ELSEVIER

Information Processing Letters 55 (1995) 123–127

Information Processing Letters

A scheme to construct distance-three codes using latin squares, with applications to the n-cube

Pranava K. Jha [a,*], Giora Slutzki [b]

[a] Department of Computer Engineering, Delhi Institute of Technology: Delhi, Kashmere Gate, Delhi 110 006, India
[b] Department of Computer Science, Iowa State University, Ames, IA 50011, USA

Communicated by D. Gries; received 9 November 1994; revised 8 March 1995

Keywords: Combinatorial problems; Distance-three codes; n-cube; Latin squares; Independent domination number

1. Introduction

Let B_n denote the set of n-bit *binary strings*, and let Q_n denote the graph of the *n-cube* where $V(Q_n) = B_n$ and where two vertices are adjacent iff their *Hamming distance* is exactly one. A subset C of B_n is called a *code*, and the elements of C are referred to as *codewords*. C is said to be a *linear code* if the codeword obtained from component-wise sum (modulo 2) of any two elements of C is again in C; otherwise it is a nonlinear code. By a *distance-three code* is meant a code in which the Hamming distance between any two distinct codewords is at least three. Distance-three codes possess the capability to correct one error and detect two or fewer errors.

It is known that if n is of the form $2^k - 1$, then B_n admits of a partition into equal-size sets $V_0, \ldots, V_n$ such that each V_i is a distance-three code and is maximal with respect to this property (see e.g. [5] or [3].) The main contribution of this paper is a scheme that systematically constructs a large family of such partitions by means of *latin squares*. In a somewhat similar study, Sloane and Seidel [6] earlier employed conference matrices to construct a family of nonlinear codes with high minimum distance.

We derive sharp bounds on the *domination number* and the *independent domination number* of the n-cube. Indeed, our upper bound on each of the two invariants of Q_n is within twice the optimal. These corollaries are important in their own right, since the general problem of determining any of these two invariants is known to be NP-hard. In fact, independent domination number is, in general, not even approximable in polynomial time within a factor of $n^{1-\varepsilon}$ for any $\varepsilon > 0$ unless P = NP, cf. [1].

By a graph is meant a finite, simple, undirected graph. Let $G = (V, E)$ be a graph, and let $S \subseteq V$. S is said to be an *independent set* if all elements of S are mutually nonadjacent in G. An independent set that is maximal with respect to the independence property is called a *maximal independent set*. S is said to be a *dominating set* if

* Corresponding author. Email: pkj@dit.ernet.in.

0020-0190/95/$09.50 © 1995 Elsevier Science B.V. All rights reserved
SSDI 0020-0190(95)00066-6

every vertex of G that is not in S is adjacent to some vertex of S. It is easy to see that S is a maximal independent set iff it is an independent set as well as a dominating set. The *domination number dom*(G) of G is defined to be the size of a smallest dominating set. A maximal independent set of smallest size is called a *minimum independent dominating set* (*mids*), and its cardinality is referred to as *independent domination number*, denoted by *idom*(G).

For two binary strings x and y, let $x \cdot y$ denote concatenation of x and y, and for two sets X and Y of binary strings, let $X \cdot Y = \{x \cdot y \mid x \in X \wedge y \in Y\}$. A subset S of B_n is said to be closed under *bitwise complementation* if $a_0 \cdots a_{n-1} \in S$ implies $\bar{a}_0 \cdots \bar{a}_{n-1} \in S$, where $\bar{0} = 1$ and $\bar{1} = 0$.

It is straightforward to see that Q_n is a bipartite graph with 2^n vertices and $n2^{n-1}$ edges. The following two lemmas are relevant.

Lemma 1.1. $2^n/(n+1) \leqslant dom(Q_n) \leqslant idom(Q_n)$.

Proof. It suffices to settle the lower bound on $dom(Q_n)$. Note that every vertex of Q_n is adjacent to n other vertices and hence dominates a total of $n+1$ vertices including itself. Thus, in order to dominate all 2^n vertices of Q_n, we need to select a minimum of $2^n/(n+1)$ vertices. □

Lemma 1.2. *Let* $n = 2^k - 1$ *where* $k \geqslant 2$, *and let* S *be a vertex subset of* Q_n *such that* $|S| = 2^n/(n+1)$. S *is a minimum independent dominating set of* Q_n *iff for any two distinct elements* x *and* y *of* S, $d_H(x, y) \geqslant 3$.

Proof. Let n, k and S be as in the statement of the lemma. First suppose that $d_H(x, y) \geqslant 3$ for any two distinct elements x and y of S. Thus, no two distinct elements of S have a common neighbor, so a vertex of Q_n that is not in S is adjacent to at most one element of S. Consequently, S dominates a total of $|S| \cdot (n+1) = 2^n$ vertices of Q_n, that is, all of them. By Lemma 1.1, S is a minimum independent dominating set of Q_n.

For the converse, note that if $x, y \in S$ and $d_H(x, y) < 3$, then S (which is of size $2^n/(n+1)$) cannot even be a dominating set of Q_n. □

An $r \times r$ *latin square* is defined to be a square matrix M over the set $\{0, \ldots, r-1\}$ such that every row and every column of M contains each element of $\{0, \ldots, r-1\}$ exactly once. For instance, the following cyclic matrix is a latin square.

$$\begin{pmatrix} 0 & 1 & 2 & \ldots & r-1 \\ 1 & 2 & 3 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ r-1 & 0 & 1 & \ldots & r-2 \end{pmatrix}$$

Section 2 consists of the main result while Section 3 contains certain corollaries, which lead to sharp bounds on $dom(Q_n)$ and $idom(Q_n)$.

2. Main result

Throughout this section, let $n = 2^k - 1$, where $k \geqslant 1$. We present a scheme, called *CubePartition*, that inducts on k and builds a partition of B_{2n+1} from that of B_n. The trick is to employ an $(n+1) \times (n+1)$ latin square and exploit its structure to construct mutually disjoint distance-three codes.

procedure *CubePartition*;

(* For $n = 2^k - 1$, inductively construct a partition of B_n into $n+1$ equal-size distance-three codes *)

begin

1. If $n = 1$, the partition is unique: return $\{\{0\}, \{1\}\}$.
2. If $n = 3$, the partition is unique: return $\{\{000, 111\}, \{001, 110\}, \{010, 101\}, \{011, 100\}\}$.
3. We have $n = 2^k - 1$, where $k \geqslant 2$. Suppose $\{V_0, \ldots, V_n\}$ is a partition of B_n into equal-size distance-three codes. Thus, each V_i is of size $2^n/(n+1) = r + 1$ (say). Let $V_i = \{v_{i,0}, \ldots, v_{i,r}\}$, $0 \leqslant i \leqslant n$.
4. Let $C_i = \{v_{i,0} \cdot b_{i,0}, \ldots, v_{i,r} \cdot b_{i,r}\}$ and $D_i = \{v_{i,0} \cdot \bar{b}_{i,0}, \ldots, v_{i,r} \cdot \bar{b}_{i,r}\}$, $0 \leqslant i \leqslant n$, where $b_{i,0} = 0$ (resp. 1) if the number of 1's in $v_{i,j}$ is even (resp. odd), and $\bar{b}_{i,j} = 1 - b_{i,j}$.
   (* Sets $C_0, \ldots, C_n$, $D_0, \ldots, D_n$ form a partition of B_{n+1}. *)
   (* Elements of C_i (resp. D_i) are of even (resp. odd) parity. *)

5. Let $T=(t_{i,j})$ be an $(n+1)\times(n+1)$ latin square.
6. Return the sets $W_0,\dots,W_{2n+1}$, where $2n+1=2^{k+1}-1$, and the W_i are constructed as follows:

$$W_i=\begin{cases} C_0\bullet V_{t_{i,0}}\cup\cdots\cup C_n\bullet V_{t_{i,n}}, \\ \quad 0\leqslant i\leqslant n, \\ D_0\bullet V_{t_{i-n-1,0}}\cup\cdots\cup D_n\bullet V_{t_{i-n-1,n}}, \\ \quad n+1\leqslant i\leqslant 2n+1 \end{cases}$$

end. (* *CubePartition* *)

We now prove that sets $W_0,\dots,W_{2n+1}$, obtained above, constitute a well-defined partition of B_{2n+1} into equal-size distance-three codes.

Proposition 2.1. *Let $W_0,\dots,W_m$ be sets obtained at the termination of procedure CubePartition, where $m=2n+1$ and $n=2^k-1$.*
(1) $|W_i|=2^m/(m+1)$, $0\leqslant i\leqslant m$.
(2) Each element of W_i is a binary string of length m.
(3) For $i\neq j$, $W_i\cap W_j=\emptyset$.
(4) For distinct $x, y\in W_i$, $d_H(x, y)\geqslant 3$.

Proof. (1) follows from the fact that the sets $C_0\bullet V_{t_{i,0}}, \dots, C_n\bullet V_{t_{i,n}}$ (resp. the sets $D_0\bullet V_{t_{i-n-1,0}}, \dots, D_n\bullet V_{t_{i-n-1,n}}$) are mutually disjoint, where $0\leqslant i\leqslant n$ (resp. $n+1\leqslant i\leqslant m$). (2) is obvious while (3) is a consequence of the structure of a latin square and the facts that (a) the sets $C_0,\dots,C_n$, $D_0,\dots,D_n$ are mutually disjoint and (b) the sets $V_0,\dots,V_n$ are mutually disjoint.

We prove (4) by induction on k. The basis is trivially true. Let x, y be distinct elements of W_i, where $0\leqslant i\leqslant n$. Then for some $a, b, c, d\in\{0,\dots,n\}$ we have $x\in C_a\bullet V_b$ and $y\in C_c\bullet V_d$, where $b=t_{i,a}$ and $d=t_{i,c}$, and $(t_{i,j})$ is a latin square as in Step 5 of the procedure. We may write $x=x_1\cdot x_2$ and $y=y_1\cdot y_2$, where $x_1\in C_a$, $x_2\in V_b$, $y_1\in C_c$ and $y_2\in V_d$. Since x, y are distinct, it cannot happen that $x_1=y_1$ and $x_2=y_2$. First suppose that $x_1=y_1$ and $x_2\neq y_2$. That $x_1=y_1$ implies $a=c$, and hence $b=d$. Consequently, x_2, y_2 are distinct elements of V_b. By induction hypothesis, $d_H(x_2, y_2)\geqslant 3$, and hence $d_H(x, y)\geqslant 3$. Argument is similar for the case when $x_1\neq y_1$ and $x_2=y_2$. Next suppose that $x_1\neq y_1$ and $x_2\neq y_2$. There are two subcases: $a=c$ and $a\neq c$. If $a=c$, then $b=d$, and hence x_1, y_1 (resp. x_2, y_2) are distinct elements of C_a (resp. V_b), and the claim is immediate. On the other hand, if $a\neq c$, then $b\neq d$, and we must have $d_H(x_1, y_1)\geqslant 2$ and $d_H(x_2, y_2)\geqslant 1$. (Note that two distinct binary strings that are of the same parity must have a Hamming distance of at least two.) It follows that $d_H(x, y)\geqslant 3$. The argument is similar for the case when x, y are distinct elements of W_i, where $n+1\leqslant i\leqslant m$. □

At Step (6) of procedure *CubePartition*, sets $W_0,\dots,W_{2n+1}$ may alternatively be defined as follows:

$$W_i=\begin{cases} V_0\bullet C_{t_{i,0}}\cup\cdots\cup V_n\bullet C_{t_{i,n}}, \\ \quad 0\leqslant i\leqslant n, \\ V_0\bullet D_{t_{i-n-1,0}}\cup\cdots\cup V_n\bullet D_{t_{i-n-1,n}}, \\ \quad n+1\leqslant i\leqslant 2n+1 \end{cases}$$

The resulting partition will, in general, be different from that obtained earlier.

If $k\geqslant 2$, then each of the sets constructed by procedure *CubePartition* is closed under bitwise complementation. In other words, if a vertex x of the n-cube is in a particular set W_i, then the antipodal (that is, diametrically opposite) vertex of x is also in W_i. This is seen by the following inductive proof. For $k=2$ (and hence $n=3$), this is clearly true. Suppose that $\{V_0,\dots,V_n\}$ is a partition of B_n as in Step (3) of the procedure and that each V_i is closed under bitwise complementation. It is easy to see that each of the sets $C_0,\dots,C_n$, $D_0,\dots,D_n$ will also have this property. Further, if two sets X and Y obey this closure property, then so do $X\cup Y$ and $X\bullet Y$. The relevance of this observation may be seen from the fact that a code that is closed under the above operation and that does not contain the zero vector is necessarily nonlinear.

Let $\{V_0,\dots,V_n\}$ be a partition of B_n as in Step (3) of *CubePartition*, and let M_1 and M_2 be two distinct $(n+1)\times(n+1)$ latin squares. These latin squares may or may not lead to distinct partitions of B_{2n+1}. In particular, if the set of rows of M_2 is a permutation of the set of rows of M_1, then the resulting partitions will not be different. On the other hand, if there is no such relationship between M_1 and M_2, then the corresponding partitions will be different.

Our scheme may not generate all possible distance-three codes. To demonstrate this, we present a partition of Q_7 that cannot be obtained by means of this procedure. For convenience, let us use decimal (rather than binary) notation for the vertices of Q_7, that is, $V(Q_7) = \{0, \dots, 127\}$. Eight sets that constitute one such partition are as follows.

$\{0, 11, 21, 30, 38, 45, 51, 56,$
$71, 76, 82, 89, 97, 106, 116, 127\}$,

$\{1, 10, 20, 31, 39, 44, 50, 57,$
$70, 77, 83, 88, 96, 107, 117, 126\}$,

$\{2, 9, 23, 28, 36, 47, 49, 58,$
$69, 78, 80, 91, 99, 104, 118, 125\}$,

$\{3, 8, 22, 29, 37, 46, 48, 59,$
$68, 79, 81, 90, 98, 105, 119, 124\}$,

$\{4, 15, 17, 26, 34, 41, 55, 60,$
$67, 72, 86, 93, 101, 110, 112, 123\}$,

$\{5, 14, 16, 27, 35, 40, 54, 61,$
$66, 73, 87, 92, 100, 111, 113, 122\}$,

$\{6, 13, 19, 24, 32, 43, 53, 62,$
$65, 74, 84, 95, 103, 108, 114, 121\}$,

$\{7, 12, 18, 25, 33, 42, 52, 63,$
$64, 75, 85, 94, 102, 109, 115, 120\}$.

The reader may verify that these sets have the desired characteristics. That this partition cannot be obtained by our scheme follows from the observations that (i) there is a unique partition of B_3, that is, $\{\{0, 7\}, \{1, 6\}, \{2, 5\}, \{3, 4\}\}$ and (ii) no 4×4 latin square coupled with this partition can yield a partition of B_7 in which the elements 0 (decimal) and 11 (decimal) appear in the same subset. Interestingly enough, all the above sets are also closed under bitwise complementation.

3. Corollaries

Recall Lemmas 1.1 and 1.2, and note that procedure *CubePartition* may be viewed as a scheme for a vertex decomposition of Q_n into minimum independent dominating sets, where n is of the form $2^k - 1$. In this section, we discuss cube decomposition into maximal independent sets for the case when n is not of the foregoing form, and obtain bounds on $dom(Q_n)$ and $idom(Q_n)$.

Assuming that $n \neq 2^k - 1$, let r be the largest integer such that $n > r$ and $r = 2^k - 1$, that is, $r + 1 = 2^{\lfloor \log_2(n+1) \rfloor}$. Obtain a partition $\{V_0, \dots, V_r\}$ of $V(Q_r)$ by means of procedure *CubePartition*. Next, let $\{A_0, A_1\}$ be a partition of $V(Q_{n-r})$ such that A_0 (resp. A_1) is the set of binary strings of even (resp. odd) parity. Thus, $|A_0| = |A_1| = 2^{n-r-1}$. For $0 \leqslant i \leqslant (r-1)/2$, let

$$W_{ii} = A_0 \bullet V_{2i} \cup A_1 \bullet V_{2i+1} \quad \text{and}$$
$$W_{2i+1} = A_0 \bullet V_{2i+1} \cup A_1 \bullet V_{2i}.$$

That the sets $W_0, \dots, W_r$ are equal-size maximal independent sets of Q_n, and constitute a partition of $V(Q_n)$ follows from the following five claims, which may be argued as in the proof of Proposition 2.1.

(1) $|W_i| = 2^n/(r+1)$, $0 \leqslant i \leqslant r$.
(2) Each element of W_i is a binary string of length n.
(3) For $i \neq j$, $W_i \cap W0_j = \emptyset$.
(4) For distinct $x, y \in W_i$, $d_H(x, y) \geqslant 2$.
(5) W_i is a dominating set of Q_n, $0 \leqslant i \leqslant r$.

It follows from the discussions of the preceding section and of the present section that for all $n \geqslant 1$, the n-cube admits of a vertex decomposition into maximal independent sets each of which is of size $2^n / 2^{\lfloor \log_2(n+1) \rfloor}$. This conclusion and Lemma 1.1 yield the following bounds on $dom(Q_n)$ and $idom(Q_n)$:

$$\frac{2^n}{n+1} \leqslant dom(Q_n) \leqslant idom(Q_n) \leqslant \frac{2^n}{2^{\lfloor \log_2(n+1) \rfloor}}.$$

Note that the upper bound on $idom(Q_n)$ is the least power of two that is at least $2^n/(n+1)$. Observe also that the lower bound and the upper bound are within a factor of two, and for n of the form $2^k - 1$, they coincide and hence yield the exact value. This partially answers a question raised by Harary et al [2] with respect to the determination of $dom(Q_n)$. Certain amplifications of these issues appear in [4]. Exact determination of $dom(Q_n)$ and $idom(Q_n)$ is open.

Acknowledgments

The authors are grateful to Dr. Jonathan D.H. Smith for his interest, encouragement and perceptive comments. They are also thankful to Dr. David Gries and the anonymous referees whose observations led to a substantial improvement in the presentation of the paper.

References

[1] M.M. Halldórsson, Approximating the minimum maximal independence number, *Inform. Process. Lett.* **46** (1993) 169–172.

[2] F. Harary, J.P. Hayes and H.-J. Wu, A survey of the theory of hypercube graphs, *Comput. Math. Appl.* **15** (1988) 277–289.

[3] D.G. Hoffman, D.A. Leonard, C.C. Lindner, K.T. Phelps, C.A. Rodger and J.R. Wall. *Coding Theory: The Essentials* (Marcel Dekker, New York, 1991).

[4] P.K. Jha, Hypercubes, median graphs and products of graphs: Some algorithmic and combinatorial results, Ph.D. Dissertation, Iowa State, 1990.

[5] V. Pless, *Introduction to the theory of error-correcting codes* (John Wiley & Sons, New York, 2nd ed., 1989).

[6] N.J.A. Sloane and J.J. Seidel, A new family of nonlinear codes obtained from conference matrices, *Ann. New York Acad. Sci.* **175** (1970) 363–365.