

## Distance Distribution of Nodes in Star Graphs \*

L. Wang, S. Subramanian and S. Latifi

Department of Electrical Engineering  
University of Nevada  
Las Vegas, NV 89154-4026

P. K. Srimani

Department of Computer Science  
Clemson University  
Clemson, SC 29634-0974

August 29, 2005

### Abstract

The purpose of the paper is to provide an answer to a long standing problem to compute the distance distribution among the nodes in a star graph, i.e., to compute the exact number of nodes at a distance  $k$  from the identity node in a star graph where  $k$  varies from 0 to the diameter of the graph. A star graph is a Cayley graph like the hypercubes; for a hypercube  $Q_n$ , there are exactly  $\binom{n}{r}$  nodes at a distance  $r$  from the identity node where  $r$  varies from 0 to  $n$ .

**Keywords:** Cayley Graph, diameter, distance distribution, Star graph.

## 1 Introduction

A suitable interconnection network is an integral part of any distributed computing system. The network is usually modeled by a symmetric (undirected) graph where the nodes (vertices) denote the processing elements and the edges (arcs) denote the bidirectional communication channels. Interconnection topologies are evaluated in terms of low degree, small diameter, high fault tolerance, low fault diameter etc. One of the most efficient interconnection network has been the well known binary  $n$ -cubes or hypercubes; they have been used to design various commercial multiprocessor machines and they have been extensively studied [SS88, Lei92, Ho91]. The hypercubes have been shown to be a member of a larger family of graphs, called Cayley graphs [LJD93]. Another member family of graphs, called the star graphs [AK89, AK87] have also been investigated in the literature; star graphs enjoy most of the desirable properties [RS93, DT94, RLS96, LS96] of the hypercubes at considerably less cost; they accommodate more nodes with less interconnection hardware and less communication delay. It has also been shown [QAM94, FA91, MS92] that many parallel algorithms can be efficiently mapped on star graphs. While it has been shown that the star graphs outperforms the hypercubes in terms of the diameter and average distance between nodes, the exact distance distribution of nodes in a star graph is not known, although the corresponding problem for hypercubes has been solved [SS88]. Our purpose in the present paper is to compute the exact number of nodes at a distance  $k$  from the identity node in a star graph where  $k$  varies from 0 to the diameter of the graph (a star graph is a vertex symmetric graph like the hypercubes [DT94]); there are exactly  $\binom{n}{r}$  nodes at a distance  $r$  from the identity node in a hypercube  $Q_n$  where  $r$  varies from 0 to  $n$ . Thus, we seem to provide an answer to a long standing problem to compute the distance distribution among the star graph nodes.

---

\***Address for Correspondence:** Pradip K Srimani, Department of Computer Science, Clemson University, Clemson, SC 29634-0974, Email: srimani@cs.clemson.edu, Voice: (864) 656-7552, Fax: (864) 656-0145

## 2 Basic Concepts

The  $n$ -dimensional star graph,  $S_n$ , is an edge and node symmetric graph containing  $n!$  nodes and  $(n-1)n!/2$  edges (links). Each node is labeled by a distinct permutation of the set of integers  $\{1, \dots, n\}$ . Two nodes are joined with a link labeled  $i$  if and only if the label of one can be obtained from the label of the other by swapping the first digit (conventionally the leftmost) and the  $i$ th digit where  $1 < i \leq n$ . For instance, in  $S_5$ , two nodes 12345 and 42315 are neighbors and joined via a link labeled 4; Figure 1 shows a 4-star,  $S_4$ . The identity permutation is denoted by  $I = (1234\dots)$ .

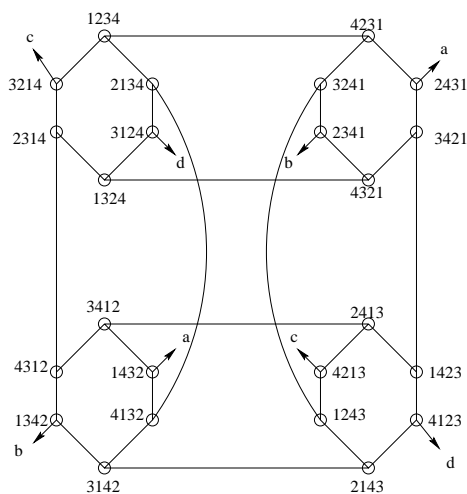


Figure 1: A 4-Star

Star graphs have been extensively studied for their applications and structural properties. We summarize some of the properties that are relevant to our discussion.

- The star graphs are members of Cayley group graphs. A star graph  $S_n$  has  $n-1$  generators,  $g_2, g_3, \dots, g_n$ , where  $g_i$  swaps the first symbol with the  $i$ -th symbol of any permutation. Each generator is its own inverse, i.e., the star graph is symmetric. Also, the star graph  $S_n$  is a  $(n-1)$ -regular graph with  $n!$  nodes and  $n!(n-1)/2$  edges.
- Since star graphs are vertex symmetric [AK89], we can always view the distance between any two arbitrary nodes as the distance between the source node and the identity permutation by suitably renaming the symbols representing the permutations. Hence, in our subsequent discussion about a path from a source node to a destination node, the destination node is always assumed to be the identity node  $I$  without any loss of generality. Also, the distance of a node (permutation) is defined to be its distance to the identity node.
- It is easy to see that any permutation of  $n$  elements can also be specified in terms of its unique cycle structure with respect to the identity permutation  $I$ . For example,  $345216 = (135, 24, 6)$ . The maximum number of cycles in a permutation of  $n$  elements is  $n$  and the minimum number is 1. When a cycle has only one symbol, that symbol is in its correct position in the permutation with respect to the identity permutation. For the rest of the paper we omit the singletons if the number of symbols in the permutation is understood from the context. The length of a cycle is defined to be the number of symbols present in the cycle.
- The diameter of  $S_n$  is given by  $\Delta_n = \lfloor 3(n-1)/2 \rfloor$ .

- If  $D(u)$  denotes the distance of the node  $u$  (to the identity permutation) and  $\Pi(u)$  denotes the cycle representation of the node (permutation)  $u$ ,  $\Pi = \{\pi_1, \pi_2, \dots, \pi_k\}$  ( $\pi_i$ 's are the individual cycles in the node), then it has been shown in [AK89] that the distance  $D(u)$  for a given node (permutation)  $u$  is given by:

$$D(u) = \begin{cases} c + m & \text{when "1" is the first symbol} \\ c + m + 2 & \text{otherwise} \end{cases}$$

where  $c$  is the number of cycles of length at least 2 in any permutation  $u$  and  $m$  is the total number of symbols in those  $c$  cycles of the permutation  $u$ .

Given an arbitrary node  $u = \{\pi_1, \pi_2, \dots, \pi_k\}$ , the algorithm to find the shortest path to the destination node (the identity permutation) is described in [AK87]; it has also been shown that in general this shortest path is not unique, i.e., there may be multiple shortest paths to the destination from a given node  $u$ . In order to generate a path from  $u$  to  $I$ , generators are applied in a particular sequence. In order that the path is minimal in length, the moves (generating the next vertex in the sequence) must be restricted to two possible choices : for any node  $u = \{\pi_1, \pi_2, \dots, \pi_k\}$ , **either (1)** exchange  $\pi_1^1$  with  $\pi_1^2$  thereby reducing the length of the cycle  $\pi_1$  by 1, **or (2)** exchange  $\pi_1^1$  with  $\pi_j^\ell$  (where  $2 \leq j \leq k$  and  $1 \leq \ell \leq |\pi_j|$ ) thereby merging the two cycles  $\pi_1$  and  $\pi_j$  into one cycle  $(\pi_j^\ell, \pi_j^{\ell+1}, \dots, \pi_j^{|\pi_j|}, \pi_j^1, \dots, \pi_j^{\ell-1}, \pi_1^1, \dots, 1)$ . When a path from a node  $u$  to the destination is said to follow the *shortest routing scheme* if the moves are restricted to these two types.

**Definition 1** The number of nodes that are at a distance  $k$  from the identity node  $I$  is denoted by  $\mathcal{N}_k^n$ , where  $0 \leq k \leq \lfloor 3(n-1)/2 \rfloor$ .

**Definition 2** Consider the cycle structure of an arbitrary node  $u$  in  $S_n$ . A cycle that contains the symbol "1" is called a  $\mathcal{CO}$  and a cycle that does not contain the symbol "1" is called a  $\mathcal{CN}$ . A  $\mathcal{CO}$  (or  $\mathcal{CN}$ ) of length  $r$  is denoted by  $r\text{-}\mathcal{CO}$  (or  $r\text{-}\mathcal{CN}$ )

**Example:** Consider the node  $u = 3241$  in  $S_4$  with the cycle structure  $u = (314, 2)$ ; the cycle 314 is a 3- $\mathcal{CO}$  (the cycle contains the symbol "1" and its length is 3). Consider the node  $u = 1324$  in  $S_4$  with the cycle structure  $u = (1, 23, 4)$ ; the cycle 23 is a 2- $\mathcal{CN}$ .  $\square$

**Definition 3** The number of nodes in  $S_n$  that are at a distance  $k$  from the identity node  $I$  and that have exactly  $j$  many  $\mathcal{CN}$ s (of any length) is denoted by  $\Psi_k^j$  ( $j$  denotes the number of  $\mathcal{CN}$ s in a node at a given distance  $k$ ). Thus, we have

$$\mathcal{N}_k^n = \sum_{j=0}^n \Psi_k^j$$

Consider an arbitrary node  $u$  in  $S_n$  at a distance  $k$ ,  $0 \leq k \leq \lfloor 3(n-1)/2 \rfloor$ . In order to compute  $\mathcal{N}_k^n$ , we consider two possible cases.

#### Case 1: Node $u$ does not contain any $\mathcal{CN}$

**Lemma 1** Node  $u$ , with no  $\mathcal{CN}$  and at a distance  $k$ , has only one  $r\text{-}\mathcal{CO}$ , where  $r = k + 1$ ; further, such a node  $u$  exists only when  $1 \leq k \leq n - 1$ .

**Proof :** The first part follows from the fact that a node can have at most one, i.e., an  $r\text{-}\mathcal{CO}$  (Definition 2) and that for a  $r\text{-}\mathcal{CO}$  it takes exactly  $r - 1$  moves to resolve the cycle completely (the shortest routing algorithm). The second part follows from the fact that the maximum length of a cycle is  $n$ .  $\square$

**Lemma 2** *The number of nodes  $u$  in  $S_n$  at a distance  $k$ , which does not contain any  $\mathcal{CN}$ , is given by*

$$\Psi_k^0 = \begin{cases} \frac{(n-1)!}{(n-k-1)!} & 1 \leq k \leq n-1 \\ 0 & \text{otherwise} \end{cases}$$

**Proof :** Node  $u$  has only one  $(k+1)$ - $\mathcal{CO}$ ; in addition to the symbol “1” in the cycle, the rest  $k$  symbols are to be arranged taken from a total of  $n-1$  symbols to form the distinct cycles. This, coupled with Lemma 1, leads to the claim.  $\square$

## Case 2: Node $u$ does contain at least one $\mathcal{CN}$

In this case, the node  $u$  may or may not have other  $\mathcal{CN}(s)$  or a  $\mathcal{CO}$ .

**Lemma 3** *When  $k \leq 2$ , there is no node  $u$  with at least one  $\mathcal{CN}$  and at a distance  $k$  (i.e.,  $\Psi_k^j = 0$ ,  $k = 1, 2$ ,  $j \geq 1$  for all  $n$ ). When  $k$ ,  $3 \leq k \leq \lfloor 3(n-1)/2 \rfloor$ , the length  $r$  of the  $r$ - $\mathcal{CN}$  has the property  $2 \leq r \leq \min\{k-1, n-1\}$ .*

**Proof :** The first part follows from the fact that the minimum length of a  $r$ - $\mathcal{CN}$  is 2 and the minimum distance of any node with a 2- $\mathcal{CN}$  is 3 (see the shortest routing algorithm). The second part follows from the fact that the maximum length of a  $\mathcal{CN}$  is  $n-1$  (since there are only  $n-1$  distinct symbols other than “1”) and the minimum distance of any node with a  $r$ - $\mathcal{CN}$  is  $r+1$ .  $\square$

**Lemma 4** *Consider a node  $u$  with at least one  $r$ - $\mathcal{CN}$ ,  $2 \leq r \leq n-1$ . The number of such distinct  $r$ - $\mathcal{CN}$  is given by  $(r-1)! \binom{n-1}{r}$ .*

**Proof :** A  $r$ - $\mathcal{CN}$  cannot contain the symbol “1” and hence we can choose  $r$  symbols from  $n-1$  symbols in  $\binom{n-1}{r}$  and each such choice of  $r$  symbols can be arranged in  $(r-1)$  ways to form distinct cycles.  $\square$

Consider a node  $u$  (at a distance  $k$ ,  $3 \leq k \leq \lfloor 3(n-1)/2 \rfloor$ ) with one  $r$ - $\mathcal{CN}$ ,  $2 \leq r \leq \min\{k-1, n-1\}$  (the node may or may not have other cycles in its permutation). Our approach is to enumerate how many distinct such  $r$ - $\mathcal{CN}$  we can form and thereby compute the number of distinct nodes in  $S_n$  at a distance  $k$ . We apply the shortest routing algorithm on this node first by combining the specified  $r$ - $\mathcal{CN}$  cycle with the cycle containing the symbol “1” (this cycle containing “1” may be a singleton cycle) and then making moves so as to put the symbols in this  $r$ - $\mathcal{CN}$  into their respective correct positions. This will take exactly  $r+1$  moves; i.e., a node has been reached with two properties: (1) node  $v$  is at a distance  $k-r-1$  (from  $I$ ); and (2) the symbols in the  $r$ - $\mathcal{CN}$  of  $u$  have been put in their respective correct position in the node  $v$ . Thus, node  $v$  can be viewed as a node at a distance  $k-r-1$  in a star graph  $S_{n-r}$ . Consider the node  $u = 54231 = (15, 234)$  in  $S_5$  with one 3- $\mathcal{CN}$  and one 2- $\mathcal{CO}$ ; the node is at a distance  $k = 5$  from the identity node. Here,  $r = 3$  and hence after exactly 4 shortest routing moves from node  $u$ , we arrive at the node  $v = 52341 = (15, 2, 3, 4)$  which is at a distance 1 from the identity node in  $S_2$  (the subgraph induced in  $S_5$  by the nodes 1 and 5).

We make the following observations:

- The number of all possible distinct such  $r$ - $\mathcal{CN}$  is given by  $(r-1)! \binom{n-1}{r}$ , for each  $r$ .
- Consider the node  $v$  (as described above) for a given node  $u$ . Note that node  $v$  must be the identity node  $I$  if  $r = k-1$  (since the node  $u$  is at a distance  $k$  from  $I$ ).
- For  $r < k-1$ , there are exactly  $\mathcal{N}_{k-r-1}^m$  such nodes  $v$  in a star graph  $S_m$  where  $m = n-r$ . We note that  $\mathcal{N}_{k-r-1}^m = \sum_{j \geq 0} \Psi_j^{k-r-1}$ . If any node  $v$  does have a  $\mathcal{CN}$ , we need be careful about not counting its contribution more than once.

| $n$ | $n!$ | $k$ | $\Psi_k^0$ | $\Psi_k^1$ | $\Psi_k^2$ | $\mathcal{N}_k^n$ |
|-----|------|-----|------------|------------|------------|-------------------|
| 4   | 24   | 0   | 1          | 0          | 0          | 1                 |
| 4   | 24   | 1   | 3          | 0          | 0          | 3                 |
| 4   | 24   | 2   | 6          | 0          | 0          | 6                 |
| 4   | 24   | 3   | 6          | 3          | 0          | 9                 |
| 4   | 24   | 4   | 0          | 5          | 0          | 5                 |
| 5   | 120  | 0   | 1          | 0          | 0          | 1                 |
| 5   | 120  | 1   | 4          | 0          | 0          | 4                 |
| 5   | 120  | 2   | 12         | 0          | 0          | 12                |
| 5   | 120  | 3   | 24         | 6          | 0          | 30                |
| 5   | 120  | 4   | 24         | 20         | 0          | 44                |
| 5   | 120  | 5   | 0          | 26         | 0          | 26                |
| 5   | 120  | 6   | 0          | 0          | 3          | 3                 |

Table 1: Distance Spectrum of Nodes in  $S_n$ ,  $n = 4, 5$

**Lemma 5** For  $j \geq 1$ , we have

$$\Psi_k^j = \sum_{r=2}^{\min\{k-1, n-1\}} (r-1)! \binom{n-1}{r} \times \frac{1}{j} \Psi_{k-r-1}^{j-1} \quad j \geq 1$$

**Proof :** The proof follows from the earlier discussion after noting that the factor of  $j$  in the denominator eliminates the double counting of nodes.  $\square$

Combining Lemma 2 and Lemma 5, we can state the following theorem to compute the number of nodes in  $S_n$  at all possible distances.

**Theorem 1** The total number of nodes at a distance  $k$ ,  $0 \leq k \leq \lceil 3(n-1)/2 \rceil$  from the identity node in a star graph  $S_n$  is given by

$$\mathcal{N}_k^n = \sum_{j=0}^n \Psi_k^j$$

As an example of Theorem 1, Table 1 lists the results in the cases  $n = 4$  and 5.

### 3 Acknowledgement

The last author, Srimani's work was partially supported by a NSF grant # ANI-0218495.

### References

- [AK87] S. B. Akers and B. Krishnamurthy. The star graph: an attractive alternative to n-cube. In *Proceedings of International Conference on Parallel Processing (ICPP-87)*, pages 393–400, St. Charles, Illinois, August 1987.

- [AK89] S. B. Akers and B. Krishnamurthy. A group-theoretic model for symmetric interconnection networks. *IEEE Transactions on Computers*, 38(4):555–566, April 1989.
- [DT94] K. Day and A. Tripathi. A comparative study of topological properties of hypercubes and star graphs. *IEEE Transactions on Parallel and Distributed Systems*, 5(1):31–38, January 1994.
- [FA91] P. Fragopoulou and S. G. Akl. Parallel algorithm for computing Fourier transforms on the star graph. In *Proceedings of the International Conference on Parallel Processing*, volume III, pages 100–106, St. Charles, Illinois, 1991.
- [Ho91] C. T. Ho. Optimal broadcasting on simd hypercubes without indirect addressing capability. *Journal of Parallel and Distributed Computing*, 13(2):246–245, 1991.
- [Lei92] F. T. Leighton. *Introductions to Parallel Algorithms and Architectures: Arrays, Trees and Hypercubes*. Morgan Kaufman, 1992.
- [LJD93] S. Lakshmirarahan, J. S. Jwo, and S. K. Dhall. Symmetry in interconnection networks based on Cayley graphs of permutation groups: a survey. *Parallel Computing*, 19:361–407, 1993.
- [LS96] S. Latifi and P. K. Srimani. Transposition networks as a class of fault-tolerant robust networks. *IEEE Transactions on Computers*, 45(2):230–238, February 1996.
- [MS92] V. E. Mendia and D. Sarkar. Optimal broadcasting on the star graph. *IEEE Transactions on Parallel and Distributed Systems*, 3(4):389–396, July 1992.
- [QAM94] K. Qiu, S. G. Akl, and H. Meijer. On some properties and algorithms for the star and pancake interconnection networks. *Journal of Parallel and Distributed Computing*, 22(1), July 1994.
- [RLS96] Y. Rouskov, S. Latifi, and P. K. Srimani. Conditional fault diameter of the star graph networks. *Journal of Parallel and Distributed Computing*, 33:91–97, 1996.
- [RS93] Y. Rouskov and P. K. Srimani. Fault diameter of star graphs. *Information Processing Letters*, 48(5), December 1993.
- [SS88] Y. Saad and M. H. Shultz. Topological properties of hypercubes. *IEEE Transactions on Computers*, 37(7):867–872, July 1988.