Application of Perfect Difference Sets to the Design of Efficient and Robust Interconnection Networks

Behrooz Parhami
Dept. of Electrical & Computer Engineering
University of California
Santa Barbara, CA 93106-9560, USA
parhami@ece.ucsb.edu

Mikhail Rakov
Dept. of Computer Science
University of California
Santa Barbara, CA 93106-5110, USA
mikhrak@aol.com

Abstract

In this paper, we focus on deriving low-diameter networks, beginning with $D = 2$*, the next best value to that of the complete network, and proceeding to somewhat larger (constant) values leading to more economical networks. We show that perfect difference networks (PDNs), which are based on the mathematical notion of perfect difference sets, offer a diameter of* 2 *in an asymptotically optimal manner. In other words, PDNs allow* $O(d^2)$ *nodes when nodes are of degree* d*, or, equivalently, have a node degree that grows as the square-root of the network size. The symmetry and rich connectivity of PDNs lead to balanced communication traffic and good fault tolerance. Multidimensional PDNs offer a tradeoff between cost and performance in the sense that for any constant number* q *of dimensions, a* q*-dimensional PDN has diameter* $D = 2q$ *and node degree that grows as the* $(2q)$*th root of* n*.*

1. Introduction

Low latency, high bandwidth, energy efficiency, and robustness are some of the properties that are sought in networks for parallel and distributed computing. Given that network performance parameters depend not only on the network architecture but also on a number of factors relating to applications and their data exchange characteristics, the challenge in interconnection network design is finding the right match between communication needs of applications on one side and capabilities and limitations inherent in each architecture on the other. This, in turn, explains the proliferation of implemented and proposed connectivities, sometimes characterized as the sea of interconnection networks [Parh99].

Ideally, each node is directly connected to every other node, thus allowing one-hop communication between any pair of nodes. This connectivity pattern is modeled by the n-node *complete graph* K_n. Physically, however, complete-graph connectivity is difficult to provide for large systems that are of practical interest, due to both the high cost of nodes with many communication channels and lack of scalability for system growth. At the other extreme from K_n, the simplest possible physical connectivity pattern is that of n-node *ring* R_n. Here, each node has only two communication channels. The number of links is n, as is the aggregate network bandwidth with unit-capacity links. Data exchange is direct only between each node and one or two neighbor(s) and indirect in all other cases. The ring architecture is an example of *bipartite graph*, with even- and odd-numbered nodes constituting the two parts and links connecting nodes that are not in the same part. The complete graph is not bipartite, but $K_{n/2,n/2}$ can be defined (for even n) in which a node in one part is connected to all nodes in the other. This leads to 1- or 2-hop connectivity between nodes.

Intermediate architectures between K_n and R_n can be obtained in a variety of ways, providing tradeoffs in cost and performance. Network cost is affected, among other things, by the (maximum) node degree d, while indicators of network performance include *diameter* D and *bisection bandwidth* B. The degree-diameter product dD is sometimes used as a composite measure of cost-performance or cost-effectiveness. Many of these intermediate architectures can be viewed as *chordal rings* [Arde81], rings to which *bypass links* or *chords* have been added to reduce the network diameter, or richly connected graphs from which certain links are systematically removed via *pruning* [Kwai98], [Parh99a] so as to reduce the node degree and, thereby, network cost. Other mechanisms for deriving cost-effective interconnection networks from other networks include cross-product composition, recursive substitution, and hierarchical composition. Cross product of networks will be discussed in Section 4 along with the notion of swapped networks as an example of hierarchical composition. These combining strategies lead to families of networks that are all based on the same component networks and thus share a number of common topological and performance parameters.

One of the main foci of interconnection network research over the past two decades has been the exploration of the network design space, with particular emphasis on deriving networks with sublogarithmic diameters that can provide some of the desirable properties of the hypercube. Emphasis on sublogarithmic-degree networks was justified by considerations of VLSI area and packaging, including pin limitations [Parh00]. Perfect difference networks which form the primary focus of this paper, provide us with design points closer to the K_n extreme. They offer benefits of full connectivity at a much lower cost. The part of design space that falls between the hypercube and K_n has been of little interest in architectures with wired connectivity, but becomes more practical, and thus interesting, with wireless/optical links.

2. Mathematical Preliminaries

Given that the complete graph K_n (diameter $D = 1$) is impractical for large n, it is quite natural to consider the best topology for $D = 2$, the next most desirable network diameter. Based on Moore bounds, a degree-d directed graph with $D = 2$ can have no more than $d^2 + d + 1$ nodes. The corresponding upper bound $n = d^2 + 1$ for undirected graphs isn't much different [Parh99]. Perfect difference sets provide the mathematical tools for achieving this optimum number of nodes, asymptotically, within the framework of perfect difference networks or PDNs. Figure 1 shows the place of PDN in the spectrum of network choices with regard to diameter.

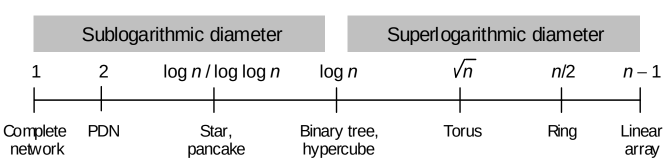

Fig. 1. The spectrum of interconnection networks in terms of diameter for size n.

Perfect difference sets were first discussed in 1938 by J. Singer. The formulation was in terms of points and lines in a finite projective plane. The theory of finite projective planes is highly developed [Hall67], but these mathematical notions are not required to understand the exposition that follows. We first present a theorem that forms the basis of the definition of perfect difference sets, and then proceed with the definition itself. All of the results in this section are from [Sing38].

Theorem 1: A sufficient condition that there exist $\delta + 1$ integers $s_0, s_1, \ldots, s_\delta$, having the property that their $\delta^2 + \delta$ differences $s_i - s_j$, $0 \le i \ne j \le \delta$, are congruent, modulo $\delta^2 + \delta + 1$, to the integers $1, 2, \ldots, \delta^2 + \delta$ in some order is that δ be a power of a prime. ■

Definition 1: Perfect difference set (PDS) – A set $\{s_0, s_1, \ldots, s_\delta\}$ of $\delta + 1$ integers having the property that their $\delta^2 + \delta$ differences $s_i - s_j$, $0 \le i \ne j \le \delta$, are congruent, modulo $\delta^2 + \delta + 1$, to the integers $1, 2, \ldots, \delta^2 + \delta$ in some order is a *perfect difference set* of order δ. Perfect difference sets are sometimes called *simple difference sets*, given that they correspond to the special $\lambda = 1$ case of difference sets for which each of the possible differences is formed in exactly λ ways. ■

Note that a PDS need not contain an integer outside the interval $[0, \delta^2 + \delta]$, because any integer outside the interval can be replaced by another integer in the interval without affecting the defining property of the PDS. The following is easily proven.

Theorem 2: Given a PDS $\{s_0, s_1, \ldots, s_\delta\}$ of order δ, the set $\{as_0 + b, as_1 + b, \ldots, as_\delta + b\}$, where a is relatively prime to $\delta^2 + \delta + 1$, also forms a perfect difference set. ■

By definition, any perfect difference set contains a pair of integers s_u and s_v such that $s_v - s_u \equiv 1 \bmod \delta^2 + \delta + 1$. By theorem 2, and the observation that preceded it, subtracting s_u from all integers in such a PDS yields another PDS that contains 0 and 1.

Definition 2: Normal PDS – A PDS $\{s_0, s_1, \ldots, s_\delta\}$ is reduced if it contains the integers 0 and 1. A reduced PDS is in normal form if it satisfies $s_i < s_{i+1} \le \delta^2 + \delta$, $0 \le i < \delta$. ■

Definition 3: Equivalent PDSs – Two different PDSs are equivalent iff they have the same normal form $\{0, 1, s_2, \ldots, s_\delta\}$. ■

Henceforth, we deal only with normal-form PDSs, some examples of which appear in Table 1. Several important properties of PDSs are noted in the following paragraphs.

Theorem 1 guarantees that a PDS exists for any n of the form $\delta^2 + \delta + 1$, where $\delta = p^h$ for a prime number p. It is suspected, though not yet proven for arbitrarily large n, that PDSs do not exist for other values of n [Guy94]. However, practically speaking, this is not alarming, given that primes and their powers are quite abundant, both in the range of practical interest and asymptotically; e.g., there are 197 primes and powers of primes under 1000.

For some values of δ, there exist more than one PDS. For example, we have the following PDSs of order 3: $\{0, 1, 3, 9\}$ and $\{0, 1, 4, 6\}$. It is easily verified that all numbers in the interval $[1, 12]$ can be formed as mod-13 difference of numbers in each of the sets above. Multiple difference sets of the same order lead to alternate interconnection network designs.

A PDS of order $\delta = p^h$, where p is a prime number, represents a set of n points and n lines in the 3D Euclidian space such that each point is on $\delta + 1$ lines and each line contains $\delta + 1$ points. This geometric interpretation leads to a PDS of order $\delta = p^h$ being generated from an irreducible degree-3 polynomial in GF(p^h); for details, see [Sing38].

Table 1. Perfect difference sets of orders up to 16. Note that the values of δ shown are powers of prime numbers and $n = \delta^2 + \delta + 1$.

δ	*n*	Example PDS of order δ in normal form
2	7	0, 1, 3
3	13	0, 1, 3, 9
4	21	0, 1, 4, 14, 16
5	31	0, 1, 3, 8, 12, 18
7	57	0, 1, 3, 13, 32, 36, 43, 52
8	73	0, 1, 3, 7, 15, 31, 36, 54, 63
9	91	0, 1, 3, 9, 27, 49, 56, 61, 77, 81
11	133	0, 1, 3, 12, 20, 34, 38, 81, 88, 94, 104, 109
13	183	0, 1, 3, 16, 23, 28, 42, 76, 82, 86, 119, 137, 154, 175
16	273	0,1,3,7,15,31,63,90,116,127,136,181,194,204,233,238,255

Besides the design of interconnection networks, discussed here, perfect difference sets have many other applications in error control coding, block designs (which are related to orthogonal Latin squares), and signal encoding to ensure negligible autocorrelation. These applications may be characterized by their need for provision of distance, variety, and/or orthogonality, or for avoiding coincidence, all of which are facilitated by unique differences offered by a PDS.

3. Perfect Difference Networks

Consider the normal-form PDS $\{0, 1, s_2, \ldots, s_\delta\}$ of order δ. We can construct a direct interconnection network with $n = \delta^2 + \delta + 1$ nodes based on this PDS as follows [Rako98], [Rako98a], [Rako01], [Parh03].

Definition 4: Perfect difference network (PDN) based on the PDS $\{0, 1, s_2, \ldots, s_\delta\}$ – There are $n = \delta^2 + \delta + 1$ nodes, numbered 0 to $n - 1$. Node i is connected via directed links to nodes $i \pm 1$ and $i \pm s_i$ (mod n), for $2 \le i \le \delta$. Because all index expressions in this paper are evaluated modulo n, henceforth we will delete the qualifier "mod n." The preceding connectivity leads to a chordal ring of in- and out-degree $d = 2\delta$ and diameter $D = 2$. Because for each link from node i to node j, the reverse link (j, i) also exists, the network corresponds to an undirected graph (see Figs. 2-3 for examples depicting the first two entries of Table 1). ■

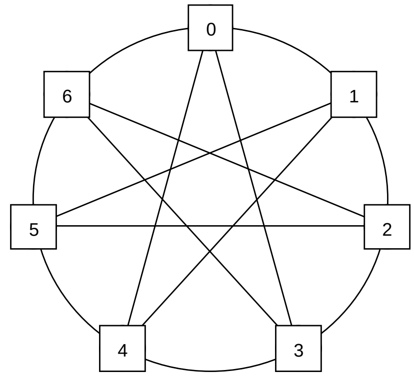

Fig. 2. PDN with $n = 7$ nodes based on the perfect difference set {0, 1, 3}.

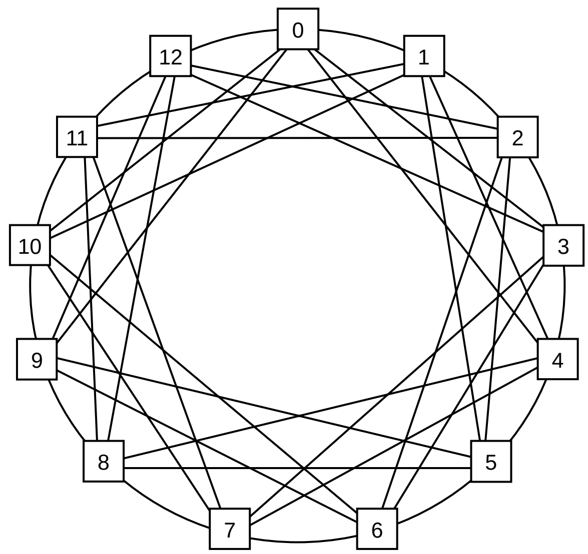

Fig. 3. PDN with $n = 13$ nodes based on the perfect difference set {0, 1, 3, 9}.

Every normal-form PDS contains 1 as a member. Therefore, PDNs based on normal-form PDSs are chordal rings. In the terminology of chordal rings, the links connecting consecutive nodes i and $i + 1$ are *ring links*, while those that connect nonconsecutive nodes i and $i + s_i$, $2 \le i \le \delta$, are *skip links* or *chords*. The link connecting nodes i and $i + s_i$ is referred to as the *forward skip link* of node i and *backward skip link* of node $i + s_i$. Similarly, the ring link between nodes i and $i + 1$ is a *forward (backward) ring link* for i ($i + 1$).

There is an alternate way in which we can define an interconnection structure based on the normal-form PDS $\{0, 1, s_2, \ldots, s_\delta\}$ of order δ [Rako00]. This scheme was briefly discussed in [Beut98], but the filing of the patent in [Rako98] and the second author's prior work leading to it predate [Beut98].

Definition 5: Bipartite PDN based on the PDS $\{0, 1, s_2, \ldots, s_\delta\}$ – There are $n = \delta^2 + \delta + 1$ host nodes, numbered 0 to $n - 1$, and similarly numbered switch nodes. Each host node i is connected via a pair of directed links to each of the switch nodes $i + 1$ and $i + s_i$, for $2 \le i \le \delta$. The preceding connectivity leads to a bipartite network, with host and switch nodes constituting the two parts. Both nodes and switches have in- and out-degrees $\delta + 1$. The host-to-host diameter of the network is $D = 2$. All host-to-host shortest paths are of length 2, leading to the average interhost distance $\Delta = 2$. Again, the bipartite network can be drawn as an undirected graph. ■

An example bipartite PDN for $n = 7$, based on the PDS {0, 1, 3}, is depicted in Fig. 4. One advantage of a bipartite PDN over a basic PDN is that its node degree is reduced from 2δ to $\delta + 1$ through the use of n switches, with each switch being a $(\delta + 1) \times (\delta + 1)$ communication node with full-crossbar or partial connection capability. The bipartite PDN can be viewed as simply a method for implementing the basic PDN. This is easily understood by drawing boxes around similarly numbered host and switch nodes in Fig. 4 to form the nodes of a basic PDN.

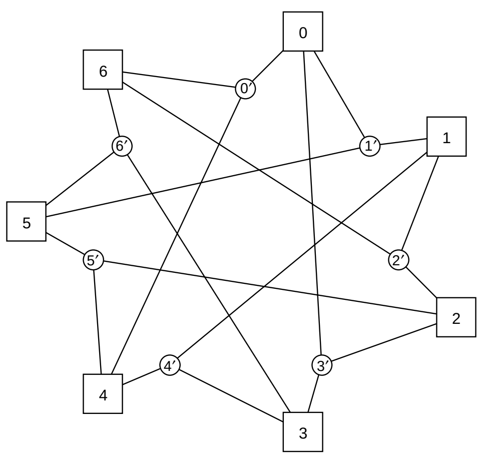

Fig. 4. Bipartite PDN with 7 hosts (squares) and 7 switches (circles), based on the PDS {0, 1, 3}.

It is also possible to interpret the bipartite PDN as a $2n$-node, degree-$(\delta + 1)$ network by simply viewing all nodes in Fig. 4 as host nodes. The resulting network has a diameter of 3. This is easily seen as follows. The host nodes replacing the original switch nodes are denoted by primed indices. Each such primed node is directly connected to several unprimed nodes and any pair of unprimed nodes are connected by a shortest path of length no greater than 2. As an example, node 0 in Fig. 4 is not connected to node 2′ by any path of length 2 or less, but there are several paths of length 3: 0 1′ 1 2′, 0 3′ 2 2′, 0 0′ 6 2′. These paths are node- and edge-disjoint. In general, there would be $\delta + 1$ such paths through all switches connected to the source node, given that the interswitch diameter is also 2.

4. PDN-Based Composite Networks

The perfect difference network, with its $O(n^{1/2})$ node degree and small constant diameter, in both its basic and bipartite forms, falls between the hypercube and complete graph in the design space of Fig. 1, offering performance close to the latter, at a much lower cost. If further cost reduction is desired, networks of smaller node degrees can be built based on the PDN concept. These networks fall in the space between hypercube and PDN in Fig. 1, offering somewhat lower performance than the latter at reduced cost, thus allowing cost-performance tradeoffs in numerous configurations. A wide variety of networks can be obtained through cross-product composition. For example, the q-cube (q-dimensional *binary hypercube*) is $K_2 \times K_2 \times \ldots \times K_2$. We thus define multidimensional PDN as the cross product of several component PDNs.

Definition 6: Product graph – The (*cross*) *product* of q graphs, $G_i = (V_i, E_i)$, $0 \le i \le q - 1$, denoted as $G = G_{q-1} \times G_{q-2} \times \ldots \times G_0$, is a graph with $n = n_{q-1} \times n_{q-2} \times \ldots \times n_0$ nodes, each labeled with a distinct q-digit *mixed-radix* integer $x_{q-1}x_{q-2} \ldots x_0$ in the range 0 to $n - 1$, so that nodes x and y are connected iff their labels differ in one and only one digit, say $x_j \ne y_j$, *and* x_j is connected to y_j in G_j [Yous95]. ■

Theorem 3: Topological properties of product networks – The node degree, average internode distance, and diameter of $G = G_{q-1} \times G_{q-2} \times \ldots \times G_0$ are the sums of the respective parameters for the q component networks [Yous95]. ■

Definition 7: Multidimensional PDN – Consider the q PDNs $H_0, H_1, \ldots, H_{q-1}$ based on their respective PDSs of orders $\delta_0, \delta_1, \ldots, \delta_{q-1}$. The product network $H_{q-1} \times H_{q-2} \times \ldots \times H_0$ is a qD, or q-dimensional, PDN. Nodes of a qD PDN are labeled by q-tuples $(x_{q-1}x_{q-2} \ldots x_0)$, where x_i belongs to the node set of H_i, $0 \le i < q$. When the q component PDNs H_i are identical, the resulting network H^q is a PDN-based *power network*. ■

For concreteness, we limit our discussion to 2D PDN-based power network H^2, depicted in Fig. 5, where node connections within rows or columns are removed to avoid clutter. The statements that follow are easily generalizable to higher dimensions and nonidentical component PDNs. Nodes in row i (column j) of H^2 are linked exactly as in an n-node PDN. Therefore, the total number of links in Fig. 5 is a factor of $2n$ greater than the number of links in its n-node basis PDN. Hence, increasing the number of nodes by a factor of n using 2D PDN has led to a factor $O(n)$ increase in link multiplicity. By contrast, had we opted for an $O(n^2)$-node PDN, its $O(n^3)$ links would have been a factor of $O(n^{3/2})$ higher that the corresponding number for an n-node PDN.

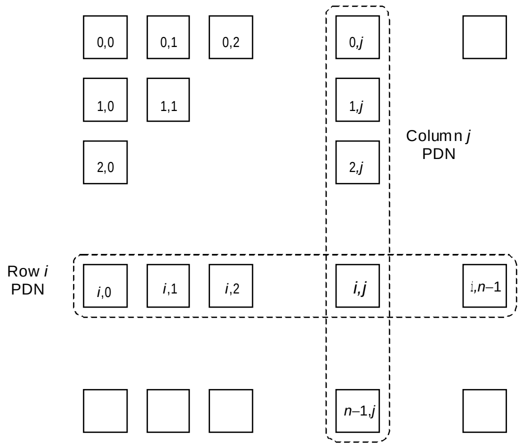

Fig. 5. The structure of the 2D PDN power network H^2.

Based on the properties of product graphs, the diameter of H^2 is 4 and its node degree is 4δ, where δ is the order of the PDS defining H. If each row/column PDN has n nodes, the 2D PDN power network will have $N = n^2$ nodes of degree $O(n^{1/2})$. Thus, node degree of H^2 grows as the fourth root of its size N. For example, a PDN with roughly 10^6 nodes requires node degree of about 2000, whereas a 2D PDN power network of the same size can be built of nodes with degrees that are about 16 times smaller. A consequence of this slower growth of node degree is that H^2 has a much more favorable degree-diameter product than a simple PDN of comparable size. Asymptotically, the dD factors are 16δ for H^2 versus $4(\delta^2 + \delta + 1)$ for the equivalent H', with the former being better except for $\delta = 2$.

Because H^2 is a power network, all algorithmic properties of power networks are applicable to it. For example, routing in H^2 can be accomplished via generalized "row/column" routing where a message is first routed in the "row" PDN and then in the "column" PDN, or vice versa. Broadcasting is done similarly. Any sorting algorithm for a square mesh that uses row and column sorts as its basic steps can be adapted to H^2 by emulating a linear-array sorting algorithm on the row and column PDNs. Hence, derivation of efficient algorithms for PDNs leads directly to a number of corresponding algorithms for H^2 with no additional effort.

A way of building composite networks with more favorable properties based on the PDN concept is via the "swap" connectivity [Yeh96], [Parh05].

Definition 8: Swapped networks – The *swapped network* G^{swapped}, based on the n-node nucleus graph G, is a graph with n copies of G numbered 0 to $n - 1$, so that nodes i in copy j is connected to node j of copy i for all $i \neq j$ and $0 \leq i, j \leq n - 1$ [Yeh96]. ■

Theorem 4: Properties of swapped networks [Yeh96] – If G has node degree d and diameter D, the node degree and diameter of G^{swapped} are $d + 1$ and $2D + 1$. ■

Consider a swapped network based on an n-node PDN. This network has n^2 nodes of degree $2\delta + 1$, with a network diameter of 5. The bisection width of this swapped network is upper bounded by $n^2/4$ (collapsing the clusters produces a K_n network) and is thus considerably smaller than that of a similar size PDN which has a bisection width of order n^3 [Parh03]. Routing, broadcasting, or total-exchange algorithms are quite simple for this network but under heavy traffic, the communication performance is likely to be lower than that of a PDN due to the smaller bisection.

5. Algorithms and Fault Tolerance

Efficiency of certain key communication algorithms play an important role in the usefulness of any network. The most important among these are point-to-point communication (one node sending a message to another node), one-to-all broadcasting (a source node sending a message to every other node), all-to-all broadcasting, and total exchange (every node sending a unique message to every other node).

Algorithm 1: Routing in a PDN – Routing from source node x to destination node y in H is tantamount to determining an intermediate node k such that $k - x = s_i$ and $k - y = s_j$, for some pair of elements s_i and s_j in the PDS on which H is based. The problem thus reduces to the determination of s_i, viz. the first link to be traversed on route to node y, since both k and s_j are then uniquely determined. The value of s_i can be obtained based on table lookup or by means of calculations relating to the mathematical notion of finite projective planes [Sing38]. ■

Algorithm 2: Broadcasting in a PDN – Broadcasting in H is a 2-phase process, given that the network diameter is 2. In the first phase, the initiating node x sends the broadcast message to node $x + s_i$ for each nonzero member s_i of the PDS on which H is based. At the end of this phase, which involves δ message transmissions, $\delta + 1$ nodes are aware of the broadcast message. In the second phase, each node y that already has the broadcast message (including the initiator) sends the message to nodes $y - s_j$, except for the node from which the broadcast message was received. In this step, x sends δ messages, while each of the δ intermediate sources sends $\delta - 1$ messages, for a total of δ^2 messages. Broadcasting is thus completed in the minimum possible number $\delta^2 + \delta$ of message transmissions. With the single-port communication model, where a node can send but one message in each time step, the total broadcast time is 2δ steps. All-port communication, on the other hand, leads to 2 time steps for broadcasting; this model is less realistic for networks with large node degrees and will not be considered further in this paper. ■

Algorithm 3: All-to-all broadcasting in a PDN – All-to-all broadcasting involves each node sending a message to all other nodes in the network; hence, n distinct messages must be sent, with each one going to $n - 1$ destinations. Each node follows Algorithm 2, with single-port communication, independently. First, the broadcast message of a node x is sent to all its neighbors $x + s_i$ in δ steps. These messages do not conflict with each other because the neighbors $u + s_i$ and $v + s_i$ for distinct nodes u and v are distinct. At the end of this phase, each node has already received δ of the expected $\delta^2 + \delta$ broadcast messages. A node y now sends its own message to all nodes $y - s_j$, and each of its received δ messages to the same nodes, except the one from which the message originated. These steps require $\delta + \delta(\delta - 1) = \delta^2$ transmissions; again, there is no conflict. All-to-all broadcasting is thus completed in $\delta^2 + \delta = n - 1$ transmission steps, which is the minimum possible. ■

Algorithm 4: Total exchange in a PDN – Total exchange (also known as gossiping or all-to-all personalized communication) involves each node sending a distinct message to each of the other nodes in the network; hence, $n(n - 1)$ distinct messages must be transmitted, with each one having a single destination. Distinct messages must go from each node x to all nodes $x + s_i - s_j$. As s_i and s_j assume all values in the PDS, all distinct nodes are covered. In phase 1, requiring δ message transmissions, each node sends its messages to nodes $x + s_i$; i.e., nodes that can be reached with $s_j = 0$. In the second phase, again involving δ message transmissions, nodes $x - s_j$ are addressed. The remaining $\delta(\delta - 1)$ destinations for each source must be reached in 2 steps. All nodes step through possible s_i and s_j ($s_i \neq s_j$) values in unison, with node x sending a message to node $x + s_i$, requesting that it be forwarded to node $x + s_i - s_j$. This process needs $2\delta(\delta - 1)$ message transmissions and involves no conflict. The algorithm requires $2\delta^2$ message transmissions in all, which is optimal in view of each node sending $n - 1$ messages and each message traveling the average internode distance of $2\delta^2/(n - 1)$ hops. The preceding algorithm and analysis assume that a node can send or receive a message in each time step, but not both at the same time. If simultaneous transmission and reception of messages is possible within a node, then the $2\delta(\delta - 1)$ steps above can be reduced to $\delta(\delta - 1) + 1$ through pipelining, leading to a total exchange time of $\delta^2 + \delta + 1 = n$ steps which is very nearly the best possible. It is easily proven that in the course of this algorithm, all PDN links carry the same number of messages; hence, message traffic is fully balanced. ■

Like the communication processes described in Algorithms 1-4, one can develop algorithms for common parallel computations that involve both data transfer steps and local node operations. An alternative to developing parallel algorithms from scratch is to show that a new connectivity is capable of emulating a well-known architecture efficiently, thus allowing available algorithms to run on the new architecture in a step-by-step emulation mode. The following result shows that PDN can emulate a complete network with asymptotically optimal slowdown.

Theorem 5: A PDN of order δ can emulate a complete network of the same size with a slowdown factor of at most $2\delta + 2$. This emulation is asymptotically optimal.

Proof: As a result of our ability to construct a set of $n(n - 1)/2$ paths between all pairs of nodes such that no more than $\delta + 1$ paths pass through any one link, the bisection width of an n-node PDN is shown to be of order $n^{3/2}$. We say that the congestion of this routing scheme is $c = \delta + 1$. If we use these routing paths for communication between nodes in our PDN, any communication step in the associated complete graph will be slowed down by a factor no greater than $\delta + 1$ due to link congestion and by a factor of at most 2 due to routing distance increasing from 1 to 2. Asymptotic optimality of this emulation is a direct consequence of the bisection widths of the complete graph and PDN being of orders n^2 and $n^{3/2}$, respectively. ■

Any computer system utilizing a large number of nodes and links must be robust if the failure of a very small subset of the many components is not to lead to total system crash. One aspect of network robustness is its survivability, which requires lack of vulnerable spots along with ability to withstand limited failures or attacks [Hobb91]. Intuitively, lack of vulnerable spots is synonymous with “blandness” which means that careful study of the network will not reveal parts that are particularly attractive as targets of a malicious attack. Ability to withstand limited attacks implies richness of connectivity so that disrupting certain nodes and/or links will not disconnect the network. PDNs are certainly both bland and richly connected.

Another aspect of robustness is the amount by which the shortest path between two nodes increases in length when node and/or link failures occur. Consider, for example, the effects of removing a single link of a PDN. Does this action increase the shortest distance between any pair of nodes? Unfortunately, it does. Note that whenever the shortest distance between a pair of nodes is 2, there are two link-disjoint shortest paths so that the removal of a single link never increases the length of the shortest path. When the two nodes are directly connected, however, removal of the link connecting them may increase the length of the shortest path to 3. One way around this problem is using a 0-free PDS, defined as one that does not contain 0. Such a PDS guarantees the existence of two node- and link-disjoint paths of length 2 between any pair of nodes in the associated PDN. A canonical 0-free PDS can be obtained by adding 1 to all elements of a canonical PDS. The 0-free version of the PDN in Fig. 2 turns into the complete graph K_7, but generally, the node degree increases from 2δ to $2\delta + 2$, which is still $O(n^{1/2})$.

6. Conclusion

We have introduced PDNs and the mathematical underpinnings that make them desirable as robust, high-performance communication and parallel processing networks. Some topological properties of PDNs were discussed and various routing algorithms for them were presented. It was shown that an n-node PDN can emulate the complete network K_n with optimal slowdown and balanced message traffic. Although other interconnection architectures with topological and performance characteristics similar to PDNs exist, we view PDNs as worthy additions to the repertoire of computer system designers. Alternative network topologies offer additional design points that can be exploited to accommodate the needs of new and emerging technologies. Further study is needed to resolve some open questions and to derive cost/performance comparisons for PDNs.

Whereas PDNs are interesting and important as asymptotically optimal diameter-2 interconnection structures, it is much more likely that hybrid or composite networks involving PDNs as component structures prove useful for practical applications. We introduced multidimensional and swapped PDNs as examples that lead to constant-diameter networks with lower cost and performance than pure PDNs. Various hierarchical or multilevel combinations of PDNs with other networks are also possible.

Possible generalizations of the perfect difference concept may lead to more efficient networks. For example, given an interest in 2-hop routing, we do not need to restrict ourselves to differences; sums can also be used. A natural question then is whether the use of difference/sum sets can lead to smaller sets (lower degree) or larger networks with the same node cost. As an example, the set {0, 2, 5, 6} leads to the mod-15 sums and differences {2, 5, 6, 7, 8, 11} and {1, 2, 3, 4, 5, 6, 9, 10, 11, 12, 13, 14}, which together cover all integers in [1, 14].

References

[Arde81] Arden, B.W. and H. Lee, "Analysis of Chordal Ring Networks," *IEEE Trans. Computers*, Vol. 30, No. 4, pp. 291-295, April 1981.

[Beut98] Beutelspacher, A. and U. Rosenbaum, *Projective Geometry: From Foundations to Applications*, Cambridge, 1998, Section 2.8, pp. 81-88.

[Guy94] Guy, R.K., *Unsolved Problems in Number Theory*, Springer, 2nd ed., 1994, pp. 118-121.

[Hall67] Hall, M., *Combinatorial Theory*, Blaisdell, 1967.

[Hobb91] Hobbs, A.M., "Network Survivability," in *Applications of Discrete Mathematics*, ed. by J.G. Michaels and K.H. Rosen, McGraw-Hill, 1991, pp. 332-353.

[Kwai98] Kwai, D.-M. and B. Parhami, "Pruned Three-Dimensional Toroidal Networks," *Information Processing Letters*, Vol. 68, No. 4, pp. 179-183, 1998 November 30.

[Parh99] Parhami, B., *Introduction to Parallel Processing: Algorithms and Architectures*, Plenum Press, 1999.

[Parh99a] Parhami, B. and D.-M. Kwai, "Periodically Regular Chordal Rings," *IEEE Trans. Parallel and Distributed Systems*, Vol. 10, No. 6, pp. 658-672, June 1999. (Printer errors corrected in Vol. 10, No. 7, pp. 767-768, July 1999.)

[Parh00] Parhami, B. and D.-M. Kwai, "Challenges in Interconnection Network Design in the Era of Multiprocessor and Massively Parallel Microchips," *Proc. Int'l Conf. Communications in Computing*, pp. 241-246, June 2000.

[Parh05] Parhami, B., "Swapped Interconnection Networks: Topological, Performance, and Robustness Attributes," *J. Parallel and Distributed Computing*, to appear.

[Rako98] Rakov, M., "Method of Interconnecting Nodes and a Hyperstar Interconnection Structure", US Patent 5 734 580, issued March 1998.

[Rako98a] Rakov, M., "Hyperstar: A New Interconnection Topology," *J. China Universities of Posts and Telecommunications*, Vol. 5, No. 2, pp. 10-18, December 1998.

[Rako99] Rakov, M., "Multidimensional Hyperstar and Hyperhub Interconnection Methods and Structures," US Patent Application No. 09/410 175, filed September 1999.

[Rako00] Rakov, M., "Hyperstar and Hyperhub Optical Networks Interconnection Methods and Structures," US Patent Application No. 09/634 129, filed August 2000.

[Rako01] Rakov, M. and J. Mackall, "Method of Interconnecting Functional Nodes and a Hyperstar Interconnection Structure," US Patent 6 330 706, issued December 2001.

[Sing38] Singer, J., "A Theorem in Finite Projective Geometry and Some Applications to Number Theory," *Trans. American Mathematical Society*, Vol. 43, pp. 377-385, 1938.

[Yeh96] Yeh, C.-H. and B. Parhami, "Swapped Networks: Unifying the Architectures and Algorithms of a Wide Class of Hierarchical Parallel Processors," *Proc. Int'l Conf. Parallel and Distributed Systems*, June 1996, pp. 230-237.

[Yous95] Youssef, A., "Design and Analysis of Product Networks," *Proc. Symp. Frontiers of Massively Parallel Computation*, February 1995, pp. 521-528.