



HAL
open science

Extension of universal cycles for globally identifying colorings of cycles

Pierre Coupechoux

► **To cite this version:**

Pierre Coupechoux. Extension of universal cycles for globally identifying colorings of cycles. *Discrete Mathematics*, 2017, 340 (7), pp.1456-1466. 10.1016/j.disc.2017.02.002 . hal-01704385

HAL Id: hal-01704385

<https://laas.hal.science/hal-01704385>

Submitted on 8 Feb 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Extension of universal cycles for globally identifying colorings of cycles

Pierre Coupechoux^a

^aLAAS-CNRS, Université de Toulouse, CNRS, Toulouse, France

Abstract

In 1998, Karpovsky, Chakrabarty and Levitin introduced identifying codes to model fault diagnosis in multiprocessor systems [1]. In these codes, each vertex is identified by the vertices belonging to the code in its neighborhood. There exists a coloring variant as follows: a globally identifying coloring of a graph is a coloring such that each vertex is identified by the colors in its neighborhood. We aim at finding the maximum length of a cycle with such a coloring, given a fixed number of colors we can use. Parreau [2] used Jackson's work [3] on universal cycles to give a lower bound of this length. In this article, we will adapt what Jackson did, to improve this result.

Keywords: Graph Theory, Identifying codes, Identifying coloring, Cycles, Universal Cycles

2010 MSC: 05C15, 05C38

1. Introduction

In this article, we will build a globally identifying coloring of a cycle. This coloring is a variant of identifying codes.

Definition 1 (Identifying code). *An identifying code C of a graph $G = (V, E)$ is a subset of V such that:*

$$(1) \forall x \in V, C \cap N[x] \neq \emptyset$$

$$(2) \forall x, y \in V, x \neq y, C \cap N[x] \neq C \cap N[y]$$

where $N[x]$ is the closed neighborhood of x , that is to say the neighbors of x plus x itself.

Definition 2. *The minimum size of an identifying code of a graph G is denoted by $\gamma_{id}(G)$.*

These codes were introduced by Karpovsky, Chakrabarty and Levitin [1] to solve the problem of fault diagnosis of multiprocessor systems. The processors can be seen as the nodes of a graph, while the edges are the communication links between processors. Each processor may run a software routine which can detect an error in the processor itself or one of its neighborhood. The subset of processors

Email address: pierre.coupechoux@laas.fr (Pierre Coupechoux)

which will run this routine must be a identifying code of the graph. The condition 1 of the Definition 1 expresses that if an error occurs, we want to detect it. The condition 2 allows to localize where the problem happened; the error is detected by some processors running the routine, and this subset of processors is unique. The aim is to have the minimum number of processors running the routine, that is to minimize $|C|$.

Remark that these codes can only exist if there are no twin vertices in the graph. Indeed, if there are two vertices x and y such that $N[x] = N[y]$, then it is impossible to satisfy the condition 2 of the Definition 1. Therefore we will only consider graph without twin vertices.

We know some bounds of $\gamma_{id}(G)$.

Proposition 1 ([1]). *For a graph G with n vertices, $\gamma_{id}(G) \geq \log_2(n + 1)$.*

There exist several variants of identifying codes. For instance we could expect to identify subsets of vertices rather than vertices [4, 5, 6]. We could also require to identify vertices at distance at most r : these are called r -identifying codes [7, 8, 9]. In this paper we focus on an extension of identifying codes to colorings.

Definition 3 (Coloring). *A coloring c of a graph $G = (V, E)$ is a function from the set of vertices V to a set of colors, represented by \mathbb{N} .*

Given a subset of vertices $S \subseteq V$, $c(S) = \{c(x) : x \in S\}$ is the set of the colors of vertices in S .

Definition 4 (Globally identifying coloring). *A globally identifying coloring of a graph $G = (V, E)$ is a coloring c of this graph, such that:*

$$\forall x, y \in V, x \neq y, \quad c(N[x]) \neq c(N[y])$$

Remark that the coloration does not need to be proper. That is to say there can be two neighbors with the same color.

Definition 5. *The minimum number of colors of a globally identifying coloring of a graph G is denoted by $\chi_{id}(G)$.*

There exist links between identifying codes and globally identifying colorings.

Theorem 1. *For a graph G , $\chi_{id}(G) \leq \gamma_{id}(G) + 1$.*

Proof. Let C be an identifying code of size of G . We can build a globally identifying coloring of G as follows: all the vertices in C have a unique color in $\{1, \dots, |C|\}$; the others have the color 0. This coloring uses $|C| + 1$ colors and is indeed a globally identifying coloring. Let x and y be two vertices of G . The code C is identifying so $C \cap N[x] \neq C \cap N[y]$. Thus there exists a vertex z in the symmetric

difference of $C \cap N[x]$ and $C \cap N[y]$, that is z is in the code and in the neighborhood of only one vertex among x and y . The vertex z has a unique color, so $c(N[x]) \neq c(N[y])$ as $c(z)$ is in the symmetric difference. \square

This bound is tight for some graphs. For example, as we can see in Figure 1, $\gamma_{id}(P_3) = 2$ and $\chi_{id}(P_3) = 3$.

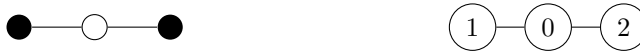


Figure 1: An identifying code and a globally identifying coloring of a P_3 .

These globally identifying colorings also exist in a local variant.

Definition 6 (Locally identifying coloring). *A locally identifying coloring of a graph $G = (V, E)$ is a coloring c of this graph, such that:*

$$\forall (x, y) \in E, \quad c(N[x]) \neq c(N[y])$$

In a locally identifying coloring, we only need to identify adjacent vertices.

Definition 7. *The minimum number of colors of a locally identifying coloring of a graph G is denoted by $\chi_{lid}(G)$.*

Cycles have been investigated for identifying codes [8] and r -identifying codes [7, 10], as well as for locally identifying colorings [11].

Theorem 2 ([11]). *Let $n \geq 4$ be an integer, and C_n be the cycle of length n . Then:*

- $\chi_{lid}(C_n) = 3$ if $n \equiv 0 \pmod{4}$
- $\chi_{lid}(C_n) = 5$ if $n = 5$ or $n = 7$
- $\chi_{lid}(C_n) = 4$ else.

Theorem 2 gives us the minimum number of colors we need to build a locally identifying coloring of a cycle, but there are no equivalent yet with a globally identifying coloring.

Definition 8 (Identified cycle). *An identified cycle is a cycle C and a globally identifying coloration c of C .*

In this paper, we will use a number of colors $L \geq 6$. Smaller values of L will be considered at the end of this article. We want to build a cycle as long as possible which admits a globally identifying coloring with at most L colors. In her PhD dissertation, Parreau [2] noticed that universal cycles are

identified cycles. Thus, Jackson’s construction for universal cycles can be used to get identified cycles. In this paper we will focus on Jackson’s construction in order to modify and improve it to get directly identified cycles.

Definition 9 ([3]). *A k -universal cycle of $\binom{L}{k}$ is a cycle where the vertices are labeled by numbers from 0 to $L - 1$ and which contains each k combination of $\binom{L}{k}$ as a unique sub-sequence of k consecutive vertices.*

A 3-universal cycle of $\binom{L}{3}$ can be seen as an identified cycle, because the closed neighborhood of a vertex is a sub-sequence of 3 consecutive vertices.

Theorem 3 ([3]). *If $\gcd(n, 3) = 1$ and $n \geq 8$, there is a universal cycle of $\binom{L}{3}$.*

Corollary 1 ([2]). *If $\gcd(L, 3) = 1$ and $L \geq 8$, the maximum length of an identified cycle with L colors is at least $\binom{L}{3}$.*

Proof. Let C be an universal cycle of $\binom{L}{3}$. We build a coloration of the same cycle simply by taking the label of the vertex as its color. As each sub-sequence of 3 labels of consecutive vertices is unique, the coloration is a globally identifying coloring. \square

These 3-universal cycles do not consider repetition of a label in a 3 sub-sequence of consecutive vertices (by definition, all the labels in a 3 sub-sequence are distinct). Thus, this result may not be perfect in the case of globally identifying coloring, as we may have the same color used up to 3 times in a 3 sub-sequence in an identified cycle. The example in Figure 2 shows this. We can verify all the vertices are identified by the colors in their neighborhood. For example the top vertex only have the color 0 in its neighborhood, and it is the only vertex to have this property.

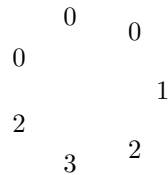


Figure 2: Example of a globally identifying coloring of a cycle of length 7 with 4 colors.

Jackson also built k -universal cycles with repetitions, where the labels in k sub-sequences do not need to be different. However, for $k = 3$, these cycles could not be seen as globally identifying colorings, because they include all 3 combinations with repetition of $\binom{L}{3}$. For example, $(0, 0, 1)$ and $(0, 1, 1)$ are two of these 3 combinations but they would not lead to a globally identifying coloring since both vertices would be surrounded by the set of colors $\{0, 1\}$.

Lemma 1 ([2]). *An upper bound of the size of an identified cycle with at most L colors is $\frac{L^3+5L}{6}$.*

Proof. Each vertex x of the cycle has exactly two neighbors, so $N[x]$ is of cardinality 3. Therefore $c(N[x])$ is of cardinality at most 3 (a single color, two different colors, or three different colors). These colors can take L different values, so the number of possibilities of $c(N[x])$ is

$$\binom{L}{3} + \binom{L}{2} + \binom{L}{1} = \frac{L^3 + 5L}{6}$$

□

2. Types and transition digraph

In order to build bigger identified cycles than those obtained from Jackson's construction as in Corollary 1, we expand the method Jackson used in his proof, including some repetitions. The main idea is that when we look at one vertex, what is really important is its neighborhood, so a sub-sequence of 3 colors. We will sort these sub-sequences into equivalent classes. Two different sub-sequences in the same class will be distinguished by the color of the first of the three vertices. Then we will build some paths which do not contain two 3 sub-sequences of colors in the same class. Then we will merge these paths to obtain a cycle (solving problems that may occur).

We note a 3 sub-sequence of colors (of three consecutive vertices) by (a, b, c) . From the point of view of globally identifying coloring, we look at the set $\{a, b, c\}$, regardless of the order and repetition ((0, 0, 1), (0, 1, 0), (0, 0, 1), and (1, 0, 1) are different 3 sub-sequences but give the color set $\{0, 1\}$ to the middle vertex).

As Jackson did in his proof, we will assign a *type* to each 3 sub-sequence, using the differences between the colors of the sub-sequence. The type of a sub-sequence is its equivalent class.

Definition 10. *Given a 3 sub-sequence (a, b, c) , we consider the differences $d_1 = b - a$, $d_2 = c - b$ and $d_3 = a - c$ modulo L . We use only two of these three differences to define the type of the sub-sequence, denoted $\langle t, u \rangle$, according to these rules :*

- *If the three are different, we keep the lowest two, starting with the following of the greatest one. That means, if d_2 is the greatest one, we keep d_3 and d_1 ; that is to say the type is $\langle d_3, d_1 \rangle$*
- *Else there is at least two differences equal to a difference d . We keep in this case twice d ; that is to say the type is $\langle d, d \rangle$.*

These two differences are the type of (a, b, c) , denoted by $\langle t, u \rangle$.

Remark that when we have the type of a 3 sub-sequence, we can retrieve the three differences (unordered), as the sum of the three is equal to L . We arbitrary choose to consider only the 3 sub-sequence so that the first two differences match the two chosen in the definition of the type. We ignore

$\langle 0, 0 \rangle$	$\langle 0, 1 \rangle$	$\langle 0, 2 \rangle$
$\langle 0, 3 \rangle$	$\langle 1, 0 \rangle$	$\langle 1, 1 \rangle$
$\langle 1, 2 \rangle$	$\langle 1, 3 \rangle$	$\langle 2, 0 \rangle$
$\langle 2, 1 \rangle$	$\langle 2, 2 \rangle$	$\langle 3, 0 \rangle$
$\langle 3, 1 \rangle$	$\langle 3, 3 \rangle$	$\langle 4, 4 \rangle$

Figure 3: List of existing types for $L = 8$

all other permutation of the three same values. For example, if we have the type $\langle 1, 2 \rangle$ and the initial color 0, this give the 3 sub-sequence $(0, 1, 3)$ as $0 + \mathbf{1} = 1$ and $1 + \mathbf{2} = 3$. All other permutation of $(0, 1, 3)$, such as $(1, 0, 3)$ or $(1, 3, 0)$, will not appear in the following construction. Each 3 sub-sequence we consider is therefore exactly represented by its type and an initial color (to give of the first of the three vertices). This is for this reason we made the arbitrary choice.

As previously stated, two different 3 sub-sequences can have the same set of colors (for example, $(0, 0, 1)$ and $(0, 1, 1)$ have color set $\{0, 1\}$).

Definition 11. *Given a number of colors L , we say the type $\langle u, v \rangle$ **exists** if there exists 3 colors a, b and c such that the sub-sequence (a, b, c) is of type $\langle u, v \rangle$.*

Lemma 2. *The type $\langle j, k \rangle$ exists if and only if one of the two statements holds :*

1. $j = k$ and $j \leq \frac{L}{2}$
2. $2j < L - k$ and $2k < L - j$

Proof. Assume the type $\langle j, k \rangle$ exists. Then, there exists a 3 sub-sequence $(i, i + j, i + j + k)$ which has type $\langle j, k \rangle$. The consecutive differences are j, k and $L - j - k$. These differences give a type $\langle j, k \rangle$ only if one of the statements is valid. Now, if $j = k$ and $j < \frac{L}{2}$, $\{0, j, 2j\}$ is indeed of type $\langle j, k \rangle$. If $j = k$ and $j = \frac{L}{2}$ then $\{0, 0, \frac{L}{2}\}$ fits. Finally, if $2j < L - k$ and $2k < L - j$, then $\{0, j, j + k\}$ is of type $\langle j, k \rangle$ because the third difference $L - j - k$ is the greatest of the three. \square

Remark that if the type $\langle j, k \rangle$ exists (that means the 3 sub-sequence $(0, j, j + k)$ has type $\langle j, k \rangle$), then the type $\langle k, j \rangle$ also exists ($(0, k, k + j)$ has type $\langle k, j \rangle$). Figure 3 shows all the existing types for $L = 8$.

Given L and a type, we can enumerate how many color sets are of this type.

Lemma 3. *Let $L \in \mathbb{N}$ and type $\langle u, v \rangle$ be an existing type.*

1. *If L is even, the type $\langle \frac{L}{2}, \frac{L}{2} \rangle$ represents $\frac{L}{2}$ different color sets.*

2. If L is a multiple of 3, the type $\langle \frac{L}{3}, \frac{L}{3} \rangle$ represents $\frac{L}{3}$ different color sets.
3. The types $\langle 0, j \rangle$ and $\langle j, 0 \rangle$ represent the same L color sets.
4. In all other cases, when type $\langle j, k \rangle$ exists, it represents L different colors sets, and these sets can not be obtained from another type.

When a type represents less than L different color sets, we say this type is **incomplete**.

- Proof.*
1. The sets represented by this type only have two different colors, i and $i + \frac{L}{2}$. Among all possible initial values i , only $\frac{L}{2}$ gives different sets, because for i in $[\frac{L}{2}, \dots, L - 1]$, we obtain the set $\{i, i + \frac{L}{2}\}$ which is equal to the set $\{i - \frac{L}{2}, i\}$ obtained from $i - \frac{L}{2}$.
 2. For i in $[0, \dots, \frac{L}{3} - 1]$, the choice of the initial value as $i, i + \frac{L}{3}$ or $i + 2\frac{L}{3}$ gives the same 3 colors set $\{i, i + \frac{L}{3}, i + 2\frac{L}{3}\}$.
 3. The 3-tuples which lead to type $\langle 0, j \rangle$ (respectively $\langle j, 0 \rangle$) are of the form $\{i, i, i + j\}$ (respectively $\{i, i + j, i + j\}$). In both cases, they form a 2 colors set $\{i, i + j\}$.
 4. The 3-tuple starting with i is $\{i, i + j, i + j + k\}$. $j \neq 0$ and $k \neq 0$ so $i \neq i + j$ and $i + j \neq i + j + k$. The only possibility to have $i = i + j + k$ would be if $j + k = L$, but in this case, $j = k = \frac{L}{2}$ (or else it would not be a type $\langle j, k \rangle$), which is a previous case of this lemma.

□

Figure 4 shows some of existing types for $L = 8$, with, for each type, the only 3 sub-sequence with initial color 0 and its color set. As we saw in Lemma 3, types $\langle 0, 2 \rangle$ and $\langle 2, 0 \rangle$ have different 3 sub-sequences which have the same color sets.

Type	3 sub-sequence with initial color 0	color set
$\langle 4, 4 \rangle$	(0, 4, 0)	{0, 4}
$\langle 0, 2 \rangle$	(0, 0, 2)	{0, 2}
$\langle 2, 0 \rangle$	(0, 2, 2)	{0, 2}
$\langle 1, 2 \rangle$	(0, 1, 3)	{0, 1, 3}
$\langle 2, 1 \rangle$	(0, 2, 3)	{0, 2, 3}

Figure 4: Examples of some types with the 3 sub-sequences with initial color 0 and their color set, for $L = 8$.

We use the types previously defined to build a particular digraph as follows.

Definition 12. Given a set of types, the transition digraph $G = (V, E)$ of this set verifies:

- V is the set of all integers that appear in at least one type.

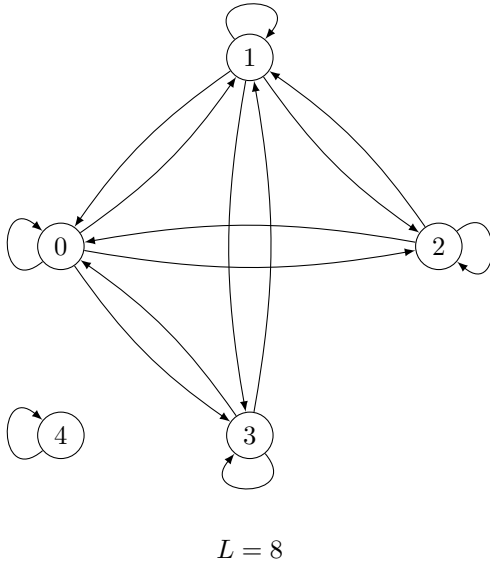


Figure 5: Transition digraph of all types for $L = 8$.

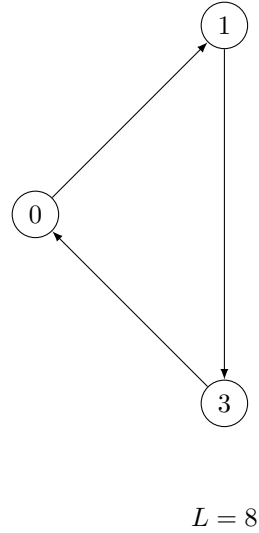


Figure 6: Transition digraph from E_L^1 for $L = 8$.

- $E = \{(j, k) : \langle j, k \rangle \text{ is an existing type}\}$

Figure 5 shows the transition digraph obtained from all the types of Figure 3.

A path on the transition digraph matches a sequence of types, from which we can build a sequence of colors, choosing an initial color for the first type. The initial colors of the other types will be determined by the previous type. For example, the path starting at vertex 0, followed by vertices 1, 2 and 1 gives us the types $\langle 0, 1 \rangle$, $\langle 1, 2 \rangle$ and $\langle 2, 1 \rangle$. Starting from arbitrary color 0, this matches the sequence of colors 00134. The first 3 sub-sequence $\langle 0, 0, 1 \rangle$ is indeed of type $\langle 0, 1 \rangle$, $\langle 0, 1, 3 \rangle$ is of type $\langle 1, 2 \rangle$, and $\langle 1, 3, 4 \rangle$ is of type $\langle 2, 1 \rangle$.

The approach to get an identified cycle is to choose an Eulerian circuit of a transition digraph. Since each type corresponds to an arc of this graph, then each type will appear exactly once. Then, starting with an arbitrary color i_0 , we will get a sequence of colors starting with i_0 and finishing with i_1 . Then we will repeat this process starting from the previous ending point i_1 , and iterate as long as we don't finish with i_0 . Then we will have a sequence of colors we can see as a cycle by "merging" the start and the end. This process will be explained in subsection 3.2, and Figures 9 and 10 show the result for $L = 9$.

However, several problems may occur in this approach. The transition digraph may not be Eulerian, and there may be redundant or incomplete types. The graph shown in Figure 5 is not Eulerian (as it is not connected). In order to have an Eulerian transition digraph, we will choose which types we keep, and which ones we discard. We will start with types that give an Eulerian graph of types, and then add other types which do not change this property. Moreover, we do not want to have incomplete

types (less than L sub-sequences of this type, such as type $\langle \frac{L}{2}, \frac{L}{2} \rangle$ according to Lemma 3) or redundant types (different types with different sub-sequences, but same color sets ; such as types $\langle 0, 1 \rangle$ and $\langle 1, 0 \rangle$, according to Lemma 3).

Initially, we choose the following types, for a given number of colors L :

$$\langle 0, k \rangle, \langle k, m_k \rangle, \text{ and } \langle m_k, 0 \rangle, \text{ with} \\ m_k = \left(\left\lfloor \frac{L}{2} \right\rfloor - k \right), \forall k \in \mathbb{N} \text{ where } 0 < 2k < \left\lfloor \frac{L}{2} \right\rfloor. \quad (1)$$

We denote the list of these types by E_L^1 . Remark that there are respectively $\frac{L}{4} - 1$, $\frac{L-1}{4} - 1$, $\frac{L-2}{4}$ and $\frac{L-3}{4}$ values of k in equation 1 for $L \equiv 0, 1, 2, 3 \pmod{4}$. All these types exist, according to Lemma 2.

Lemma 4. *The transition digraph obtained from E_L^1 is Eulerian.*

Proof. An Eulerian circuit on this graph is $0, 1, m_1, 0, 2, m_2, 0, \dots, 0$. Indeed, the arcs of this circuit are $(0, 1)$, $(1, m_1)$, $(m_1, 0)$, \dots , $(0, i)$, (i, m_i) , $(m_i, 0)$, \dots . These are exactly the arcs obtained from E_L^1 . \square

The Figure 6 shows an example of this graph for $L = 8$.

Then we add to E_L^1 all pairs of types $\langle j, k \rangle$ and $\langle k, j \rangle$ for $j \neq 0$, $k \neq 0$ and $j \neq k$ which do not appear in E_L^1 . We denote by E_L^2 this new list of types.

Lemma 5. *The graph of type obtained from E_L^2 is Eulerian.*

Proof. This graph can be seen as the graph obtained from E_L^1 to which we add some pairs of symmetric arcs (and vertices if needed). Thus the new graph is balanced (each vertex has its outdegree equals its indegree). Let us show it is strongly connected. Let $\langle j, k \rangle$ and $\langle k, j \rangle$ be two types we add. Assume $j < k$. There are 3 cases :

- If $2j < \lfloor \frac{L}{2} \rfloor$, then the vertex j was already present, through the type $\langle j, m_j \rangle$.
- If $2j = \lfloor \frac{L}{2} \rfloor$, then $2k > \lfloor \frac{L}{2} \rfloor$. Let $k' = \lfloor \frac{L}{2} \rfloor - k$. $2k' < 2 \lfloor \frac{L}{2} \rfloor - \lfloor \frac{L}{2} \rfloor$, that is $2k' < \lfloor \frac{L}{2} \rfloor$. Moreover, $m_{k'} = k$, so vertex k already exists, through type $\langle k', k \rangle$.
- If $2j > \lfloor \frac{L}{2} \rfloor$, then let $j' = \lfloor \frac{L}{2} \rfloor - j$. As in the previous case, we can show that $m_{j'} = j$ and that vertex j was already in the graph.

In all cases, at least one of the vertices j and k was already present in the graph obtained from E_L^1 . Thus the strong connectivity of the graph stands when we add types $\langle j, k \rangle$ and $\langle k, j \rangle$ in the graph. \square

Then we add to E_L^2 all types $\langle j, j \rangle$ for $0 \leq j < \lfloor \frac{L}{2} \rfloor$ and $j \neq \frac{L}{3}$. We denote by E_L^3 this new list. Figure 7 shows the transition digraph obtained from E_L^3 for $L = 8$.

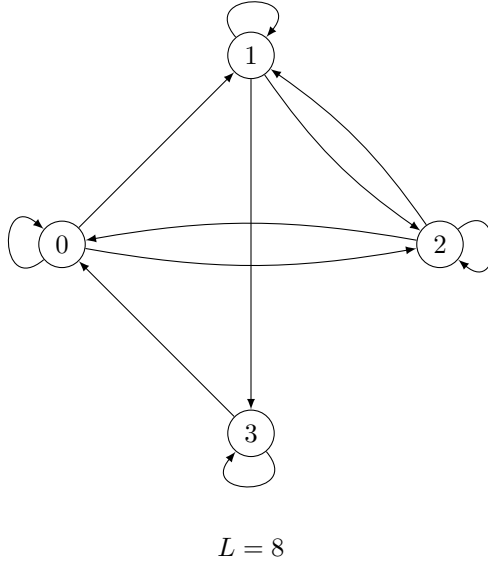


Figure 7: Transition digraph from E_L^3 for $L = 8$.

Lemma 6. *The graph of types obtained from E_L^3 is Eulerian.*

Proof. This graph can be seen as the graph obtained from E_L^2 to which we add some loops, thus the graph is still balanced. If $j \leq 1$, we know the vertex j already exists. If $2 \leq j < \frac{L}{2}$, then the types $\langle 1, j \rangle$ and $\langle j, 1 \rangle$ exist, according to Lemma 2, so the type $\langle 1, j \rangle$ is in E_L^1 if $j = m_1$ or else in E_L^2 , so the vertex j is already present. \square

3. Construction of the identified cycle

3.1. Construction of identified cycles by merging paths

Since the graph obtained from E_L^3 is Eulerian, there exists an Eulerian circuit. We can exhibit one as follow. We start from the circuit seen in the proof of Lemma 4. For all arcs added by E_L^2 , we add the pair of arcs (j, k) and (k, j) for $j < k$ and $j \neq \lfloor \frac{L}{2} \rfloor$ the first time we come through vertex j (for $L = 8$, adding arcs $(1, 2)$ and $(2, 1)$ in the circuit 013 results in 01213), and the pair of arcs (k, j) and (j, k) for $j < k$ and $j = \lfloor \frac{L}{2} \rfloor$ the first time we come through vertex k . Finally, we add all the arcs of the form (j, j) the first time we come through vertex j .

Examples of circuit we obtain for different values of L are shown in Figure 8.

As previously stated, we will use this Eulerian circuit to build a sequence of types. We conveniently choose the beginning of the circuit just before the loop edge $(1, 1)$. Given an initial value and this circuit, we can build a colored path which follows the types on the circuit (while this path is not closed to form a cycle, we will not consider the end vertex, which still misses one neighbor). For example, for

L	Circuit of transition digraph obtained from E_L^3
8	0 0 1 1 2 2 1 3 3
9	0 0 1 1 2 2 1 3 2 3
10	0 0 1 1 2 2 1 3 3 1 4 4 0 2 3
11	0 0 1 1 2 2 1 3 3 1 4 4 0 2 4 2 3

Figure 8: Examples of Eulerian circuits of transition digraph obtained from E_L^3 for different values of L .

$L = 8$ and the circuit exposed in Figure 8 (112213300), with an initial color of 0, we get the colored path 0124672555. We can verify the first 3 sub-sequence 012 is of type $\langle 1, 1 \rangle$, the second one 124 is of type $\langle 1, 2 \rangle$, and so on. The only missing type $\langle 0, 1 \rangle$ – the last one – will appear later. We denote by n the total difference between the first and the last colors (modulo L). In this example, $n = 5 - 0 = 5$.

Now, we take the ending color as the new initial value to create a new colored path, which can be merged with the previous one. These two paths contained the same types, in the same order, but with different initial values. The operation of merging consists here of concatenation of the first and second paths, where the first vertex of second path has been removed. As we chose an Eulerian circuit, this operation is possible. In the example above, merging the two paths 0124672555 and 5671347222 results in 0124672555671347222. This adds the last type of the cycle ($\langle 0, 1 \rangle$) of the first path, which had been previously removed.

We can repeat this process, as long as the ending value we obtain is different from the first initial value. Then, we can cycle to the first vertex, what will add the type $\langle 0, 1 \rangle$ to the last path. On the end, there are two possibilities. If $\gcd(n, L) = 1$, then all values of $0, L - 1$ were reached as initial colors of the paths, before getting 0 again. In this case, all 3 sub-sequences represented by a type of E_L^3 are present exactly once in the cycle. Else, if $\gcd(n, L) = k (\neq 1)$, we only reach $\frac{L}{k}$ initial colors, the multiples of n modulo L .

The second case occurs for instance with $L = 9$. By construction, we obtain $n = 6$, so $\gcd(n, L) = 3$, and the initial colors successively are 0, 6 and 3. In order to get a larger cycle, we can move some vertices in sub-sequences of type $\langle 1, 1 \rangle$ as follows.

3.2. Unification of the cycles when $\gcd(n, L) \neq 1$

Denote by P_i the path obtained, starting with initial value i , and iterating the previous algorithm until it ends with color i . We consider k independent paths like this (with no identical 3 sub-sequences) among the L possibilities, namely $P_0 = P_L, P_{L-1}, \dots, P_{L-k+1}$. Denote by P'_i the path obtained from P_i , removing the first vertex. This path contains the same 3 sub-sequences as P_i , except for the first one $(i, i+1, i+2)$, of type $\langle 1, 1 \rangle$. This path starts with initial color $i+1$ and ends with color i . Thus, we can merge the path P_i and P'_{i-1} , as illustrated in Figure 10. We can use this to merge P_0 with P'_{L-1} ,

Differences	+1	+1	+2	+2	+1	+3	+2	+3	+0	+0
$i_0 = 0$	1	2	4	6	7	1	3	6	6	(6)
$i_1 = 6$	7	8	1	3	4	7	0	3	3	(3)
$i_2 = 3$	4	5	7	0	1	4	6	0	0	(0)
$i'_0 = 8$	0	1	3	5	6	0	2	5	5	(5)
$i'_1 = 5$	6	7	0	2	3	6	8	2	2	(2)
$i'_2 = 2$	3	4	6	8	0	3	5	8	8	(8)
$i''_0 = 7$	8	0	2	4	5	8	1	4	4	(4)
$i''_1 = 4$	5	6	8	1	2	5	7	1	1	(1)
$i''_2 = 1$	2	3	5	7	8	2	4	7	7	(7)

Figure 9: Paths obtained from the Eulerian circuit for $L = 9$.

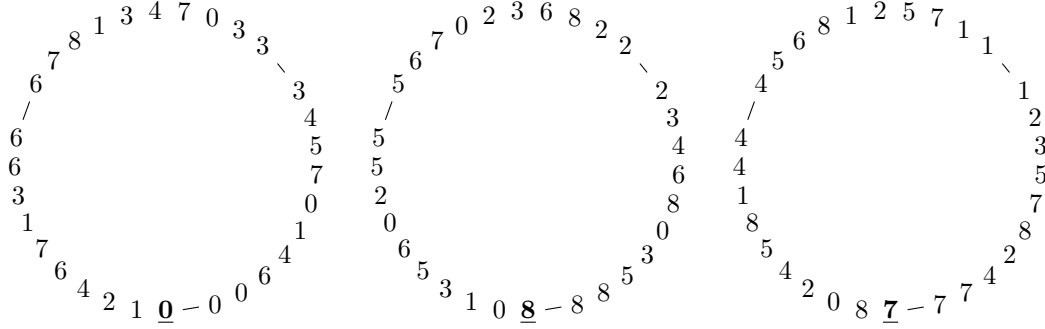


Figure 10: Cycles obtained by merging the paths of Figure 9 for $L = 9$.

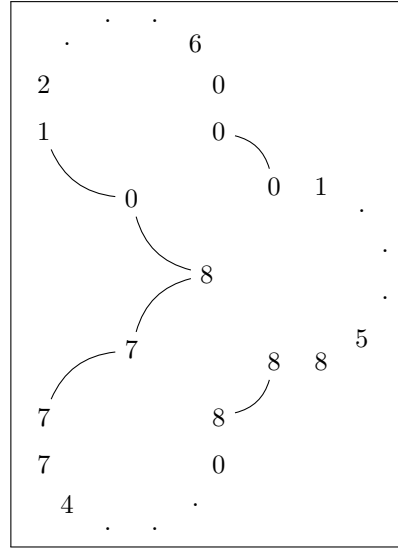
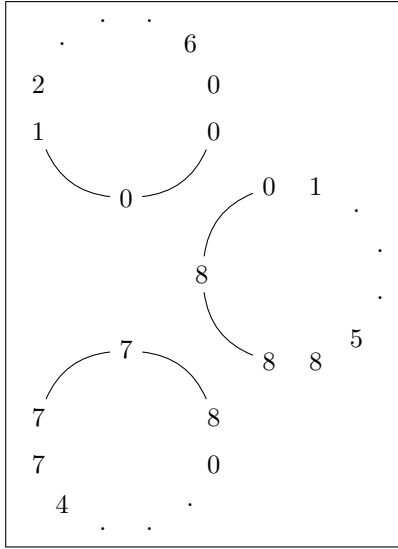
$P'_{L-2}, \dots, P'_{L-k+1}$ in this order. This new path starts with initial color 0 and ends with $L - k + 1$. We can append to this path all the colors $L - k + 2, \dots, L - 1, 0$, in this order. By doing this, we add the 3 sub-sequences of type $\langle 1, 1 \rangle$ we previously removed. Figure 11 shows this merging for $L = 9$. We can verify that each vertex has exactly the same colors in its neighborhood, so there is no two identical 3 sub-sequence.

In all cases, we obtain a cycle that contains all 3 sub-sequences represented by any type in E_L^3 . Figure 12 shows the cycle we obtain for $L = 9$.

3.3. Odd number of colors

When L is odd, we can improve the number of vertices in the cycle, using redundant types $\langle 0, \frac{L-1}{2} \rangle$, $\langle \frac{L-1}{2}, 0 \rangle$ and type $\langle \frac{L-1}{2}, \frac{L-1}{2} \rangle$. We insert one 3 sub-sequence of each of the first two types, and $L - 1$ of the third type. We build a path with these 3 sub-sequences as follows:

- The first vertex is 0, and the first type is $\langle 0, \frac{L-1}{2} \rangle$ (so the first three values are 0, 0 and $\frac{L-1}{2}$).



3 independent cycles, as shown in Figure 10.

Shifting 1 vertex per cycle.

Figure 11: Merging of cycles obtained in subsection 3.2 for $L = 9$.

```

0 1 2 4 6 7 1 3 6 6 6 7 8 1 3 4 7 0 3 3 3 4
8                                     5
7 7 7 4 2 8 7 5 3 2           3 1 0 0 0 6 4 1 0 7
                                     1   5
4 5 6 8 1 2 5 7 1 1           6 0 2 5 5 5 6 7 0 2
4                                     3
4 1 8 5 4 2 0 8 8 8 5 3 0 8 6 4 3 2 2 2 8 6

```

Underlined vertices are the first vertices of P_8 and P_7 , removed in P'_8 and P'_7 . **Bold** ones show the passage in a "new cycle" (in the sense of Figure 10).

Figure 12: Cycle obtained in subsection 3.2 for $L = 9$.

- We append $L - 1$ values to form triplets of type $\langle \frac{L-1}{2}, \frac{L-1}{2} \rangle$.
- We append the value 0 to the path, to form the type $\langle \frac{L-1}{2}, 0 \rangle$.

By this construction, we have $L - 1$ 3 sub-sequences of types $\langle \frac{L-1}{2}, \frac{L-1}{2} \rangle$, with initial colors being the multiples of $\frac{L-1}{2}$. As $\gcd(L, \frac{L-1}{2}) = 1$, all these 3 sub-sequences are different. The third color of the last 3 sub-sequence is 0, because the initial color of the path is 0 and we add L times $\frac{L-1}{2}$ to this color to have $L - 1$ 3 sub-sequence of type $\langle \frac{L-1}{2}, \frac{L-1}{2} \rangle$. Thus, as we appended 0, the last 3 sub-sequence is of type $\langle \frac{L-1}{2}, 0 \rangle$.

We can verify that the two 3 sub-sequences of type $\langle 0, \frac{L-1}{2} \rangle$ and $\langle \frac{L-1}{2}, 0 \rangle$ do not represent the same set of colors: the first one is $\{0, \frac{L-1}{2}\}$, and the second one is $\{0, \frac{L+1}{2}\}$.

This path starts and ends with two zeros. Thus we can insert it in the previously obtained cycle in place of the two last colors of the 3 sub-sequence $(0, 0, 0)$ (which is indeed in the cycle from subsection 3.2, at the end of the first cycle of subsection 3.1).

For example, with $L = 9$, the path is 004837261500. We replace 0001 in the cycle of Figure 12 by 00048372615001.

4. Results

4.1. Case $L \geq 6$

Theorem 4. *The method explained in subsections 3.1, 3.2 and 3.3 builds, with L colors, an identified cycle of length:*

$$\frac{L^3 + 5L}{6} - \left\{ \begin{array}{ll} \frac{L+2}{4}L & \text{if } L \equiv 0 \pmod{4} \\ \frac{L+3}{4}L - 1 & \text{if } L \equiv 1 \pmod{4} \\ \frac{L^2}{4} & \text{if } L \equiv 2 \pmod{4} \\ \frac{L+1}{4}L - 1 & \text{if } L \equiv 3 \pmod{4} \end{array} \right\} - \left\{ \begin{array}{l} \frac{L}{3} \text{ if } L \equiv 0 \pmod{3} \\ 0 \text{ else} \end{array} \right\}$$

Proof. The total number of color sets is $\binom{L}{1} + \binom{L}{2} + \binom{L}{3} = \frac{L^3 + 5L}{6}$. We count the missing ones in the cycle obtained at subsection 3.3.

1. The type $\langle j, k \rangle$ for $j, k \neq 0$ and $j \neq k$ does not appear in E_L^3 if and only if $\langle k, j \rangle$ was already in E_L^1 (else we add $\langle j, k \rangle$ and $\langle k, j \rangle$ in E_L^3). There are respectively $\frac{L}{4} - 1$, $\frac{L-1}{4} - 1$, $\frac{L-2}{4}$ and $\frac{L-3}{4}$ such types for $L \equiv 0, 1, 2, 3 \pmod{4}$, according to construction of E_L^1 . Each of these types represents L color sets. Since all the types are different, all the represented color sets are different.
2. If L is even, then the type $\langle \frac{L}{2}, \frac{L}{2} \rangle$ is missing. This type represents $\frac{L}{2}$ color sets.

3. If $L \equiv 0 \pmod{3}$, then the type $\langle \frac{L}{3}, \frac{L}{3} \rangle$ is missing. It represents $\frac{L}{3}$ color sets.
4. According to Lemma 2, there are $\frac{L}{2}$ different types $\langle 0, k \rangle$ when L is even, $\frac{L-1}{2}$ when L is odd. There are respectively $\frac{L}{2} - 2$, $\frac{L-1}{2} - 2$, $\frac{L-2}{2}$ and $\frac{L-3}{2}$ types of the form $\langle 0, k \rangle$ or $\langle j, 0 \rangle$ ($j, k \neq 0$) for $L \equiv 0, 1, 2, 3 \pmod{4}$, by construction of E_L^1 . In E_L^3 , the type $\langle 0, 0 \rangle$ is always added. Thus the number of values of i such that types $\langle 0, i \rangle$ and $\langle i, 0 \rangle$ are both missing is 1 when L equals 0 or 1 modulo 4, 0 else (for example, if $L \equiv 0 \pmod{4}$, there are $(\frac{L}{2} - 2) + 1$ types out of $\frac{L}{2}$). Each of these pair of types represents L color sets (Lemma 3). The number of missing color sets of type $\langle 0, k \rangle$ or $\langle j, 0 \rangle$ is therefore L if $L \equiv 0, 1 \pmod{4}$, and 0 if $L \equiv 2, 3 \pmod{4}$.
5. When L is odd, we added $L + 1$ color sets during the last step, in subsection 3.3.

The types considered in point 1 to 4 are obviously different. They are all types represented in E_L^3 , on the contrary of types in point 5, so all the types from point 1 to 5 are different, and so are the color sets represented by these types.

For example, for $L \equiv 0 \pmod{4}$, we know that there are $\frac{L-4}{4}L$ color sets missing from point 1, $\frac{L}{2}$ from point 2 and L from point 4. Thus $\frac{L+2}{4}L$ color sets are missing. If $L \equiv 0 \pmod{3}$, there are $\frac{L}{3}$ additionally missing color sets from point 3. □

4.2. Small values of L

We perfectly know the size of identified cycles of maximal length for $L < 6$.

For $L = 1$, the only identified cycle is obviously a single vertex. There exists no identified cycle with $L = 2$ colors. This is quite straightforward: if such a cycle exists, it has at least 2 vertices (in order to have 2 colors), and no more than 3 vertices ($\frac{2^3+5*2}{6} = 3$). These two cycles have twin vertices, and can not admit a globally (or locally) identifying coloring. The following propositions deal with the cases $L = 3$ and $L = 4$, which are less straightforward.

Proposition 2. *There exists no identified cycle with $L = 3$ colors.*

Proof. There exist $7(= \frac{3^3+5*3}{2})$ color sets with at most 3 colors : $\{0\}$, $\{1\}$, $\{2\}$, $\{0, 1\}$, $\{0, 2\}$, $\{1, 2\}$ and $\{0, 1, 2\}$. The singleton color sets are represented in an identified cycle by three consecutive vertices of the same color. Remark that a 3 sub-sequence of three consecutive vertices of the same color can not follow such another 3 sub-sequence: if we have 6 consecutive vertices x_1 to x_6 , with $c(x_{i \leq 3}) = c_1$ and $c(x_{i > 3}) = c_2$, then x_3 and x_4 have the same colors in their neighborhood. That means that there is at least one vertex between each 3 sub-sequence representing a single color. Therefore, there is at most one such 3 sub-sequence (if there were at least 2, the cycle would have at least 8 vertices). This implies that only 5 of the 7 previously enumerated color sets can be represented in an identified cycle.

Now let us suppose the remaining singleton color set (for example $\{0\}$) is represented. There exists a 3 sub-sequence of consecutive vertices having the color 0. The two others vertices must have colors 1 and 2 (to have the 3 colors in the cycle). However, this does not lead to an identified cycle. So only 4 color sets can exist in an identified cycle (which is consequently of size at most 4). Finally, remark that you need 4 colors to build a globally identifying coloring on a cycle of length 4: if $c(x_i) = c(x_{i+1})$, then $c(N[x_{i+2}]) = c(N[x_{i+3}])$, and if $c(x_i) = c(x_{i+2})$, then $c(N[x_{i+1}]) = c(N[x_{i+3}])$. As we saw that there exists no identified cycle of length 3, this concludes the proof. \square

Proposition 3. *The size of an identified cycle of maximal length for $L = 4$ is 10.*

This result has been obtained by an exhaustive enumeration of all the cycles with 4 colors and at most 14 vertices. This enumeration has been made by a computer search. Note that we could use arguments similar to those in proof of Proposition 2. The cycle is 0121333200.

The last case, $L = 5$ colors, shows that the bound $\frac{L^3+5*L}{6}$ can be reached.

Proposition 4. *There exists an identified cycle of length 25 for $L = 5$ vertices. All the color sets of at most 3 colors among 5 are represented in this cycle.*

Proof. An example of such a cycle is 0124440133340222341112300. This cycle is obtained with the method described in this article, but with a slight modification in the definition of a type. If we use the Definition 10, the types are $\langle 0, 0 \rangle$, $\langle 1, 1 \rangle$, $\langle 2, 2 \rangle$, $\langle 0, 1 \rangle$ (or $\langle 1, 0 \rangle$) and $\langle 0, 2 \rangle$ (or $\langle 2, 0 \rangle$). Here we choose to replace the type $\langle 2, 2 \rangle$ by the type $\langle 1, 2 \rangle$. This choice only change the order of a 3 sub-sequence of this type. For example, instead of the 3 sub-sequence $(0, 2, 4)$ (obtained with initial color 0 and type $\langle 2, 2 \rangle$), we have the 3 sub-sequence $(4, 0, 2)$ (with initial color 4 and type $\langle 1, 2 \rangle$). With this definition, the transition digraph obtained from E_L^3 is Eulerian. That means there exists a circuit of this digraph which visits all edges exactly once. Using this circuit, we obtain the identified cycle previously exposed. This identified cycle is indeed optimal, as $25 = \frac{5^3+5*5}{6}$. \square

This proof is a possible idea to improve the results given in Theorem 4.

References

- [1] M. G. Karpovsky, K. Chakrabarty, L. B. Levitin, On a new class of codes for identifying properties in graphs, IEEE transactions on information theory 44.
- [2] A. Parreau, Problèmes d'identification dans les graphes, Ph.D. thesis, Institut Fourier (2012).
- [3] B. W. Jackson, Universal cycles of k -subsets and k -permutations, Discrete Mathematics 117 (1993) 141–150.

- [4] A. Frieze, R. Martin, J. Moncel, M. Ruszinkó, C. Smyth, Codes identifying sets of vertices in random networks, *Discrete Mathematics* 307 (2007) 1094–1107.
- [5] I. Honkala, T. Laihonen, On the identification of sets of points in the square lattice, *Discrete & Computational Geometry* 29 (1) (2002) 139–152.
- [6] I. Honkala, T. Laihonen, S. Ranto, On codes identifying sets of vertices in hamming spaces, *Designs, Codes and Cryptography* 24 (2) (2001) 193–204.
- [7] N. Bertrand, I. Charon, O. Hudry, A. Lobstein, Identifying and locating-dominating codes on chains and cycles, *European Journal of Combinatorics* 25 (2004) 969–987.
- [8] S. Gravier, J. Moncel, A. Semri, Identifying codes of cycles, *European Journal of Combinatorics* 27 (2006) 767–776.
- [9] I. Honkala, O. Hudry, A. Lobstein, On the number of optimal identifying codes in a twin-free graph, *Discrete Applied Mathematics* 180 (2015) 111–119.
- [10] V. Junnila, T. Laihonen, Optimal identifying codes in cycles and paths, *Graphs and Combinatorics* 28 (4) (2012) 469–481.
- [11] F. Foucaud, I. Honkala, T. Laihonen, A. Parreau, G. Perarnau, Locally identifying colourings for graphs with given maximum degree, *Discrete Mathematics* 312 (2012) 1832–1837.