AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **32** (2005), Pages 177–195

Possible cardinalities for identifying codes in graphs

Irène Charon   Olivier Hudry   Antoine Lobstein

CNRS & ENST
46, rue Barrault
75634 Paris Cedex 13
France
{charon, hudry, lobstein}@infres.enst.fr

Abstract

Consider a connected undirected graph $G = (V, E)$, a subset of vertices $C \subseteq V$, and an integer $r \geq 1$; for any vertex $v \in V$, let $B_r(v)$ denote the ball of radius r centered at v, i.e., the set of all vertices linked to v by a path of at most r edges. If for all vertices $v \in V$, the sets $B_r(v) \cap C$ are all nonempty and different, then we call C an r-identifying code.

It is known that the cardinality of a minimum r-identifying code in any connected undirected graph G having a given number, n, of vertices lies in the interval $[\lceil \log_2(n+1) \rceil, n-1]$, and that the values $\lceil \log_2(n+1) \rceil$ and $n-1$ can be achieved. Here, we prove that any inbetween value can also be reached.

1 Introduction

Given a connected undirected graph $G = (V, E)$ and an integer $r \geq 1$, we define $B_r(v)$, the *ball* of radius r centered at $v \in V$, by

$$B_r(v) = \{x \in V : d(x, v) \leq r\},$$

where $d(x, v)$ denotes the number of edges in any shortest path between v and x, and $S_r(v)$, the *shell* of radius r centered at $v \in V$, by

$$S_r(v) = \{x \in V : d(x, v) = r\}.$$

Whenever $d(x, v) \leq r$, we say that x and v *r-cover* each other (or simply *cover* if there is no ambiguity). A set $X \subseteq V$ covers a set $Y \subseteq V$ if every vertex in Y is covered by at least one vertex in X.

A *code* C is a nonempty set of vertices, and its elements are called *codewords*. For each vertex $v \in V$, we denote by

$$K_{C,r}(v) = C \cap B_r(v)$$

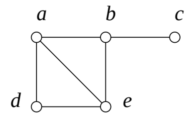

Figure 1: A graph G admitting no 1-identifying code.

the set of codewords which r-cover 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*, or *separated*, by code C. This can be expressed as

$$C \cap (B_r(v_1) \Delta B_r(v_2)) \neq \emptyset,$$

where Δ stands for the symmetric difference.

A code C is called *r-identifying*, or *identifying*, if the sets $K_{C,r}(v), v \in V$, are all nonempty and distinct [9]. It is called *r-locating-dominating*, or *locating-dominating*, if the same is true for all $v \in V \setminus C$ [7]. In other words, in the first case all vertices must be covered and pairwise separated by C; in the latter case only the noncodewords need to be covered and separated.

Remark 1. For given graph $G = (V, E)$ and integer r, there exists an r-identifying code $C \subseteq V$ if and only if

$$\forall v_1, v_2 \in V \; (v_1 \neq v_2), \; B_r(v_1) \neq B_r(v_2).$$

Indeed, if for all $v_1, v_2 \in V$, $B_r(v_1)$ and $B_r(v_2)$ are different, then $C = V$ is r-identifying. Conversely, if for some $v_1, v_2 \in V$, $B_r(v_1) = B_r(v_2)$, then for any code $C \subseteq V$, we have $K_{C,r}(v_1) = K_{C,r}(v_2)$. For instance, there is no r-identifying code in a complete graph. See also Example 1 below.

Remark 2. For given graph $G = (V, E)$ and integer r, an r-locating-dominating code always exists (simply take $C = V$), and any r-identifying code is r-locating-dominating.

Example 1. Consider the graph G in Figure 1. We see that $B_1(a) = \{a, b, d, e\}$, $B_1(b) = \{a, b, c, e\}$, $B_1(c) = \{b, c\}$, $B_1(d) = \{a, d, e\}$, $B_1(e) = \{a, b, d, e\}$; consequently, because $B_1(a) = B_1(e)$, there is no 1-identifying code in G (cf. Remark 1 above). On the other hand, $C = \{a, b\}$ is 1-locating-dominating, since the sets $K_{C,1}(c) = \{b\}$, $K_{C,1}(d) = \{a\}$, and $K_{C,1}(e) = \{a, b\}$, are all nonempty and different.

Definition 1. A graph is said to be *r-identifiable* if it admits at least one r-identifying code.

The motivations come, for instance, from fault diagnosis in multiprocessor systems. Such a system can be modeled as a graph where vertices are processors and edges are links between processors. Assume that at most one of the processors is malfunctioning and we wish to test the system and locate the faulty processor. For this purpose, some processors (constituting the code) will be selected and assigned the task of testing their neighbourhoods (i.e., the vertices at distance at most r). Whenever a selected processor (= a codeword) detects a fault, it sends an alarm signal, saying that one element in its neighbourhood is malfunctioning. We require that we can uniquely tell the location of the malfunctioning processor based only on the information which ones of the codewords gave the alarm, and in this case an identifying code is what we need.

If the selected codewords are assumed to work without failure, or if their only task is to test their neighbourhoods (i.e., they are not considered as processors anymore) and we assume that they perform this simple task without failure, then we shall search for locating-dominating codes. These codes can also be considered for modeling the protection of a building, the rooms of which are the vertices of a graph.

Locating-dominating codes were introduced in [7], identifying codes in [9], and they constitute now a topic of their own: both were studied in a large number of various papers, investigating particular graphs or families of graphs (such as planar graphs, certain infinite regular grids, or the n-cube), dealing with complexity issues, or using heuristics such as the noising methods for the construction of small codes. See, e.g., [3], [4], [5], [8], [10], and references therein, or [11].

In this paper, we concentrate on identifying codes and continue the investigation started in [6]; it is known that, for all $r \geq 1$, the cardinality of a minimum r-identifying code in any connected, undirected, r-identifiable graph G having a given number, n, of vertices lies in the interval $[\lceil \log_2(n+1) \rceil, n-1]$, and that the values $\lceil \log_2(n+1) \rceil$ and $n-1$ can be achieved, provided that n is large enough, see [6]. Here, we prove that any in-between value can also be reached. Note that allowing disconnected graphs would only make things easier.

2 Previous Results and Constructions

We first give the (trivial) lower bound.

Theorem 1 *Let $r \geq 1$ and $n \geq 1$ be two integers; let $G = (V, E)$ be a connected, undirected, r-identifiable graph with n vertices. If $C \subseteq V$ is r-identifying, then*

$$|C| \geq \lceil \log_2(n+1) \rceil.$$

The upper bound is slightly more difficult to establish.

Theorem 2 *(Theorem 2 in [6]) Let $r \geq 1$ and $n \geq 3$ be two integers; let $G = (V, E)$ be a connected, undirected, r-identifiable graph with n vertices. If $C \subseteq V$ is a minimum r-identifying code, then*

$$|C| \leq n - 1.$$

The following three theorems state that, for n large enough with respect to r, these bounds can be achieved.

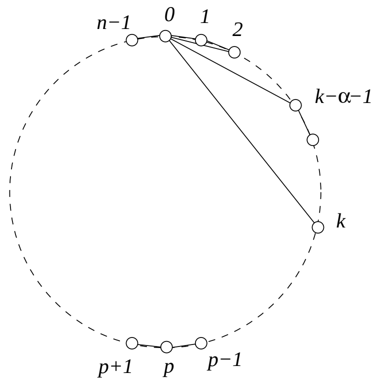

Figure 2: A partial representation of the graph $G^{(n)}_{even}$.

Theorem 3 *(Theorem 3 in [6]) Let $r \geq 1$ and n be integers such that $n \geq 2^{2r+1}$. There exists a connected graph with n vertices admitting an r-identifying code with size $\lceil \log_2(n+1) \rceil$.*

Theorem 4 *(cf. Theorem 5 in [6]) Let r be a fixed integer, $r \geq 1$. For all even n, $n \geq 3r^2$, there exists a connected graph $G^{(n)}_{even}$ with n vertices, such that any minimum r-identifying code in $G^{(n)}_{even}$ contains $n-1$ vertices.*

Theorem 5 *(cf. Theorem 6 in [6]) Let r be a fixed integer, $r \geq 1$. For all odd n, $n \geq 3r^2+1$, there exists a connected graph $G^{(n)}_{odd}$ with n vertices, such that any minimum r-identifying code in $G^{(n)}_{odd}$ contains $n-1$ vertices.*

Now what interests us in the previous two theorems is how the graphs $G^{(n)}_{even}$ and $G^{(n)}_{odd}$ are constructed when $r \geq 2$, since we shall make use of them subsequently (in Section 4).

In the even case, we write $n = 2p$ and $p = kr - \alpha$, with $1 \leq \alpha \leq r$.

Then $G^{(n)}_{even} = (V^{(n)}_{even}, E^{(n)}_{even})$ is the following graph (cf. Figure 2):

$$V^{(n)}_{even} = \{0, 1, \ldots, n-1\},$$

$$E^{(n)}_{even} = \{\{i, i+j \bmod n\} : i \in V^{(n)}_{even}, j \in \{1, 2, \ldots, k-\alpha-1, k\}\}.$$

Actually, in view of Section 4.2, all the reader needs to remember about $G^{(n)}_{even}$ is the following:

- Any subset of $V^{(n)}_{even}$ with $n-1$ elements is a minimum r-identifying code in $G^{(n)}_{even}$.
- For all $i \in V^{(n)}_{even}$,
  - there is a symmetry with respect to the diameter $[i, i+p \bmod n]$.
  - $B_r(i) = V^{(n)}_{even} \setminus \{i+p \bmod n\} \neq B_{r-1}(i)$.
  - $S_{r+1}(i) = \{i+p \bmod n\}$.

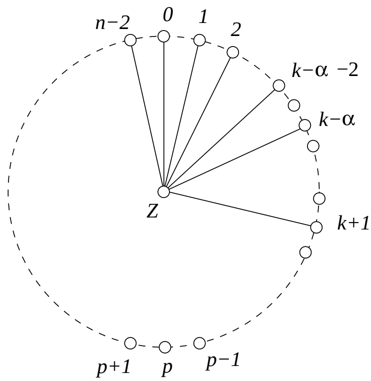

Figure 3: A partial representation of the graph $G^{(n)}_{odd}$.

– For any j between 0 and r, $S_j(i) \neq \emptyset$.

When n is odd, we add one vertex, Z, to $G^{(n-1)}_{even}$, we delete some edges in $E^{(n-1)}_{even}$, and add some edges between Z and vertices in $V^{(n-1)}_{even}$ (cf. Figure 3), but again, all the reader needs to remember about $G^{(n)}_{odd}$ is the following:

- $V^{(n)}_{odd} \setminus \{Z\}$ is the only minimum r-identifying code in $G^{(n)}_{odd}$.
- There is a symmetry with respect to the diameter $[0, p]$.
- There is an edge between 0 and Z.
- $B_r(Z) = V^{(n)}_{odd}$.
- For all $i \in V^{(n)}_{odd} \setminus \{Z\}$,
  - $B_r(i) = V^{(n)}_{odd} \setminus \{i + p \bmod n\} \neq B_{r-1}(i)$.
  - $S_{r+1}(i) = \{i + p \bmod n\}$.
  - For any j between 0 and r, $S_j(i) \neq \emptyset$.

In both cases, odd and even, we see that "the circle" ($V^{(n)}_{odd} \setminus \{Z\}$ or $V^{(n)}_{even}$) has an even number of points, and that any point i lying on it r-covers all vertices but its diametrical opposite $\text{opp}(i) = i + p \bmod n$, and $(r+1)$-covers $\text{opp}(i)$. As a consequence, for i and j on the circle, $i \neq j$, $B_r(i) \Delta B_r(j) = \{\text{opp}(i), \text{opp}(j)\}$, and $B_r(i) \Delta B_r(Z) = \{\text{opp}(i)\}$.

3 The Case $r = 1$

We treat separately the case $r = 1$, which is easier, and show, in this section, that any value between $n - 1$ and $\lceil \log_2(n+1) \rceil$ can be reached by the size of a minimum 1-identifying code, for a well-chosen connected graph with n vertices. Moreover, in passing, we treat the case of trees and of bipartite graphs.

First, we recall an easy result, with its short proof.

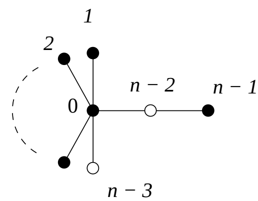

Figure 4: The tree G of Theorem 7, with a minimum 1-identifying code. Codewords are in black.

Theorem 6 *For all $n \geq 3$, there exists a tree G with n vertices, such that any minimum 1-identifying code in G contains $n-1$ vertices.*

Proof. Consider the "star", i.e., the tree consisting of n vertices 0, 1, ..., $n-1$, and $n-1$ edges $\{0,i\}$, $1 \leq i \leq n-1$. It is immediate to check that taking for codewords any set of $n-1$ vertices is necessary and sufficient to obtain a 1-identifying code, except for $n=3$, where only $\{1,2\}$ is 1-identifying. △

Next, we divide the interval from $n-2$ to $\lceil \log_2(n+1) \rceil$ into three (possibly overlappping) parts: $n-2$ (Theorem 7), from $n-3$ to $\lceil 3(n+1)/7 \rceil$ (Theorem 8), and from $\lfloor n/2 \rfloor$ to $\lceil \log_2(n+1) \rceil$ (Theorem 9).

The tree described in the following theorem is represented in Figure 4.

Theorem 7 *Consider the tree $G=(V,E)$ defined by*

$$V = \{0,1,2,\ldots,n-3,n-2,n-1\},$$

$$E = \{\{0,i\} : 1 \leq i \leq n-2\} \cup \{\{n-2,n-1\}\},$$

with $n = |V| \geq 5$. Any minimum 1-identifying code in G contains $n-2$ elements.

Proof. Let C be a 1-identifying code in G. Because $n-2$ and $n-1$ must be 1-separated, $0 \in C$. Then at least $n-4$ of the $n-3$ vertices 1, 2, ..., $n-4$, $n-3$, must be taken in C, and one more codeword is necessary, to 1-cover $n-1$. On the other hand, $\{0,1,\ldots,n-4,n-1\}$ is a possible 1-identifying code. △

Theorem 8 *For $n \geq 6$, for any integer c between $n-3$ and $\lceil 3(n+1)/7 \rceil$, there exists a tree G with n vertices in which any minimum 1-identifying code has size c.*

Proof. Our tree G is constructed using one tree with minimum 1-identifying code containing $3/7$ of the vertices, and one tree where all but between three and six vertices are necessary. Varying the respective sizes of these two subgraphs, we can attain all values between $n-3$ and $\lceil 3(n+1)/7 \rceil$.

Let $G_i = (V_i, E_i)$, $i \geq 1$, be defined as follows:

$$V_i = \{v_{i,1}, v_{i,2}, v_{i,3}, v_{i,4}, v_{i,5}, v_{i,6}, v_{i,7}\},$$

$$E_i = \{\{v_{i,1}, v_{i,2}\}, \{v_{i,3}, v_{i,4}\}, \{v_{i,5}, v_{i,6}\}, \{v_{i,1}, v_{i,3}\}, \{v_{i,3}, v_{i,5}\}, \{v_{i,5}, v_{i,7}\}\},$$

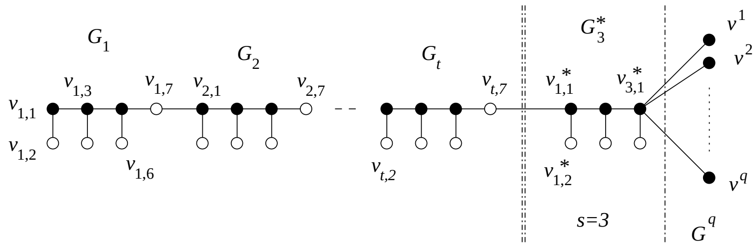

Figure 5: The tree G of Theorem 8, for $s = 3$, with a minimum 1-identifying code. Codewords are in black.

cf. Figure 5. Next, consider, for $s \in \{3, 4, 5, 6\}$, the tree $G_s^* = (V_s^*, E_s^*)$, with:

$$V_s^* = \{v_{i,1}^*, v_{i,2}^* : 1 \leq i \leq s\},$$

$$E_s^* = \{\{v_{i,1}^*, v_{i+1,1}^*\} : 1 \leq i \leq s-1\} \cup \{\{v_{i,1}^*, v_{i,2}^*\} : 1 \leq i \leq s\}$$

(cf. Figure 5), and the graph $G^q = (V^q, E^q)$, with q disconnected vertices ($q \geq 0$): $V^q = \{v^1, v^2, \ldots, v^q\}$, $E^q = \emptyset$.

Finally, for $t \geq 0$, $G_{t,s,q} = (V_{t,s,q}, E_{t,s,q})$ is given by

$$V_{t,s,q} = (\cup_{i=1}^t V_i) \cup V_s^* \cup V^q,$$

$$E_{t,s,q} = \cup_{i=1}^t E_i \cup (\cup_{i=1}^{t-1} \{\{v_{i,7}, v_{i+1,1}\}\}) \cup \{\{v_{t,7}, v_{1,1}^*\}\} \cup E_s^* \cup \ (\cup_{i=1}^q \{\{v_{s,1}^*, v^i\}\})$$

(see Figure 5). The tree $G_{t,s,q}$ contains $n = 7t + 2s + q$ vertices, and a 1-identifying code in $G_{t,s,q}$ requires at least $3t + s + q$ codewords: first, each G_i must obviously contain at least three codewords, then $G_s^* \setminus \{v_{s,1}^*, v_{s,2}^*\}$ clearly contains at least $s - 1$ codewords, and finally, the set $\{v_{s,1}^*, v_{s,2}^*\} \cup V^q$, which is a star, needs at least $q + 1$ codewords; on the other hand,

$$C = (\cup_{i=1}^t \{v_{i,1}, v_{i,3}, v_{i,5}\}) \cup (\cup_{j=1}^s \{v_{j,1}^*\}) \cup V^q$$

is 1-identifying and has size $3t + s + q$ (this might be untrue for $s < 3$).

Now, for fixed $n \geq 6$ and c verifying $n - 3 \geq c \geq \lceil 3(n + 1)/7 \rceil$, we want to find integers t, s, q such that:

$$t \geq 0, q \geq 0, 6 \geq s \geq 3,$$

$$n = 7t + 2s + q, \tag{1}$$

$$c = 3t + s + q. \tag{2}$$

Combining (1) and (2), we have:

$$t = \frac{n - c - s}{4}. \tag{3}$$

First we choose s in $\{3, 4, 5, 6\}$ such that t is an integer; then q is an integer, and all that remains to be proved is that both t and q are nonnegative.

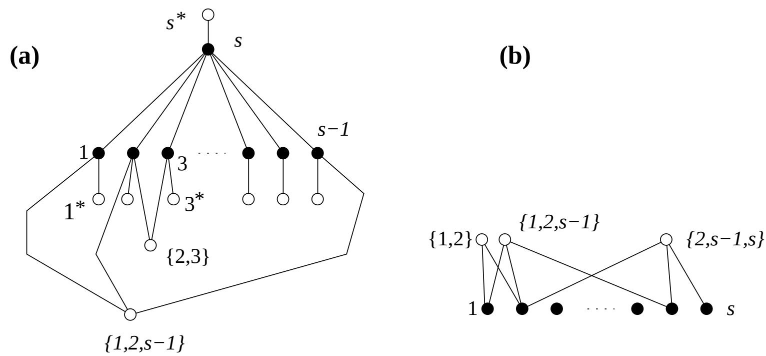

Figure 6: (a) The bipartite graph $G_{s,q}$, with $n = 2s + 2$. (b) A partial representation of a bipartite graph when $n > 2^{s-1} + s$. In both cases, a minimum 1-identifying code of size s is given, with codewords in black.

By (2) and (3), $q \leq -1$ implies that $c \leq (3n + s - 4)/7$. Since $c \geq (3n + 3)/7$ and $s \leq 6$, this is impossible.

By (3), $t \leq -1$ implies that $n - c \leq s - 4$. Since $c \leq n - 3$ and $s \leq 6$, this is not possible either.

Therefore, we see that for any couple (n, c) with $n \geq 6$, $n - 3 \geq c \geq \lceil 3(n+1)/7 \rceil$, we can find integers $t \geq 0$, $s \in \{3, 4, 5, 6\}$, $q \geq 0$, such that (1) and (2) hold, i.e., we can construct a tree with n vertices admitting a 1-identifying code of size c. △

Note that in the construction of the previous theorem, if we allow s to be equal to one or two, with $t = 0$, we find the constructions of Theorems 6 and 7, respectively (with $q \geq 1$).

Theorem 9 *For $n \geq 6$, for any integer c between $\lfloor n/2 \rfloor$ and $\lceil \log_2(n+1) \rceil$, there exists a connected bipartite graph G with n vertices in which any minimum 1-identifying code has size c.*

Proof. First, we consider the following graph $G_s = (V_s, E_s)$, with $s \geq 3$:

$$V_s = \{1, 2, \ldots, s-1, s, 1^*, 2^*, \ldots, s^*\},$$

$$E_s = \{\{i, i^*\} : 1 \leq i \leq s\} \cup \{\{i, s\} : 1 \leq i \leq s-1\},$$

cf. Figure 6(a). Because of the vertices of type i^*, any 1-identifying code in G_s has size at least s, and $C = \{1, 2, \ldots, s\}$ is suitable.

The construction of G_s can be seen as follows: first, a set of s vertices, from 1 to s; then, the addition of s vertices, associated to the s one-element subsets of $\{1, 2, \ldots, s\}$, each of these vertices being linked to one and only one vertex in $\{1, 2, \ldots, s\}$.

This gives the idea of adding vertices and edges to G_s: let q be an integer lying between 0 and $2^{s-1} - (s-1) - 1$. To G_s we add q vertices, associated to q distinct, nonempty, non-singleton, subsets of $\{1, 2, \ldots, s-1\}$, and each of these vertices is linked in a very natural way to G_s, by linking it to the elements in $\{1, 2, \ldots, s-1\}$ which belong to the subset it is associated to. There are precisely $2^{s-1}-s$ possibilities. The graph $G_{s,q}$ thus constructed is bipartite and has $n = 2s+q$ vertices ($6 \leq 2s \leq n \leq 2^{s-1}+s$), see Figure 6(a). Its very construction shows that C is still a minimum 1-identifying code in $G_{s,q}$. Therefore, $|C| \leq n/2$ and $|C| \geq \lceil \log_2(n-|C|)\rceil + 1$, and all in-between integer values can be achieved, according to the values taken by $s \geq 3$ and q.

The only remaining case is when $2^s - 1 \geq n > 2^{s-1} + s$. In this case, we can exploit the construction illustrated by Figure 6(b), still with the same idea as before: s vertices $1, \ldots, s$, and the use of q appropriate subsets (distinct, nonempty, non-singleton) of $\{1, 2, \ldots, s\}$, with $n = q+s$, $2^{s-1}+1 \leq q \leq 2^s - 1 - s$. The graph will be connected (because q is large), bipartite, and, since $\lceil \log_2(n+1)\rceil \geq s$, Theorem 1 shows that any 1-identifying code contains at least s elements, and clearly $C = \{1, 2, \ldots, s\}$ suits.

Therefore, for any couple (n, c) with $\lfloor n/2 \rfloor \geq c \geq \lceil \log_2(n+1)\rceil$, by taking $s = c$ and accordingly $q = n-2c$ or $q = n-c$, we can construct a connected bipartite graph with n vertices admitting a 1-identifying code of size c. △

The cases $3 \leq n \leq 5$ can be studied easily; since trees are bipartite graphs, the following theorem is the consequence of Theorems 6–9.

Theorem 10 *For $n \geq 3$, for any integer c between $n-1$ and $\lceil \log_2(n+1)\rceil$, there exists a connected bipartite graph with n vertices in which any minimum 1-identifying code has size c.*

The following theorem, the consequence of Theorems 6–8, cannot be improved with respect to c, since it is known (see [1] or [2]) that in a tree with n vertices, a 1-identifying code contains at least $\lceil 3(n+1)/7 \rceil$ elements.

Theorem 11 *For $n \geq 3$, for any integer c between $n-1$ and $\lceil 3(n+1)/7 \rceil$, there exists a tree with n vertices in which any minimum 1-identifying code has size c.*

4 Achieving All Values

In this section we prove our main result: given an integer $r \geq 2$ and an integer n sufficiently large with respect to r, for any integer c between $n-2$ and $\lceil \log_2(n+1)\rceil + 1$, there exists a connected graph with n vertices admitting a minimum r-identifying code of size c. Or the other way round: given an integer $r \geq 2$ and an integer c sufficiently large with respect to r, for any integer n between $c+2$ and $2^{c-1}-1$, there exists a connected graph with n vertices admitting a minimum r-identifying code of size c.

Therefore, we fix $r \geq 2$ and c (how large c must be will be specified later; let us only say here that c is of order r^2). We will proceed in several steps: first from $c+2$ to $c+r+1$, second from $c+r+2$ to $c+2r^2+4r+1$, third from $c+2r^2+4r+2$ to $2^{c-3r^2}+3r^2+r$, and finally from $(r+1)c+1$ to $2^{c-1}-1$ (Lemma 4(ii) will guarantee that, for c large enough, there is no gap between the last two intervals).

Our idea is the following. We start with a "dense" graph (needing as many as (#vertices −1) codewords) and, in order to gradually diminish the density, we combine it first with chains of different lengths, then with a "sparse" graph (needing down to $\lceil \log_2(\#\text{vertices} + 1)\rceil$ codewords), varying their respective sizes. The final step uses the sparse graph only. Small technical problems arise because a dense graph has size at least $3r^2$, therefore we cannot act on its size as freely as we would like to.

4.1 Five Easy Lemmas

Let $V_q = V \cup W_q$, with $V = \{v_0, v_1, \ldots, v_r\}, W_q = \{w_1, w_2, \ldots, w_q\}$, $1 \le q$, and let $G_q = (V_q, E_q)$ be a chain:

$$E_q = \{\{v_0, v_1\}, \{v_1, v_2\}, \ldots, \{v_{r-1}, v_r\}, \{v_r, w_1\}, \{w_1, w_2\}, \ldots, \{w_{q-1}, w_q\}\}.$$

We say that a code $C_q \subseteq V_q$ is V-semi-identifying in G_q if it contains V and r-covers and pairwise r-separates all vertices in W_q; we denote by $\gamma_q + r + 1$ the minimum size of such a code ($\gamma_q \ge 0$).

Lemma 1

$$(i)\ \textit{For } 1 \le q \le r, \quad \gamma_q = 0. \tag{4}$$

$$(ii)\ \textit{For } r+1 \le q \le 2r, \quad \gamma_q \le q - r. \tag{5}$$

$$(iii)\ \textit{For } 2r+1 \le q, \quad \gamma_q \le \left\lfloor \frac{q-2r-1}{2} \right\rfloor + 2r + 1. \tag{6}$$

$$(iv)\ \textit{For } 1 \le q, \quad \gamma_q \le r + \frac{q+1}{2}. \tag{7}$$

Proof. Case (i) is trivial and (iv) is a consequence of (i)–(iii).

In case (ii), $C_q = V \cup \{w_i : 1 \le i \le q - r\}$ is V-semi-identifying in G_q.

In case (iii), $C_q = V \cup \{w_i : i \text{ even }, 1 \le i \le q - 2r - 1\} \cup \{w_i : q - 2r \le i \le q\}$ is V-semi-identifying in G_q. Note that, since inequality (7) meets our subsequent requirements, we did not try here to optimize C_q. △

Lemma 2

$$\textit{For any } q \ge 1, \quad 0 \le \gamma_{q+1} - \gamma_q \le 1. \tag{8}$$

Proof. 1) $\gamma_q \le \gamma_{q+1}$: by Lemma 1(i), we can assume that $q \ge r+1$. Let C_{q+1} be a minimum V-semi-identifying code in G_{q+1}: $|C_{q+1}| = \gamma_{q+1}$. Because w_{q+1} and w_q are separated by C_{q+1}, necessarily $w_{q-r} \in C_{q+1}$.

If $w_{q+1} \notin C_{q+1}$, then $C_{q+1} \subseteq V_q$ is V-semi-identifying in G_q, and $\gamma_q \le \gamma_{q+1}$.

If $w_{q+1} \in C_{q+1}$, then we consider $C_q = C_{q+1} \setminus \{w_{q+1}\}$; two things can prevent C_q from being V-semi-identifying in G_q:

– either some vertex in W_q which was covered by w_{q+1} is not covered by C_q; this is not possible, since $w_{q-r} \in C_q$;

– or w_{q-r} and w_{q-r+1}, which were separated by w_{q+1}, are not separated by any element in C_q. But if $r+1 \le q \le 2r$, then w_{q-r} and w_{q-r+1} are separated by a codeword in V; so necessarily $2r+1 \leq q$, in which case $C_q \cup \{w_{q-2r}\}$ separates w_{q-r} and w_{q-r+1}, is V-semi-identifying in G_q, and has size γ_{q+1}.

2) $\gamma_{q+1} \leq \gamma_q + 1$: by Lemma 1(i), we can assume that $r \leq q$. Let C_q be a minimum V-semi-identifying code in G_q. If $q = r$, then w_q and w_{q+1} are separated by v_r (which implies that w_{q+1} is separated from all vertices by C_q), and with at most one additional codeword covering w_{q+1}, we obtain a V-semi-identifying code in G_{q+1}. So we assume that $r+1 \leq q$.

If $w_{q-r} \in C_q$, then w_q and w_{q+1} are separated by C_q, and with at most one additional codeword covering w_{q+1}, we obtain a V-semi-identifying code in G_{q+1}. If $w_{q-r} \notin C_q$, then the codeword which covers w_q also covers w_{q+1}, and by taking w_{q-r} as a codeword, we obtain a V-semi-identifying code in G_{q+1}. △

Lemma 3 *The set* $\{q - \gamma_q : q \geq 1\}$ *is equal to* $\mathbb{N}^*$.

Proof. By (8), for $q \geq 1$, $(q+1-\gamma_{q+1}) - (q - \gamma_q) = 0$ or 1; by (4), $1 - \gamma_1 = 1$; by (7), $q - \gamma_q \geq (q-1)/2 - r$. Therefore the sequence $(q - \gamma_q)$ starts at one, possibly increases by one only, and is bounded from below by a quantity going to infinity with q. △

Lemma 4 *Let* $r \geq 2$.

$$(i) \text{ For } x \geq 2r+3, \quad 2^x \geq (x+1)(r+1). \tag{9}$$

$$(ii) \text{ For } x \geq 5r^2+5r+1, \quad 2^{x-3r^2} + 3r^2 + r \geq x(r+1). \tag{10}$$

Proof. (i) Let $f(x) = 2^x - (x+1)(r+1)$; $f'(x) = 2^x \ln 2 - (r+1) \geq 0$ when $x \geq 2r+3, r \geq 2$, and $f(2r+3) = 2^{2r+3} - (2r+4)(r+1)$ is nonnegative for $r \geq 2$.

(ii) Let $g(x) = 2^{x-3r^2} + 3r^2 + r - x(r+1)$; similarly, $g'(x) \geq 0$ when $x \geq 5r^2+5r+1, r \geq 2$, and $g(5r^2+5r+1)$ is nonnegative for $r \geq 2$. △

Let $s \geq 2r+3$ and $G(s) = (V(s), E(s))$ be the following graph, with $(r+1)s+1$ vertices (cf. Figure 7):

$$V(s) = \{x_{i,j} : 1 \leq i \leq s, 1 \leq j \leq r\} \cup X(s) \text{ with } X(s) = \{z_i : 1 \leq i \leq s+1\},$$

$$E(s) = \{\{x_{i,1}, x_{i+1,1}\} : 1 \leq i \leq s-1\} \cup \{\{x_{s,1}, x_{1,1}\}\} \cup$$
$$\{\{x_{i,j}, x_{i,j+1}\} : 1 \leq i \leq s, 1 \leq j \leq r-1\} \cup \{\{z_i, z_j\} : 1 \leq i < j \leq s+1\} \cup$$
$$\{\{z_i, x_{j,r}\} : 1 \leq i \leq s, 1 \leq j \leq s, i \neq j\} \cup \{\{z_{s+1}, x_{j,r}\} : 1 \leq j \leq s\}.$$

Lemma 5 *For* $r \geq 2$, $s \geq 2r+3$, *the only minimum* r*-identifying code in* $G(s)$ *is* $C = \{x_{i,1} : 1 \leq i \leq s\}$.

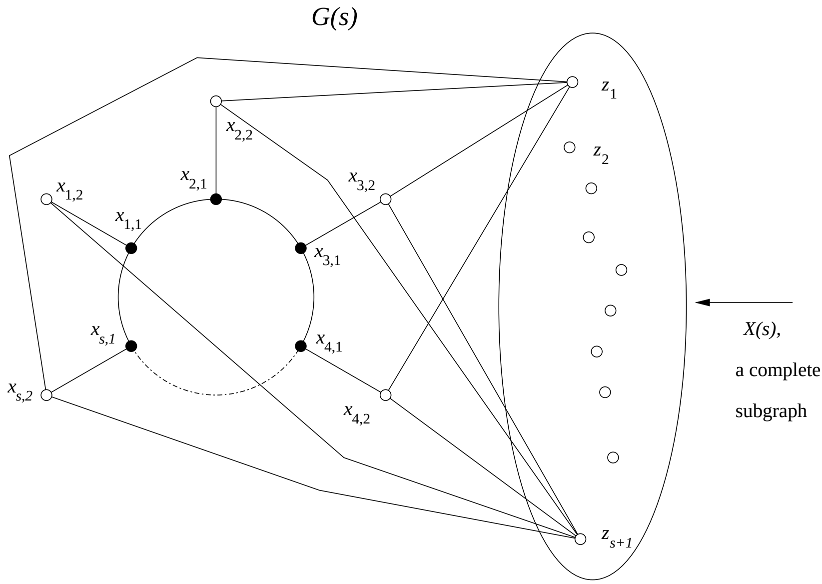

Figure 7: The case $r = 2$: a partial representation of $G(s)$.

Proof. For $1 \leq i \leq s$, z_i is r-covered by all vertices but $x_{i,1}$, and z_{s+1} is covered by all vertices, so $B_r(z_i)\Delta B_r(z_{s+1}) = \{x_{i,1}\}$. Therefore, any r-identifying code in $G(s)$ contains $\{x_{i,1} : 1 \leq i \leq s\}$. On the other hand, it is easy to check that this set is suitable: for a given j, the vertices $x_{i,j}$ are covered by different sets of codewords of size $1 + 2(r - j + 1) \leq 2r + 1$; for $1 \leq i \leq s$, the vertices z_i are covered by different sets of codewords of size $s - 1 \geq 2r + 2$, and z_{s+1} is covered by all s codewords. $\triangle$

Remark 3. Still with the same identifying code $C = \{x_{i,1} : 1 \leq i \leq s\}$, we can have more vertices and construct a graph $G_{s,q}$ with up to $2^s - 1$ vertices (cf. the proof of Theorem 9): let q be an integer between 0 and $2^s - 1 - |V(s)| = 2^s - (r + 1)s - 2$. We set $X(s, q) = X(s) \cup \{w_j : 1 \leq j \leq q\}$ and construct new edges in the following way: first, $X(s, q)$ is a complete subgraph; next, to all vertices w_j we associate distinct subsets of C, $K(w_j)$, which are nonempty and not equal to any of the sets $K_{C,r}(y), y \in V(s)$. There are exactly $2^s - 1 - |V(s)|$ such subsets available. Then every vertex w_j is linked to the vertex $x_{i,r}$ whenever $x_{i,1} \in K(w_j)$: obviously, the set of codewords r-covering w_j is $K(w_j)$, i.e., $K(w_j) = K_{C,r}(w_j)$, and we see that, by construction, C is still the only minimum r-identifying code in the new graph $G_{s,q}$.

4.2 From $c + 2$ to $c + r + 1$

Let $n_0 = c+1$ and $G^{(n_0)} = (V^{(n_0)}, E^{(n_0)})$ be either $G^{(n_0)}_{even}$ or $G^{(n_0)}_{odd}$, and let $\mathcal{V}^{(n_0)} = V^{(n_0)}$ or $V^{(n_0)} \setminus \{Z\}$, according to the parity of n_0 (cf. Section 2); in both cases, $\mathcal{V}^{(n_0)}$ is the set of vertices of an even cycle, and for a set $X \subseteq V^{(n_0)}$, we denote by opp(X)

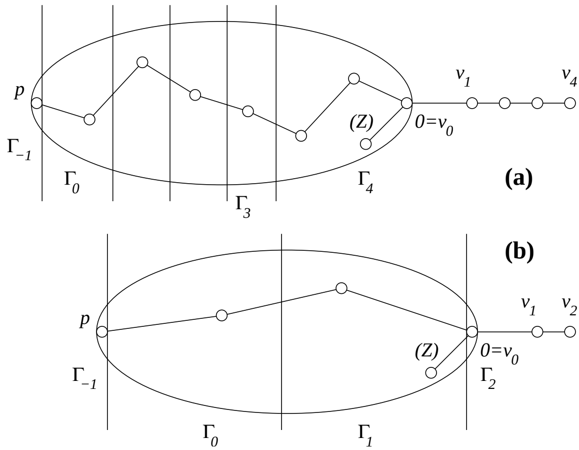

Figure 8: A representation of $G^{(n_0)}$ based on the distances to the vertices v_i: (a) $r = 6, \ell = 4$ ($\Gamma_{-1} = \{p\}$); (b) $r = \ell = 2$ ($\Gamma_{-1} = \{p\}, \Gamma_2 = \{0\}$).

the set of points which are the diametrical opposites of the points in X, with the convention that if $Z \in X$, then $Z \in \mathrm{opp}(X)$.

Let $\ell \in \{1, 2, \ldots, r\}$; to $V^{(n_0)}$ we add ℓ vertices, $v_1, v_2, \ldots, v_\ell$, and construct a graph $G_{n_0+\ell} = (V_{n_0+\ell}, E_{n_0+\ell})$ in the following way:

$$V_{n_0+\ell} = V^{(n_0)} \cup \{v_1, v_2, \ldots, v_\ell\},$$

$$E_{n_0+\ell} = E^{(n_0)} \cup \{\{0, v_1\}, \{v_1, v_2\}, \ldots, \{v_{\ell-1}, v_\ell\}\}.$$

For notational reasons, we set $v_0 = 0$. We now prove that in $G_{n_0+\ell}$, which contains $n = n_0 + \ell$ vertices, with $c+2 = n_0+1 \leq n \leq n_0+r = c+r+1$, there is a minimum r-identifying code of size c, provided that $c \geq 3r^2$ (so that n_0 fulfills the assumptions in Theorems 4 and 5).

We need further notation (see Figure 8 for an illustration). For i between 0 and $\ell - 1$, let $\Gamma_i = S_r(v_i) \cap \mathcal{V}^{(n_0)}$; let $\Gamma_{-1} = \{p\}$ (we remind the reader that the vertex $p = \lfloor n/2 \rfloor$ is the diametrical opposite of the vertex 0), and $\Gamma_\ell = B_r(v_\ell) \cap \mathcal{V}^{(n_0)}$. When n_0 is odd, we add Z to Γ_ℓ if $\ell < r$, and to $\Gamma_{\ell-1}$ if $\ell = r$ (because there is an edge between 0 and Z, we choose to put Z in a set where there are vertices at distance one from 0). Note that all points in Γ_i, $-1 \leq i \leq \ell - 1$, are at distance $r - i$ from 0, and all points in Γ_ℓ are within distance $r - \ell$ from 0. This shows that the $\ell + 2$ sets Γ_i, $-1 \leq i \leq \ell$, partition the set $V^{(n_0)}$; the sets $\mathrm{opp}(\Gamma_i)$ also partition $V^{(n_0)}$, but because of Z, they might be less regular with respect to distances.

Let C be an r-identifying code in $G_{n_0+\ell}$. We want to prove that necessarily $|C| \geq c$.

Lemma 6 *For any i between -1 and ℓ, there is at most one noncodeword in opp(Γ_i).*

Proof. Assume on the contrary that there is an i such that two elements in $\mathrm{opp}(\Gamma_i)$, a_i and b_i, do not belong to C.

a) Neither of them is Z; since $\mathrm{opp}(a_i)$, $\mathrm{opp}(b_i) \in \Gamma_i$, either $\mathrm{opp}(a_i)$ and $\mathrm{opp}(b_i)$ are at the same distance from 0 (hence, they are at the same distances from v_1, $\ldots$, v_ℓ), or they are within distance r from v_ℓ; therefore, the symmetric difference of $B_r(\mathrm{opp}(a_i))$ and $B_r(\mathrm{opp}(b_i))$ is equal to $\{a_i, b_i\}$, its intersection with C is empty, and so $\mathrm{opp}(a_i)$ and $\mathrm{opp}(b_i)$ are not r-separated by C, a contradiction.

b) One of them, say a_i, is equal to Z; because both Z and $\mathrm{opp}(b_i)$ are either at the same distance from 0 or within distance r from v_ℓ, the symmetric difference of $B_r(\mathrm{opp}(b_i))$ and $B_r(Z)$ is equal to $\{b_i\}$, which again shows that Z and $\mathrm{opp}(b_i)$ are not r-separated by C. $\triangle$

Lemma 7 *For any i, j with $-1 \leq i < j \leq \ell$, if there is one noncodeword in opp(Γ_i) and one noncodeword in opp(Γ_j), then there is at least one codeword among $v_{i+1}, \ldots, v_j$.*

Proof. Assume that there exists i and j ($-1 \leq i < j \leq \ell$) such that one vertex a_i in $\mathrm{opp}(\Gamma_i)$ and one vertex b_j in $\mathrm{opp}(\Gamma_j)$ do not belong to C.

a) Neither of them is Z; the symmetric difference of $B_r(\mathrm{opp}(a_i))$ and $B_r(\mathrm{opp}(b_j))$ is equal to $\{a_i, b_j, v_{i+1}, \ldots, v_j\}$, and so at least one of $v_{i+1}, \ldots, v_j$ must be a codeword.

b) $a_i = Z$; since $Z \in \Gamma_\ell$ if $\ell < r$, or $Z \in \Gamma_{\ell-1}$ if $\ell = r$, and $i < j \leq \ell$, necessarily $\ell = r$, $i = \ell - 1$ and $j = \ell$. Then the symmetric difference of $B_r(\mathrm{opp}(b_j))$ and $B_r(Z)$ is equal to $\{b_\ell, v_\ell\}$, and v_ℓ must be a codeword.

c) $b_j = Z$; the symmetric difference of $B_r(\mathrm{opp}(a_i))$ and $B_r(Z)$ is equal to $\{a_i, v_{i+1}, \ldots, v_{\ell-1}\}$ if $j = \ell - 1$ (when $\ell = r$), or to $\{a_i, v_{i+1}, \ldots, v_\ell\}$ if $j = \ell$ (when $\ell < r$). In both cases, at least one vertex of type v must be a codeword. $\triangle$

Now let s be the number of noncodewords in $V^{(n_0)}$. By Lemma 6, and because the sets $\mathrm{opp}(\Gamma_i)$ partition $V^{(n_0)}$, $s \leq \ell + 2$. If $s = 0$ or 1, then $|C| \geq n_0 - 1 = c$ and we are done. So we can assume that $2 \leq s \leq \ell + 2$.

Also, there are s indices, $i_1, \ldots, i_s$ ($-1 \leq i_1 < \ldots < i_s \leq \ell$) such that each of $\mathrm{opp}(\Gamma_{i_1})$, $\ldots$, $\mathrm{opp}(\Gamma_{i_s})$ contains exactly one noncodeword.

Then, by Lemma 7, the $s - 1$ sets $\{v_{i_1+1}, \ldots, v_{i_2}\}$, $\{v_{i_2+1}, \ldots, v_{i_3}\}$, $\ldots$, $\{v_{i_{s-1}+1}, \ldots, v_{i_s}\}$ each contain at least one codeword, which would lead to at least $(n_0 - s) + (s - 1) = n_0 - 1 = c$ codewords. The last point consists in showing that these $s - 1$ codewords of type v belong to $v_1, \ldots, v_\ell$ and do not interfere with $v_0 = 0$, which is a vertex in $V^{(n_0)}$. If 0 is not a codeword, we are done.

If 0 is a codeword, then, since $\mathrm{opp}(\Gamma_{-1}) = \{0\}$, there is no noncodeword in $\mathrm{opp}(\Gamma_{-1})$, and $i_1 \neq -1$, i.e., $v_{i_1+1} \neq 0$.

We can therefore conclude that C contains at least c codewords.

On the other hand, in $G_{n_0+\ell}$ we can construct an r-identifying code C with exactly c elements in the following way: as codewords, we take all elements v_i, $1 \leq i \leq \ell$, and in $\mathcal{V}^{(n_0)}$ all vertices but $\ell + 1$, choosing exactly one noncodeword in each of the $\ell + 1$ sets $\mathrm{opp}(\Gamma_i)$, $0 \leq i \leq \ell$ (therefore, $0 \in C$). We leave the checking to the reader. Therefore the following theorem has been proved.

Theorem 12 *Let $r \geq 2$ and $c \geq 3r^2$. For n between $c+2$ and $c+r+1$, there exists a connected graph G with n vertices, such that any minimum r-identifying code in G contains c elements.*

Remark 4. The inequality $|C| \geq c$ stems only from the necessary covering and separating of vertices in $V^{(n_0)}$. Therefore, if $\ell = r$ and if the graph G_{n_0+r} is extended into a new graph $G^* = (V_{n_0+r} \cup V', E_{n_0+r} \cup E')$ where the only edges between V_{n_0+r} and V' are between v_r and V', then, because no vertex in V' is within distance r from a vertex in $V^{(n_0)}$, any r-identifying code C^* in G^* is such that $|C^* \cap V_{n_0+r}| \geq c$.

Consider in particular the case when $V' = W_q = \{w_1, \ldots, w_q\}$ and $E' = \{\{v_r, w_1\}, \{w_1, w_2\}, \ldots, \{w_{q-1}, w_q\}\}$ (cf. Section 4.1); then, since we can construct a minimum identifying code C in G_{n_0+r} which contains all vertices v_i $(0 \leq i \leq r)$ and since clearly this is the most favourable situation for the vertices in W_q, any r-identifying code in G^* has cardinality at least $c + \gamma_q$; actually, we can reach exactly $c + \gamma_q$, since C^*, the union of C with a $\{v_0, \ldots, v_r\}$-semi-identifying code C' in $\{v_0, \ldots, v_r\} \cup W_q$, is identifying in G^*: all vertices in V_{n_0+r} are r-covered and pairwise r-separated by C, all vertices in W_q are r-covered and pairwise r-separated by C', and, because $0 \in C^*$, w_1 is separated by C^* from v_r, which guarantees that any vertex in W_q is separated by C^* from any vertex in V_{n_0+r}.

4.3 From $c + r + 2$ to $c + 2r^2 + 4r + 1$

We shall use the previous construction and Remark 4 to prove the following.

Theorem 13 *Let $r \geq 2$ and $c \geq 5r^2 + 5r + 1$. For n between $c + r + 2$ and $c + 2r^2 + 4r + 1$, there exists a connected graph G with n vertices, such that any minimum r-identifying code in G contains c elements.*

Proof. Let j be such that $r + 2 \leq j \leq 2r^2 + 4r + 1$; then $j - r - 1 \geq 1$, and by Lemma 3, there exists an integer m such that $m - \gamma_m = j - 1 - r$. By (7), $j \geq r + 1 + (m-1)/2 - r = (m+1)/2$, or: $m \leq 2j - 1$.

Let $n_0 = c + j - m - r$. Thanks to the assumptions on c and j, we have:

$$\begin{aligned} n_0 &\geq (5r^2 + 5r + 1) + j - (2j - 1) - r \\ &= 5r^2 + 4r + 2 - j \\ &\geq 5r^2 + 4r + 2 - (2r^2 + 4r + 1) \\ &= 3r^2 + 1. \end{aligned}$$

Therefore, n_0 satisfies the conditions of Theorems 4 and 5. We mentioned in the proof of Lemma 1 that we did not try to optimize inequality (7); we can see here that any general improvement on (7) would change only marginally the condition involving c in the statement of Theorem 13.

Next, consider the graph $G_{n_0+r} = (V_{n_0+r}, E_{n_0+r})$ constructed in the previous section with $\ell = r$. To V_{n_0+r} we add m vertices, $v_{r+1}, v_{r+2}, \ldots, v_{r+m}$, where m is such that $m - \gamma_m = j - 1 - r$, and construct a graph $G_{n_0+r+m} = (V_{n_0+r+m}, E_{n_0+r+m})$ in the following way:

$$V_{n_0+r+m} = V_{n_0+r} \cup \{v_{r+1}, v_{r+2}, \ldots, v_{r+m}\},$$

$$E_{n_0+r+m} = E_{n_0+r} \cup \{\{v_r, v_{r+1}\}, \{v_{r+1}, v_{r+2}\}, \dots, \{v_{r+m-1}, v_{r+m}\}\}.$$

The number of vertices in G_{n_0+r+m} is equal to $n_0+r+m = (c+j-m-r)+r+m = c+j$, with $r+2 \le j \le 2r^2+4r+1$. By Remark 4 we have immediately the size of a minimum r-identifying code in G_{n_0+r+m}: it is
$n_0 - 1 + \gamma_m = (c+j-m-r) - 1 + (m-(j-1-r)) = c.$ △

4.4 From $c + 2r^2 + 4r + 2$ to $2^{c-3r^2} + 3r^2 + r$

Theorem 14 *Let $r \ge 2$ and $c \ge 3r^2 + 2r + 3$. For n between $c + 2r^2 + 4r + 2$ and $2^{c-3r^2} + 3r^2 + r$, there exists a connected graph G with n vertices, such that any minimum r-identifying code in G contains c elements.*

Proof. Let $c_1 \in [2r+3, c-3r^2]$, and consider the sparse graph $G_{c_1,q}$ in Remark 3, with $0 \le q \le 2^{c_1} - (r+1)c_1 - 2$: $G_{c_1,q}$ has $n_1 = (r+1)c_1 + 1 + q$ vertices, $(r+1)c_1 + 1 \le n_1 \le 2^{c_1} - 1$, and admits a minimum r-identifying code C_1 of size c_1, consisting of the vertices $x_{i,1}, 1 \le i \le c_1$.

Let $c_2 = c - c_1 \ge 3r^2$, $n_2 = c_2 + 1$, and G_{n_2+r} be the dense graph constructed in Section 4.2 with $\ell = r$: this graph has $n_2 + r$ vertices and admits a minimum identifying code C_2 of size $n_2 - 1$ containing the vertices v_0, v_1, …, v_r.

Now construct a new graph G by taking the union of $G_{c_1,q}$ and G_{n_2+r} and adding the set of edges

$$\{\{v_r, y\} : y \in X(c_1, q)\}.$$

Since v_r cannot distinguish the vertices in $X(c_1, q)$, it is now quite clear that $C_1 \cup C_2$ is a minimum r-identifying code in G, of size $c = c_1 + c_2$. On the other hand, G has $n_1 + n_2 + r$ vertices, and, setting $n = n_1 + n_2 + r$,

$$n = n_1 + (c_2 + 1) + r = n_1 + (c - c_1) + 1 + r;$$

therefore,

$$(r+1)c_1 + 1 + (c - c_1) + 1 + r \le n \le 2^{c_1} - 1 + (c - c_1) + 1 + r.$$

Or, setting $\min(c_1) = r(c_1 + 1) + c + 2$ and $\max(c_1) = 2^{c_1} + c - c_1 + r$:

$$\min(c_1) \le n \le \max(c_1),$$

with $2r + 3 \le c_1 \le c - 3r^2$. Which values of n can be achieved? We see that the functions min(.) and max(.) increase with c_1; moreover, by (9), we know that $2^{c_1} \ge (c_1+1)(r+1)$, which proves that $\max(c_1)+1 \ge \min(c_1+1)$. From this, we can conclude that $n_1 + n_2 + r$ achieves all values between $\min(2r+3) = c + 2r^2 + 4r + 2$ and $\max(c - 3r^2) = 2^{c-3r^2} + 3r^2 + r$, provided that $c \ge 3r^2 + 2r + 3$. △

4.5 From $(r+1)c + 1$ to $2^{c-1} - 1$

Theorem 15 *Let $r \ge 2$ and $c \ge 2r+3$. For n between $(r+1)c+1$ and $2^{c-1} - 1$, there exists a connected graph G with n vertices, such that any minimum r-identifying code in G contains c elements.*

Proof. Use Lemma 5 and Remark 3, with $c = s$, to construct a graph with a number of vertices between $(r+1)c+1$ and $2^c - 1$, admitting a minimum r-identifying code of size c. △

4.6 Recapitulatory

By Lemma 4(ii), for $c \geq 5r^2 + 5r + 1$ and $r \geq 2$, we have $2^{c-3r^2} + 3r^2 + r \geq c(r+1)$; therefore there is no gap between the interval $[(r+1)c+1, 2^{c-1} - 1]$ and the previous interval $[c + 2r^2 + 4r + 2, 2^{c-3r^2} + 3r^2 + r]$. Gathering Theorems 12–15 and their conditions on c, plus Theorems 3–5 and 10, one obtains the following.

Theorem 16 *Let $r \geq 1$ and $c \geq 5r^2 + 5r + 1$. For n between $c+1$ and $2^c - 1$, there exists a connected graph G with n vertices, such that any minimum r-identifying code in G contains c elements.*

5 Conclusion

We conclude with some open problems and one conjecture.

The graphs we consider here are still finite, unoriented and connected. Let $r \geq 1$, $n \geq 3$ be integers, and $\mathcal{G}$ be a family of graphs G (e.g., trees, multipartite graphs, planar graphs, series-parallel graphs, ...); in particular we denote by $\mathcal{T}$ the family of trees and by $\mathcal{B}$ the family of bipartite graphs. We define the following parameters:
– $\underline{c}(r, n, \mathcal{G})$ (respectively, $\overline{c}(r, n, \mathcal{G})$) is the smallest (respectively, largest) size of a minimum r-identifying code among the r-identifiable graphs $G \in \mathcal{G}$ having n vertices.
– $\underline{c}(r, n)$ (respectively, $\overline{c}(r, n)$) is the smallest (respectively, largest) size of a minimum r-identifying code among all the r-identifiable graphs G having n vertices.

Using Theorems 1 and 2, we have, for any family $\mathcal{G}$ and for any $n \geq 3$ for which the parameters are defined,

$$\lceil \log_2(n+1) \rceil \leq \underline{c}(r, n) \leq \underline{c}(r, n, \mathcal{G}) \leq \overline{c}(r, n, \mathcal{G}) \leq \overline{c}(r, n) \leq n - 1.$$

Open Problem 1. Study $\underline{c}(r, n)$, $\overline{c}(r, n)$; study $\underline{c}(r, n, \mathcal{G})$, $\overline{c}(r, n, \mathcal{G})$ for certain families of graphs.

By Theorems 3–5, we know that there exist values of n such that

$$\underline{c}(r, n) = \lceil \log_2(n+1) \rceil \tag{11}$$

$$\text{or } \ \overline{c}(r, n) = n - 1. \tag{12}$$

We denote by $E(r)$ (respectively, $F(r)$) the set of values of n such that (11) (respectively, (12)) holds. Again by Theorems 3–5, we know that

$$[2^{2r+1}, +\infty) \subseteq E(r), \ \ [3r^2 + 1, +\infty) \subseteq F(r). \tag{13}$$

Open Problem 2. Study $E(r)$ and $F(r)$. In particular, is there a function $e(r)$ (respectively, $f(r)$) such that $E(r) = [e(r), +\infty)$ (respectively, $F(r) = [f(r), +\infty)$)? what happens when $n \notin E(r)$ or $n \notin F(r)$ (i.e., for “small” values of n)?

By Theorems 10, 11 and the remark in-between, we have

$$\underline{c}(1,n,\mathcal{B}) = \lceil \log_2(n+1) \rceil, \tag{14}$$

$$\overline{c}(1,n,\mathcal{B}) = n-1, \tag{15}$$

$$\underline{c}(1,n,\mathcal{T}) = \lceil 3(n+1)/7 \rceil, \tag{16}$$

$$\text{and } \overline{c}(1,n,\mathcal{T}) = n-1, \tag{17}$$

and, if we denote by $E(1,\mathcal{B})$ (respectively, $F(1,\mathcal{B})$, $E(1,\mathcal{T})$, $F(1,\mathcal{T})$) the set of values of n such that (14) (respectively, (15), (16), (17)) holds, then we also have, still by Theorems 10 and 11, that

$$E(1) = F(1) = E(1,\mathcal{B}) = F(1,\mathcal{B}) = E(1,\mathcal{T}) = F(1,\mathcal{T}) = [3,+\infty).$$

It can also be proved that $E(2) = [6,+\infty)$, $F(2) = [5,+\infty)$, and $[2^{2r},+\infty) \subseteq E(r)$, improving the first inclusion in (13).

Open Problem 3. Study other values of r, other families of graphs (with the same questions as in Open Problem 2).

Finally, Theorems 10, 11 and 16, which give results about intermediate values, suggest the following.

Conjecture. *For any $r \geq 1$ and $n \geq 3$,*
(i) for any integer c between $\underline{c}(r,n)$ and $\overline{c}(r,n)$, there is a graph G with n vertices admitting a minimum r-identifying code with c elements;
(ii) for some families $\mathcal{G}$ of graphs such as trees, multipartite graphs, planar graphs, ..., for any integer c between $\underline{c}(r,n,\mathcal{G})$ and $\overline{c}(r,n,\mathcal{G})$, there is a graph $G \in \mathcal{G}$ with n vertices admitting a minimum r-identifying code with c elements.

References

[1] N. Bertrand, Codes identifiants et codes localisateurs-dominateurs sur certains graphes, Mémoire de stage de maîtrise (under O. Hudry's supervision), ENST, June 2001.

[2] N. Bertrand, I. Charon, O. Hudry, A. Lobstein, Identifying codes on trees, *Australas. J. Combin.*, to appear.

[3] U. Blass, I. Honkala, S. Litsyn, Bounds on identifying codes, *Discrete Math.* **241** (2001), 119–128.

[4] I. Charon, O. Hudry, A. Lobstein, Identifying codes with small radius in some infinite regular graphs, *Electron. J. Combin.* **9(1)** (2002), R11.

[5] I. Charon, O. Hudry, A. Lobstein, Minimizing the size of an identifying or locating-dominating code in a graph is NP-hard, *Theoret. Comput. Sci.* **290(3)** (2003), 2109–2120.

[6] I. Charon, O. Hudry, A. Lobstein, Extremal cardinalities for identifying and locating-dominating codes in graphs, Rapport Interne Télécom Paris-2003D006, August 2003. Also submitted to *Discrete Math.*

[7] C. J. Colbourn, P. J. Slater, L. K. Stewart, Locating dominating sets in series parallel networks, *Congr. Numer.* **56** (1987), 135–162.

[8] I. Honkala, T. Laihonen, S. Ranto, On strongly identifying codes, *Discrete Math.* **254** (2002), 191–205.

[9] M. G. Karpovsky, K. Chakrabarty, L. B. Levitin, On a new class of codes for identifying vertices in graphs, *IEEE Trans. Inform. Theory* **44** (1998), 599–611.

[10] S. Ranto, I. Honkala, T. Laihonen, Two families of optimal identifying codes in binary Hamming spaces, *IEEE Trans. Inform. Theory* **48** (2002), 1200–1203.

[11] http://www.infres.enst.fr/~lobstein/bibLOCDOMetID.html

(Received 15 Sep 2003)