On the identification of vertices using cycles

Petri Rosendahl*
Department of Mathematics and
Turku Centre for Computer Science TUCS
University of Turku
20014 Turku, Finland
perosen@utu.fi

Submitted: Apr 3, 2002; Accepted: Jan 22, 2003; Published: Jan 29, 2003
MR Subject Classifications: 05C38, 94C15

Abstract

A set of cycles $C_1, \dots, C_k$ in a graph G is said to identify the vertices v if the sets $\{j : v \in C_j\}$ are all nonempty and different. In this paper, bounds for the minimum possible k are given when G is the graph $\mathbb{Z}_p^n$ endowed with the Lee or Hamming metric or G is a complete bipartite graph.

1 Introduction

For basic notions of graph theory we use the terminology of [2]. Especially, we will use the following notions.

A *walk* in a graph is a finite non-null sequence $W = v_0 v_1 \cdots v_n$ of vertices such that v_i and v_{i+1} are adjacent for all $i = 0, \dots, n-1$. We also say that W is a walk *from* v_0 *to* v_n. The number n is the *length of* W. If $v = v_i$ for some i, we say that W *visits* the vertex v and write (by a slight abuse of notation) $v \in W$. If $v_0 = v_n$ and the length of W is positive, we say that the walk W is *closed*.

If all the vertices in a walk $W = v_0 v_1 \cdots v_n$ are distinct, then W is called a *path*. If W is closed, $n \geq 3$, and $v_i \neq v_j$ whenever $i \neq j$ and $i, j = 0, \dots, n-1$, then W is a *cycle*. A path of length one is essentially nothing but an edge.

A *Hamilton cycle* of a graph G is a cycle which visits every vertex of G. A graph is said to be *Hamiltonian* if it contains a Hamilton cycle (as a subgraph).

Definition 1. A collection $A_1, A_2, \dots, A_k$ of subsets of a set S is called identifying if the sets $\{i : x \in A_i\}$ $(x \in S)$ are nonempty and different.

*Research supported by the Academy of Finland under grant 44002.

An obvious test whether or not $A_1, \ldots, A_k$ is an identifying collection of subsets of an s-element set S, is provided by a $k \times s$ matrix $\mathbf{A}$ whose columns are indexed with elements of S and whose rows are characteristic vectors of the subsets A_i. Namely, $A_1, \ldots, A_k$ is identifying if and only if the columns of $\mathbf{A}$ are nonzero and distinct. Even if the sets A_i are required to have some kind of structure, $\mathbf{A}$ can sometimes be used to solve existence questions (see Theorem 12).

The identification problem can be interpreted as a (graph) coloring problem in the following way. Suppose that we want to colour the vertices of a graph, each with a different colour. How many colours are needed initially, when the same vertex may be colored with several different colours and mixtures of different colours are assumed to produce a new colour?

For arbitrary sets we have the following obvious lower bound for the number of identifying subsets.

Theorem 2. *An identifying collection of subsets of an s-element set contains at least $\lceil \log_2(s+1) \rceil$ subsets.*

The lower bound given in Theorem 2 can always be attained if there is no requirement for the structure of the identifying subsets.

Lemma 3. *For an s-element set S there is an identifying collection $A_1, \ldots, A_k$ of subsets such that $k = \lceil \log_2(s+1) \rceil$ and $|A_i| = \lceil s/2 \rceil$ for all $i = 1, \ldots, k$.*

Proof. Clearly there is an identifying collection of $\lceil \log_2(s+1) \rceil$ subsets. Assume that A_1, for instance, contains less than $\lceil s/2 \rceil$ elements. Then the corresponding matrix $\mathbf{A}$ has less columns beginning with 1 than columns beginning with 0. Thus there is a column beginning with 0 whose $k-1$ last bits do not appear as $k-1$ last bits of any other column. Now the first bit of this column can be changed to 1. The same argument holds for zeros too (except that we have to avoid the zero column). $\square$

The study of identifying collections of subsets is motivated by fault diagnosis in multiprocessor architectures, i.e., these collections are used for locating a malfunctioning processor. As usual, a multiprocessor architecture is represented as a graph. Each vertex corresponds to a processor and each edge represents a link between two processors. In the simplest variant we assume that at most one of the processors is malfunctioning, and we wish to identify it (or to find that none of them is malfunctioning). We use the following scheme. Let S be the set of vertices. We choose the identifying subsets $A_1, \ldots, A_k$ of S. The set of processors in each A_i is checked and we get YES/NO answers telling whether or not any problems were detected in A_i. Based on these k YES/NO answers we are able to identify the malfunctioning processor (or to tell that there is none). It is natural to pose various constraints on the sets A_i, e.g., to require that that they are balls with respect to the graphic distance, or that they are cycles or closed walks. For more details, see [8], [7] and [4]. This problem is also related to fault-tolerant message routing in multiprocessor architectures, see [9] and [10]. Yet another kind of routing problem and multiprocessor diagnosis technique is treated in [1].

Identification problems have been studied in various graphs, see e.g. [3] and [5] for further references.

In Section 2 we consider vertex identification in the p-ary Lee space $\mathbb{Z}_p^n$ and in the Hamming space $\mathbb{Z}_p^n$. The identifying sets are required to consist of vertices of a cycle.

In Section 3 we construct a vertex identifying collection of cycles in the complete bipartite graph $K_{m,m}$ and also consider the more general case $K_{m,n}$.

2 Cycles for identification in the Hamming and Lee spaces

By a Lee space we mean the set $\mathbb{Z}_p^n$ endowed with the Lee metric. In other words two vertices are adjacent if and only if they differ by $\pm 1 \mod p$ in exactly one coordinate and the other coordinates coincide. The smallest cardinality of a vertex identifying collection (if there is one) of cycles in the p-ary n-dimensional Lee space is denoted by $V_L(p,n)$. Respectively, in the Hamming space, this number is denoted by $V_H(p,n)$. Note that $V_L(p,1)$, $V_H(p,1)$, and $V_H(2,2)$ are not defined.

In the case of vertex identification, constructions proved for the Lee space are also valid in the Hamming space, i.e., $V_H(p,n) \leq V_L(p,n)$, since the edge set of the former is a subset of the edge set of the latter. Note also that when $p = 2$ or $p = 3$, then the Hamming and Lee spaces are the same, and thus in this case $V_H(p,n) = V_L(p,n)$.

We note here that the graphs $\mathbb{Z}_p^n$ are always Hamiltonian, with respect to both Lee and Hamming metric, when $p^n \geq 3$. This can be easily proved by induction.

We begin with an example.

Example 4. By writing down the corresponding matrix, the following five cycles C_i are seen to be vertex identifying in $\mathbb{Z}_3^3$. The trivial lower bound $\lceil \log_2 28 \rceil$ given in Theorem 2 is thus attained in this case. Hence $V_L(3,3) = V_H(3,3) = 5$.

C_1	C_2	C_3	C_4	C_5
000	001	011	001	021
200	201	010	101	121
202	221	020	102	120
222	222	021	100	100
122	122	121	110	101
022	121	122	210	201
012	120	022	200	200
212	100	222	220	202
210	200	212	221	212
110	202	202	211	211
111	002	201	212	111
101	012	221	222	112
001	011	220	202	102
011	001	210	201	122
010		211	001	022
000		011		021

Theorem 5. *If $p > 2$, then $V_L(p, m+n) \leq V_L(p, m) + V_L(p, n)$ for all $m, n \geq 2$. If $p = 2$, the same is valid for all $m, n \geq 3$.*

Proof. Let $C = u_1 u_2 \ldots u_k u_1$ be any cycle in $\mathbb{Z}_p^n$. First we show that there is a cycle through every vertex of the set

$$\mathbb{Z}_p \times C = \{(x, y) \in \mathbb{Z}_p^{n+1} : x \in \mathbb{Z}_p, y \in C\}.$$

Consider the cycles C_i where $i = 0, 1, \ldots, p-1$ and $C_i = \{(i, y) \in \mathbb{Z}_p^{n+1} : y \in C\}$. We form the required cycle as follows. First we visit every vertex and edge of C_0 except for the last edge. Being now at the vertex $(0, u_k)$ we move to vertex $(1, u_k)$. Now we visit every vertex of C_1 counterclockwise except for $(1, u_1)$ and move to C_2, which is traveled clockwise and without $(2, u_1)$. We proceed in this way, i.e., every other cycle is traveled counterclockwise and vertices (i, u_1) are not visited for $1 < i < p-1$, and finally arrive at the cycle C_{p-1}. This time we visit also the vertex $(p-1, u_1)$ and then complete the cycle with vertices $(p-2, u_1)$,...,$(1, u_1)$ and $(0, u_1)$.

Inductively we see that we can form cycles also from vertices of the set

$$\mathbb{Z}_p^l \times C = \{(x, y) \in \mathbb{Z}_p^{n+l} : x \in \mathbb{Z}_p^l, y \in C\},$$

for every $l \geq 1$.

Suppose that $D_1, \ldots, D_s$ are vertex identifying cycles in $\mathbb{Z}_p^m$ and $E_1, \ldots, E_t$ are vertex identifying cycles in $\mathbb{Z}_p^n$. Then the $s+t$ cycles $D_i \times \mathbb{Z}_p^n$ and $\mathbb{Z}_p^m \times E_i$ are vertex identifying in $\mathbb{Z}_p^{m+n}$. Indeed, every vertex lies in some cycle and if u and v lie in the same set of cycles then their first m coordinates must coincide and their last n coordinates must coincide too. $\square$

From now on the vertices of $\mathbb{Z}_p^n$ are denoted by boldface letters, e.g., $\mathbf{x}$. The i-th coordinate of a vertex is then specified by a subscript, e.g., x_i.

In the Hamming space $\mathbb{Z}_p^n$ we have the following theorem. Note that this says that if p is a power of two, then the trivial lower bound given in Theorem 2 can be achieved.

Theorem 6. *Assume that $n \geq 2$. If p is a power of two, then excluding the case $p = 2$ and $n = 2$, we have $V_H(p, n) = n \log_2 p + 1$. If p is not a power of two, then $V_H(p, n) \leq n\lceil \log_2 p \rceil$.*

Proof. Let $k = \lceil \log_2 p \rceil$. We choose k subsets A_s of $\mathbb{Z}_p$ such that the sets $\{s : x \in A_s\}$ are distinct for $x \in \mathbb{Z}_p$. Furthermore, they are required to be nonempty if p is not a power of two. We construct the sets C_{ij}, where $i = 1, \ldots, n$ and $j = 1, \ldots, k$, by the rule:

C_{ij} consists of those vertices whose i-th component is in the set A_j.

If p is a power of two then there is an element $x \in \mathbb{Z}_p$ which lies in none of the sets A_s and in this case we add to our collection any cycle which visits the vertex $(x, x, \ldots, x)$.

Clearly every vertex lies in at least one of the constructed subgraphs. Moreover, if $\mathbf{x} \neq \mathbf{y}$ then $x_i \neq y_i$ for some i, and thus C_{ij} contains either $\mathbf{x}$ or $\mathbf{y}$, but not both, for some j.

Lastly, we have to show that there is a cycle consisting of the vertices of C_{ij}. To begin with, we note that there is a cycle C through every vertex of $\mathbb{Z}_p^{n-1}$. Thus the required cycle can be constructed in the same way as in the proof of Theorem 5. □

Remark 7. It is proved in [7] and [6] that $V_H(2, n) = n+1$ for all $n \geq 3$. The construction is briefly as follows. For $i = 1, \ldots, n$ take a cycle consisting of those vertices which have 0 as the i-th coordinate (e.g. Gray code of dimension $n-1$), and then take any cycle which contains the all-one vector. It is straightforward to show that the set of these $n+1$ cycles is identifying.

Theorem 8. *We have $V_L(4, n) = 2n + 1$, for all $n \geq 2$.*

Proof. By Theorem 2 $V_L(4, n) \geq 2n + 1$. We map the elements of $\mathbb{Z}_4^n$ onto $\mathbb{Z}_2^{2n}$ by the Gray map $0 \mapsto 00$, $1 \mapsto 01$, $2 \mapsto 11$ and $3 \mapsto 10$. Since two elements of the Lee space $\mathbb{Z}_4^n$ are adjacent if and only if their images are adjacent in the Hamming space $\mathbb{Z}_2^{2n}$, we get from the remark above (and also by Theorem 6) that $V_L(4, n) \leq 2n + 1$. □

In the sequel the identity $\lceil \log_2(s + 1) \rceil = \lfloor \log_2 s \rfloor + 1$ is used. To see this, choose a k such that $2^k \leq s < 2^{k+1}$ and observe that both sides equal $k + 1$.

It is proved in [7] and [6] that

$$V_L(p, 2) \leq 2\lceil \log_2(p + 1) \rceil + 1,$$

for all $p \geq 4$. We give an upper bound in the case $n \geq 3$.

Theorem 9. *In the Lee space $\mathbb{Z}_p^n$, where $p \geq 4$ and $n \geq 3$, we have $\lfloor n \log_2 p \rfloor + 1 \leq V_L(p, n) \leq n\lceil \log_2(p + 1) \rceil + 2$.*

Proof. Theorem 2 gives the lower bound.

To prove the upper bound we give a construction. The idea of this construction is briefly as follows. The cycles $C_{i,j}$ are used to decide the i-th coordinate of the vertex $\mathbf{v}$ which is to be identified. The auxiliary cycles B_i are needed in the construction (and they will not be in the final collection of cycles). The cycles D and E are needed in identification of the vertices in the cycles B_i, since the cycles $C_{i,j}$ may leave two possibilities for $\mathbf{v}$.

Throughout the proof we will illustrate the construction in the case $p = 4$ and $n = 3$. For notational reasons we write $\mathbb{Z}_p = \{1, 2, \ldots, p\}$.

First of all, let $k = \lceil \log_2(p+1) \rceil$ and let $\mathbf{C}$ be the $k \times p$ matrix whose i-th column is the binary representation of $i - 1$. In particular, $\mathbf{C}$ does not contain the all-one column. In our example, we have

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

For $i = 2, \ldots, n$, let B_i be the cycle which starts from the vertex 1^n and the next vertices are obtained by increasing every other time the $(i-1)$-st coordinate by one and every other time the i-th coordinate by one, the $(i-1)$-st coordinate being increased first. In addition, we construct the cycle B_1. It starts from the vertex 131^{n-2}, then the first coordinate increases by one. Next the second coordinate increases by one, then the first, and so on. Note that the cycles B_i and B_j, $i, j \geq 3$ are disjoint except for the vertex 1^n. In addition the cycles B_1 and B_2 are disjoint (here we need the assumption $p \geq 4$). Hence, in our example, B_2 consists of vertices 111, 211, 221, 321, 331, 431, 441, 141 and 111. Similarly, B_1 consists of 131, 231, 241, 341, 311, 411, 421, 121 and 131. It should be noted that B_3 (and more generally B_i for $i \geq 3$) can be obtained from B_2 by permuting the coordinates.

From the matrix $\mathbf{C}$ we form the cycles $C_{i,j}$, where $i = 1, \ldots, n$ and $j = 1, \ldots, k$. The cycle $C_{i,j}$ is obtained using the auxiliary cycle B_i and j-th row of $\mathbf{C}$ as follows. We start from the vertex 1^n except that in case $i = 1$ we start from the vertex 131^{n-2}. Then we move along B_i to the first vertex which has 2 as its i-th coordinate, i.e., to the vertex $1^{i-2}221^{n-i}$ if $i > 1$ and to the vertex 231^{n-2} if $i = 1$. If the second element of the j-th row of $\mathbf{C}$ is one, then the hyperplane $x_i = 2$ is traveled through in such a way that we end at the vertex $1^{i-2}321^{n-i}$ which lies in B_i (in the case $i = 1$ the last vertex is 241^{n-2} which lies in B_1). This is possible by symmetry and the fact that $\mathbb{Z}_p^{n-1}$ is Hamiltonian. Next we move along B_i to the vertex $1^{i-2}331^{n-i}$ which lies in B_i (if $i = 1$ we move to 341^{n-2}). If the second element of the j-th row of $\mathbf{C}$ was zero, we move along the cycle B_i to $1^{i-2}321^{n-i}$ and $1^{i-2}331^{n-i}$ (and to 241^{n-2} and 341^{n-2} if $i = 1$) without going through the hyperplane $x_i = 2$.

Similarly, for all t, the hyperplane $x_i = t$ is traveled through if and only if the t-th element of the j-th row is one, otherwise we move along the cycle B_i to the next plane $x_i = t + 1$. More formally, denote $\mathbf{C} = (c_{ij})$. Then $\mathbf{x}$ lies in $C_{i,j}$ if and only if (i) $\mathbf{x}$ lies in B_i or (ii) $x_i = t$ and $c_{jt} = 1$.

In our example we have $C_{1,1} = B_1$, $C_{2,1} = B_2$ and $C_{3,1} = B_3$. The cycle $C_{1,2}$ consists now of vertices 131, 231, 241, 341, 342, 343, 344, 314, 313, 312, 322, 323, 324, 334, 333, 332, 331, 321, 311, 411, 412, 413, 414, 424, 423, 422, 432, 433, 434, 444, 443, 442, 441, 431, 421, 121 and 131. So it contains all the vertices which have 3 or 4 as the first coordinate and in addition it contains the vertices of B_1. As another example we give the cycle $C_{2,3}$. It consists of the vertices 111, 211, 221, 222, 223, 224, 324, 323, 322, 422, 423, 424, 124, 123, 122, 121, 421, 321, 331, 431, 441, 442, 443, 444, 144, 143, 142, 242, 243, 244, 344, 343, 342, 341, 241, 141 and 111. As it should, it contains all the vertices of B_2 and vertices which have either 2 or 4 as the second coordinate. Note that the cycle $C_{3,3}$ is now easily obtained from $C_{2,3}$ by permuting the coordinates.

Let D be the cycle 1^n, 121^{n-2}, 221^{n-2}, 231^{n-2}, 331^{n-2}, ..., 1^n. It has the important property that every other vertex of both B_1 and B_2 belong to it.

Lastly, we construct a cycle E, which will include every other vertex of each of the cycles $B_3, ..., B_n$. To begin with, we construct a path P_i for $i = 3, ..., n$. The path P_i starts with the vertex whose $(i-1)$-st and i-th coordinates are 2 and other coordinates are 1. The next vertices are obtained by increasing every second time the i-th coordinate and every second time the $(i-1)$-st coordinate, the i-th coordinate being increased first. The path P_i ends in the vertex whose $(i-1)$-st and i-th coordinates are p and the other coordinates are 1. In our example the path P_3 consists of 122, 123, 133, 134 and 144. Now we use P_i as the building blocks of the cycle E as follows. Firstly, we visit the vertices of P_3. Next, if $n > 3$, we change the fourth coordinate to 2, decrease in steps of one the third coordinate to 2 and change the second coordinate to 1. We are now at the vertex 11221^{n-4}, which lies in P_4. Now we travel the vertices of P_4. If $n > 4$, we repeat the process above. This process is repeated until we are at the last vertex of P_n, i.e., $1^{n-2}pp$. Now we move to vertices $p1^{n-3}pp$, $p1^{n-2}p$ and $p1^{n-1}$. The first coordinate is now changed to 1 and thus the all-one vertex is in the cycle E. Now we change the first coordinate to 2 and then move to the vertex 2221^{n-3} via 221^{n-2}. Now we change the first coordinate to 1 and the cycle E is complete. The cycle E contains now every second vertex of $B_3, \ldots, B_n$ and some vertices which are in none of the $B_3, \ldots, B_n$.

In our example $n = 3$ and the cycle E consists of the vertices 122, 123, 133, 134, 144, 444, 414, 411, 111, 211, 221, 222 and 122. As an other example, consider the case $n = 4$ and $p = 4$. Now the cycle E consists of the vertices 1221, 1231, 1331, 1341, 1441, 1442, 1432, 1422, 1122, 1123, 1133, 1134, 1144, 4144, 4114, 4111, 1111, 2111, 2211, 2221 and 1221.

We claim that the cycles $C_{i,j}$ together with the cycles D and E form an identifying collection. Clearly the number of the constructed cycles is as claimed.

First we have to show that every vertex lies in at least one of the constructed cycles. So, let $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}_p^n$. Suppose for the moment that $x_i = t \neq 1$ for some i. Since the matrix $\mathbf{C}$ has a row, say the j-th row, which has 1 as its t-th element we conclude that $\mathbf{x}$ lies in $C_{i,j}$. The vertex 1^n lies in E, for example.

We observe here that if $\mathbf{v} \notin B_i$, then the i-th coordinate of $\mathbf{v}$ can be decided by testing for which $j = 1, \ldots, k$ we have $\mathbf{v} \in C_{i,j}$.

Let $\mathbf{v} \in \mathbb{Z}_p^n$. We use the cycles D, $C_{1,j}$ and $C_{2,j}$ to determine the first and second coordinates of $\mathbf{v}$ in the following way. If $\mathbf{v}$ lies in every cycle $C_{1,j}$, then it lies in the cycle B_1 ($\mathbf{C}$ does not contain the all-one column) and thus cannot lie in B_2. Thus we can determine v_2 using the cycles $C_{2,j}$. This leaves two possibilities for v_1, and since $\mathbf{v} \in B_1$, two possibilities for $\mathbf{v}$. Furthermore, these are consecutive in the cycle B_1. Since D contains every second vertex of B_1, we can decide which one $\mathbf{v}$ is. If $\mathbf{v}$ lies in B_2 we determine v_1 first and proceed similarly. If $\mathbf{v}$ lies in neither B_1 nor B_2 there is no difficulty.

Suppose now that the coordinates $v_1, ..., v_t$, where $t \geq 2$, have been determined. We have to determine the coordinate v_{t+1}. Firstly, if $\mathbf{v}$ lies in all the cycles $C_{t+1,j}$ then it lies in B_{t+1}. Since we know v_t, there is only two possibilities for v_{t+1}, and thus also for $\mathbf{v}$. These are consecutive in the cycle B_{t+1} and thus we can use E to decide which one $\mathbf{v}$ is. Secondly, if $\mathbf{v}$ does not lie in all of the cycles $C_{t+1,j}$, then there is no difficulty. □

To cover the case $p = 3$ we give the following bounds.

Theorem 10. $\lfloor 3n \log_2 3 \rfloor + 1 \leq V_L(3, 3n) = V_H(3, 3n) \leq 5n$.

Proof. This follows from Theorem 5 and the fact $V_L(3, 3) = V_H(3, 3) = 5$ proven in Example 4. □

It is not difficult to see that $V_L(3, 2) = V_H(3, 2) = 4$. Together with the preceding theorem we obtain an upper bound for all $V_L(3, n)$ and $V_H(3, n)$.

3 On the identification of vertices in bipartite graphs

A graph G is a complete bipartite graph if the vertex set of G can be partitioned as $V = V_1 \cup V_2$ and two vertices v_1 and v_2 are adjacent if and only if $v_1 \in V_1$ and $v_2 \in V_2$. If $|V_1| = m$ and $|V_2| = n$, this graph is denoted by $K_{m,n}$.

The identification problem for complete bipartite graphs using connected subgraphs solves trivially.

Theorem 11. *For the complete bipartite graph $K_{m,n}$ the smallest cardinality of an identifying collection of connected subgraphs is $\lceil \log_2(m + n + 1) \rceil$ for all $m, n \geq 1$.*

Proof. The number of vertices is $m + n$, and from Theorem 2 we get the lower bound.

If $m = n = 1$ there is nothing to prove. By symmetry we may assume now that $m \geq 2$ and $n \geq 1$. We have to show that the identifying subsets of $V = V_1 \cup V_2$ can be chosen in such a way that they contain vertices from both V_1 and V_2. Let $u, v \in V_1$ and $w \in V_2$ and consider the corresponding matrix $\mathbf{A}$ defined in the Introduction. We take the columns corresponding to u and v to be $011 \cdots 1$ and $101 \cdots 1$. The column corresponding to w is chosen to be $11 \cdots 1$. The remaining columns can be chosen freely among the remaining nonzero vectors. □

In the next theorem we use the observation that there is a cycle through every vertex of a subset S of the vertex set of $K_{m,n}$ if and only if S contains equally many vertices from V_1 and V_2 and at least two from both.

Theorem 12. *For the complete bipartite graph $K_{m,m}$ $(m \geq 3)$ the smallest cardinality of an identifying set of cycles is* *(i)* $\lfloor \log_2 m \rfloor + 3$*, if* m *is of the form* $m = 2^k - 1$*,*

(ii) $\lfloor \log_2 m \rfloor + 2$*, otherwise.*

Proof. From Theorem 2, we have a lower bound $\lceil \log_2(2m+1) \rceil = \lfloor \log_2 2m \rfloor + 1 = \lfloor \log_2 m \rfloor + 2$.

The rest of the proof is now given in several steps. First we give a construction which proves an upper bound which equals the lower bound if m is even. If m is odd, this bound is $\lfloor \log_2 m \rfloor + 3$. Secondly, we show that the lower bound is not reachable if $m = 2^k - 1$ for some k. Lastly, we assume that m is odd but not of the form $m = 2^k - 1$, and give a construction which equals the lower bound.

If m is even, there is a very simple construction. For every vertex in V_1 we choose a counterpart in V_2. Next we choose an identifying collection S of $\lceil \log_2(m+1) \rceil = \lfloor \log_2 m \rfloor + 1$ subsets of V_1 such that each of these subsets contains at least two vertices (this can be done by Lemma 3), and from V_2 this determines subsets consisting of counterpart vertices. The union of a subset in S and its counterpart subset can be made connected with one cycle. Since S is an identifying set of V_1, the collection of cycles constructed so far is identifying except that it does not separate a vertex from its counterpart. Since m is even all such pairs can be separated with one cycle.

In the odd case a similar construction works also, but this time we need two cycles to separate vertices of V_1 from their counterparts in V_2. This gives an upper bound.

To prove (i), let $m = 2^k - 1$ and $C_1, \dots, C_{k+1}$ be a collection of subsets of the vertex set of $K_{m,m}$. Let $\mathbf{A}$ be the corresponding matrix (as in the Introduction). Suppose that the columns of $\mathbf{A}$ are distinct and nonzero. If $\mathbf{A}$ contained all $2^{k+1} - 1$ nonzero vectors of length $k+1$ as columns, then every C_i would contain an even number of vertices. However, as is the case, leaving one vector out forces at least one C_i to have an odd number of vertices. Thus one C_i cannot be a cycle.

Assume finally that m is odd and $2^{k-1} < m < 2^k - 1$. We arrange the vertices $v_1, \dots, v_{2m}$ in such a way that $v_i \in V_1$ if and only if i is odd, and construct a suitable $(k+1) \times 2m$ matrix. To do this we take $0100\cdots0$, $0110\cdots0$, $0010\cdots0$ and $100\cdots0$ as the first four columns. We will not use $110\cdots0$, $1110\cdots0$, $1010\cdots0$ and $00\cdots0$ as columns; these vectors can be discarded since $2m$ and 2^{k+1} differ by at least four. Now we choose $m-2$ different other vectors beginning with 1 and use these as columns from 5 to $m+2$. The remaining $m-2$ columns are the same as the columns just chosen but this time beginning with zero and written in reverse order, i.e., the columns corresponding to vertices v_{m+2-i} and v_{m+3+i}, where $i = 0, \dots, m-3$, differ only in the first coordinate. Now it is easy to see that the constructed matrix satisfies the requirements, i.e., there are as many odd and even vertices in every row and the columns are distinct and nonzero. $\square$

Example 13. The matrix for $K_{11,11}$ is

$$\left(\begin{array}{cccc|ccccccccc|ccccccccc}
0&0&0&1&1&1&1&1&1&1&1&1&1&0&0&0&0&0&0&0&0&0\\
1&1&0&0&0&0&0&0&0&0&1&1&1&1&1&1&0&0&0&0&0&0\\
0&1&1&0&0&0&0&1&1&1&0&0&0&0&0&0&1&1&1&0&0&0\\
0&0&0&0&0&1&1&0&1&1&0&1&1&1&1&0&1&1&0&1&1&0\\
0&0&0&0&1&0&1&1&0&1&1&0&1&1&0&1&1&0&1&1&0&1
\end{array}\right).$$

Note the symmetry around the second vertical line.

The methods given above do not apply if $n < m$, and in this case the exact cardinality of the smallest identifying collection of cycles may be hard to find. As an example we treat the case $n = 3$.

Theorem 14. *Let $k \geq 4$. The smallest cardinality of an identifying collection of cycles in $K_{2k-2,3}$ and $K_{2k-1,3}$ is k.*

Proof. We may assume that in the corresponding matrix $\mathbf{A}$ the columns corresponding to V_2 are the last three columns. We refer to the last three columns as the right part of $\mathbf{A}$ and the remaining columns as the left part.

First we give a lower bound. Suppose that k cycles is enough. We note that a cycle can now contain at most three vertices from both V_1 and V_2. Moreover, we must have at least two cycles which have only two vertices from V_2. Otherwise it would be impossible to separate the vertices in V_2. Thus the left part of $\mathbf{A}$ can contain at most $2 \times 2 + 3 \times (k-2) = 3k - 2$ 1's. There are only k binary vectors of weight one and hence we can choose at most $((3k-2)-k)/2 = k-1$ vectors of weight two. Thus at most $k + (k-1) = 2k-1$ vertices of V_1 can be separated from each other. This proves that at least k cycles are needed to identify the vertices of $K_{2k-1,3}$ and $K_{2k-2,3}$.

Actually, the argument above suggests the way in which we construct the set of identifying cycles. In the case of $K_{2k-1,3}$ the first k columns are the vectors of Hamming weight one, i.e., 100...0, 010...0, ..., 00...01. The next $k-1$ columns are 1100...0, 0110...0, 00110...0, ..., 00....011. The columns corresponding to V_2 are now 11...10, 11...1 and 011...1. In the case of $K_{2k-2,3}$ the first $k-1$ columns are 10...01, 010...0, ..., 00...010 (instead of the two vectors 10...0 and 00...01 we use their sum). The next $k-1$ columns are again 1100...0, 0110...0, ..., 00...011. The last three columns are the same as in the case $K_{2k-1,3}$. The validity of these selections for $\mathbf{A}$ is easily checked. $\square$

Example 15. The corresponding matrices for $K_{10,3}$ and $K_{11,3}$ are

$$\left(\begin{array}{cccccccccc|ccc}
1&0&0&0&0&1&0&0&0&0&1&1&0\\
0&1&0&0&0&1&1&0&0&0&1&1&1\\
0&0&1&0&0&0&1&1&0&0&1&1&1\\
0&0&0&1&0&0&0&1&1&0&1&1&1\\
0&0&0&0&1&0&0&0&1&1&1&1&1\\
1&0&0&0&0&0&0&0&0&1&0&1&1
\end{array}\right)$$

and

$$\left(\begin{array}{ccccccccccc|ccc}
1&0&0&0&0&0&1&0&0&0&0&1&1&0\\
0&1&0&0&0&0&1&1&0&0&0&1&1&1\\
0&0&1&0&0&0&0&1&1&0&0&1&1&1\\
0&0&0&1&0&0&0&0&1&1&0&1&1&1\\
0&0&0&0&1&0&0&0&0&1&1&1&1&1\\
0&0&0&0&0&1&0&0&0&0&1&0&1&1
\end{array}\right).$$

It is easy to prove that the smallest identifying collection of cycles in $K_{4,3}$ consist of three cycles. In $K_{5,3}$ the smallest identifying collection of cycles consists of four cycles.

References

[1] U. Blass, I. Honkala, M.G. Karpovsky, S. Litsyn, Short dominating paths and cycles in the binary hypercube, Annals of Combinatorics 5 (2001), pp. 51-59.

[2] J.A. Bondy, U.S.R. Murty, Graph theory with applications. North-Holland, New York, 1976.

[3] I. Charon, O. Hudry, A. Lobstein: Identifying codes with small radius in some infinite regular graphs, Electronic Journal of Combinatorics 9 (1) 2002.

[4] K. Chakrabarty, M.G. Karpovsky, L.B. Levitin: Fault isolation and diagnosis in multiprocessor architectures with point-to-point connections, Fault tolerant parallel and distributed systems, Kluwer 1998, pp. 285-301.

[5] G. Cohen, I. Honkala, A. Lobstein and G. Zémor: New Bounds for Codes Identifying Vertices in Graphs, Electronic Journal of Combinatorics 6 (1) 1999.

[6] I. Honkala, M.G. Karpovsky, S. Litsyn, Cycles identifying vertices and edges in binary hypercubes and 2-dimensional tori, Discrete Applied Mathematics, to appear.

[7] I. Honkala, M.G. Karpovsky, S. Litsyn, On the identification of vertices and edges using cycles, Proc. AAECC-14, S. Boztas, I.E. Shparlinski (eds.) (2001), pp. 308-314.

[8] M.G. Karpovsky, K. Chakrabarty, L.B. Levitin, On a new class of codes for identifying vertices in graphs, IEEE Transactions on Information Theory 44 (1998), pp. 599-611.

[9] L. Zakrevski, M.G. Karpovsky, Fault-tolerant message routing in computer networks, Proc. Int. Conf. on Parallel and Distributed Processing Techniques and Applications (1999), pp. 2279-2287.

[10] L. Zakrevski, M.G. Karpovsky, Fault-tolerant message routing for multiprocessors, Parallel and Distributed Processing, J. Rolim (ed.), Springer (1998), pp. 714-731.