Discrete Applied Mathematics 37/38 (1992) 553-562
North-Holland

Application of coding theory to interconnection networks

Gilles Zémor and Gérard D. Cohen

Département d'Informatique, C215-4, ENST, 46, rue Barrault, 75634 Paris Cedex 13, France

Received 21 June 1989
Revised 27 March 1990

Abstract

Zémor, G. and G.D. Cohen, Application of coding theory to interconnection networks, Discrete Applied Mathematics 37/38 (1992) 553-562.

We give a few examples of applications of techniques and results borrowed from error-correcting codes to problems in graphs and interconnection networks. The degree and diameter of Cayley graphs with vertex set $(Z/2Z)^r$ are investigated. The asymptotic case is dealt with in Section 2. The robustness, or fault tolerance, of the n-cube interconnection network is studied in Section 3.

A few words on codes and graphs

Network theory deals with the estimation of graphs parameters such as the diameter, maximal degree, robustness, etc. These concepts have coding-theoretical counterparts in several situations: for instance, Cayley graphs can be associated with codes to yield the following correspondences:

diameter ↔ covering radius,

degree ↔ length,

robustness ↔ dual distance.

The aim of this paper is to show how coding theory can interpret some network-theoretical problems, to use the correspondence to give bounds on the parameters of an interconnection network, and to suggest a few constructions.

We have tried to make this paper self-contained with respect to coding theory. However for general background, the interested (or distressed) reader can consult [11].

A (binary linear) $[n, k, d]$ code C is a linear subspace of dimension k of F^n, the n-dimensional vector space over $F = \{0, 1\}$.

F^n is endowed with the Hamming metric $d_H(.,.)$ which counts the number of different coordinates between two vectors, and d (or $d(C)$) is the *minimum distance* of C, i.e., the minimum distance between any two distinct elements (codewords) of C.

Labeling vertices of a graph with elements of F^n for some n and drawing edges between pairs of vertices at a given (Hamming) distance is convenient for many purposes: it yields natural routing algorithms (see e.g. [15]), and good classes of graphs (de Bruijn, Kautz, odd graphs, etc.). For example (see [4]), Petersen and Akers graphs are odd graphs obtained by taking for vertices all n-tuples of weight $(n+1)/2$ and joining two vertices by an edge if their Hamming distance is $n-1$, for $n=5$ and $n=7$ respectively.

Unless stated otherwise, codes are binary and linear hereafter and log is to the base 2.

Two elements $x=(x_i)$ and $y=(y_i)$ in F^n are *orthogonal* if $\langle x, y\rangle := \sum x_i y_i = 0$ (in F).

The *dual* code of C:

$$\bar{C} := \{x \in F^n : \langle x, c\rangle = 0 \text{ for all } c \text{ in } C\}$$

is a linear $[n, n-k]$ code. Its dimension $r := n-k$ is called the *redundancy* of C.

The elements in a basis of $\bar{C}$ can be written as rows of an $(n-k)\times n$ matrix called a *parity-check matrix* H. Then:

$$c \in C \text{ iff } cH^t = 0. \tag{1}$$

Let G be an undirected connected graph with v vertices and m edges. Identify any subset of edges with an element of F^m (its characteristic function). Then there are two classical codes associated with G (see e.g. [13]). Although they are not optimal, we shall give them here as an illustration. The *circuit code* C_1 is the set of circuits and disjoint unions of circuits. Its minimum distance is clearly the girth "g" of G: C_1 has parameters $[m, m-v+1, g]$. Similarly, the *cutset code* C_2 is defined as the set of cutsets and disjoint unions of cutsets. Its dimension is $v-1$. Since a cutset and a circuit intersect in an even number of edges, the cutset code is the dual of the circuit code. Its minimum distance is the edge-connectivity λ of G. Thus C_2 has parameters $[m, v-1, \lambda]$.

Example. For the famous Petersen graph, the circuit code and cutset code have respectively the following parameters: [15, 6, 5] and [15, 9, 3].

There are a few general bounds in coding theory relating n, k and d. Let us simply mention a good general one, due to McEliece, Rodemich, Rumsey and Welch (see [11]):

$$k/n \leq h(1/2 - \sqrt{d(1-d/n)/n}),$$

where $h(x) := -x\log x - (1-x)\log(1-x)$ is the entropy function.

1. Cayley graphs

1.1. Introduction

Interconnection networks should have a high degree of regularity, so it is only natural to look for them among Cayley graphs. Given a group G, and a generating subset S of G, let us denote by (G,S) the associated Cayley graph. Recall that (G,S) is a graph having the elements of G for vertices and that there is an edge from g to g' iff $g'=gs$ where $s\in S$. If $S=S^{-1}$, then (G,S) is an undirected graph. Note that having chosen for S a generating subset makes (G,S) a connected graph.

Cayley graphs can be particularly interesting because they have high symmetry - they are vertex-transitive - and because it is also very easy to ensure that they have optimum connectivity. For instance we have the following result:

Theorem 1.1. *If for any nontrivial subgroup H of G, we have: $|SH|-|S|\geq|H|$ and $|HS|-|S|\geq|H|$, then (G,S) has optimum connectivity (i.e., equal to the degree Δ of the graph).*

For a proof of this result, see [17]. (It is also the object of a forthcoming paper [18].) For a survey on Cayley graphs and networks see [7] and also [1].

Remark 1.2. From the purely degree and diameter point of view, one should turn to Cayley graphs on noncommutative groups. This is done in [6], where record-breaking constructions are obtained using computers and considering groups of matrices over finite fields. However, since less refined structures can also be useful (e.g. for solving problems such as routing) and since our main concern is to show how coding theory can be brought in, we will restrict our attention to the commutative case, and more precisely to the case when G is the group of binary vectors, of length r, $G=F^r$. Note that any (G,S) is undirected in this instance. Our approach will be the following: what can coding theory tell us about the graph-theoretical properties of (G,S)?

1.2. The code associated with (G,S)

In this section, we introduce the main tool coding theory can provide for the study of (G,S). Let $S=\{s_1,\ldots,s_n\}$, $0\notin S$ (so there are no loops) and $|S|=n$.

We have the following obvious:

Fact 1.3. *The degree of the graph (G,S) is:*

$$\Delta(G,S)=n=|S|.$$

Taking the elements s_i of S as column vectors we can form an $r\times n$ matrix $H=[s_1,\ldots,s_n]$. Next we use H as a parity-check matrix to define a code $C(S)$. In other words $C(S)$ is defined as the set of words orthogonal to every row of H.

Note that the parity-check matrix H thus defined puts Cayley graphs over $(Z/2Z)^r$ into one-to-one correspondence with projective codes (i.e., codes with distance at least 3, or equivalently with distinct columns in their parity-check matrix) after identifying codes which differ only by a permutation of columns of H.

$C(S)$ is therefore a linear code of length n and dimension $n-r$ (since S is a generating subset); the purpose of defining $C(S)$ is that its code-theoretic properties can tell us a great deal about the graph-theoretic properties of (G,S).

The crucial link between the two structures lies in the following fact:

Fact 1.4. *There is a one-to-one correspondence between the words of weight m of $C(S)$ and the subsets T of S such that*:

$$|T|=m \quad \text{and} \quad \sum_{s\in T} s=0.$$

To see this just associate any word w of F^n to its *support* supp(w), that is to the subset J of $[1,n]$ corresponding to the nonzero coordinates of w, then apply the definition of $C(S)$ to see that:

$$w\in C(S) \quad \text{iff} \quad \sum_{i\in \mathrm{supp}(w)} s_i=0. \tag{1'}$$

Given a code C and a vector w of F^n we need to define the (Hamming) distance between w and C; this is: $d_H(w,C)=\min_{c\in C} d_H(w,c)$.

The quantity $d_H(w,C)$ can be interpreted in terms of the set S:

Fact 1.5. $d_H(w,C)=\min\{t\colon \exists T\subset S,\ |T|=t \text{ and } \sum_{s\in T} s=\sum_{i\in \mathrm{supp}(w)} s_i\}$.

To see this, first note that:

$$d_H(w,C)=\min_{c\in C} |\mathrm{supp}(c+w)|$$

and that, by (1'):

$$c\in C \quad \text{iff} \quad \sum_{i\in \mathrm{supp}(c+w)} s_i=\sum_{i\in \mathrm{supp}(w)} s_i.$$

Next we introduce another parameter of a code C: we call the *covering radius* of a code C of length n, and denote it by $\varrho(C)$, the maximum distance between C and an arbitrary vector w of F^n:

$$\varrho(C)=\max_{w\in F^n} d(w,C).$$

Denote by $D(G,S)$ the diameter of the graph (G,S), i.e.,

$$D(G,S)=\max_{g,g'\in G} d(g,g'),$$

where $d(.,.)$ denotes here the graph distance. Since (G,S) is a Cayley graph, its

diameter is simply the maximum distance between 0 and an arbitrary element x of G, i.e.,

$$D(G,S)=\max_{x\in F^r} d(0,x).$$

Now $d(0,x)$ is, by definition of (G,S), given by:

$$d(0,x)=\min\{t:\ \exists T\subset S,\ |T|=t \text{ and } \sum_{s\in T} s=x\}.$$

So, Fact 1.5 tells us that:

$$d_H(w,C)=d\left(0,\ \sum_{i\in \mathrm{supp}(w)} s_i\right). \tag{2}$$

Since S generates $G=F^r$, $\sum_{i\in \mathrm{supp}(w)} s_i$ ranges over all F^r, when w ranges over F^n. Hence:

$$\max_{w\in F^n} d_H(w,C)=\max_{x\in F^r} d(0,x).$$

In other words, we have proved:

Fact 1.6. *The diameter of* (G,S) *equals the covering radius of* $C(S)$:

$$D(G,S)=\varrho(C(S)).$$

The covering radius of codes has been extensively studied; see for example [8]; the next sections show how this knowledge can be applied to the study of the diameter of (G,S).

1.3. Networks associated with perfect codes

Good interconnection networks should have a small diameter. For (F^r,S) to have this property, we should turn to coding theory for linear codes with a good (small) covering radius. Let us look at some examples.

Given a code C of length n, for every vector w of F^r there is a codeword c such that $d(w,c)\le\varrho(C)$. A code C is called *perfect* if the c verifying the above inequality is unique, i.e.,

$$\forall w\in F^n,\ \exists!\, c\in C \text{ such that } d_H(w,c)\le\varrho(C).$$

Obviously perfect codes have good covering properties, so do they yield interesting Cayley graphs? Unfortunately the class of perfect codes is very small: there are only three types of (binary linear) perfect codes.

(1) The Hamming codes: they correspond to the set $S=F^r\setminus\{0\}$. This gives us the rather uninteresting complete graph over 2^r vertices.

(2) The repetition codes: when r is an even integer, those correspond to a set S consisting of all the vectors (of length r) of weight 1, plus the all-one vector. The corresponding graph is the r-cube (over 2^r vertices) where an edge is added between every pair of diametrically opposed vertices. Its diameter is $r/2$.

(3) There is just another (binary) perfect code, the Golay code: its length is 23, its dimension 12, its minimal distance 7, and its covering radius is 3. This means that the associated Cayley graph is on F^{11}, with degree $\Delta = 23$ and diameter $D = 3$.

Recall the Moore bound which upperbounds the number of vertices v of a regular graph of degree Δ and diameter D:

$$v \leq M(\Delta, D) := (\Delta(\Delta-1)^D - 2)/(\Delta - 2).$$

In our case $M(23,3) = 11662$ which leaves some space for improvement since the "Golay" graph is on 2048 vertices.

On the other hand, among "commutative" Caley graphs, the Golay does very well, since we can improve the Moore bound to obtain:

Proposition 1.7. *The number of vertices $v = |G|$ of a Cayley graph (G, S) on a commutative group G, with degree $\Delta = |S|$ and diameter D is upperbounded by:*

$$|G| \leq \sum_{i=0}^{D} \binom{\Delta - 1 + i}{i} := M^*(\Delta, D).$$

Proof. To see this, count all vertices at distance i from the zero element and note that they cannot exceed $\binom{\Delta-1+i}{i}$, i.e., the number of choices of i elements, not necessarily different, among $\Delta = |S|$. □

We have $M^*(23,3) = 2600$, so the "Golay" graph is near optimal among commutative Cayley graphs. Its connectivity is, by the way, easily seen to be optimal, i.e., equal to the degree, 23.

1.4. Diameter of networks obtained from codes

Next we address the question: how good a diameter does an arbitrary Cayley graph over F^r have?

Here again coding theory gives us good bounds on $D(G, S)$. Let us sketch the way this is achieved. We shall need to state a few more facts from coding theory: denote by $k(n, d)$ the maximum dimension of a linear code of length n and Hamming distance d and by $[\cdot]$ the integer part.

The following is well known:

Proposition 1.8. $k(n, 3) = n - 1 - [\log n]$.

The next proposition is due to Godlewski [9]:

Proposition 1.9. *If C is an $[n, k, d]$ code, its covering radius ϱ satisfies:*

$$k + k(\varrho, d) \leq k(n, d).$$

Proof. Let C be an $[n,k,d]$ code with covering radius ϱ, and z be such that $d(z,C)=w(z)=\varrho$. Assume w.l.o.g. that $\mathrm{supp}(z)=\{1,2,\dots,\varrho\}$. Consider $C'[\varrho,k(\varrho,d),d]$ an optimal code built on $\mathrm{supp}(z)$ and $(C'\mid \mathbf{0})$ the $[n,k(\varrho,d),d]$ code obtained by appending $\mathbf{0}\in F^{n-\varrho}$ to all words in C'. Then for any c' in $(C'\mid \mathbf{0})$, one clearly has $d(c',C)=d(c',\mathbf{0})\geq d$ and the direct sum $C\oplus(C'\mid \mathbf{0})$ is an $[n,k+k(\varrho,d),d]$ code containing C. □

The code $C(S)$ associated with (G,S) is an $[n=|S|,n-r,3]$ linear code, so the above proposition gives us, for any connected Cayley graph (G,S) with $G=F^r$:

$$n-r+\varrho-1-[\log\varrho]\leq n-1-[\log n],$$

i.e.,

$$[\log n]\leq r-\varrho+[\log\varrho].$$

Remember that (G,S) has degree $\Delta=|S|=n$, so

$$\Delta\leq 2^{r-\varrho+[\log\varrho]+1}.$$

Hence we have the following relation (between degree Δ, diameter D and $|G|$):

Proposition 1.10. $|G|\geq \Delta 2^{D-1}/D$.

The latter gives a good general upperbound on D, that seems difficult to achieve without coding theory.

2. Asymptotic results

Due to information-theoretic limitations (see the famous work of Shannon), a substantial amount of coding theory deals with the asymptotical behavior of codes.

Most results are unfortunately nonconstructive. We shall rephrase here a few classical theorems in terms of networks. We set $k/n=R$ (the rate), $d/n=\delta$ (normalized distance), $\varrho/n=\theta$ (normalized covering radius). On $[0,1/2]$, the entropy function $h(\cdot)$ is strictly increasing and we shall consider its inverse $h^{-1}(\cdot)$. The following statements are valid for n large enough and $k=[nR]$, with R fixed (see [11, Chapter 17]).

Proposition 2.1. *There exist codes lying above the Varshamov-Gilbert bound, i.e., $[n,nR,n\delta]$ codes satisfying*

$$\delta\geq h^{-1}(1-R). \tag{3}$$

Proposition 2.2. *There exist codes $[n,nR]$ with covering radius θn satisfying*

$$\theta=h^{-1}(1-R). \tag{4}$$

Remarks. (1) One can sketch the proof of Proposition 2.1 in the following way: define a code to be maximal if it is not properly contained in a code with the same minimum distance. For a maximal code C, clearly $\varrho \leq d-1$ holds. Otherwise, pick a vector x at distance d from C, and construct $C' := C \cup \{x+C\}$, contradicting maximality. Thus the trivial covering bound:

$$|C| \sum_{i=0}^{\varrho} \binom{n}{i} \geq 2^n$$

becomes

$$\sum_{i=0}^{d-1} \binom{n}{i} \geq 2^{n(1-R)}.$$

One now obtains (3) using the following approximations for the sums of binomial coefficients (see, [13, Appendix A])

$$\sum_{i=0}^{\lambda n} \binom{n}{i} \cong 2^{nh(\lambda)}, \quad \text{for } 0 \leq \lambda \leq 1/2.$$

(2) As for Proposition 2.2, the covering bound gives in the same way $\theta \geq h^{-1}(1-R)$. The reverse inequality is obtained by "constructing" the code with a greedy algorithm.

(3) In fact both propositions are true for almost all codes (see [5]). Hence for almost all codes:

$$\theta \cong \delta \cong h^{-1}(1-R).$$

Returning to the representation of codes with Cayley graphs (by means of the parity-check matrix), we obtain graphs with

$$N = 2^{n(1-R)} \text{ vertices},$$

$$\Delta = n,$$

$$D \cong nh^{-1}(1-R).$$

This yields for example the following relations:

$$D \cong (1-R)^{-1} h^{-1}(1-R) \log N,$$

$$N = 2^{\Delta h(D/N)}.$$

3. Robustness of the n-cube

A large class of parallel algorithms, including sorting and routing, can be efficiently implemented with processors interconnected in a network such as the n-cube, the shuffle-exchange, etc. Whenever processors become faulty, one wants to estimate the efficiency of the surviving network, i.e., the subgraph induced by

nonfaulty processors and links. In the case of the n-cube, the dimension of the largest nonfaulty subcube will be the relevant parameter for most basic algorithms.

Following [3], we shall denote by $f(n,s)$ the minimum size of a set R of vertices that must be removed (faulty processors) to make any $(n-s)$-dimensional subcube K faulty (i.e., miss a vertex). That is a covering problem:

$$\text{Find } \min|R|: \quad R \cap K \neq \emptyset \text{ for all } (n-s)\text{-dimensional } K.$$

In that setting, if at most $f(n,s)-1$ processors are faulty, there exists a surviving cube of dimension $n-s$, and the "slow-down factor" (see [3]) for performing the algorithm will be 2^s. Let us rephrase the problem in $\{0,1\}$-matrix terms: $R \subset F^n$, with $|R|=r$ is the set of rows of an $r \times n$ matrix M with the following property:

For any ordered s-tuple of columns $(i_1, i_2, \ldots, i_s)$ and any binary s-tuple: $(e_1, e_2, \ldots, e_s) \in F^s$, there exists a row u of M $(m_1, m_2, \ldots, m_n)$ s.t. $m_{ui} = e_j$ for $j = 1, 2, \ldots, s$.

For example, if $n=4$, $s=2$, take $R = \{(0000), (1110), (1101), (1011), (0111)\}$. Indeed the matrix

$$M = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}$$

has the required property: any two columns (say the first and the third) contain as rows the four possible 2-tuples (say rows 1, 2, 3, 5 for 00, 11, 10, 01 respectively).

Obviously any subcube with dimension $n-s=2$, obtained by fixing $s=2$ components will intersect R.

For example, if $K = \{(x_1, x_2, x_3, x_4): x_1 = 0 \text{ and } x_2 = 1\}$, then $K \cap R = (0111)$.

This problem is equivalent to the one of s-surjectivity, occurring in universal testing of combinatorial devices (see e.g. [14,16]). The dual form of it, namely fixing r and maximizing n, has been considered in [10] under the name "k-independent families", where the authors prove the following:

$$f(n,2) = \log n + 1/2 \log\log n + O(1).$$

For s fixed and n large enough:

$$2^{s-1}(s-1)\log n \leq f(n,s) \leq 2^s \log\left(2^s \binom{n}{s}\right).$$

The upper bound is nonconstructive.

Some easy values are: $f(n,n) = 2^n$, $f(n, n-1) = 2^{n-1}$, $f(n, n-2) = [2^n/3]$ (see [14]).

Let us notice that a stronger property than s-surjectivity can be obtained directly from coding: take for R the codewords of the dual $\bar{C}$ of a code C with distance $s+1$. Then any s-tuple of columns of M has rank s (otherwise there would exist a codeword with weight s in C). Hence every binary s-tuple appears in exactly 2^{n-k-s}

rows. In that case, M is called an *orthogonal array of strength* s (see [11]). Unfortunately this is too demanding, and apart for small values of n and s or $n-s$, constructions obtained from codes are not good. Let us simply mention a construction due to Alon [2], based on Justesen codes, giving

$$r = c_s \log n,$$

where c_s is huge but does not depend on n.

Acknowledgement

We thank M. Fellows for many valuable comments on a first version of this work.

References

[1] S.B. Akers and B. Krishnamurty, A group-theoretic model for symmetric interconnection networks, IEEE Trans. Comput. 38 (1989) 555–566.
[2] N. Alon, Explicit construction of exponential size families of k-independent sets, Discrete Math. 58 (1985) 191–193.
[3] B. Becker and H.U. Simon, How robust is the n-cube? Inform. and Comput. 77 (1988) 162–178.
[4] J.C. Bermond, C. Delorme and J.J. Quisquater, Strategies for interconnection networks: some methods from graph theory, J. Parallel Distribut. Comput. 3 (1986) 433–449.
[5] V.M. Blinovskii, Bounds on efficiency of coverings by linear codes, Problems Inform. Transmission, to appear (in Russian).
[6] L. Campbell, G.E. Carlsson, M.J. Dinneen, V. Faber, M.R. Fellows, M.A. Langston, J.W. Moore, A.P. Mullhaupt and H.B. Sexton, Small diameter symmetric networks from linear groups, IEEE Trans. Comput. 40 (1992).
[7] G. Carlsson, J. Cruthirds, H. Sexton and C. Wright, Interconnection networks based on a generalization of cube-connected cycles, IEEE Trans. Comput. 34 (1985) 769–772.
[8] G.D. Cohen, M. Karpovsky, H.F. Mattson Jr and J.R. Schatz, Covering radius – survey and recent results, IEEE Trans. Inform. Theory 31 (1985) 328–344.
[9] P. Godlewski, WOM-codes construits à partir des codes de Hamming, Discrete Math. 65 (1987) 237–243.
[10] D.J. Kleitmann and J. Spencer, Families of k-independent sets, Discrete Math. 6 (1973) 255–262.
[11] F.J. MacWilliams and N.J.A. Sloane, The Theory of Error-Correcting Codes (North-Holland, Amsterdam, 1977).
[12] H.B. Mann, Addition Theorems (Wiley, New York, 1965).
[13] W.W. Peterson and E.J. Weldon Jr, Error-Correcting Codes (MIT Press, Cambridge, MA, 2nd ed., 1972).
[14] C. Roux, k-propriétés dans des tableaux de n colonnes: k-surjectivité et k-permutivité, Thèse de Doctorat, Paris VI (1987).
[15] E. Shamir and A. Schuster, Communication aspects of networks based on geometric incidence relations, Theoret. Comput. Sci. to appear.
[16] D.T. Tang and C.L. Chen, Iterative exhaustive pattern generation for logic testing, IBM J. Res. Develop. 28 (1984) 212–219.
[17] G. Zémor, Problèmes combinatoires liés à l'écriture sur des mémoires, Ph.D. Dissertation, Telecom, Paris (1989).
[18] G. Zémor, A generalisation to noncommutative groups of a theorem of Mann, Discrete Math., to appear.