Des. Codes Cryptogr.
DOI 10.1007/s10623-012-9638-x

Sphere coverings and identifying codes

David Auger · Gérard Cohen · Sihem Mesnager

Received: 8 November 2011 / Revised: 26 January 2012 / Accepted: 10 February 2012

Abstract In any connected, undirected graph $G = (V, E)$, the *distance* $d(x, y)$ between two vertices x and y of G is the minimum number of edges in a path linking x to y in G. A *sphere* in G is a set of the form $S_r(x) = \{y \in V : d(x, y) = r\}$, where x is a vertex and r is a nonnegative integer called the *radius* of the sphere. We first address in this paper the following question: What is the minimum number of spheres with fixed radius $r \geq 0$ required to cover all the vertices of a finite, connected, undirected graph G? We then turn our attention to the Hamming Hypercube of dimension n, and we show that the minimum number of spheres *with any radii* required to cover this graph is either n or $n + 1$, depending on the parity of n. We also relate the two above problems to other questions in combinatorics, in particular to identifying codes.

Keywords Sphere coverings · Identifying codes · Hamming spaces

Mathematics Subject Classification 05C70 · 05D05 · 94B25 · 94C12

This is one of several papers published in *Designs*, *Codes and Cryptography* comprising the "Special Issue on Coding Theory and Applications".

D. Auger
INRIA Saclay and University of Paris XI, Paris, France
e-mail: auger@gmx.fr

G. Cohen (✉)
Telecom-Paristech, UMR 5141, CNRS, Paris, France
e-mail: cohen@telecom-paristech.fr

S. Mesnager
Department of Mathematics, University of Paris VIII and University of Paris XIII, CNRS UMR 7539 LAGA, Paris, France
e-mail: smesnager@univ-paris8.fr

Published online: 22 March 2012

1 Introduction

We define identifying codes in a connected, undirected graph $G = (V, E)$, in which a *code* is simply a nonempty subset of vertices. These definitions can help, in various meanings, to unambiguously determine a vertex. The motivations may come from processor networks where we wish to locate a faulty vertex under certain conditions, or from the need to identify an individual, given its set of attributes. Then we turn our attention to the Hamming Hypercube of dimension n, and we show that the minimum number of spheres *with any radii* required to cover this graph is either n or $n + 1$, depending on $n \mod 2$. We also relate the two above problems to other questions in combinatorics, in particular to identifying codes. In G we define the usual distance $d(v_1, v_2)$ between two vertices $v_1, v_2 \in V$ as the smallest possible number of edges in any path between them. For an integer $r \geq 0$ and a vertex $v \in V$, we define $B_r(v)$ the *ball* (resp. $S_r(v)$ the *sphere*) of radius r centred at v, as the set of vertices within (resp. at) distance r from v. Whenever two vertices v_1 and v_2 are such that $v_1 \in B_r(v_2)$ (or, equivalently, $v_2 \in B_r(v_1)$), we say that they *r-cover* each other. Similarly, if v_1 and v_2 are such that $v_1 \in S_r(v_2)$ (or, equivalently, $v_2 \in S_r(v_1)$), we say that they *exactly r-cover* each other. A set $X \subseteq V$ (exactly) r-covers a set $Y \subseteq V$ if every vertex in Y is (exactly) r-covered by at least one vertex in X. The elements of a code $C \subseteq V$ are called *codewords*. For each vertex $v \in V$, we denote by $K_{C,r}(v) = C \cap B_r(v)$ the set of codewords r-covering v. Analogously, we denote by $X_{C,r}(v) = C \cap S_r(v)$ the set of codewords exactly r-covering v. Two vertices v_1 and v_2 with $K_{C,r}(v_1) \neq K_{C,r}(v_2)$ are said to be *r-separated* by code C, and any codeword belonging to exactly one of the two sets $B_r(v_1)$ and $B_r(v_2)$ is said to *r-separate* v_1 and v_2;

A code $C \subseteq V$ is called *r-identifying* [6] if all the sets $K_{C,r}(v)$, $v \in V$, are nonempty and distinct. In other words, every vertex is r-covered by at least one codeword, and every pair of vertices is r-separated by at least one codeword.

2 Identifying and covering by spheres

2.1 Mediating codes

It is proved in [3] (Corollary 4) that identifying codes give special coverings by spheres. In fact, a weaker property than identification, namely mediation, that we now define, will already be sufficient for that purpose. A code $C \subseteq V$ is called *r-mediating* if every vertex is r-covered by at least one codeword, but the property that $K_{C,r}(v_1)$ and $K_{C,r}(v_2)$ be distinct is only required for *adjacent* vertices v_1 and v_2. This implies in fact that for any two *adjacent* vertices v_1 and v_2, there exists a codeword c with $v_i \in S_r(c)$ and $v_j \in S_{r+1}(c)$, with $\{i, j\} = \{1, 2\}$. For $L \subset [0, n]$, define an *L-shell* by: $S_L(v) = \{x \in V : d(x, v) \in L\}$. Thus $B_r(v) = S_{[0,r]}(v)$, $S_r(v) = S_{\{r\}}(v)$.

Proposition 1 *If C is r-mediating, then $\cup_{c \in C}\{S_{\{r,r+1\}}(c)\} = V$. In words, V is covered by the $L = \{r, r + 1\}$-shells centered at codewords.*

Proof Suppose indirectly C mediating and the existence of a vertex v uncovered by such shells. Then for all $c \in C$, $d(v, c) \leq r - 1$ or $\geq r + 2$. Thus, $K_{C,r}(v) \subset B_{r-1}(v)$. Consider any v' adjacent to v; then $K_{C,r}(v) \subset B_r(v')$ by the triangle inequality and thus $K_{C,r}(v) \subset K_{C,r}(v')$. Since $K_{C,r}(v) \neq K_{C,r}(v')$ by the mediation property, there exists a $c \in C$ with $d(c, v') = r$ and $d(c, v) = r + 1$, a contradiction. □

2.2 Lower bounds for sphere coverings

A special kind of sphere covering is studied in [4], *exact r-step domination*. This corresponds to the requirement that any vertex is exactly r-covered by a *unique* codeword: $|X_{C,r}(v)| = 1$, for every $v \in V$. It is proved in [4] that every such code has size at least $\log_2 r + 1$. The proof extends in fact trivially to the relaxed case of sphere covering:

Proposition 2 *If C is a covering of V by r-spheres, then $|C| \geq \log_2 r + 1$.*

We need a few more definitions and easy facts. The *diameter* $\Delta(G) = \Delta$ of a graph G is the maximum distance between two vertices. The *radius* $\rho(G) = \rho$ is the minimum integer such that $B_\rho(v) = V$ for some $v \in V$; such a v is called a *center*. If C is r-identifying, then $r \leq \rho \leq \Delta \leq 2\rho$, with a unique center in case of equality $r = \rho$.

Consider a maximal path $\mathcal{P}$ of length Δ in G, and a codeword $c \in C$, a r-sphere covering. We show that c cannot cover too many vertices of $\mathcal{P}$ and deduce a lower bound on $|C|$.

Proposition 3 $|S_r(c) \cap \mathcal{P}| \leq 2r + 1$.

Proof Denote by $[v^1, v^{\Delta+1}]$ the vertices of $\mathcal{P}$, identified with $[1, \Delta + 1]$. Let $i \in \mathcal{P}$ be the "smallest" vertex r-covered by c, and j the "largest". Note that we do not necessarily have that $[i, j] \in \mathcal{P}$;thus $|S_r(c) \cap \mathcal{P}| \leq j - i + 1$. Since $d(c, i) = d(c, j) = r$, by the triangle inequality $d(i, j) = j - i \leq 2r$. □

Corollary 4 *A r-sphere covering C of a graph satifies: $|C| \geq \Delta/(2r + 1) \geq \rho/(2r + 1)$.*

2.3 A construction

An example of exact r-domination is given in [4] with the following parameters:

$$\Delta = 9, r = 6, |C| = 4 = 2r/3.$$

From this example, we can easily construct, for an infinite number of r's (multiples of 6), a graph inheriting an exact r-dominating code (thus a r-sphere covering) C with $|C| = 2r/3$.

3 Covering the hamming space by spheres

We now focus on the binary Hamming space of dimension n, also called the binary n-cube, which is a regular bipartite graph. We need to give some specific definitions and notation. We consider the n-cube as the set of binary row-vectors of length n, denote it by $G = (F^n, E)$ with $F = \{0, 1\}$ and $E = \{\{x, y\} : d(x, y) = 1\}$, the usual graph distance $d(x, y)$ between two vectors x and y being called here the *Hamming distance* — it simply consists of the number of coordinates where x and y differ. A sort of converse of Proposition 1 is proved in [5]

Proposition 5 *If $0 < r \leq n - 2$ and C_0 is such that $\cup_{c\in C_0}\{S_{\{r,r+1\}}(c)\} = F^n$, then $C := \cup_{c\in C_0}\{S_1(c)\}$ is r-identifying.*

The following result is proved in [1,2]:

Theorem 6 *If C is a covering by L- shells, then $|C| \geq n/|L|$.*

Note that this bound is generally weaker than the trivial *sphere-covering* bound: $|C| \geq 2^n/|S_L|$, unless L is centered around $n/2$ (in which case $|S_L| \approx 2^n$).

Corollary 7 *A r- mediating code has size at least $n/2$.*

Proof By Proposition 1, such a code is L-covering with $|L| = 2$. □

We now present a generalization of the previous theorem to the case where each codeword c^i is surrounded by its own L_i-shell (we allow multisets for codes).

For $x = (x_i)$, $y = (y_i) \in F^n$, it is easy to see that

$$d(x, y) = \sum_{i=1}^{n}(x_i + y_i - 2x_i y_i).$$

Theorem 8 *Consider $k \geq 1$ vertices $x^1, x^2, \dots, x^k$ (not necessarily distinct) of F^n and k non-negative radii $r_1, r_2, \dots, r_k$ such that*

$$F^n = \bigcup_{j=1}^{k} S_{r_j}(x^j).$$

Then $k \geq n$ if n is even, and $k \geq n + 1$ if n is odd.

Let us denote by U the set of all $y \in \{-1, 1\}^n$ and $\{1, 2, \dots n\}$ by $[n]$. A vector $y \in U$ is said to be *even* if its number of -1 is even, otherwise it is *odd*. We shall need the following (Lemma 1 from [1]).

Lemma 9 *Let $P(y_1, \dots y_n)$ be a n-multilinear function over the reals with degree strictly less than $\frac{n}{2}$, i.e.*

$$P(y_1, \dots, y_n) = \sum_X \lambda_X \prod_{i \in X} y_i$$

where the sum is taken over all subsets X of $[n]$ of size $|X| < \frac{n}{2}$. Suppose that $P(y) = 0$ for all even $y \in U$ (or similarily for all odd $y \in U$), then $P = 0$.

Proof of the theorem For $x \in F^n$, consider the vector $\bar{x} \in U$ with $\bar{x}_i = 1$ if $x_i = 0$ and $\bar{x}_i = -1$ if $x_i = 1$; thus $\bar{x}_i = 1 - 2x_i$ so for $x, y \in F^n$ we have $d(x, y) = \frac{1}{2}(n - \sum_{i=1}^{n} \bar{x}_i \bar{y}_i)$. Let us call a vertex $v \in F^n$ even if $\sum_{i=1}^{n} v_i$ is even, otherwise odd. With the previous notation, x is even if and only $\bar{x}$ is even. Now if $v, w \in F^n$ then $d(v, w)$ is even if and only if v and w have the same parity. Hence for even $v \in F^n$ we have $d(v, x^j) - r_j$ even if and only if x_j and r_j have the same parity : let us denote by J the set of $j \in \{1 \dots k\}$ with this property. We then have

$$\prod_{j \in J}(d(v, x^j) - r_j) = 0$$

for all even $v \in F^n$, and so

$$Q(y) = \prod_{j \in J}(n - 2r_j - < \bar{x}^j, y >)/2$$

vanishes over all even $y \in U$. Moreover, $Q(y) \neq 0$ if $y \in U$ is odd. Using the fact that $(-1)^2 = 1$, we can expand Q and simplify all squares of variables in the expansion of Q, to obtain a multilinar polynomial P with $P(y) = Q(y) = 0$ for all even $y \in U$, and $P(y) = Q(y) \neq 0$ for all odd $y \in U$. Using the lemma, we see that the degree of P is at least $\frac{n}{2}$: we conclude that $|J| \geq \frac{n}{2}$. The same argument holds if we consider the set K of $j \in [k]$ such that x^j and r_j do not have the same parity : we have $|K| \geq \frac{n}{2}$. Puting these facts together we see that $k \geq n$ if n is even, and $k \geq n + 1$ if n is odd. □

The bounds given in the theorem are tight : indeed, for any vertex x we have

$$F^n = \bigcup_{i=0}^{n} S_{\{i\}}(x).$$

If n is even then

$$F^n = \bigcup_{i=1}^{n-1} S_{\{i\}}(x) \cup S_{n/2}(y)$$

where y is any vertex satisfying $d(x, y) = n/2$.

Corollary 10 *Let $C = \{c^i\}$ be a covering by L_i-shells, then $\Sigma_i |L_i| \geq n$.*

4 Open problems

In the general case, we have the following extension of Corollary 10:

Conjecture Let $C = \{c^i\}$ be a covering of a graph G by L_i-shells, then $\Sigma_i |L_i| = \Omega(\rho(G))$.

Worth studying is the following specialization of identifying codes to *exact identification*: C is a covering of V by r-spheres and furthermore all the sets $X_{C,r}(v)$, $v \in V$, are nonempty and distinct.

Also, it would be interesting to narrow the gap between lower and upper bounds for coverings of graphs by r-spheres.

Acknowledgements The authors are thankful to the referees for their constructive remarks which led to a more concise and accurate version of the paper.

References

1. Alon N., Bergmann E., Coppersmith D., Odlyzko A.: Balancing sets of vectors. IEEE Trans. Inf. Theory **34**(1), 128–130 (1988).
2. Cohen G., Honkala I., Litsyn S., Lobstein A.: Covering codes. Elsevier, Amsterdam (1997).
3. Exoo G., Junnila V., Laihonen T., Ranto S.: Upper bounds for binary identifying codes. Adv. Appl. Math. **42**, 277–289 (2009).
4. Hersh P.: On exact n-step domination. Discret. Math. **205**, 235–239 (1999).
5. Honkala I., Lobstein A.: On identifying codes in binary Hamming spaces. J. Comb. Theory Ser. A **99**, 232–243 (2002).
6. Karpovsky M.G., Chakrabarty K., Levitin L.B.: On a new class of codes for identifying vertices in graphs. IEEE Trans. Inf. Theory **44**(2), 599–611 (1998).