Let $G = (V, E)$ be a simple connected graph and $S = \{w_1, \cdots , w_t\} \subseteq V$ an ordered subset of vertices. The metric representation of a vertex $u \in V$ with respect to $S$ is the $t$-vector $r(u\mid S) = (d_G(u, w_1), \cdots , d_G(u, w_t))$, where $d_G(u, v)$ represents the length of a shortest $u - v$ path in $G$. The set $S$ is called a resolving set for $G$ if $r(u\mid S) = r(v\mid S)$ implies $u = v$ for every $u, v \in V$. The smallest cardinality of a resolving set is the metric dimension of $G$. In this article we propose, to the best of our knowledge, a new problem in Graph Theory that resembles to the aforementioned metric dimension problem. We call $S$ a $k$-antiresolving set if $k$ is the largest positive integer such that for every vertex $v \in V - S$ there exist other $k - 1$ different vertices $v_1, \cdots , v_{k-1} \in V - S$ with $r(v\mid S) = r(v_1\mid S) = \cdots = r(v_{k-1}\mid S)$, i.e., $v$ and $v_1, \cdots , v_{k-1}$ have the same metric representation with respect to $S$. The $k$-metric antidimension of $G$ is the minimum cardinality among all the $k$-antiresolving sets for $G$.

In this article, we introduce a novel privacy measure, named $(k, \ell)$-anonymity and based on the $k$-metric antidimension problem, aimed at evaluating the resistance of social graphs to active attacks. We, therefore, propose a true-biased algorithm for computing the $k$-metric antidimension of random graphs. The success rate of our algorithm, according to empirical results, is above 80% and 90% when looking for a $k$-antiresolving basis and a $k$-antiresolving set respectively. We also investigate theoretical properties of the $k$-antiresolving sets and the $k$-metric antidimension of graphs. In particular, we focus on paths, cycles, complete bipartite graphs and trees.

**Keywords:** graph; privacy; active attack; resolving set; $k$-antiresolving set; $k$-metric antidimension

**AMS Subject Classification Numbers:** 05C12; 91D30; 05C82; 05C85.
1 Introduction

Social networking services are widely used in modern society as illustrated by the Alexa’s Top 500 Global Sites statistics\(^1\) where facebook and linkedin rank 2nd and 11th respectively in 2014. Such popularity has enabled governments and third-party enterprises to massively collect social network data, which eventually can be released for mining and analysis purposes.

The power of social network analysis is questionless. It might uncover previously unknown knowledge such as community-based problem, media use, individual engagement, amongst others. Sociology is a trivial example of a field that certainly benefits from social graphs publication. Many other fields (e.g., economics, geography, or political science) and systems (e.g., service-oriented systems, advertisers, or recommended systems) improve their decisions, processes, and services, based on users interaction.

However, all these benefits are not cost-free. An adversary can compromise users privacy using the published social network, which results in the disclosure of sensitive data such as e-mails, instant messages, or relationships. A simple and popular approach to prevent this privacy problem is anonymization by means of removing potential identifying attributes. Doing so, aggregate knowledge still can be inferred (e.g., connectivity, distance, or node degrees) while the “who” information has been removed. In practice, however, this naive approach is not enough for protecting users’ privacy.

What makes social network anonymization a challenging problem is the combination of the adversary’s background knowledge with the released structure of the network. Considering a social network as a simple graph, in which individuals are represented by vertices and their bidirectional relationships by edges, the adversary’s background knowledge about a victim may take many forms, e.g., vertex degrees, connectivity, or local neighborhood. This structural knowledge, together with the released graph, is often enough to perform passive attacks where the users and their relationships are re-identified \([16]\).

Other privacy attacks exist. In 2007, Backstrom et al. [1] introduced active attacks based on the creation and insertion in the network of attacker nodes controlled by the adversary. The attacker nodes could be either new accounts with pseudonymous or spoofed identities (Sybil nodes), or legitimate users in the network who collude with the adversary. Attacker nodes establish links with other nodes in the network (also between them self) aiming at creating a sort of fingerprint in the network. Once the social graph is released, the adversary just need to retrieve such a fingerprint (the attacker nodes) and use it as a hub to re-identify other nodes in the network. Backstrom et. al. proved that \(O(\sqrt{\log n})\) attacker nodes in the network can compromise the privacy of arbitrary targeted nodes with high probability, which makes active attack particularly dangerous.

1.1 Contribution and Plan of the Article

Several active attacks to social graphs have been proposed. They could even target random nodes in the network as recently shown in \([18]\). However, to the best of our knowledge, no privacy measure aimed at evaluating the resistance of a social graph to this kind of attack exists. The lack of such a measure prevents the development of privacy-preserving

\(^1\)http://www.alexatop.com/topsites
methods with theoretically proven privacy guarantees.

In this article we first define the adversary’s background knowledge concerning a node \( u \) and a subset \( S \) of attacker nodes as the metric representation of \( u \) with respect to \( S \). We then propose a privacy measure based on a new problem in Graph Theory: the \( k \)-metric antidimension. The privacy metric, called \((k,\ell)\)-anonymity, can be applied to real-life social graphs in order to measure their resistance to active attacks. Recognizing the hardness of the \( k \)-metric antidimension problem (is NP-complete for \( k = 1 \) [12]), we propose a true-biased algorithm whose computational complexity and success rate can be adjusted. The empirical results show that our algorithm finds \( k \)-antiresolving sets in random graphs of order at most 100 with a success rate above 90%. Finally, we provide theoretical results on the \( k \)-metric antidimension of graphs, such as paths, cycles, complete bipartite graphs and trees.

This article is structured as follows. Section 2 briefly reviews the literature on privacy-preserving publication of social network data. Section 3 presents the metric representation as a reasonable definition of the adversary’s background knowledge. It also proposes the concept of \( k \)-antiresolving sets as the basis for the privacy measure \((k,\ell)\)-anonymity. In Section 4 we present a true-biased algorithm for computing the \( k \)-metric antidimension of a graph, and evaluate the proposed algorithm through experiments. Preliminary results (mathematical properties) on the new problem (the \( k \)-metric antidimension) are provided in Sections 5 and 6 (this last one specifically addresses the case of tree graphs). Finally, Section 7 draws conclusions and future work.

2 Related work

A social graph \( G = (V, E) \) is a simple graph where \( V \) represents the set of social actors and \( E \subseteq V \times V \) their relationships. Both vertices and edges could be enriched with attribute values such as weights representing trustworthiness or labels providing meaning. We consider, however, social network data in its most “simplest” form, \textit{i.e.}, a simple graph without further annotation.

Privacy breaches in social networks are mainly categorized in identity disclosure or link disclosure [23]. To perform such attacks adversaries rely on background knowledge, which is usually defined as structural knowledge such as vertex degrees [14] or neighborhoods [30]. The assumptions on the adversary’s background knowledge determine the type of privacy attacks and the corresponding countermeasures.

Privacy-preserving methods for the publication of social graphs are normally based on the well-known concept \( k \)-anonymity [19] adapted to graphs. \( k \)-anonymity, initially proposed for microdata, aims at ensuring that no record in a database can be re-identified with probability higher than \( 1/k \). To do so, identifying attributes should be obviously removed, and any combination of non-identifying attribute values should not be unique in the database. In practice, not all the attributes need to be combined, because they do not belong to the adversary’s knowledge. This leads to the concept of quasi-identifier, that is, an attribute that can be found in external source of information and, combined with other quasi-identifiers, can uniquely identify a record in the database.

Even though graphs can be represented in tabular form and, thus, graph \( k \)-anonymity can be defined in terms of quasi-identifying attributes [21], graph \( k \)-anonymity is typically
defined in terms of structural properties of the graph rather than on attributes. For instance in [10], the adversary’s background knowledge is defined as a *knowledge query* \( Q(x) \) evaluated for a given target node of the original graph \( G \). The knowledge query \( Q(x) \) allows the creation of a candidate set consisting of \( \{ y \in V | Q(x) = Q(y) \} \). In other words, all the nodes in the network matching the query \( Q(.) \) are equally likely to be the target node \( x \). This simple concept is the basis of several passive attacks and privacy-preserving methods in the publication of social graphs [14, 30, 32].

Other privacy notions based on entropy rather than on \( k \)-anonymity have been proposed [5]. This type of privacy measure is better suited for methods based on random addition, deletion, or switching of edges. The perturbation could be made in such a way that the number of edges or the degree of the vertices are preserved [27, 28]. However, empirical results obtained in [26, 28] suggest that random obfuscation poorly preserves the topological features of the network.

Passive attacks to social networks can be combined with active attacks. In addition to structural knowledge, in an active attack the adversary manages to control a subset of nodes (attacker nodes) of the original graph \( G \) [1]. The attacker nodes aim at creating links with their victims by either identity theft or cloning of existing users profiles [4]. They also establish links between them self so as to build a subgraph \( H \) of attacker nodes with the following properties: i) \( H \) can be efficiently identified in \( G \) and ii) \( H \) does not have a non-trivial automorphisms. Once \( H \) has been identified, the adversary is able to re-identify neighbor nodes of \( H \) [1] or even arbitrary nodes in the network [16, 18].

Performing active attacks is not easy, given that there exist several detection mechanisms of attacker or Sybil nodes in a network [29]. However, such defenses strongly depend on assumptions on the topological structure of the social network, which does not hold in many real-world scenarios [15]. Actually, recent works aim at mitigating, instead of preventing, the impact of Sybil attacks [22]. Furthermore, a group of users who collude in order to breach the privacy of other users in the network can be also regarded as attacker nodes.

Other types of active attacks exist. For instance, the maximal vertex coverage (MVC) attack consists in attacking a few nodes so as to delete as many edges of the network as possible. In this attack, the attacker tries to convince some users to leave the social network in order to reduce the number of residual social ties. Metrics to quantify the impact of MVC attacks have been studied in [13]. MVC is not a privacy attack, though.

While there exist several published active attacks to social graphs, there does not exist yet a rational privacy metric for evaluating the resistance of social graphs to this type of privacy attack. To overcome this problem, in this article we introduce \((k, \ell)\)-anonymity; a privacy notion based on \( k \)-anonymity and the metric representation of nodes in a graph. Note that, privacy notions with the same name has been already proposed. For instance, Feder and Nabar proposed \((k, \ell)\)-anonymity where \( \ell \) represents the number of common neighbors of two nodes [8]. This notion was later generalized by Stokes and Torra in [21]. In our privacy notion, however, \( \ell \) represents an upper bound on the \( k \)-metric antidimension of the graph.

Interested readers could refer to [17, 23, 31] for further reading on privacy-preserving publication of social graphs.
3  \textit{k}-antiresolving sets

In this section we first present the metric representation as a reasonable definition of the adversary’s background knowledge. We then propose the concept of \textit{k}-antiresolving sets and, finally, the privacy measure \((k, \ell)\)-anonymity.

3.1 Adversary’s background knowledge

Vulnerabilities in an anonymized social graphs are better understood once the adversary’s knowledge has been properly modeled. This knowledge can be acquired from public information sources and through malicious actions. In practice, the adversary could even be a close friend, which makes the publication of social network where users cannot re-identify themselves a reasonable privacy goal.

Adversary’s background information in passive attacks is typically modeled as structural knowledge on the network such as vertex degrees, connectivity, or local neighborhood. This is a sort of \textit{global} view on the network defined in terms of \textit{knowledge queries} in [10]. However, adversaries provided with attacker nodes in a network are undoubtedly more powerful. In addition to the \textit{global} view, they have a \textit{local} view determined by the relationship of the attacker nodes with the network.

We aim in this article at proposing a conservative model for this \textit{local} view. We claim that the distance between an attacker node and a legitimate node is a reasonable representation of this kind of knowledge. This is particularly true for neighbor nodes [1], and could also be extended to other nodes in the network [16, 18]. An adversary could also estimate such a distance by considering the presence of groups in social networks, such as family members, friends, work colleges, or job-related contacts\(^2\), together with other kind of external information. More precisely, given a social graph \(G\) and a set of attacker nodes \(S\), we define the adversary’s background knowledge about a target node \(u\) as the \textit{metric representation} of \(u\) (see Definition 1).

\textbf{Definition 1} (Metric representation). Let \(G = (V,E)\) be a simple connected graph and \(d_G(u,v)\) be the length of the shortest path between the vertices \(u\) and \(v\) in \(G\). For an ordered set \(S = \{u_1, \cdots, u_t\}\) of vertices in \(V\) and a vertex \(v\), we call \(r(v \mid S) = (d_G(v, u_1), \cdots, d_G(v, u_t))\) the \textit{metric representation} of \(v\) with respect to \(S\).

The concept of metric representation leads to other two useful concepts as follows.

\textbf{Definition 2} (Resolving set and metric dimension). [9, 20] Let \(G = (V,E)\) be a simple connected graph. A set \(S \subset V(G)\) is said to be a resolving set for \(G\) if any pair of vertices of \(G\) have different metric representations with respect to \(S\). A resolving set of the smallest possible cardinality is called a metric basis, and its cardinality the metric dimension of \(G\).

It is worth mentioning that the definitions above has been already motivated by problems related to unique recognition of an intruder position in a network [20], where resolving sets were called locating sets. The name “resolving set” is due to Harary and Melter [9], who introduced the concept independently in 1976. The concept of resolving sets has been

\(^2\)In the scientific community we could mention https://www.researchgate.net and http://genealogy.math.ndsu.nodak.edu
recently extended to $k$-resolving sets \([7, 24]\). Those problems, however, are different in nature to the problem introduced in this article.

### 3.2 $(k, \ell)$-anonymity

$(k, \ell)$-anonymity is a privacy measure that evolves from the adversary’s background knowledge defined previously. It is based on the concept of $k$-antiresolving set defined as follows.

**Definition 3** ($k$-antiresolving set). Let $G = (V, E)$ be a simple connected graph and let $S = \{u_1, \cdots, u_t\}$ be a subset of vertices of $G$. The set $S$ is called a $k$-antiresolving set if $k$ is the greatest positive integer such that for every vertex $v \in V - S$ there exist at least $k - 1$ different vertices $v_1, \cdots, v_{k-1} \in V - S$ with $r(v | S) = r(v_1 | S) = \cdots = r(v_{k-1} | S)$, i.e., $v$ and $v_1, \cdots, v_{k-1}$ have the same metric representation with respect to $S$.

The following concepts derive from Definition 3, whose study is one of the goals of this article.

**Definition 4** ($k$-metric antidimension and $k$-antiresolving basis). The $k$-metric antidi-\(mension of a simple connected graph $G = (V, E)$ is the minimum cardinality amongst the $k$-antiresolving sets in $G$ and is denoted by $\text{adim}_k(G)$. A $k$-antiresolving set of cardinality $\text{adim}_k(G)$ is called a $k$-antiresolving basis for $G$.

Intuitively, if $S$ is a $k$-antiresolving set, it cannot uniquely re-identify other nodes in the network with probability higher than $1/k$. However, it might exist other subset of vertices in $G$ that forms a $k'$-antiresolving set where $k' < k$. This is, indeed, considered in our privacy measure below.

**Definition 5** ($(k, \ell)$-anonymity). We say that a graph $G$ meets $(k, \ell)$-anonymity with respect to active attacks if $k$ is the smallest positive integer such that the $k$-metric antidi-\(mension of $G$ is lower or equal than $\ell$.

In Definition 5 the parameter $k$ is used as a privacy threshold, whilst $\ell$ is an upper bound on the expected number of attacker nodes in the network. The later can be obtained through statistical analysis. It has been shown that attacker nodes are difficult to enrol in a network without been detected \([29]\). Therefore, the number of attacker nodes might be considered substantially lower than the total number of nodes in the network. Definition 5 can be better understood through the following example result.

**Theorem 1.** For every $n > 0$ and $0 < \ell < n$, the graph $K_n$ meets $(n - \ell, \ell)$ - anonymity.

**Proof.** Since all the vertices in $K_n$ are connected, every subset $S$ of vertices of $K_n$ is an $(n - |S|)$-antiresolving set. Therefore, the $k$-metric antidimension of $K_n$ is $n - k$.

According to Definition 5, the $k$-metric antidimension should be lower or equal than $\ell$, which implies that $k \leq n - \ell$. Moreover, $k$ should be the smallest positive integer satisfying the previous condition. Therefore, $K_n$ holds $(n - \ell, \ell)$-anonymity. \(\square\)

**Corollary 2.** A social graph $K_n$ guarantees that a user cannot be re-identified with probability higher than $\frac{1}{n-\ell}$ by an adversary controlling at most $\ell$ attacker nodes.
These simple and intuitive results obtained in Theorem 1 and Corollary 2 shows the role of the privacy measure \((k, \ell)\)-anonymity in privacy-preserving publication of social graphs. Before releasing a social graph \(G\), the goal is to find \(k\) such that \(G\) holds \((k, \ell)\)-anonymity. To do so, theoretical results and efficient algorithms on the \(k\)-metric antidimension of a graph need to be investigated.

4 Computing the \(k\)-metric antidimension

In this section we address the problem of computing the \(k\)-metric antidimension of a graph. This problem becomes the well-known metric dimension problem for \(k = 1\), which has been proven to be NP-Complete [12]. We remark that the extension of metric dimension (the \(k\)-metric dimension) problem is also known as NP-complete [24]. For these reasons and due to the nature of the problem, we conjecture that the \(k\)-metric antidimension problem is NP-Complete for \(k > 1\) as well, and leave its proof for future work.

4.1 A true-biased algorithm

A true-biased algorithm is always correct when it returns true, it might fail with some small probability when its output is false. True-biased algorithms normally are Monte Carlo algorithms with deterministic running time and randomized behavior. The algorithm we introduce resembles to a Monte Carlo algorithm in the sense that is deterministic and has the true-biased property, however, it is not randomized.

The mathematical foundation of our algorithm requires the introduction of notation as follows. For a given subset of vertices \(X \subseteq V(G)\), we denote \(\sim_X: V(G) \times V(G)\) to the symmetric, reflexive and transitive relation satisfying that \(u \sim_X v \implies r(u|X) = r(v|X)\). The set of equivalent classes created by \(\sim_X\) over the subset of vertices \(V(G) - X\) is denoted as \(C_X\). We deliberately abuse notation and use \(\sim_v\) and \(C_v\) instead of \(\sim_{\{v\}}\) and \(C_{\{v\}}\) for every vertex \(v \in V(G)\).

**Proposition 3.** Let \(S \subseteq V(G)\) and \(S' \subseteq S\):

- \(u \sim_S v \implies u \sim_{S'} v\)
- \(\forall X \in C_S\) there exists \(X' \in C_{S'}\) such that \(X \subseteq X'\)
- \(\forall X \in C_S\) and \(\forall X' \in C_{S'}\), \(X \cap X' \neq \emptyset \implies X \subseteq X'\)

**Lemma 4.** Let \(S\) be a \(k\)-antiresolving set and let \(S' \subseteq S\). Let \(Y = \{X \in C_{S'} : |X| < k\}\), then \(S' \cup (\bigcup_{y \in Y} y) \subseteq S\).

**Proof.** By Proposition 3, for every \(X \in C_S\) there exists \(X' \in C_{S'}\) such that \(X \subseteq X'\), which implies that \(|X'| \geq |X| \geq k\) due to the definition of \(k\)-antiresolving set. Consequently, \(|X'| < k\) implies that there does not exist \(X \in C_S\) such that \(X \subseteq X'\), meaning that there does not exist \(X \in C_S\) such that \(X \cap X' \neq \emptyset\) according to Proposition 3. Therefore, \(X' \cap (V(G) - S) = \emptyset\) and thus \(X' \subseteq S\). \(\square\)
In the spirit of Lemma 4, let \( f : V(G) \to V(G) \) be the function defined recursively as follows:

\[
f(S) = \begin{cases} 
  f(S \cup (\bigcup_{y \in Y} y)), & \text{if } Y = \{X \in C_S : |X| < k\} \text{ is not empty}, \\
  S, & \text{otherwise}.
\end{cases}
\] (1)

According to Lemma 4, if \( S \) is a subset of a \( k \)-antiresolving set, so is \( f(S) \). We therefore give some useful properties of the function \( f \) in Theorem 5 below.

**Theorem 5.** The function defined in Equation 1 satisfies the following properties.

1. \( f(f(S)) = f(S) \)
2. \( S' \subseteq S \implies f(S') \subseteq f(S) \)
3. \( \forall S' \subseteq S, f(S) = f(f(S - S') \cup f(S')) \)
4. \( S' \subseteq f(S) \implies f(S') \subseteq f(S) \)

**Proof.** The first property comes straightforwardly from Equation 1. In order to prove the second property, let \( S' \subseteq S \) and \( u \in f(S') \). If \( u \in S \), then \( u \in f(S) \) by definition. Let us thus assume that \( u \notin S \). Given that \( u \in f(S') \), there exist \( X' \in C_{S'} \) such that \( |X'| < k \) and \( u \in X' \). Let \( X \in C_S \) such that \( u \in X \). Note that, such an \( X \) exists because \( u \notin S \).

According to Proposition 3, since \( X \cap X' \neq \emptyset \) and \( S' \subseteq S \), then \( X \subseteq X' \), which means that \( |X| < k \) and that \( X \subseteq f(S) \), which proves the second property.

The third property can be proven by using the first property. Given that \( S - S' \subseteq S \) and \( S' \subseteq S \), then \( f(S - S') \subseteq f(S) \) and \( f(S') \subseteq f(S) \), hence, \( f(f(S - S') \cup f(S')) \subseteq f(S) \).

Similarly, \( S - S' \subseteq f(S - S') \) and \( S' \subseteq f(S') \) by definition, which implies that \( S \subseteq f(S - S') \cup f(S') \). Again, applying the first property we obtain that \( f(S) \subseteq f(f(S - S') \cup f(S')) \).

The two results lead to \( f(S) = f(f(S - S') \cup f(S')) \).

Finally, the last property is proven as follows. If \( S' \subseteq f(S) \), then \( f(S') \subseteq f(f(S)) \) by applying the second property. The proof is concluded by simply considering the first property. \( \square \)

The function \( f(.) \) is the basis of Algorithm 1, which aims to find a \( k \)-antiresolving set in a graph. Algorithm 1 is an optimized version supported by Theorem 5 of the following algorithm. Consider all the subsets \( S \subseteq V(G) \) with cardinality lower than or equal to \( m \). If \( f(S) \) is a \( k \)-antiresolving set for some set \( S \subseteq V(G) \) where \( |S| \leq m \), then a positive output is provided. If not, a proof that a \( k \)-antiresolving set does not exist is found when \( f(S) = V(G) \) for every \( S \subseteq V(G) \) such that \( |S| = m \). Note that, this impossibility result comes from the monotonicity of the function \( f \), i.e., \( S' \subseteq S \implies f(S') \subseteq f(S) \). Any other case leads to the unknown state where neither a proof nor a disproof of the existence of a \( k \)-antiresolving set can be found.

Algorithm 1 can be considered a true-biased algorithm if the unknown state is regarded as a negative result. Its computational complexity is clearly exponential in terms of \( m \); the set \( C_h \) has cardinality \( N^{2^h} \) in the worst case. However, \( m \) can be tuned with respect to the number of vertices in the graph in order to balance properly false negatives versus computational cost.
Algorithm 1 Given a positive integer $k$, this algorithms outputs: i) true if it finds a $k$-antiresolving set, ii) false if such a set does not exist, iii) unknown when neither a $k$-antiresolving set nor a proof that such a set does no exist was found.

Require: A graph $G$, an integer value $m$ to control the exponential explosion, and the integer value $k$.

Let $V(G) = \{v_1, \ldots, v_N\}$

Let $C_1 = \{f(\{v_1\}), \ldots, f(\{v_N\})\}$

if $\exists S \in C_1$ that is a $k$-antiresolving set then return true

for $h = 2$ to $m$ do

Let $C_h$ be an empty set

for $i = 1$ to $|C_{h-1}|$ do

Let $S_i$ be the $i$th element of $C_{h-1}$

for $j = i + 1$ to $|C_{h-1}|$ do

Let $S_j$ be the $j$th element of $C_{h-1}$

if $S_i \not\subseteq S_j$ and $S_j \not\subseteq S_i$ then

$S = f(S_i \bigcup S_j)$

if $S$ is a $k$-antiresolving set then return true

Add $S$ to $C_h$

if $\forall S \in C_m, S = V(G)$ then return false

else return unknown

In order to show the feasibility of our algorithm, we ran experiments by considering $m \in \{1, 2, 3\}$ and random graphs with at most 100 vertices. A random graph is created by choosing integer values uniformly distributed in the interval $[1, 100]$ as the number of vertices $N$, the number of edges also distributes uniformly in the interval $[0, N \times (N-1)/2]$, and the edges are added randomly to the graph. For each $m \in \{1, 2, 3\}$ and each $k = \{1, 2, 3, 4, 5, 6, 7, 8\}$, we executed 10 000 times Algorithm 1 aiming at providing statistically sounds data about its accuracy.

Figure 1(a) shows the performance of Algorithm 1 in terms of success rate for different values of $m$. Algorithm 1 is considered to succeed if it outputs either true or false. According to Figure 1(a), Algorithm 1 performs poorly for $m = 1$, while it achieves a success rate near to 90% when $m = 2$. As expected, the most computationally demanding version of Algorithm 1, that is, when $m = 3$, is the best in terms of success rate.

Algorithm 1 can be easily adapted to find a $k$-antiresolving basis rather than a $k$-antiresolving set by using Proposition 6 below. We therefore executed similar experiments to the ones described above, however, this time with Algorithm 1 adapted to find a $k$-antiresolving basis. Figure 1(b) shows the results of these experiments. The success rate of this new algorithm naturally downgrades with respect to the original Algorithm 1, yet it is above 80% for $m = 2$ and $m = 3$.

Proposition 6. Let $S$ be the subset of smaller cardinality in $V(G)$ such that $f(S)$ is a $k$-antiresolving set. Then, $f(S)$ is a $k$-antiresolving basis, if $\forall S' \subseteq V(G)$ such that $|S'| = |S|$ it follows that $|f(S)| \leq |f(S')|$.

Given the decrease in success rate of our algorithm as $k$ grows shown by Figures 1(a) and 1(b), it makes sense to wonder whether this decreasing tendency continues for $k > 8$. 
Figure 1: Figure 1(a) and Figure 1(b) depicts the success rates of our algorithm when executed to find a \( k \)-antiresolving set and a \( k \)-antiresolving basis respectively. The considered values of \( m \) are \( \{1, 2, 3\} \), while \( k \) varies from 1 to 8.

However, our algorithm has 100% of success rate if \( k \) is equal to the order of the graph. Indeed, the output of Algorithm 1 in that case would be \text{false} for every graph, because \( \forall u \in V(G) \forall X \in C_u(|X| < |V(G)|) \) and, thus, all the nodes in the graph should be contained in a \( k \)-antiresolving set according to Lemma 4.

We end this section by showing the ratio between the different outputs of Algorithm 1. We consider the two versions of Algorithm 1, namely the one looking for a \( k \)-antiresolving set and the one looking for a \( k \)-antiresolving basis. Due to lack of space, we consider two settings only: one where \( k = 4 \) and \( m = 2 \) (see Figures 2(a) and 2(b)); another one where \( k = 8 \) and \( m = 2 \) (see Figures 2(c) and 2(d)). According to the results, an 8-antiresolving basis is not only harder to find than a 4-antiresolving basis, but also less likely to exist. A similar conclusion can be drawn for the problem of finding a \( k \)-antiresolving set.

5 Mathematical properties on the \( k \)-metric antidimension of graphs

In the next two sections we provide some primary theoretical results on the \( k \)-metric antidimension problem. We focus on giving mathematical properties that, supported by the results in Section 4, can determine or bound the \( k \)-metric antidimension for some families of graphs. In particular in this section, we study the \( k \)-metric antidimension of cycles, paths, complete bipartite graphs, and other graph families satisfying some specific conditions. To do so, we first observe some basic properties of \( k \)-antiresolving sets, which will be further used.

Observation 1.

(i) A 1-antiresolving set is also a resolving set.

(ii) There does not exist \( k > 1 \) such that all the vertices of a graph form a \( k \)-antiresolving set.
Figure 2: Four Pie charts showing the ratio of true, false, and unknown results provided by Algorithm 1 on two different inputs: \((k = 4, m = 2)\) and \((k = 8, m = 2)\). Figure 2(a) and 2(c) are devoted to the version of Algorithm 1 aimed at finding a \(k\)-antiresolving set, Figure 2(b) and 2(d) considers the version that looks for a \(k\)-antiresolving basis.
There does not exist any $n$-antiresolving set in a graph of order $n$.

Not for every graph $G$ of order $n$ and every $k < n$, there exist a $k$-antiresolving set in $G$. For instance, if $G$ is a path graph, for $k \geq 3$ there does not exist a $k$-antiresolving set in $G$.

In order to continue with our study we need to introduce some terminology and notation. For a graph $G$ and a vertex $v \in V(G)$, the set $N_G(v) = \{u \in V : uv \in E(G)\}$ is the open neighborhood of $v$ and the set $N_G[v] = N_G(v) \cup \{v\}$ is the closed neighborhood of $v$. Two vertices $x, y$ are called (false) true twins if $(N_G(x) = N_G(y)) \quad N_G[x] = N_G[y]$. In this sense, a vertex $x$ is a twin if there exists $y \neq x$ such that they are either true or false twins. The diameter of $G$ is defined as $D(G) = \max_{u,v \in V} \{d_G(u,v)\}$.

As mentioned before, not for every integer $k$ it is possible to find a $k$-antiresolving set in a graph $G$. Thus, it is desirable to analyze first the interval of suitable values for $k$ satisfying that $G$ contains at least one $k$-antiresolving set. According to Definition 3 we present the following concept, which is relevant in the study of the $k$-metric antidimension of graphs.

**Definition 6** ($k$-metric antidimensional graph). A simple connected graph $G = (V,E)$ is $k$-metric antidimensional, if $k$ is the largest integer such that $G$ contains a $k$-antiresolving set.

### 5.1 $k$-metric antidimensional graphs

In order to study the $k$-metric antidimension of graphs, we first focus into obtaining the values of $k$ for which a given graph is $k$-metric antidimensional. First notice that any graph $G$ is always $k$-metric antidimensional for some $k \geq 1$, and a natural upper bound for $k$ which makes that $G$ would be $k$-metric antidimensional is clearly the maximum degree of the graph, since the number of vertices at distance one from any vertex is at most the maximum degree of the graph.

**Observation 2.** If $G$ is a connected $k$-metric antidimensional graph of maximum degree $\Delta$, then $1 \leq k \leq \Delta$.

Since the maximum degree of a graph is at most the order of the graph minus one, a particular case of the above result is the next one.

**Remark 7.** If $G$ is any connected $k$-metric antidimensional graph of order $n$, then $1 \leq k \leq n - 1$. Moreover, $G$ is $(n - 1)$-metric antidimensional if and only if $G$ has maximum degree $n - 1$.

**Proof.** The upper bound is a particular case of Remark 2. Now, it is straightforward to observe that if $v$ is a vertex of $G$ of degree $n - 1$, then for every vertex $u, w \neq v$ it follows that $r(u\{v\}) = r(w\{v\}) = 1$, i.e., every vertex different from $v$ has the same metric representation with respect to $\{v\}$. Thus, $\{v\}$ is a $(n - 1)$-antiresolving set, since there exists no $n$-antiresolving sets in $G$. Thus, $G$ is $(n - 1)$-metric antidimensional.

On the contrary, we assume that $G$ is $(n - 1)$-metric antidimensional. Hence, if $S$ is a $(n - 1)$-antiresolving set, then $|S| = 1$. Thus, the only possibility is that $S$ is formed by a single vertex and that every vertex is adjacent to it. □
To continue with our study we need some extra notation. The eccentricity $\varepsilon(v)$ of a vertex $v$ in a connected graph $G$ is the maximum graph distance between $v$ and any other vertex $u$ of $G$. Given a vertex $u$ of a graph $G$, we consider the following local parameter. For every $i \in \{1, ..., \varepsilon(u)\}$, let $d_i(u) = \{v \in V(G) : d(v, u) = i\}$. Now, for every $u \in V(G)$, let

$$\phi(u) = \min_{1 \leq i \leq \varepsilon(u)} \{|d_i(u)|\}.$$ 

and for any graph $G$, let $\phi(G) = \max_{v \in V(G)} \{\phi(v)\}$.

**Theorem 8.** Any connected graph $G$ is $k$-metric antidimensional for some $k \geq \phi(G)$.

**Proof.** Assume $x$ is a vertex of degree at least two in $G$ such that $\phi(G) = \phi(x)$. Thus, for any vertex $y \neq x$, there exist at least $\phi(x) - 1$ vertices $v_1, v_2, ..., v_{\phi(G)-1}$ in $V(G) - \{x, y\}$ such that $d(y, x) = d(v_1, x) = ... = d(v_{\phi(G)-1}, x)$. Moreover, since $\phi(G) = \phi(x)$, there exists at least one vertex $y'$ such that there are exactly $\phi(G) - 1$ different vertices satisfying the above mentioned. So, $\{x\}$ is a $\phi(x)$-antiresolving set and $G$ is $k$-metric antidimensional for some $k \geq \phi(G)$.

If a graph $G$ is 1-metric antidimensional, then its 1-antiresolving sets are standard resolving sets as defined in [9, 20] and this has been studied in several scientific articles (e.g., [2, 3, 6, 11, 25]). According to that, in this work we are interested mainly in those graphs being $k$-metric antidimensional for some $k \geq 2$. An example of a graph being 1-metric antidimensional is for instance the path graph of even order.

### 5.2 Graphs that are $k$-metric antidimensional for some $k \geq 2$

To begin with the description of some families of graphs being $k$-metric antidimensional for some $k \geq 2$ we define the radius and the center of a graph as follows. The radius $r(G)$ of $G$ is the minimum eccentricity of any vertex in $G$. The center of $G$ is the set $S$ of vertices of $G$ having eccentricity equal to the radius of $G$.

**Remark 9.** If the center of a graph $G$ is only one vertex, then $G$ is $k$-metric antidimensional for some $k \geq 2$.

**Proof.** Let $v$ be the center of $G$. Hence, there exist two diametral vertices $u, w$ such that $d_G(v, u) = d_G(v, w) = \varepsilon(v) = r(G)$. Since $v$ has eccentricity $r(G)$, there is no vertex $z \neq u, w$ in $G$ such that $d_G(v, z) > d_G(v, u) = d_G(v, w)$. Thus, it follows that for any vertex $x \neq v$ there exists at least a vertex $y$ belonging to the $u - w$ path such that $d_G(x, v) = d_G(y, v)$. Therefore, $\{v\}$ is a $k$-antiresolving set in $G$ for some $k \geq 2$.

If a path graph has odd order, then its center is formed by only one vertex. Also, for every vertex of any path, there exists at most other different vertex having equal distance to a third vertex of the path. Thus, it is clear the following consequence of the Remark above.

**Corollary 10.** If a path $P_n$ has odd order, then it is 2-metric antidimensional.

Another example of 2-metric antidimensional are the cycle graphs as we next see.
Remark 11. Any cycle graph $C_n$ is 2-metric antidimensional.

Proof. We assume first that $n$ is odd. Let $v$ be any vertex of $C_n$. Hence, for any vertex $x \neq v$ of $C_n$, there exists only one $y \neq x, v$ such that $d_{C_n}(x,v) = d_{C_n}(y,v)$. Thus, $\{v\}$ is a 2-antiresolving set in $C_n$. Assume now that $n$ is even and let $\{u,w\}$ be any two diametral vertices of $C_n$. We observe that for any vertex $x \neq u,w$ of $C_n$, there exists only one $y \neq x,u,w$ such that $d_{C_n}(x,u) = d_{C_n}(y,u)$ and $d_{C_n}(x,w) = d_{C_n}(y,w)$. Thus, $\{u,w\}$ is a 2-antiresolving set in $C_n$.

On the other hand, there does not exists $k > 2$ such that $C_n$ contains a $k$-metric antiresolving set, since for any vertex of $C_n$, there exists at most another different vertex having equal distance to a third vertex of the path. Therefore, $C_n$ is 2-metric antidimensional. □

If the vertices of a set $S$ are pairwise twins in a graph $G$, then it is clear that they have the same distance to every other vertex $x \notin S$. So, $V(G) - S$ is a $|S|$-antiresolving set for $G$. Hence, the following result.

Observation 3. If the vertices of a set $S$ are pairwise twins in a graph $G$, then $G$ is $k$-metric antidimensional for some $k \geq |S|$.

Complete bipartite graph\(^3\) are special kind of graphs, since they have a bipartition of the vertex set in which all the vertices belonging to one of the sets of the bipartition are pairwise twin vertices. Let $K_{r,t}$ be a complete bipartite graph. Next we analyze the suitable values $k$ making a complete bipartite graph $k$-metric antidimensional.

Remark 12. Any complete bipartite graph $K_{r,t}$ with $r \geq t$ is $r$-metric antidimensional.

Proof. Let $U$ and $V$ be the two disjoint sets of $K_{r,t}$ with $|U| = r$ and $|V| = t$. Notice that $U$ (respectively $V$) is a set of pairwise twin vertices. Thus, by Observation 3 we have that $K_{r,t}$ is $k$-metric antidimensional for some $k \geq |U| = r$. Suppose that $k \geq r + 1$ and let $S$ be a $k$-antiresolving set for $K_{r,t}$. Since every vertex of $K_{r,t}$ has distance either one or two to any other vertex of $K_{r,t}$ it is not possible to find $k - 1$ vertices having the same distance to every vertex of $S$, a contradiction. So, $k = r$ and the proof is complete. □

5.3 The $k$-metric antidimension of graphs

In this subsection we compute the $k'$-metric antidimension of some graphs which were already described to be $k$-metric antidimensional for some value $k \geq k'$. It is clear that the first natural bound which follows for the $k$-metric antidimension of a graph of order $n$ is $\text{adim}_k(G) \leq n - k$. Such a bound is tight. It is achieved, for instance, for the complete bipartite graphs $K_{r,t}$ as we can see at next by taking the case $t < k \leq r$.

Proposition 13. Let $r, t$ be two positive integers with $r \geq t$.

1. If $t < k \leq r$, then $\text{adim}_k(K_{r,t}) = r + t - k$.

2. If $1 < k \leq t$, then $\text{adim}_k(K_{r,t}) = r + t - 2k$.

\(^3\)A graph $G$ is complete bipartite if its vertex set can be divided into two disjoint sets $U$ and $V$ such that every vertex in $U$ is adjacent to every vertex in $V$ and no more.
Proof. From Proposition 12 we know that $K_{r,t}$ is a $r$-metric antidimensional graph. Let $U$ and $V$ be the two partite sets of $K_{r,t}$ with $|U| = r$ and $|V| = t$. We assume first that $t < k \leq r$. Let $A \subseteq U$ with $|A| = k$ and let be $S = (V \cup U) - A$. Notice that if $k = r$, then $A = U$ and so, $S = V$. Since any vertex $v \notin S$ (or equivalently $v \in A$) is adjacent to every vertex of $V$ and it has distance two to every vertex in $U - A$, we have that all the vertices of $A$ have the same metric representation with respect to $S$. As $|A| = k$, it follows that $S$ is a $k$-antiresolving set and $\text{adim}_k(K_{r,t}) \leq r + t - k$. Now, suppose $\text{adim}_k(K_{r,t}) < r + t - k$ and let $S'$ be a $k$-antiresolving set for $K_{r,t}$. So, we have either one of the following situations.

- There exist more than $k$ vertices of $U$ not in $S'$. Hence, for any vertex $u \in U - S'$ there exist at least $k$ vertices not in $S'$ which, together with $u$, have the same metric representation with respect to $S'$. So, $S'$ is not a $k$-antiresolving set, but a $k'$-antiresolving set for some $k' \geq k + 1$, a contradiction.

- There exists at least one vertex of $V$ not in $S'$. It is a direct contradiction, since $|V| = t < k$.

Therefore, we obtain that $\text{adim}_k(K_{r,t}) = r + t - k$.

On the other hand, we assume that $1 < k \leq r$. Let $X \subseteq U$ with $|X| = k$, let $Y \subseteq V$ with $|Y| = k$ and let $Q = (V - Y) \cup (U - X)$. Hence, for any vertex $v \notin Q$ (or equivalently $v \in X \cup Y$), there exist exactly $k - 1$ vertices, such that all of them, together with $v$, have the same metric representation with respect to $Q$. Thus, $Q$ is a $k$-antiresolving set and $\text{adim}_k(K_{r,t}) \leq r + t - 2k$.

Now, suppose that $\text{adim}_k(G) < r + t - 2k$ and let $Q'$ be a $k$-antiresolving set in $K_{r,t}$. Hence, either there exist more than $k$ vertices of $U$ not in $Q'$ or there exist more than $k$ vertices of $V$ not in $Q'$. As above, in any of both possibilities we obtain that $Q'$ is not $k$-antiresolving set, but a $k'$-antiresolving set for some $k' \geq k + 1$, a contradiction. As a consequence, we obtain that $\text{adim}_k(K_{r,t}) = r + t - 2k$. \qed
odd and $a$ is any vertex of $C_n$, then we can check that for any vertex $b \neq a$, there exists exactly one vertex $c \neq a, b$, such that $b, c$ have the same metric representation with respect to $\{a\}$. Thus, $\text{adim}_2(C_{2n+1}) = 1$. 

\section{The particular case of trees}

Let $T$ be a tree and let $u$ be a vertex of $T$ of degree at least two. Let $v$ be a neighbor of $u$. A $v$-branch of $T$ at $u$ is the subtree $T_{u,v}$ obtained from the union of all length maximal paths beginning in $u$, passing throughout $v$ and finishing at a vertex of degree one in $T$. Given a $y$-branch $T_{x,y}$ at $x$, we say that $\xi(T_{x,y})$ is the eccentricity of the vertex $x$ in the $y$-branch $T_{x,y}$. Two branches $T_{x,y_1}$ and $T_{x,y_2}$ at $x$ are $\xi_x$-equivalent if $\xi(T_{x,y_1}) = \xi(T_{x,y_2})$. For every vertex $x$ of $T$, let $\xi(x)$ represents the maximum number of pairwise $\xi_x$-equivalent branches at $x$ and let $l_\xi(x)$ equals the length of any $\xi_x$-equivalent branch. Now, for any tree $T$, we define the following parameter:

$$
\xi(T) = \max_{x \in V(T) : \delta(x) \geq 2} \{\xi(x)\}.
$$

An example which helps to clarify the above definitions is given in Figure 3. There we have a tree $T$ satisfying the following. The vertex $v_5$ has 4 branches: $T_{v_5,v_6}$, $T_{v_5,v_{10}}$, $T_{v_5,v_{15}}$ and $T_{v_5,v_4}$. For instance $V(T_{v_5,v_{15}}) = \{v_5, v_{15}, v_{16}, v_{17}, v_{11}, v_{18}, v_{14}, v_{13}\}$. We observe that $\xi(T_{v_5,v_6}) = 2$, $\xi(T_{v_5,v_{10}}) = 1$, $\xi(T_{v_5,v_{15}}) = 3$ and $\xi(T_{v_5,v_4}) = 4$. So, $v_5$ has no $\xi_5$-equivalent branches and $\xi(v_5) = 0$. Similarly, it can be noticed that $v_3$ and $v_{15}$ are the only vertices of $T$ which have equivalent branches. That is, $T_{v_3,v_2}$ and $T_{v_3,v_9}$ are $\xi_3$-equivalent, since $\xi(T_{v_3,v_2}) = \xi(T_{v_3,v_9}) = 2$. Thus $\xi(v_3) = 2$ and $\ell_\xi(v_3) = 2$. Analogously, $\xi(T_{v_{15},v_{11}}) = \xi(T_{v_{15},v_{10}}) = \xi(T_{v_{15},v_{14}}) = 2$ and $\xi(v_{15}) = 3$, $\ell_\xi(v_{15}) = 2$. Therefore $\xi(T) = 3$.

![Figure 3: A 3-metric antidimensional tree $T$.](image)

Now, for the particular case of trees, we next use the definition of $\phi(G)$ already presented in Section 5. As an example, for the tree of Figure 3 we have that, for instance, $\phi(v_3) = 3$, $\phi(v_4) = 3$ and $\phi(v_5) = 2$. Also, some calculations give that $\phi(T) = 3$.

Now, with the definitions above we present the following result.

**Theorem 15.** Any tree $T$ is $k$-metric antidimensional for some $k \geq \max\{\phi(T), \xi(T)\}$.

**Proof.** From Theorem 8 it follows that $k \geq \phi(T)$. Now, let $x$ be a vertex of degree at least two in $T$ such that $\xi(T) = \xi(x)$. Hence, there exist $\xi(T)$ disjoint paths beginning in $x$, passing throughout a vertex $y_j$ (neighbor of $x$), and ending in a vertex $w_j$ of degree one
in $T$ with $j \in \{1, ..., \xi(T)\}$. Moreover, every $y_j$-branch $T_{x,y_j}$ does not contain any other vertex further away from $x$ than $w_j$.

We consider now the set

$$A = V(T) - \left( \bigcup_{i=1}^{\xi(T)} V(T_{x,y_i}) \right) \bigcup \{x\}.$$  

Notice that for any vertex $u \not\in A$, there exist at least $\xi(T) - 1$ different vertices $v_1, v_2, ..., v_{\xi(T)} - 1$ in $V(T) - A$ such that $d(u, z) = d(v_1, z) = ... = d(v_{\xi(T)} - 1, z)$ for every $z \in A$. Moreover, since $\xi(T) = \xi(x)$, it follows that there exists a vertex $u'$ such that there are exactly $\xi(T) - 1$ different vertices satisfying the above mentioned. Thus, $A$ is a $\xi(T)$-antiresolving set and, as a consequence, $T$ is $k$-metric antidimensional for some $k \geq \xi(T)$.

Therefore we obtain that $T$ is $k$-metric antidimensional for some $k \geq \max\{\phi(T), \xi(T)\}$ and the proof is complete.

According to Theorem 15, we conclude that the tree $T$ in Figure 3 is $k$-metric antidimensional for some $k \geq 3$, since $\phi(T) = 3$ and $\xi(T) = 3$; that tree is indeed 3-antidimensional. If we add some extra vertices to this mentioned tree, like in Figure 4, we obtain that $\phi(T) = 5$ (since $\phi(v_4) = 5$), and it remains $\xi(T) = 3$. Thus, this new tree is $k$-metric antidimensional for some $k \geq 5$, and by Remark 2 we have that $k = 5$.

Notice that $\xi(T) > \phi(T)$ holds for some trees. For instance, if we take a star graph $S_{1,n}$, $n \geq 4$, and we add an extra vertex $x$ connected by an edge with one leaf $y$ of $S_{1,n}$, then we have a tree $T$ such that $\phi(T) = 2$ (for the vertex $y$, $\phi(y) = 2$) and $\xi(T) = n - 1$.

Moreover, there are graphs in which the bound of Theorem 15 is not achieved. An example of this appears in Figure 5. There we have a tree $T$ such that $\xi(T) = 3$ ($\xi(v_7) = 3$) and $\phi(T) = 3$ ($\phi(v_{12}) = 3$). Nevertheless the set $\{v_{10}, v_{11}, v_{12}, v_{13}, v_{14}, v_{18}, v_{19}\}$ is a 4-antiresolving set.

### 6.1 The $k$-metric antidimension of trees

Once we have a lower bound for the integer $k'$ for which a given tree $T$ is $k'$-metric antidimensional, we are able to compute its $k$-metric antidimension for a suitable value $k \leq k'$. We first notice that if $T$ is 1-metric antidimensional, then any 1-metric antiresolving set
Corollary 17. For any tree $T$ such that $\phi(T) \geq 2$, $\text{adim}_{\phi(T)}(T) = 1$.

According to the results above, it remains to study the $k$-metric dimension of trees for the case in which every vertex $v$ of $T$ satisfies that $\phi(v) \neq k$. To do so, we need to introduce some notations.

We denote by $\Xi_k(T)$, for some $k \in \{2, ..., \xi(T)\}$, the set of vertices $v \in V(T)$ such that $\xi(v) \geq k$. Now, for every $v \in \Xi_k(T)$, let

$$N_{<l_\xi} (v) = \{x \in V(T_{v,u}) : u \in N(v) \text{ and } \xi(T_{v,u}) < l_\xi(v)\} - \{v\},$$

and for the set of vertices $u \in N(v)$ such that $\xi(T_{v,u}) = l_\xi(v)$, let $N_{=l_\xi}^u (v)$ be the maximum cardinality among all possible sets obtained as the union of $k$ vertex sets of the branches $T_{v,u}$ where $u \in N(v)$ minus the vertex $v$ itself. As an example, we consider the tree of Figure 3. For $k = 2$, there we have that $\Xi_2(T) = \{v_3, v_{15}\}$, $N_{<l_2} (v_3) = \{v_{12}\}$, $N_{<l_2} (v_{15}) = \emptyset$, $N_{=l_2}^u (v_3) = \{v_1, v_2, v_8, v_9\}$ and $N_{=l_2}^u (v_{15}) = \{v_{13}, v_{14}, v_{16}, v_{17}\}$ (notice that $N_{=l_2}^u (v_{15})$ can be different from this set, but it always has five vertices).

With this definition we are able to present the following result, where we analyze only those graphs being $k'$-metric antidimensional for $k' = \max \{\phi(T), \xi(T)\}$.

Theorem 18. Let $T$ be a $k'$-metric antidimensional of order $n$ with $k' = \max \{\phi(T), \xi(T)\}$. Then for any $k \leq k'$,

$$\text{adim}_k(T) \leq n - \left| \bigcup_{v \in \Xi_k(T)} N_{<l_\xi} (v) \right| - \left| \bigcup_{v \in \Xi_k(T)} N_{=l_\xi}^k (v) \right|.$$
Proof. We consider a set $S \subset V(T)$ given by

$$S = V(T) - \bigcup_{v \in \Xi_k(T)} N_{< l_k}(v) - \bigcup_{v \in \Xi_k(T)} N_{= l_k}(v).$$

In this sense, for any vertex $x \not\in S$, there exists at least $k - 1$ vertices $y_1, y_2, ..., y_{k-1}$ not in $S$ such that $d(x, w) = d(y_1, w) = ... = d(y_{k-1}, w)$ for every $w \in S$. Moreover, if there exists at least one vertex $x' \not\in S$ for which there are exactly $k - 1$ vertices not in $S$ satisfying the above mentioned, then $S$ is a $k$-metric antiresolving set and the result follows since the cardinality of $S$ is given by the formula of the theorem. On the contrary, if such a vertex does not exist, then $S$ is a $k''$-metric antiresolving set for $G$ for some $k'' \geq k$. Since in this case, $\text{adim}_{k''}(G) \geq \text{adim}_k(G)$ we obtain the result. \hfill $\square$

Consider now the example of Figure 3. According to the result above, we have that the set $S = \{v_3, v_4, v_5, v_6, v_7, v_{10}, v_{11}, v_{15}, v_{18}, v_{19}\}$ is a 2-metric antiresolving set for such a tree $T$ and $\text{adim}_2(T) \leq 10$. Nevertheless, since $\phi(v_5) = 2$, from Remark 16 we have that $\text{adim}_2(T) = 1$. Next we present a family of trees, where the bound of Theorem 18 is achieved.

We consider the family $\mathcal{F}$ of trees $T_r$ satisfying the following conditions.

- The center of $T_r$ is formed by two adjacent vertices, say $x, y$.
- $T_r$ is “rooted” in $x, y$.
- $T_r$ is a complete $r$-ary tree (each vertex of degree greater than one has $r$ children)
- Any two leaves being descendants of the same root ($x$ or $y$), have the same distance to this root.

An example of a tree $T_3$ of the family $\mathcal{F}$ is given in Figure 6.

It is straightforward to observe that $\xi(T_r) = r$ and $\phi(T_r) = r + 1$. Thus, $T_r$ is $k'$-metric antidimensional for some $k \geq r + 1$, and by Remark 2 we have that $k = r + 1$. Now on,
we compute the $r$-metric antidimension of $T_r$. According to the construction of the family $\mathcal{F}$, we see that the root vertices $x, y$ of a tree $T_r \in \mathcal{F}$ satisfy that $x, y \in \Xi_i(T_r)$. Also, $N_{\leq t}(x) = N_{\leq t}(y) = \emptyset$ and the sets $N_{> t}(x), N_{> t}(y)$ are formed by the set of all their corresponding descendants (this fact makes unnecessary to consider other vertices of $T_r$). As a consequence of this, by Theorem 18 we have that $\{x, y\}$ is a $r$-metric antiresolving set and $\text{adim}_r(T_r) \leq 2$. Since, for any non-leaf vertex $u$ of $T_r$ satisfies that $\phi(u) = 4$, we have that any singleton vertex (being not a leaf) is a $(r + 1)$-metric antiresolving set. Thus, $\text{adim}_r(T_r) \geq 2$ and we have that $\text{adim}_r(T_r) = 2$, which makes that the bound of Theorem 18 is tight.

7 Discussion and conclusions

In this article we have introduced a new problem in Graph Theory (the $k$-metric antidimension problem) that resembles to the well-known metric dimension problem. The $k$-metric antidimension is the basis of our novel privacy measure ($k, \ell$)-anonymity. This measures quantifies the level of privacy offered by an outsourced social graph against active attacks. Consequently, privacy-preserving methods for the publication of social networks ought to consider ($k, \ell$)-anonymity as one of their privacy goal.

We have proposed a true-biased algorithm aimed at finding both a $k$-antiresolving set and a $k$-antiresolving basis in a graph. The algorithm, although computationally demanding, reached a success rate above 90\% during the executed experiments when looking for a $k$-antiresolving set. We expect future experiments to be conducted over real-life social graphs so that privacy-preserving methods satisfying ($k, \ell$)-anonymity can be empirically evaluated in terms of utility and resistance to active attacks.

We have also began the study of mathematical properties of the $k$-antiresolving sets and the $k$-metric antidimension of graphs. We have studied some particular graph families like cycles, paths, complete bipartite graphs and trees. For instance, we have obtained that for any path $P_n$ of odd order, $\text{adim}_2(P_n) = 1$ and for any cycle $C_n$ it follows that $\text{adim}_2(C_n) = 1$ if $n$ is odd, and $\text{adim}_2(C_n) = 2$ if $n$ is even. Also, for every complete bipartite graph $K_{r,t}$, $\text{adim}_k(K_{r,t}) = r + t - k$ if $t < k \leq r$, and $\text{adim}_k(K_{r,t}) = r + t - 2k$ if $1 < k \leq t$. For the case of trees we have presented a tight lower bound for its $k$-metric antidimension in terms of the order of the tree and the order of some subtrees satisfying some specific conditions. We have also described an infinite family of $k$-ary trees which achieve this bound.

Finally, this article opens new and challenging open problems related to the $k$-metric antidimension of graphs and the privacy concept ($k, \ell$)-anonymity. For instance, it would be interesting to characterize the family of graphs such that they are 1-metric antidi- mensional, as well as looking for a close relationship between the $k$-metric antidimension and the $k$-metric dimension of a graph. In particular, those families of graphs that resemble to social graphs must be considered. The computational complexity of computing the $k$-metric antidimension should also be addressed. In case the problem is NP-complete, efficient heuristics and privacy-preserving methods need to be developed so as to compute the $k$-metric antidimension and, ultimately, transform a social graph into a ($k, \ell$)-anonymous graph for given values of $k$ and $\ell$. 

20
References

[1] Lars Backstrom, Cynthia Dwork, and Jon Kleinberg. Wherefore art thou r3579x?: anonymized social networks, hidden patterns, and structural steganography. The 16th International Conference on World Wide Web, pages 181–190, 2007.

[2] Robert F. Bailey and Peter J. Cameron. Base size, metric dimension and other invariants of groups and graphs. Bulletin of the London Mathematical Society, 43(2):209–242, 2011.

[3] Robert F. Bailey and Karen Meagher. On the metric dimension of grassmann graphs. Discrete Mathematics & Theoretical Computer Science, 13(4):97–104, 2011.

[4] Leyla Bilge, Thorsten Strufe, Davide Balzarotti, and Engin Kirda. All your contacts are belong to us: Automated identity theft attacks on social networks. The 18th International Conference on World Wide Web, pages 551–560, 2009.

[5] Francesco Bonchi, Aristides Gionis, and Tamir Tassa. Identity obfuscation in graphs through the information theoretic lens. Information Sciences, 275(0):232 - 256, 2014.

[6] Gary Chartrand, Linda Eroh, Mark A. Johnson, and Ortrud R. Oellermann. Resolvability in graphs and the metric dimension of a graph. Discrete Applied Mathematics, 105(1-3):99–113, October 2000.

[7] Alejandro Estrada-Moreno, Juan A. Rodríguez-Velázquez, and Ismael G. Yero. The $k$-metric dimension of a graph. arXiv:1312.6840 [math.CO].

[8] Tomás Feder, Shubha U. Nabar, and Evimaria Terzi. Anonymizing graphs. CoRR, abs/0810.5578, 2008.

[9] Frank Harary and Robert A. Melter. On the metric dimension of a graph. Ars Combinatoria, 2:191–195, 1976.

[10] Michael Hay, Gerome Miklau, David Jensen, Don Towsley, and Philipp Weis. Resisting structural re-identification in anonymized social networks. The VLDB journal, 1(1):102–114, August 2008.

[11] Teresa W. Haynes, Michael A. Henning, and Jamie Howard. Locating and total dominating sets in trees. Discrete Applied Mathematics, 154(8):1293–1300, May 2006.

[12] Samir Khuller, Balaji Raghavachari, and Azriel Rosenfeld. Landmarks in graphs. Discrete Appl. Math., 70(3):217–229, October 1996.

[13] Rong-Hua Li, Jeffrey Xu Yu, Xin Huang, Hong Cheng, and Zechao Shang. Measuring the impact of MVC attack in large complex networks. Information Sciences, 278(0):685 - 702, 2014.

[14] Kun Liu and Evimaria Terzi. Towards identity anonymization on graphs. The 2008 ACM SIGMOD International Conference on Management of Data, pages 93–106, 2008.
[15] Abedelaziz Mohaisen, Tamer AbuHmed, Hun Jeong Kang, Yongdae Kim, and Dae-Hun Nyang. Mistaking friends for foes: An analysis of a social network-based sybil defense in mobile networks. The 5th International Conference on Ubiquitous Information Management and Communication, pages 28:1–28:8, 2011.

[16] A. Narayanan and V. Shmatikov. De-anonymizing social networks. The 30th IEEE Symposium on Security and Privacy, 2009, pages 173–187, 2009.

[17] M. Netter, S. Herbst, and G. Pernul. Analyzing privacy in social networks—an interdisciplinary approach. Third International Conference on Social Computing, pages 1327–1334, Oct 2011.

[18] Wei Peng, Feng Li, Xukai Zou, and Jie Wu. A two-stage deanonymization attack against anonymized social networks. IEEE Transactions on Computers, 63(2):290–303, Feb 2014.

[19] Pierangela Samarati and Latanya Sweeney. Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression. Technical report, 1998.

[20] Peter J. Slater. Leaves of trees. Congressus Numerantium, 14:549–559, 1975.

[21] Klara Stokes and Vicenç Torra. Reidentification and k-anonymity: a model for disclosure risk in graphs. Soft Comput., 16(10):1657–1670, 2012.

[22] Bimal Viswanath, Mainack Mondal, Krishna P. Gummadi, Alan Mislove, and Ansley Post. Canal: Scaling social network-based sybil tolerance schemes. The 7th ACM European Conference on Computer Systems, pages 309–322, 2012.

[23] Xintao Wu, Xiaowei Ying, Kun Liu, and Lei Chen. A survey of privacy-preservation of graphs and social networks. Managing and Mining Graph Data, pages 421–453, 2010.

[24] Ismael G. Yero, Alejandro Estrada-Moreno, and Juan A. Rodríguez-Velázquez. The k-metric dimension of a graph: Complexity and algorithms. arXiv:1401.0342 [math.CO].

[25] Ismael G. Yero, Dorota Kuziak, and Juan A. Rodríguez-Velázquez. On the metric dimension of corona product graphs. Computers & Mathematics with Applications, 61(9):2793–2798, May 2011.

[26] Xiaowei Ying, Kai Pan, Xintao Wu, and Ling Guo. Comparisons of Randomization and K-degree Anonymization Schemes for Privacy Preserving Social Network Publishing. The 3rd Workshop on Social Network Mining and Analysis, pages 10:1–10:10, 2009.

[27] Xiaowei Ying and Xintao Wu. Randomizing Social Networks: a Spectrum Preserving Approach. SIAM International Conference on Data Mining, pages 739–750, 2008.
[28] Xiaowei Ying and Xintao Wu. On Link Privacy in Randomizing Social Networks. *Advances in Knowledge Discovery and Data Mining*, pages 28–39, 2009.

[29] Haifeng Yu, Phillip B. Gibbons, Michael Kaminsky, and Feng Xiao. Sybillimit: A near-optimal social network defense against sybil attacks. *IEEE/ACM Trans. Netw.*, 18(3):885–898, June 2010.

[30] Bin Zhou and Jian Pei. Preserving privacy in social networks against neighborhood attacks. *IEEE 24th International Conference on Data Engineering*, pages 506–515, April 2008.

[31] Bin Zhou, Jian Pei, and WoShun Luk. A brief survey on anonymization techniques for privacy preserving publishing of social network data. *SIGKDD Explor. Newsl.*, 10(2):12–22, December 2008.

[32] Lei Zou, Lei Chen, and M. Tamer Özs. K-automorphism: A general framework for privacy preserving network publication. *The VLDB Journal*, 2(1):946–957, August 2009.