Improved upper bounds for identifying codes in trees and other graphs

Florent Foucaud [1] — Université Clermont Auvergne, France
Tuomo Lehtilä [2] — University of Turku, Finland

Abstract

A dominating set C of a graph G is an identifying code if any two distinct vertices of G have distinct closed neighbourhoods within C. We prove new upper bounds for the size of the smallest (optimal) identifying codes in bipartite graphs which do not have twins of degree two or greater, and in graphs of girth at least 5. We show that $(n+\ell)/2$ is a tight upper bound for optimal identifying codes in the bipartite case, where n is the order and ℓ is the number of leaves of the graph. This bound is an improvement over two previous bounds for trees, in size and generality. Moreover, the bound is tight for several structurally different infinite families of trees. We derive from our bound a tight upper bound of $2n/3$ for twin-free bipartite graphs of order n, and give an exact characterization for graphs attaining it. Then we give an upper bound of $(5n+2\ell)/7$ for identifying codes in graphs of girth at least 5.

1 Introduction

We denote by $N(v)$ the *open neighbourhood* of vertex $v \in V(G)$ and by $N[v] = \{v\} \cup N(v)$ the *closed neighbourhood* of v. The *I-set* of vertex v is $I(v) = N[v] \cap C$ where $C \subseteq V(G)$. A set of vertices C is *dominating* if the I-set of each vertex is non-empty, i.e. $I(v) \neq \emptyset$ for each $v \in V(G)$. Furthermore, a dominating set C is an *identifying code* if any distinct vertices $v, u \in V(G)$ have distinct I-sets, that is, $I(v) \neq I(u)$. We call any vertex $v \in C$ a *codeword*. The smallest possible identifying code in graph G is called *optimal* and its cardinality is denoted by $\gamma^{\text{ID}}(G)$.

Vertex $v \in V(G)$ is a *leaf* if it has degree 1 and it is a *support vertex* if it has an adjacent leaf. We denote by $L(G)$ the set of leaves in G and by $S(G)$ the set of support vertices. The cardinalities of these sets are denoted by $|L(G)| = \ell(G)$ and $|S(G)| = s(G)$. Moreover, graph has *girth* $g(G)$ if the length of a shortest cycle in G is $g(G)$. A pair of vertices v, u is called *open twins* if $N(v) = N(u)$ and *closed twins* if $N[v] = N[u]$. A graph which does not have any open or closed twins is said to be *twin-free* and a graph without closed twins is said to be *identifiable*. Observe that if vertices v and u are closed twins, then they have $I(v) = I(u)$ for any subset of vertices C and the graph does not admit any identifying code.

Identifying codes were originally introduced over 20 years ago in [5] for multiprocessor architectures. Since then, they and related topics have been studied in numerous articles for different applications [6]. As identifying codes have been studied for over 20 years, naturally they have been considered in trees. In particular, the following two upper bounds have been presented. The bound in Theorem 1 is better when the number of leaves is small, and when there are many leaves, Theorem 2 gives the better bound.

Theorem 1 ([3, Theorem 15])**.** *Let T be a tree on $n \geq 3$ vertices. Then $\gamma^{ID}(T) \leq \frac{n+2\ell(T)-2}{2}$.*

Theorem 2 ([7, Theorem 11])**.** *Let T be a tree on $n \geq 3$ vertices. Then $\gamma^{ID}(T) \leq \frac{3n+2\ell(T)-1}{5}$. Equality holds if and only if $T = P_4$.*

[1]Research supported by the French government IDEX-ISITE initiative 16-IDEX-0001 (CAP 20-25) and by the ANR project GRALMECO (ANR-21-CE48-0004-01).

[2]Research supported by the Finnish cultural foundation and by the Academy of Finland grant 338797.

In Theorem 5, we offer a new improved upper bound $\gamma^{\mathrm{ID}}(G) \leq (n+\ell(G))/2$ for bipartite graphs which do not have any twins of degree two or greater. This new upper bound is not only more general but also an improvement even in the case of trees. This is rather surprising, considering the long history of study of identifying codes and the fact that the upper bounds presented in Theorems 1 and 2 are tight for some trees. Observe that we may present the bound in Theorem 2 as $\frac{3n+2\ell(T)-1}{5} = 3\frac{n-\ell(T)-1}{5} + \ell(T)$ and the bound in Theorem 5 as $(n-\ell(T))/2 + \ell(T)$. Now, our new bound is clearly an improvement over both previous bounds.

We also show the tight upper bound of $\gamma^{\mathrm{ID}}(G) \leq n - s(G)$ for triangle-free graphs in Lemma 1 and generalize it to every graph with the tight bound of $\gamma^{\mathrm{ID}}(G) \leq n - s(G) + 1$ in Theorem 4. Together these upper bounds give the tight upper bound $\gamma^{\mathrm{ID}}(G) \leq 2n/3$ for twin-free bipartite graphs, presented in Corollary 1. In Theorem 6, we characterize every graph attaining this bound.

Besides bipartite graphs, we also consider graphs of girth at least 5. Earlier in [1], the authors have provided the following upper bound for $\gamma^{\mathrm{ID}}(G)$ in these graphs when we have $\ell(G) = 0$.

Theorem 3 ([1])**.** *Let G be a graph of order n and girth at least 5 with minimum degree $\delta(G) \geq 2$. Then $\gamma^{ID}(G) \leq 5n/7$.*

The bound of Theorem 3 is tight for the 7-cycle. In Theorem 7, we show that when G has girth at least 5, we have $\gamma^{\mathrm{ID}}(G) \leq (5n + 2\ell(G))/7$ generalizing Theorem 3 to graphs of girth at least 5. The generalization is tight for stars and the cycle C_7. Full version of the paper is available at [2].

2 New upper bounds

We first present our new upper bounds based on the number of the support vertices. Then we give bounds using the number of leaves. After that, we combine these results to get a new upper bound as a corollary. Finally, we generalize a previous upper bound to every graph of girth at least 5.

The following lemma has been presented for total dominating identifying codes in trees in [4]. We generalize it for a larger class of graphs. In particular, it holds for bipartite graphs.

Lemma 1. *Let G be a connected graph on $n \geq 4$ vertices that is not the path P_4, such that $G - L(G)$ is identifiable or G is triangle-free. Then $\gamma^{ID}(G) \leq n - s(G)$.*

The bound in Lemma 1 is tight. We present some graphs attaining this bound in Theorem 6. Moreover, some restrictions on the structure are necessary. For example, if we consider the graph G consisting of the complete graph K_m and a single leaf added to every vertex in the clique, then we have $\gamma^{\mathrm{ID}}(G) = m + 1 = 2m - m + 1 = n - s(G) + 1$. As we show in the following theorem, this is actually a tight upper bound for every connected graph G.

Theorem 4. *Let G be a connected identifiable graph on $n \geq 3$ vertices. Then $\gamma^{ID}(G) \leq n - s(G) + 1$.*

Theorem 5 is based on the idea that if u is a non-leaf, non-codeword, then $N(u)$ forms a unique I-set. Similarly, in Theorem 7, if a non-codeword is at least 2-dominated, then it has unique I-set since there are no triangles or 4-cycles in the graph. Some restrictions on the graph structure are necessary. For example, Theorem 5 does not hold for odd cycles or for the 4-cycle C_4.

Theorem 5. *Let G be a connected bipartite graph on $n \geq 3$ vertices which does not have twins of degree two or greater. We have*

$$\gamma^{ID}(G) \leq \frac{n + \ell(G)}{2}.$$

Proof (sketch). The proof is based on first constructing two almost identifying codes C'_e and C'_o. Then we show that at least one of them has the claimed cardinality and after that one can shift some codewords to create an identifying code with the same cardinality.

We construct C'_e and C'_o in the following way: First we add every leaf to both codes. Then we choose some non-leaf vertex x and add every vertex at even distance from x to C'_e and every vertex at odd distance from x to C'_o. Either C'_e or C'_o contains at most half of the non-leaf vertices, that is, one of these codes has cardinality at most $\ell(G) + (n - \ell(G))/2$.

Moreover, a non-leaf, non-support vertex clearly has a unique I-set since either it is a non-codeword which is dominated by two codewords or it is a codeword dominated only by itself with adjacent non-codewords dominated by at least two codewords. However, we may have problems with some leaf-support vertex pairs. In particular, if there is a support vertex v which has exactly one adjacent leaf u and v is a codeword, then $I(v) = I(u) = \{v, u\}$. We can work around this problem by shifting the codeword in the leaf u to any vertex w adjacent to v. We denote the resulting codes by C_e and C_o. Now, $I(u) = \{v\}$ and $I(v) = \{v, w\}$ while $|I(w)| \geq 3$. The shifting and construction are presented in Figure 1. □

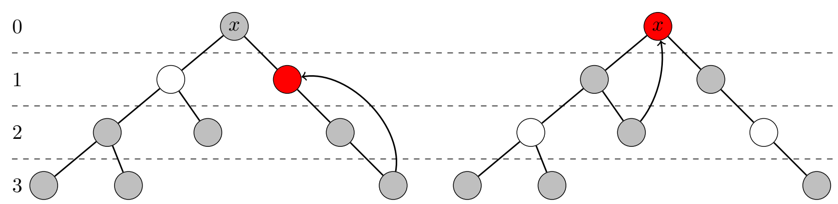

Figure 1: Gray vertices form the sets C'_e (left) and C'_o (right). Arrows and red vertices depict the shifts in forming the (non-optimal) identifying codes C_e and C_o.

Notice that Theorem 5 holds especially for trees. Moreover, the theorem is tight for quite a large class of graphs, for example, paths, even cycles (other than C_4), stars, 2-coronas presented in Theorem 6 and some other trees for which we have given examples in Figure 2.

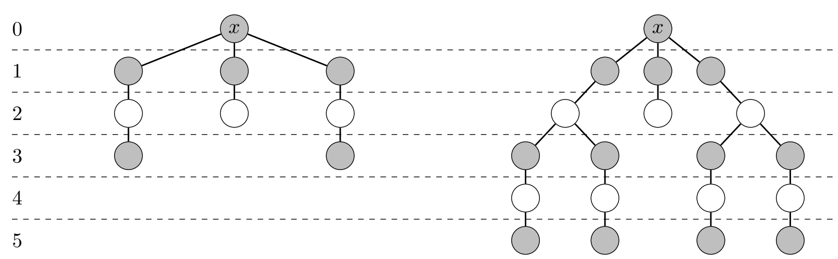

Figure 2: The gray vertices form optimal identifying codes in these two trees whose sizes are equal to the bound presented in Theorem 5.

When G is twin-free, we have $s(G) = \ell(G)$. Thus, Lemma 1 and Theorem 5 together provide the following corollary for twin-free bipartite graphs since $\gamma^{\mathrm{ID}}(G) \leq \min\{n - s(G), (n + s(G))/2\}$.

Corollary 1. *Let $G \neq P_4$ be a twin-free bipartite graph on $n \geq 3$ vertices. Then $\gamma^{ID}(G) \leq 2n/3$.*

The 2-corona $H \circ_2$ of graph H, is constructed by joining to every vertex of H a path P_2 with a single edge so that $\ell(H \circ_2) = |V(H)|$. In the following theorem, we give an exact characterization for graphs attaining the upper bound of Corollary 1.

Theorem 6. *Let G be a connected twin-free bipartite graph on n vertices. We have $\gamma^{ID}(G) = 2n/3$ if and only if G is the 2-corona $H \circ_2$ of some bipartite graph H.*

2.1 New upper bound for graphs of girth at least 5

As we have seen in Theorem 3, there exists a tight upper bound $\gamma^{\mathrm{ID}}(G) \leq 5n/7$ in graphs of girth at least 5 when $\ell(G) = 0$. Some constraint on the graph structure is necessary here since there exist graphs of lesser girth without leaves for which we have $\gamma^{\mathrm{ID}}(G) = n - 1$. In the following theorem, we generalize the result of Theorem 3 to every graph of girth at least 5. The bound does not hold for girth 4 since for the complete bipartite graph $K_{p,q}$ with $p, q \geq 3$, we have $\gamma^{\mathrm{ID}}(K_{p,q}) = n - 2$.

Theorem 7. *Let G be an identifiable graph with girth at least 5 without isolated vertices. Then $\gamma^{ID}(G) \leq (5n + 2\ell(G))/7$.*

Proof (sketch). The proof is based on induction on the number n of vertices. By Theorem 3, the bound holds when $\ell(G) = 0$ and by Theorem 5 the bound holds when we have no cycles. Thus, we may assume that there exists a vertex v which either belongs to a cycle or connects two cycles and there is a cut edge vu such that one of the components in $G - vu$ is a tree. The basic idea is to use Theorem 3 on the tree component and the induction hypothesis on the other components. After that, we may be required to do some small modifications to make sure that $I(u) \neq I(v)$. Some small tree components, such as P_2 are considered separately. □

3 Concluding remarks

We continue the study of identifying codes in bipartite graphs, trees and graphs of girth at least 5. Somewhat surprisingly, we manage to improve the known upper bounds for γ^{ID} in trees, both in generality and in size. In total, we give five new tight upper bounds for γ^{ID}.

References

[1] C. Balbuena, F. Foucaud and A. Hansberg. Locating-dominating sets and identifying codes in graphs of girth at least 5. *The Electronic Journal of Combinatorics* 22, P2.15, 2015.

[2] F. Foucaud and T. Lehtilä. Revisiting and improving upper bounds for identifying codes. Manuscript, 2022. `http://arxiv.org/abs/2204.05250`

[3] J. Gimbel, B. D. Van Gorden, M. Nicolescu, C. Umstead, and N. Vaiana. Location with dominating sets. *Congressus Numerantium*, 129-144, 2001.

[4] T. W. Haynes, M. A. Henning and J. Howard. Locating and total dominating sets in trees. *Discrete Applied Mathematics* 154(8):1293–1300, 2006.

[5] M. G. Karpovsky, K. Chakrabarty and L. B. Levitin. On a new class of codes for identifying vertices in graphs. *IEEE Transactions on Information Theory* 44:599–611, 1998.

[6] A. Lobstein. Watching systems, identifying, locating-dominating and discriminating codes in graphs: a bibliography. `https://www.lri.fr/~lobstein/debutBIBidetlocdom.pdf`

[7] H. Rahbani, N. J. Rad and S. M. MirRezaei. Bounds on the Identifying Codes in Trees. *Graphs and Combinatorics* 35:599–609, 2019.