The complexity of finding a large subgraph under anonymity constraints

Robert Bredereck[1][*], Sepp Hartung[1], André Nichterlein[1], and
Gerhard J. Woeginger[2][**]

[1] Institut für Softwaretechnik und Theoretische Informatik, TU Berlin
[2] Department of Mathematics and Computer Science, TU Eindhoven

Abstract. We define and analyze an anonymization problem in undirected graphs, which is motivated by certain privacy issues in social networks. The goal is to remove a small number of vertices from the graph such that in the resulting subgraph every occurring vertex degree occurs many times.
We prove that the problem is NP-hard for trees, and also for a number of other highly structured graph classes. Furthermore we provide polynomial time algorithms for other graph classes (like threshold graphs), and thereby establish a sharp borderline between hard and easy cases of the problem. Finally we perform a parametrized analysis, and we concisely characterize combinations of natural parameters that allow FPT algorithms.

1   Introduction

With the tremendous usage of social networks, the protection of privacy when releasing underlying data sets has become an important and active field of research [15]. If a graph contains only few vertices with some distinguished feature, then this might allow the identification (and violation of privacy) of the underlying real world entities with that particular feature. Hence in order to ensure pretty good privacy and anonymity behavior, every vertex should share its features with many other vertices. In a landmark paper[3], Liu and Terzi [11] considered in their setting the vertex degrees as feature; see Wu et al. [15] for other features considered in the literature. Correspondingly, a graph is called k-anonymous if for each vertex there are at least $k-1$ other vertices of same degree. Therein, different values of k reflect different privacy demands and the natural computational task arises to perform few changes to a graph in order to make it k-anonymous.

Liu and Terzi [11] proposed an heuristic algorithm for the task of making a graph k-anonymous by adding edges. The same variant has been studied by

[3] According to Google Scholar, accessed on June 14, 2013, since its publication in 2008 the paper has been cited 256 times.

Hartung et al. [10] from a parameterized complexity perspective. In this paper, we complement these previous studies by investigating the vertex deletion variant which is defined as follows:

DEGREE ANONYMITY BY VERTEX DELETION (ANONYM-V-DEL)
Instance: An undirected graph $G = (V, E)$; positive integers k and s.
Question: Is there subset $S \subseteq V$ of size at most s such that deleting S in G results in a k-anonymous graph?

Considering vertex deletions instead of edge additions seems to be a promising approach on practical instances, especially on social networks. Therein, the degree distribution of the underlying graphs often follow a so-called power law distribution [1] implying that there are only few high degree vertices and most vertices are of moderate degree; this suggests that only few vertices have to be removed in order to get a k-anonymous graph. For instance, consider the DBLP co-author graph (generated in Feb. 2012) with ≈ 715 thousand vertices corresponding to authors and ≈ 2.5 million edges indicating whenever two authors have a common scientific paper: This graph has maximum degree 804 but only 208 vertices are of degree larger than 208, whereas the average degree is 7. Interestingly, a heuristic that simply removes vertices violating the k-anonymous property proves that one has to remove no more than 338 vertices to make it 5-anonymous and even to make it 10-anonymous requires at most 635 vertex deletions.

While there are many different privacy models, there is a lack of algorithms with provably good performance (as explicitly observed by [3]). In this work, we will show that already the simple and highly specialized privacy model of ANONYM-V-DEL is computationally hard from the parameterized as well as from the approximation point of view. A variety of hardness results holds even in very restricted graph classes, as for instance trees, cographs, and split graphs.

One reason for this hardness is shown in the following two examples illustrating that the number s of allowed removals and the degree k of anonymity are independent of each other, and that a small change in one of these parameters might lead to a large jump of the other parameter.

Example 1. Let G be a graph on $n \geq 5$ vertices that consists of two components: a clique of size $n - 2$ and a clique of size two. This 2-anonymous graph cannot be transformed into a 3-anonymous graph by deleting only one vertex, however, deleting two vertices makes it (n−2)-anonymous. Hence, by slightly increasing s from 1 to 2 the reachable anonymity-degree jumps from $k = 2$ to $k = n - 2$.

Example 2. Let $G = (V, E)$ be a graph with vertices $X = \{x_1, \ldots, x_\ell\}$ and $Y = \{y_1, \ldots, y_\ell\}$ with an edge between x_i and y_j if $i + j > \ell$. Clearly, x_i and y_i are of degree i implying that G is 2-anonymous. Since $N(x_i) \subseteq N(x_{i+1})$ for all i, deleting any subset of Y preserves the invariant $\deg(x_1) \leq \deg(x_2) \leq \ldots \leq \deg(x_\ell)$. As the previous argument is symmetric, one can observe that to make G 3-anonymous one has to remove $2/3$ of the "jumps" in the initial sequences $\deg(x_1) < \deg(x_2) < \ldots \deg(x_\ell)$ and $\deg(y_1) < \ldots < \deg(y_\ell)$. Since removing one vertex in X (Y) removes only one jump in the sequence of X

(Y) and only one in Y (X), it follows that at least $2(\ell - 1) \cdot \frac{2}{3} \cdot \frac{1}{2} \approx \frac{2}{3}\ell = \frac{1}{3}|V|$ vertices have to be deleted in order to get a 3-anonymous graph. Summarizing, by requiring anonymity $k = 3$ instead of anonymity $k = 2$, the number of vertices needed to be removed jumps from zero to a constant fraction of vertices.

Related Work. Hartung et al. [10] studied the ANONYM-E-ADD problem as proposed by Liu and Terzi [11]. Given a graph and two positive integers k and s, ANONYM-E-ADD asks whether there exists a set of at most s edges whose addition makes the graph k-anonymous. The main result of Hartung et al. [10] is a polynomial problem kernel with respect to the parameter maximum degree Δ of the input graph. Furthermore, they showed that an heuristic algorithm proposed by Liu and Terzi [11] is optimal for ANONYM-E-ADD solutions larger than Δ^4. Chester et al. [4] investigated the computational complexity of ANONYM-E-ADD and variants with edge labels. They showed NP-hardness for the considered variants and a polynomial time algorithm for bipartite graphs.

Mathieson and Szeider [12] performed a parameterized complexity study for the problem of finding a minimum amount of graph editions in order to fulfill specified degree constraints. The graph editions considered are vertex deletion, edge insertion, edge deletions, and combinations thereof.

Our Results. Whereas every graph is trivially 1-anonymous, we will show that the combinatorial structure of 2-anonymous graphs is already rich and complicated: ANONYM-V-DEL for $k = 2$ is NP-hard, even for strongly restricted graph classes like trees, interval graphs, split graphs, trivially perfect graphs, and bipartite permutation graphs. All these hardness results are established by means of a general framework in Section 2. As a side-result, our framework implies the W[2]-hardness of various (natural) parameterized problem variants and the in-approximability of various (natural) optimization versions. Furthermore, we show that ANONYM-V-DEL is NP-hard even on graphs with maximum degree three; this result is in stark contrast with the fixed-parameter tractability of ANONYM-E-ADD with respect to the maximum degree Δ [10].

On the positive side, Section 3 presents (polynomial time) dynamic programming approaches for ANONYM-V-DEL on three graph classes: graphs of maximum degree two, P_3-free graphs, and threshold graphs. We frankly admit that these three graph classes carry an *extremely constraining* combinatorial structure: ANONYM-V-DEL is such a vicious problem that without these heavily constraining structures there remains no hope for polynomial time results. Figure 1 summarizes the considered graph classes and their containment relations.

Finally, we analyze the parametrized complexity of ANONYM-V-DEL in Section 4. Once again, ANONYM-V-DEL shows a difficult and challenging behavior: It is intractable with respect to each of the three (single) parameters s, k and Δ. Even worse, it is intractable with respect to the combined parameter (s, k). The only positive parametrized results come with the combined parameters (Δ, s) and (Δ, k). The latter result is based on bounding the number s of deleted vertices in terms of Δ and k.

Preliminaries. All graphs in this paper are undirected, loopless, and simple (that is, without multiple edges). Throughout we use n to denote the number of vertices
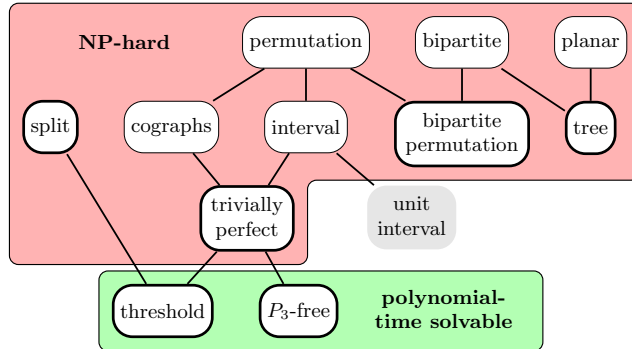
Fig. 1. The complexity landscape of Anonym-V-Del for various graph classes. The results for classes with thick frames are made in this work and they imply the results for classes with thin frames.

in the considered graph. The maximum vertex degree of a graph $G = (V, E)$ is denoted Δ_G. A vertex subset $S \subseteq V$ is called *k-deletion set* if $G[V \setminus S]$ is k-anonymous. For each vertex $v \in V$ we denote by $N_G(v)$ the set of neighbors of v and by $N_G[v] = N_G(v) \cup \{v\}$ the closed neighborhood. Correspondingly, for a vertex subset V' we set $N_G[V'] = \bigcup_{v \in V'} N_G[v]$ and $N_G(V') = N_G[V'] \setminus V'$. For $0 \le a \le \Delta$, the *block of degree a* is the set $D_G(a) \subseteq V$ of all vertices with degree a in G. Clearly, a graph is k-anonymous iff (if and only if) each block is either of size zero or at least k. We omit subscripts if the corresponding graph is clear from the context.

For the relevant notation of parameterized complexity and algorithmics we refer to the monographs of Downey and Fellows [7], Niedermeier [13]. Due to the space constraints some proofs are omitted.

2 Computational Hardness

In this section we provide NP-hardness results for Anonym-V-Del on several restricted graph classes such as trees, split graphs, and trivially perfect graphs. As a warm up, we first prove that Anonym-V-Del is NP-hard on graphs with maximum degree three. This contrasts the known fixed-parameter tractability of Anonym-E-Add with respect to the parameter maximum degree [10].

Theorem 1. *Anonym-V-Del is NP-hard on graphs with degree at most three.*

Proof. We give a reduction from the Vertex Cover problem which is known to be NP-complete even in three-regular graphs [8, GT1]. Therein, given a three-regular graph together with an integer $h \in \mathbb{N}$ the task is to decide whether there is vertex set of size at most h such that each edge has at least one endpoint in it. Given a Vertex Cover instance $(G = (V, E), h)$, start by copying G to a new graph G'. Finally, add $h + 1$ degree-zero vertices to G', set $s = h$, and $k = |V| + 1$.

If G contains a vertex cover S of size h, then deleting S in G' clearly results in an edgeless graph with $|V| + 1 = k$ vertices, implying that (G', s, k) is a yes-instance of Anonym-V-Del. In the other direction, for any k-deletion set S, since $2k > n + h + 1$ and G' contains $s + 1$ degree-zero vertices, all vertices in $G' \setminus S$ have degree zero. Thus, $S \cap V$ is a vertex cover in G. □

NP-hardness on trees. Next we show that Anonym-V-Del is NP-hard even on trees. Extracting the basic ideas of this result, subsequently we provide a generic reduction to show NP-hardness on trivially perfect graphs, bipartite permutation graphs, and split graphs. Both reductions will reduce from the NP-hard Set Cover problem, which is defined as follows [8, SP5]: Given a universe $A = \{a_1, \ldots, a_\alpha\}$, a collection $\mathcal{B} = \{B_1, \ldots, B_\beta\}$ of sets over A, and $h \in \mathbb{N}$ the task is to decide whether there is an index set $I \subseteq \{1, \ldots, \beta\}$ with $|I| \leq h$, such that $\bigcup_{i \in I} B_i = A$?

Let $(A, \mathcal{B}, h)$ be an instance of Set Cover. We assume without loss of generality that for each element $a \in A$ there exists a set $B \in \mathcal{B}$ with $a \in B$. Furthermore, we assume without loss of generality that each set $B \in \mathcal{B}$ occurs at least $h + 2$ times in $\mathcal{B}$. To reduce the amount of indices in the construction given below we introduce the function $f : A \to \mathbb{N}$ that maps an element $a_i \in A$ to $f(a_i) = \alpha + (h + 1)i$.

The reduction for trees is as follows. Set $k = 2$ and $s = h$ such that (G, k, s) is an equivalent Anonym-V-Del-instance. Graph $G = (V, E)$ is constructed as follows: For each element $a_i \in A$ add an *element gadget* consisting of a star $K_{1, f(a_i)}$ with the center vertex $v(a_i)$. Denote with $V_A = \{v(a_1), \ldots, v(a_\alpha)\}$ the set of all these center vertices.

For each set $B_j \in \mathcal{B}$ add a *set gadget* which is a tree rooted in a vertex $v(B_j)$. The root has $|B_j|$ child vertices where each element $a_i \in B_j$ corresponds to exactly one of these children, denoted by $v(a_i, B_j)$. Additionally, we add to $v(a_i, B_j)$ exactly $f(a_i)$ degree-one neighbors. Hence, the set gadget is a tree of depth three rooted in $v(B_j)$. We denote with $V_\mathcal{B} = \{v(B_1), \ldots, v(B_\beta)\}$ the set of all root vertices. Observe that, as each set $B_j \in \mathcal{B}$ occurs at least $h + 2$ times, the set gadgets are $h+2$-anonymous. Finally, to end up with one tree instead of a forest, repeatedly add edges between any degree-one-vertices of different connected components.

Observe that for each element $a_i \in A$ the only vertex of degree $f(a_i)$ is $v(a_i)$ and there are no other vertices violating the 2-anonymous property. The key point in the construction is that, in order to get a 2-anonymous graph, one has to delete vertices of $V_\mathcal{B}$: Let $a_i \in A$ be an element and $v(B_j)$ a root vertex such that $a_i \in B_j$. By construction the child vertex $v(a_i, B_j)$ of $v(B_j)$ corresponds to a_i and therefore has $f(a_i)$ child vertices. Thus, deleting $v(B_j)$ lowers the degree of $v(a_i, B_j)$ to $f(a_i)$ and, hence, $v(a_i)$ no longer violates the 2-anonymous property. Furthermore, as each set $B_j \in \mathcal{B}$ occurs at least $h+2$ times, the vertices $V_\mathcal{B}$ are 2-anonymous. Hence, given a set cover one can construct a corresponding k-deletion set of the same size and, thus, if $(A, \mathcal{B}, h)$ is a yes-instance, then (G, k, s) is a yes-instance. The proof of the converse direction which implies the following theorem will be given later, after introducing the generic reduction.

Theorem 2. Anonym-V-Del *is NP-hard on trees even if $k = 2$.*

Generic Reduction. We now generalize the reduction given in the previous paragraph. More specifically, we will define properties such that a graph G fulfilling them together with $s = h$ and $k = 2$ forms a yes-instance of ANONYM-V-DEL iff the given SET COVER instance $(A, \mathcal{B}, h)$ is a yes-instance. Based on that, we then describe the construction of a several graphs contained in different graph classes and fulfilling the properties. Formally, we require the constructed graph $G = (V, E)$ to fulfill the following:

1. For each element $a_i \in A$ there is a corresponding vertex, denoted by $v(a_i)$, in G and the vertex set $V_A = \{v(a_1), \ldots, v(a_\alpha)\}$ is exactly the set of vertices not being 2-anonymous in G.
2. For each set $B_j \in \mathcal{B}$ there is a corresponding vertex $v(B_j)$ in G and for each element $a_i \in B_j$ the vertex $v(B_j)$ has a neighbor $v(a_i, B_j)$ with $\deg(v(a_i, B_j)) = \deg(v(a_i)) + 1$.
   Set $V_{\mathcal{B}} = \{v(B_1), \ldots, v(B_\beta)\}$ and $A_{B_j} = \{v(a_i, B_j) \mid a_i \in B_j\}$.
3. The vertex subsets V_A, $V_{\mathcal{B}}$, and $A_{B_1}, \ldots, A_{B_\beta}$ are pairwise disjoint. We set $A_{\mathcal{B}} = \bigcup_{B_j \in \mathcal{B}} A_{B_j}$.
4. For each $D \subseteq V_{\mathcal{B}}$, $|D| \leq h$, the set of vertices violating the 2-anonymous property in $G[V \setminus D]$ is a subset of V_A.
5. It holds: (a) $|N[v] \cap V_A| \leq 1$ for each vertex $v \in V$, (b) $N(A_{B_j}) \cap V_{\mathcal{B}} = \{v(B_j)\}$ for all $B_j \in \mathcal{B}$, and (c) $N(V_A) \cap (V_{\mathcal{B}} \cup A_{\mathcal{B}}) = \emptyset$.
6. For each vertex $v \in V$ there is a vertex $u \in V_{\mathcal{B}}$ such that $N(v) \cap A_{\mathcal{B}} \subseteq N(u)$.
7. Any two vertices $u \in V_A$ and $v \notin A_{\mathcal{B}}$ satisfy $|\deg(v) - \deg(u)| > s$.

It is not hard to verify that the graph constructed in the reduction in the previous paragraph has the above properties. Before proving the correctness of the generic reduction we show the following observation.

Observation 1. For each $D \subseteq V_{\mathcal{B}}$, $|D| \leq h$, the set $V_A \setminus \{v(a_i) \mid \exists v(B_j) \in D : a_i \in B_j\}$ is exactly the set of vertices not being 2-anonymous in $G[V \setminus D]$.

Lemma 1. *Let G be a graph satisfying Properties 1 to 7 for a given instance $(A, \mathcal{B}, h)$ of* SET COVER*. Then $(G, 2, h)$ is a yes-instance of* ANONYM-V-DEL *if and only if $(A, \mathcal{B}, h)$ is a yes-instance of* SET COVER*.*

Proof. If there is an index set I, $|I| \leq h$, such that $\bigcup_{j \in I} B_j = A$, then by Observation 1 the set $S = \{v(B_j) \mid j \in I\} \subseteq V_{\mathcal{B}}$, $|S| = |I|$, is a k-deletion set for G. It remains to prove the reverse direction.

Let S be a k-deletion set of size at most $s = h$ for $G = (V, E)$. We form a k-deletion set S' for G such that $S' \subseteq V_{\mathcal{B}}$ and $|S'| \leq |S|$. Consider each vertex $v \in S$: If $v \in V_{\mathcal{B}}$, then add v to S' (Case 1). If $v \in N[V_A]$, then by Property 5 there is only one a_i such that $v \in N[v(a_i)]$ and we add a vertex $v(B_j) \in V_{\mathcal{B}}$ with $a_i \in B_j$ to S' (Case 2). Finally, if $v \in N[A_{\mathcal{B}}]$, then by Property 6 there is a vertex $u \in V_{\mathcal{B}}$ with $N(v) \cap A_{\mathcal{B}} \subseteq N(u)$ and we add u to S' (Case 3).

We next prove that S' is a k-deletion set for G and thus by Observation 1 the index set corresponding to the vertices in S' is a solution of size $|S'|$ to the SET COVER instance.

Assume towards a contradiction that $G[V \setminus S']$ is not 2-anonymous. Denoting by $X \subseteq V \setminus S'$ the set of vertices not being 2-anonymous, it follows from Observation 1 that $X \subseteq V_A$. Moreover, by the construction of S' (see Case 2) and Observation 1 it follows that $N[X] \cap S = \emptyset$ and thus $\deg_{G[V \setminus S]}(u) = \deg_G(u)$ for all $u \in X$. Hence, for each $u \in X$ there is a vertex $w \in V$ such that $\deg_G(u) = \deg_{G[V \setminus S]}(w)$ and thus by Property 7 it follows that $w \in N[A_{\mathcal{B}}]$. This implies a contradiction to the construction of S' because from $w \in N[A_{\mathcal{B}}]$ it follows that S' contains w's neighbor in $V_{\mathcal{B}}$ (see Case 3) and thus $u \notin X$ by Observation 1. $\square$

Using this generic reduction we now show NP-hardness on several graph classes which are defined as follows (see Brandstädt et al. [2]): *Trivially perfect graphs* are the (P_4, C_4)-free graphs, that is, they do not contain an induced path or cycle on four vertices. A graph G is a *bipartite permutation graph* if G is bipartite and does not contain an asteroidal triple (is AT-free). Three vertices of a graph form an asteroidal triple if every two of them are connected by a path avoiding the neighborhood of the third. A graph is a *split graph* if it can be partitioned into a clique and an independent set.

Theorem 3. ANONYM-V-DEL *is NP-hard on trivially perfect graphs, bipartite permutation graphs, and split graphs.*

Since SET COVER is W[2]-complete with respect to h [7] we have the following.

Corollary 1. ANONYM-V-DEL *is W[2]-hard with respect to parameter s, even if $k = 2$ and if the input graph is a tree, a bipartite permutation graph, a split graph, or a trivially perfect graph.*

Dom et al. [6] showed that SET COVER does not admit a polynomial kernel with respect to the combined parameter (α, h). Observe that in all above constructions except the one for split graphs we can bound s and Δ in a polynomial in α and h.

Corollary 2. ANONYM-V-DEL *on trees, bipartite permutation graphs or trivially perfect graphs does not admit a polynomial kernel with respect to the combined parameter (k, s, Δ).*

There are two natural optimization versions associated with ANONYM-V-DEL: in one version (called MAX ANONYM-V-DEL) the goal is to maximize the anonymity k subject to the constraint that the number s of deleted vertices does not exceed a given bound; in the other version (called MIN ANONYM-V-DEL) the goal is to minimize the number s of deleted vertices subject to the constraint that the anonymity does not go below a certain given bound. As SET COVER is NP-hard to approximate within a ratio $o(\log n)$ [14], the above reduction yields the following inapproximability result.

Corollary 3. *The optimization problem* MIN ANONYM-V-DEL *on n-vertex graphs cannot be approximated within a factor of $o(\log n)$, unless $P = NP$.*

Since the above reduction gives NP-hardness for $k = 2$, we immediately get inapproximability within a factor of two for MAX ANONYM-V-DEL.

Corollary 4. *The optimization problem* MAX ANONYM-V-DEL *cannot be approximated*

- *within a factor of $2 - \epsilon$, unless $P = NP$.*
- *within a factor of $2 - \epsilon$ in $f(s)n^{O(1)}$ time for any computable f, unless FPT=W[2].*

3 Polynomially Solvable Cases

We complement our intractability results for ANONYM-V-DEL from Theorem 2 by showing that ANONYM-V-DEL is polynomial time solvable on graphs with maximum degree two, on graphs that are disjoint unions of cliques, and on threshold graphs.

3.1 Graphs with maximum degree two

In contrast to graphs of maximum degree three (see Theorem 1), we observe that ANONYM-V-DEL is polynomial time solvable on graphs of maximum degree two. Note that a graph of maximum degree two is just a collection of paths and cycles. Given five integers d_0, d_1, d_2, x, y, it is easy to decide whether it is possible to remove x vertices from a path of length y (respectively, from a cycle of length y) such that there survive precisely d_0 vertices of degree zero, d_1 vertices of degree one, and d_2 vertices of degree two. A straight-forward dynamic programming approach based on this observation leads to the following.

Theorem 4. *On graphs of maximum degree two,* ANONYM-V-DEL *is polynomial time solvable.*

3.2 Disjoint union of cliques

Note that ANONYM-V-DEL is trivial on cliques: either the clique size is at least k, or otherwise one has to delete all the vertices. The following theorem shows that polynomial time solvability also carries over to the case where the graph is the disjoint union of several cliques. (Recall that a graph is the disjoint union of cliques if and only if it does not contain the 3-vertex path P_3 as an induced subgraph.)

Theorem 5. *On a P_3-free graph G with n vertices and maximum degree Δ,* ANONYM-V-DEL *can be solved in $O(n^2\Delta)$ time.*

3.3 A polynomial time result for threshold graphs

We recall that a graph $G(V, E)$ is a *threshold graph* if there are positive real vertex weights $w(v)$ for $v \in V$, such that $\{v_1, v_2\} \in E$ if and only if $w(v_1) + w(v_2) \geq 1$; see Chvátal and Hammer [5] and Golumbic [9] for more information. Without loss of generality we will assume throughout that the vertex weights satisfy the following conditions:

- The vertex weights are pairwise distinct, and satisfy $0 < w(v) < 1$
- Any $v_1, v_2 \in V$ satisfy $w(v_1) + w(v_2) \neq 1$; in particular $w(v_1) \neq \frac{1}{2}$

Note that the closed neighborhoods in a threshold graph are totally ordered by inclusion: whenever $w(v_1) < w(v_2)$, then $N_G[v_1] \subseteq N_G[v_2]$ and consequently $\deg(v_1) \leq \deg(v_2)$.

Lemma 2. *Let $U \subseteq V$ be a subset of vertices with $|U| \geq 2$, let $w_{\min} = \min_{u \in U} w(u)$ and $w_{\max} = \max_{u \in U} w(u)$, and let $u_0, u_1 \in U$ be the vertices with $w(u_0) = w_{\min}$ and $w(u_1) = w_{\max}$. All vertices in U have identical degree, if and only if there is no vertex $v \in V \setminus \{u_0, u_1\}$ with $1 - w_{\max} < w(v) < 1 - w_{\min}$.*

Proof. Note that all vertices in U have identical degree, if and only if $N_G[u_0] = N_G[u_1]$. The latter condition in turn holds if and only if there is no vertex v in the graph (with $v \neq u_0$ and $v \neq u_1$) that is adjacent to u_1 but not to u_0, and this is equivalent to the stated condition $1 - w_{\max} < w(v) < 1 - w_{\min}$. $\qquad\square$

Now consider some block U of constant degree in an optimal subgraph for ANONYM-V-DEL, and let $u_0, u_1 \in U$ and $w_{\min}$ and $w_{\max}$ be defined as in the lemma. The *territory* of this block is defined as the union of the two closed intervals $[w_{\min}, w_{\max}]$ and $[1 - w_{\max}, 1 - w_{\min}]$; note that these two intervals will overlap if $w_{\min} < \frac{1}{2} < w_{\max}$. The *canonical superset* $U^* \subseteq V$ consists of u_0 and u_1, together with all vertices $v \in V$ that satisfy $w_{\min} \leq w(v) \leq w_{\max}$ but not $1 - w_{\max} < w(v) < 1 - w_{\min}$. One message of Lemma 2 is that distinct blocks in an optimal subgraph must have disjoint territories. Another message of the lemma is that we may as well replace every block U by its canonical superset U^*: By adding these vertices, the degree in every block either remains the same or is uniformly increased by $|U^*| - |U|$. And if the territories of distinct blocks were disjoint before the replacement, then they will also be disjoint after the replacement. In other words, such a replacement does not violate k-anonymity but simplifies the combinatorial structure of the considered subgraph.

This suggests the following dynamic programming approach. For every real number r with $0 \leq r \leq \frac{1}{2}$, we consider the threshold graph G_r that is induced by the vertices $v \in V$ with $r \leq w(v) \leq 1 - r$; note that the only crucial values for r are the $O(n)$ values $w(v)$ and $1 - w(v)$ that fall between the bounds 0 and $\frac{1}{2}$. The goal is to compute for every graph G_r a largest k-anonymous subgraph. We start our computations with $r = \frac{1}{2}$ and work downwards towards $r = 0$.

The initialization step of the dynamic program handles subgraphs that consist of a single block whose territory contains the number $\frac{1}{2}$. Such a block will either be empty, or it is a canonical superset specified by two values $w_{\min}$ and $w_{\max}$. All in all, this only yields a polynomial number of cases to handle. In the main computation phase of the dynamic program, we consider a general graph G_r and check all possibilities for the outermost block, which is the block whose territory is farthest away from the center point $\frac{1}{2}$. Since this territory is the union of two intervals $[r, q]$ and $[1 - q, 1 - r]$, we may simply check all possibilities for the interval boundary q, and then combine the corresponding block with the (previously computed) largest k-anonymous subgraph for graph G_q. Since there

is only a linear number $O(n)$ of candidate values for q, the largest k-anonymous subgraph of G_r can be found in linear time.

Theorem 6. *On threshold graphs with n vertices,* ANONYM-V-DEL *can be solved in $O(n^2)$ time.* □

4 Parametrized results

Theorem 1, Theorem 2, and Corollary 1 show that there is no hope for fixed-parameter tractability neither for any of the individual parameters s, k or Δ nor for the combined parameter (s, k). In this subsection, we show that ANONYM-V-DEL becomes fixed-parameter tractable when considering the combined parameters (s, Δ) as well as (k, Δ). We start with a fixed-parameter algorithm for (s, Δ) and show that the minimum size of a solution is bounded by a function only depending on k and Δ.

Theorem 7. ANONYM-V-DEL *can be solved in $(s\Delta)^{O(s\Delta^2)} n^2 \log n$ time.*

Lemma 3. *For every yes-instance $(G = (V, E), k, s)$ of* ANONYM-V-DEL *with Δ denoting the maximum degree of G there is a subset $S \subseteq V$ with $|S| < 2^\Delta \Delta^3 2k$ such that $G[V \setminus S]$ is k-anonymous.*

By combining Theorem 7 and Lemma 3 we obtain fixed-parameter tractability with respect to the parameter (k, Δ): For an instance (G, k, s) of ANONYM-V-DEL apply the algorithm from Theorem 7 on $(G, k, \min\{s, 2^\Delta \Delta^3 2k\})$. The running time is bounded by $(2^\Delta \Delta^4 2k)^{O(2^\Delta \Delta^5 2k)} n^2 \log n$.

Corollary 5. ANONYM-V-DEL *is fixed-parameter tractable with respect to the combined parameter (k, Δ).*

5 Conclusion

In this paper, we have complemented the investigations of Hartung et al. [10] on the edge addition version of the degree anonymity problem to the vertex deletion version. To our surprise, there is a strong contrast in the complexity of the two problem versions: Whereas ANONYM-E-ADD admits a polynomial kernel with respect to the maximum degree [10], we proved NP-hardness of ANONYM-V-DEL on graphs with maximum degree three. Furthermore, bounding one of the input parameters s and k does not yield fpt-algorithms for ANONYM-V-DEL; however bounding the degree and bounding one of the input parameters s and k brings the problem into FPT.

Our results also provide a good view on the colorful complexity landscape of the vertex deletion version. We have shown that the problem is hard for most of the standard graph classes, and that one has to move on to highly structured classes like threshold graphs in order to get some polynomial time results. A

number of questions remained open: What is the complexity of Anonym-V-Del on claw-free graphs? What is the complexity of Anonym-V-Del on unit interval graphs? Corollary 4 does not exclude the existence of a constant-factor approximation for Max Anonym-V-Del. Are there stronger inapproximability results? Can the bounds stated in Theorem 7 and Lemma 3 be improved?

Bibliography

[1] A. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509, 1999.

[2] A. Brandstädt, V. B. Le, and J. P. Spinrad. *Graph Classes: a Survey*, volume 3 of *SIAM Monographs on Discrete Mathematics and Applications*. SIAM, 1999.

[3] S. Chester, B. Kapron, G. Srivastava, and S. Venkatesh. Complexity of social network anonymization. *Social Network Analysis and Mining*, pages 1–16, 2012. Online available.

[4] S. Chester, B. M. Kapron, G. Ramesh, G. Srivastava, A. Thomo, and S. Venkatesh. Why Waldo befriended the dummy? k-anonymization of social networks with pseudo-nodes. *Social Network Analysis and Mining*, pages 1–19, 2012. ISSN 1869-5450. Available online.

[5] V. Chvátal and P. L. Hammer. Aggregation of inequalities in integer programming. *Annals of Discrete Mathematics*, 1:145–162, 1977.

[6] M. Dom, D. Lokshtanov, and S. Saurabh. Incompressibility through colors and IDs. In *Proc. 36th ICALP*, volume 5555 of *LNCS*, pages 378–389. Springer, 2009. doi: 10.1007/978-3-642-02927-1_32.

[7] R. G. Downey and M. R. Fellows. *Parameterized Complexity*. Springer, 1999.

[8] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Freeman, 1979.

[9] M. C. Golumbic. *Algorithmic graph theory and perfect graphs*, volume 57 of *Annals of Discrete Mathematics*. Elsevier B. V., 2nd edition, 2004. First edition Academic Press, 1980.

[10] S. Hartung, A. Nichterlein, R. Niedermeier, and O. Suchý. A refined complexity analysis of degree anonymization on graphs. In *Proceedings of the 40th International Colloquium on Automata, Languages and Programming (ICALP '13)*, LNCS. Springer, 2013. Accepted for publication.

[11] K. Liu and E. Terzi. Towards identity anonymization on graphs. In *Proc. ACM SIGMOD International Conference on Management of Data*, SIGMOD '08, pages 93–106. ACM, 2008.

[12] L. Mathieson and S. Szeider. Editing graphs to satisfy degree constraints: A parameterized approach. *J. Comput. Syst. Sci.*, 78(1):179–191, 2012.

[13] R. Niedermeier. *Invitation to Fixed-Parameter Algorithms*. Oxford University Press, 2006.

[14] V. V. Vazirani. *Approximation Algorithms*. Springer, 2001.

[15] X. Wu, X. Ying, K. Liu, and L. Chen. A survey of privacy-preservation of graphs and social networks. In *Managing and mining graph data*, pages 421–453. Springer, 2010.