$.

We Assume that

1. The nodes are perfect whereas the links are imperfect.
2. The link failure is statistically independent of each other.
3. Initially all components are good and they are not repairable

B. Brief Description of the Proposed Method

The parallel computer interconnection network is first converted in to an equivalent probabilistic graph. Then the graph is decomposed in to three sub graphs through two minimal cuts. The system success is expressed in terms of certain successes of these sub graphs, and then changed into an equivalent disjoint expression, which is directly converted on a one-to-one basis into a reliability expression.

The system success is expressed as

$$S = S_1 \cup S_2$$

where

$$S_1 = \bigcup_{\alpha \in C} X \cup Y$$
and
$$S_2 = \bigcap_{\alpha \in C} Y Z_j$$

on disjointing $S$

$$S_{dis} = (S_1)_{dis} \cup (S_2)_{dis}$$

By replacing all indicator variables by their probabilities and logical sum and product operator by their arithmetic counterparts, we get the two terminal reliability of the system, i.e.,

$$TR = (S_{dis})_{\sum_{\{w \in \pi, C_{\alpha} \in C\} \rightarrow \pi_{\{i, i'\}}}}$$

C. Proposed Functions

The proposed algorithms I and II consists of the following five functions

i. Multicomputer_Decomposition

This function inputs the graph $G$, source node $(s)$ and destination node $(t)$ for which, the two terminal reliability is to be evaluated. The function partitions the graph into sub graphs $G_s$ and $G_t$ by taking two minimal cuts and finds the common node $(C)$.

Multicomputer_Decomposition $(G,s,t) /*Function for decomposing a multi computer interconnection system*/$

for $i = 1$ to $N$

for $j = 1$ to $N$

$K[i] = K[i] + A[i][j]$

Sort $K$ in ascending order


If $C[1]' = s$ and $C[1]' = t$

find the $N_C(C[1])$ for which number of 1’s is minimum:

$C[2] = N_C(C[1])$

$i = 3, j = 2$

while $G$ is not partitioned in to two sub graphs $G_i$ and $G_j$

$C[i] = K[j]$
If $C[i] = s$ and $C[i]! = t$ and $C[i]! = C[1..j-1]$ find the $N_g(C[i])$ for which number of 1’s is minimum

$C[+1] = N_g(C[i])$

Break

else

$i = i+1$

$j = j+1$

return ($C$)

ii. Compare

This function compares $G_1$ and $G_2$ on the basis of number of nodes and finds out the larger sub graph with larger number of nodes.

$Compare(G_1, G_2) /* Function for comparing two graphs*/$

if $(N(G_1) > N(G_2))$

return ($G_1$)

else

return ($G_2$)

iii. MIN Decomposition

This function inputs the graph $G$, dimension ($n$), source node ($s$) and destination node ($t$) for which, the two terminal correlation is to be evaluated. Since the multistage interconnection network is composed of a number of switching stages, the proposed function takes the above said composition of the multistage interconnection network to its advantage and takes minimal cuts for stage wise decomposition into three sub graphs.

$MIN\ Decomposition (G, n, s, t) /* Function to decompose a MIN*/$

for $i = 1$ to $|N|$ /*cardinality of $N$*/

if $(n+1 \leq i \leq 2n)$ and $I! = s$ and $I! = t$

$C[i] = N$

partition $G$ into two graphs $G_1$ and $G_2$

return ($C$)

iv. Terminal Reliability Evaluation

This function accepts common nodes $C_1$ and $C_2$ of sub graphs I, II and III. $C_1$ is common to sub graphs I and II and $C_2$ is common to sub graphs II and III. The function computes the two terminal reliability correlation.

$Terminal\ Reliability\ Evaluation\ (C_1, C_2) /* Function for reliability evaluation through double decomposition*/$

Enumerate all the paths ($P_{i,j}$) of sub graph $G_i$ from source node $s$ to $n_i$, $n_i \in C_1$, $i \in C_1$, $j \geq 1$

for $i = 1$ to $|C_1|

$p_{i,j} =$ cardinality of $P_{i,j}$

$X = \phi$

for $i = 1$ to $|C_1|

for $j = 1$ to $p_{i,j}$

$X = X \cup P_{i,j}$

Enumerate all the paths ($P_{2,i,j}$) of sub graph $G_2$ from source node $n_i$ to $n_j$, $n_i \in C_1$ & $n_j \in C_2$

For $i = 1$ to $|C_2|

$p_{2,i,j} =$ cardinality of $P_{2,i,j}$

$Y = \phi$

for $i = 1$ to $|C_2|

for $j = 1$ to $p_{2,i,j}$

$Y = Y \cup P_{2,i,j}$

Enumerate all the paths ($P_{3,i,j}$) of sub graph $G_3$ from source node $n_j$ to $n_i$, $n_j \in N_{C_2}$, $i \in C_1$, $j \geq 1$

$p_3 =$ cardinality of $P_{3,i,j}$

$Z = \phi$

for $i = 1$ to $|C_1|

for $j = 1$ to $p_3$

$Z = Z_i \cup P_{3,i,j}$

$v.\ Dis(X)$

This function accepts a Boolean expression and finds its disjoint form

$Dis(X) /* X is converted into string */$

if $X! = \omega$

$P_i \cap P_j$ /* Concatenation of $P_i$ and $X_i$ */

else

for $i = i+1$

for $j = j+1$

if $i! = j$

$X = X \cup P_i \cap \overline{P}_i$

for $j = 1$ to $i$

for $j = 1$ to $j$

if $i! = j$

$X = X \cup P_i \cap \overline{P}_i$

return (int ($X$))

D. Proposed Algorithms

Here, we propose two separate algorithms for terminal reliability evaluation of parallel computer interconnection networks. The first one is meant for multi computer while the second one accounts for multi stage interconnection networks.

Algorithm-I (Multi computer reliability evaluation)

$Multicomputer\ decomposition(G, s, t)$

$G = \text{Multicomputer}\ Decomposition\ (G, s, t); /* G is partitioned in to $G_1$ and $G_2$ */$

$g = \text{Compare}(G_1, G_2)$

$C_1 = \text{Multicomputer}\ Decomposition\ (g, s, t);$

$Terminal\ Reliability\ Evaluation\ (C_1, C_2)$

Algorithm-II (Multistage interconnection network reliability evaluation)

$MIN\ Double\ decomposition$

$C_1 = MIN\ Decomposition (G, n, s, t) /* G is partitioned in to $G_1$ and $G_2$ */$

$C_2 = \text{MIN}\ Decomposition\ (G, n, s, t) /* or $G_2$ */$

$Terminal\ Reliability\ Evaluation\ (C_1, C_2)$

Complexity: Finding the minimal paths using the proposed algorithms require $O(N^2)$ operations and dominates all other operation such as disjoining. So, the running time of proposed algorithms are $O(N^2)$ which is polynomial.
III. ILLUSTRATIONS

This section illustrates Algorithms I and II through examples 1 and 2, respectively.

The Algorithm-I is illustrated through the following example.

**Example 1**

Let us consider the network given in Fig. 1(a). In order to evaluate the terminal reliability (TR) of the network of Fig. 1(a), the network is decomposed into the three sub graphs Fig. 1(b), 1(c) and 1(d) through the cuts \( K_1 = \{a, b\} \) and \( K_2 = \{c, d\} \). The common node sets are \( C_1 = \{2, 4\} \) and \( C_2 = \{3, 4\} \). Hence the system success is

\[
S_1 = X_1Y_1 \cup X_2Y_2 \quad S_2 = \overline{Y}_Z \cap \overline{Y}_Z \quad S = S_1 \cup S_2 \tag{4}
\]

\[
X_2 = A, \quad Y_1 = D, \quad Z_3 = G \cup EF
\]

\[
X_4 = B, \quad Y_4 = C, \quad Z_4 = F \cup EG
\]

\[
S_1 = (X_1)_{\text{div}} \cap (Y_1)_{\text{div}} \cap (X_2)_{\text{div}} \cap (Y_2)_{\text{div}} \cap (X_4)_{\text{div}} \cap (Y_4)_{\text{div}} = BD \cup BD \cup CD \tag{5}
\]

\[
S_2 = (\overline{X}_Z)_{\text{div}} \cap (\overline{Y}_Z)_{\text{div}} = \overline{C}(F \cup EG)
\]

\[
S = BD \cup BD \cup CD \cup \overline{C}(G \cup EF) \cap C(F \cup EG) \tag{6}
\]

So, the two terminal reliability of the network between nodes \( s \) and \( t \) is given by

\[
TR = p_s p_t + q_s p_t p_d + p_s q_t p_d + q_s q_t p_d \tag{7}
\]

Equation (9) requires 13 multiplications and 5 sums for reliability evaluation as compared with 19 multiplications and 7 sums required by the equivalent expression in [22].

The proposed algorithm-II is illustrated through the following example.

**Example 2**

The network (Fig. 2(a)) is decomposed into the three sub graphs of Figs. 2(b)-2(d) through the cuts \( K_1 = \{x_1, x_2\} \) and \( K_2 = \{x_1, x_4\} \). The common node sets \( C_1 = \{2, 3\} \) and \( C_2 = \{3, 5\} \). Hence the system success is

\[
S_1 = (X_1)_{\text{div}} \cap (Y_1)_{\text{div}} \cap (X_2)_{\text{div}} \cap (Y_2)_{\text{div}} \cap (X_3)_{\text{div}} \cap (Y_3)_{\text{div}} \tag{8}
\]

\[
S_2 = \overline{Z}_z \cap \overline{T}_z \tag{9}
\]

\[
TR = p_s (p_t + q_t) + q_s p_t (p_t + q_t) + p_s q_t (p_t + q_t) + p_s q_t q_t p_d \tag{10}
\]

Expression (11) requires 21 multiplications and 8 sums for numerical calculation as compared with 33 multiplications and 13 sums required by the equivalent expression in [22]. The results obtained for the examples using Algorithm-I and Algorithm-II are validated with the results obtained using method [22].
This section presents the results obtained for various candidate parallel computer networks using the proposed algorithms. The computed results have been plotted for comparison and discussions.

Reliability is evaluated using the proposed algorithm-II for a some important multistage interconnection networks. The reliability is plotted against the mission time \( t \) in hours for different link failure rate \( \lambda \) (Figs. 3-5).

From Fig. 3(a), it is clear that, under link failure rate \( \lambda = 0.005 \), the TR of Extra stage cube \( (N = 8) \) becomes zero at 800 hrs. At mission time 500 hrs and link failure rate \( \lambda = 0.005, 0.003, 0.002 \) and 0.001, the reliability are found to be 0.02, 0.15, 0.3, 0.7 respectively. At mission time 1000 hrs and link failure rate \( \lambda = 0.005, 0.003, 0.002 \) and 0.001, the reliabilities are 0.0001, 0.07 and 0.47. So, it can be concluded that the Extra stage cube (ESC) does not provide a reliable parallel system under large link failure rate and for a longer period of time. Fig. 3(b) shows the TR versus time \( t \) for the Multi path omega network \((N = 16)\). For mission time 500 hrs and link failure rate \( \lambda = 0.005, 0.003, 0.002 \) and 0.001, the corresponding reliabilities are 0.06, 0.14, 2.1, and 3.8 respectively. The reliability is 0.01 at mission time 1000 hrs. So, multi path omega network may be treated as a preferable network structure than extra stage cube network as it can sustain for a longer period of time. From Fig. 3(c), it can be observed that the reliability of extra stage shuffle exchange network \((N = 8)\) becomes 0.0 at mission time beyond 500 hrs for link failure rate \( \lambda = 0.005 \), at mission time beyond 700 hrs for link failure rate \( \lambda = 0.003 \), and at mission time beyond 1000 hrs for link failure rate \( \lambda = 0.002 \). At mission time 500 hrs and for link failure rates \( \lambda = 0.005, 0.003, 0.002 \) and 0.001, the corresponding reliabilities are 0.0, 0.015, 0.1, and 0.49.

The reliability (TR) is plotted against the mission time \( t \) in hours for different link failure rate \( \lambda \) (Figs. 4(a)-4(d)). From Fig. 4(a), it can be observed that, under link failure rate \( \lambda = 0.0005 \), the two terminal reliability of Hypercube (HC) \((n = 3)\) is 0.0005 at mission time 10000 hrs, so it can not sustain for a longer period of time with a higher link failure rate. But with link failure rate \( \lambda = 0.0003, 0.0002 \) and 0.0001, the corresponding TR is 0.15, 0.35 and 0.68 at mission time 5000 hrs. and 0.03, 0.09 and 0.31 at mission time \( t = 10000 \) hrs respectively. Fig. 4(b) shows the two terminal reliability of crossed cube under different link failure rate against mission time \( t \). Under link failure rate \( \lambda = 0.0005 \), the two terminal reliability of Crossed cube (CQ) \((n = 3)\) is 0.01 at mission time 10000 hrs, which is double of that of hypercube at same mission time and same link failure rate. The TR of crossed cube under different link failure rate \( \lambda = 0.0003, 0.0002 \) and 0.0001 are 0.08, 0.39, 0.68 at mission time 5000 hrs and 0.01, 0.05, 0.12 and 0.39 at mission time 1000 hrs. From Fig. 4(c) it can be observed that under link failure rate \( \lambda = 0.0005 \), the two terminal reliability of Varietal hypercube (VH) \((n = 3)\) is 0.0 at mission time 7000 hrs, so it is not a preferable topology for a system having higher link failure rate that requires processing beyond 7000 hrs. The corresponding TR of varietal hypercube with different link failure rate \( \lambda = 0.0003, 0.0002 \) and 0.0001 are 0.05, 0.25, 0.62 at mission time 5000 and 0.0, 0.0, 0.02 and 0.22 at mission time 1000. From Fig. 4(d), it is clear that under link failure...
rate $\lambda = 0.0005$, the reliability of Exchanged hypercube
(EHC) becomes zero at mission time beyond 6000 hrs.
Under different link failure rate $\lambda = 0.0003, 0.0002$ and
0.0001, the reliability are 0.00001, 0.1 and 0.42
respectively at mission time 5000 hrs. At 10,000 hrs these
values are 0, 0.0 and 0.9. So, exchanged hypercube may
not provide a good reliable system under high link failure
rate and for longer period.

The reliabilities of Hypercube (HC), Crossed cube (CQ),
Variateal Hypercube (VH), Exchanged Hypercube (EHC)
and Cube connected cube (CCC) are plotted against
mission time t under the same link failure rate $\lambda = 0.0002$
(Fig. 5). From Fig. 5, it can be observed that the hypercube
provides better reliability than other multi computers within
mission time 3000 hrs. but as time increases, the crossed cube provides a better reliability than other multi
computer networks.

For the purpose of comparison, the three dimensional
hypercube is decomposed by methods [23] and [24]
respectively and then, the terminal reliability is computed
by the proposed function $\text{Terminal Reliability evaluation}$.
The Figure 6 compares the results obtained through the
proposed method with those of methods [23] and [24].
From Fig. 6, it can be observed that the proposed
decomposition method provides much more accurate
values of terminal reliability of the said network over its
counterparts.

Further, to show the generality of the proposed method,
the proposed method is applied to evaluate the terminal reliability of other multicomputer networks such as
Alternating Group Graph [25] (Fig. 7).

V. CONCLUSIONS

Two new algorithms for reliability evaluation of parallel
computer interconnection networks based on graph
decomposition are presented and illustrated through simple
examples. The simulated results conclude the proposed
decomposition method for evaluating the terminal reliability of parallel computer interconnection systems to
be quite simple, efficient and general. The work carried out
in this paper can be further extended to find other
reliability measures like network reliability and broadcast
reliability of parallel computer interconnection systems.

REFERENCES

collection of fault tolerant multi stage interconnection networks,"
reliability," IEEE Trans. on Computers, vol. 38, no. 11, pp. 1600-
Trans on Parallel and Distributed System, vol. 3, no. 5, pp. 513-524,
interconnection topology for large scale computer," in Proc. of Int.
Conf. on Parallel and Distributed System, pp. 703-708, 1994.
reliability and improved performance," IEEE Trans on Computer,
evaluation of super computer systems," Microelectronics and
network using hierarchical composition," IEEE Trans on Reliability,
multi stage interconnection network," IEEE Trans on Reliability,
reliability in hypercubes," IEEE Trans on Computers, vol. 44, no.7,
interconnection network," IEEE Trans on Parallel and Distributed
analysis of hypercube multi computers," Micro Electronics and
reliability of hypercube architecture," IEEE Trans. on Parallel and
evaluation of multi stage interconnection networks with multi state
elements," Micro Electronics and Reliability, vol. 36, no. 3, pp. 423-
428, Mar. 1996.
of parallel computer interconnection networks," Journal of IE,
[20] K. S. Trivedi, Probability and Statistics with Reliability Queuing:
of hypercube multi computers." IEEE Trans on Reliability, vol. 38,
[22] A. M. Rushdi and K. Abdul Aziz, "On reliability evaluation by
[23] H. L. Bodlaender, "A linear time algorithm for finding tree
decompositions of small tree width," SIAM Journal of Computing,

**R. K. Dash** received the MCA degree from F.M. University in 2001. He got his Ph.D. in Computer Science from Sambalpur University in the year 2008. He is now working as Reader and Head, CSA Department College of Engineering & Technology, Bhubaneswar, Orissa. His research areas are reliability of parallel & distributed system, analysis of & design of algorithm.

**C. R. Tripathy** received the B.Sc. (Eng.) in Electrical Engineering from Sambalpur University and M.Tech. degree in Instrumentation Engineering from Indian Institute of Technology, Kharagpur respectively. He got his Ph.D. in the field of Computer Science & Engineering from Indian Institute of Technology, Kharagpur. He is currently working as a Professor and Head, Department of Computer Sc. & Eng., VSSUT, Burla, Orissa, India. He has more than 50 publications in different national and International Journals and Conferences. His Research interest includes dependability, reliability and fault-tolerance of parallel and distributed system.