REPORTS
IN
INFORMATICS

ISSN 0333-3590

Fixed Parameter Set Splitting, Obtaining a Linear Kernel and Improving Running Time

Daniel Lokshtanov and Christian Sloper

REPORT NO 280 **August 2004**

Department of Informatics

UNIVERSITY OF BERGEN

Bergen, Norway

This report has URL http://www.ii.uib.no/publikasjoner/texrap/ps/2004-280.ps

Reports in Informatics from Department of Informatics, University of Bergen, Norway, is available at
http://www.ii.uib.no/publikasjoner/texrap/.

Requests for paper copies of this report can be sent to:

Department of Informatics, University of Bergen, Høyteknologisenteret,
P.O. Box 7800, N-5020 Bergen, Norway

Fixed Parameter Set Splitting, Obtaining a Linear Kernel and Improving Running Time.

Daniel Lokshtanov, Christian Sloper

Department of Informatics
University of Bergen
Norway

31st August 2004

Abstract

We study the problem k-SET SPLITTING from a Fixed Parameter point of view. We give a linear kernel of $2k$ elements and $2k$ sets and improve on the current best running time for the problem, giving a $\mathcal{O}^*(4^k)$ algorithm. This is done by reducing the problem to a bipartite graph problem where we use crown decomposition to reduce the graph. We show that this result also gives a good kernel for Max Cut.

Keywords: Algorithms, Graph Algorithms.

1 Introduction

The problem we study in this short note is MAXIMUM SET SPLITTING. The transformation from MAXIMUM SET SPLITTING to MAX CUT preserves the parameter and thus our kernel applies for this problem as well.

k-SET SPLITTING
INSTANCE: A collection $\mathcal{F}$ of subsets of a finite set X, and a positive integer k
PARAMETER: k
QUESTION: Is there a subfamily $\mathcal{F}' \subseteq \mathcal{F}, |\mathcal{F}'| \geq k$, and a partition of X into disjoint subsets X_0 and X_1 such that for every $S \in \mathcal{F}', S \cap X_0 \neq \emptyset$ and $S \cap X_1 \neq \emptyset$?

SET SPLITTING is a well studied problem and a decision version of the problem appears in [GJ79] as problem [SP4]. It is APX-complete [Pe94] and there have been several approximation algorithms published. The most notable are Anderson and Engebretsen [AE97] with a factor of 0.7240, and Zhang and Ling [ZL01] with a factor of 0.7499.

In parameterized algorithms there have been several results published. The first by Dehne, Fellows, and Rosamond [DFR03] give a $\mathcal{O}^*(72^k)$ FPT algorithm. Dehne, Fellows, Rosamond, and Shaw [DFRS04] then improved on this result giving a $\mathcal{O}^*(8^k)$ using a combination of the techniques *greedy localization* and *crown decomposition*.

We will show how to improve on this by reducing the problem to a bipartite graph problem, BIPARTITE COLORFUL NEIGHBORHOOD. We will use crown decomposition to reduce the graph then show that a simple greedy algorithm decides instances where $k \leq \frac{|\mathcal{F}|}{2}$. Together the two results give a linear kernel with at most $2k$ elements and at most $2k$ sets and we thus obtain a running time of $\mathcal{O}^*(4^k)$ using a trivial brute force algorithm.

2 Preliminaries

We assume that in a SET SPLITTING instance every set contains at least two elements of X. This is a natural assumption as sets of size one cannot be split in any case.

We employ the $\mathcal{O}^*$ notation introduced in [W03], which suppresses the polynomials in the running time and focus on the exponentials. Thus for a $O^*(2^k)$ algorithm there exists a constant c such that the running time is $O(2^k n^c)$.

In graphs the neighbors of a vertex v are denoted as the set $N(v)$, and the neighbors of a set $S \subseteq V$, $N(S) = \bigcup_{v \in S} N(v) - S$.

We use the simpler $G \setminus v$ to denote $G = (V \setminus v, E)$ and $G \setminus e$ to denote $G = (V, E \setminus e)$ where v and e is a vertex and an edge respectively. Likewise for sets $G \setminus V'$ denotes $G = (V \setminus V', E)$ and $G \setminus E'$ denotes $G = (V, E \setminus E')$ where V' is a set of vertices and E' is a set of edges. For a set of vertices $V' \subset V$ we use $\delta(V')$ as the lowest degree in V'.

3 Reducing to a graph problem

Recently the fixed parameter algorithms for many problems have been improved using crown decompositions. It is a common technique [FHRST04, PS04] to create an auxiliary graph model from the problem instance and then show that a reduction(using crown decomposition) in the graph model leads to reduction of the problem instance. This is essentially what we will do here, but since the crown reduction rule is our only reduction rule we think it is cleaner to transform the problem instead.

We reformulate the problem as a problem on bipartite graphs. Let $G(V_M, V_X, E)$ be bipartite graph, where V_M is a set of vertices with a vertex v_m for each set $m \in \mathcal{F}$, V_X is a set of vertices with a vertex v_x for each element $x \in X$ and let $v_x, v_m \in E$ be an edge if $x \in m$.

The problem is now reduced to color the set V_x black and white such that at least k vertices of V_M has a colorful neighborhood, i.e. at least one neighbor of each color. It is easy to see that this problem is equivalent to k-SET SPLITTING.

> BIPARTITE COLORFUL NEIGHBORHOOD (BCN)
> INSTANCE: A bipartite graph collection $G = (V_M, V_X, E)$, and a positive integer k
> PARAMETER: k
> QUESTION: Is there a two-coloring of V_X such that there exists a set $S \subseteq V_M$, $|S| \geq k$ where each element of S has a colorful neighborhood?

As mentioned we will use crown decomposition to reduce the problem. Crown decomposition is particularly well suited for use in bipartite graphs as Lemma 1 ensures us the existence of a crown decomposition in any bipartite graph.

Definition 1 *A* crown decomposition (H, C, R) *in a graph* $G = (V, E)$ *is a partitioning of the vertices of the graph into three sets* H*,* C*, and* R *that have the following properties:*

1. H (the head) *is a separator in* G *such that there are no edges in* G *between vertices belonging to* C *and vertices belonging to* R*.*
2. $C = C_u \cup C_m$ (the crown) *is an independent set in* G*.*
3. *There is an bijective mapping* $f : H \to C_m$ *(i.e. perfect matching).*

We can find the following lemma in [CFJ04].

Lemma 1 *If a graph* $G = (V, E)$ *has an independent set* $I \subseteq V(G)$ *such that* $|N(I)| < |I|$ *then a nontrivial crown decomposition* (H, C, R) *with* $C \subset I$ *for* G *can be found in time* $O(|V| + |E|)$*.*

Our main reduction rule is the following lemma that states that any crown decomposition contains a crown decomposition where the head and crown can be removed from the graph.

Lemma 2 *Given a bipartite graph* $G = (V_M, V_X, E)$ *where* $|V_M| < |V_X|$*, there exists a nontrivial crown decomposition* (H, C, R) *such that* G *has a* k*-BCN* $\iff$ $G' = (V_M \setminus H, V_X \setminus C, E)$ *has a* $(k - |H|)$*-BCN*

Proof. Since $|V_M| < |V_X|$ there exists a component $V'_M \subseteq V_M, V'_X \subseteq V_X$ where $|V'_M| < |V_X|'$. By Lemma 1 we know that this component has a crown decomposition (H', C', R') where $H' \subseteq V'_M$. We now use this crown to identify another crown (H, C, R) with the desired properties.

We assume $R \neq \emptyset$, if this is not the case we can move a vertex from C_u to R. If $C_u \cup R = \emptyset$ then $|V'_M| = |V'_X|$ contradicting that $|V'_M| < |V'_X|$.

We iteratively compute this new crown in the following manner. Let $H_0 \subseteq H'$ be the vertices of H' that have a neighbor in $V_X \setminus C$, H_0 is nonempty since $R \neq \emptyset$ and H' is a vertex separator. Let C_0 be the vertices of C that are matched to H_0.Let $H_{i+1} = N(C_i)$ and C_{i+1} be the vertices matched to H_{i+1}. Run iteratively until $H_{i+1} = H_i$ then let $H = H_i, C = \{v_x \mid N(v_x) \subseteq H\}$ and R be the remainder.

From the construction of (H, C, R) it is clear that this is a crown decomposition. We proceed to show that G has a k-BCN $\Leftrightarrow$ $G' = (V_M \setminus H, V_X \setminus C, E)$ has a $(k - |H|)$-BCN.

In one direction assume in contradiction that G has a k-BCN but that $G' = (V_M \setminus H, V_X \setminus C, E)$ has no $(k - |H|)$-BCN. Then the removed elements C must have participated in a colorful neighborhood for more than $|H|$ vertices in V_M. This is clearly impossible as $N(C) \subseteq H$.

In the other direction we have that $G' = (V_M \setminus H, V_X \setminus C, E)$ has a $(k - |H|)$-BCN. We can assume that every vertex in $V_X \setminus C$ has been colored. We can now color C such that every vertex in H has a colorful neighborhood. For every vertex $h \in H_0$ we can color the vertex matched to h different to h's neighbor in $V_X \setminus C$. Observe that after coloring C_j, all vertices in $H_{j+1} \setminus H_j$ have a neighbor in C_j. Thus we can obtain a colorful neighborhood for each vertex $h \subseteq H_{j+1} \setminus H_j$ by coloring its matched vertex appropriately. Thus every vertex in H has a colorful neighborhood and G has a k-BCN.

□

Corollary 1 *In a reduced graph* $|V_X| \leq |V_M|$

Proof. Assume in contradiction that the graph is reduced but $|V_M| < |V_X|$. By Lemma 2 this graph can be reduced, contradicting the assumption that the graph was reduced. $\square$

We have obtained the inequality $|V_X| \leq |V_M|$. We now show that we can obtain a similar relationship between $|V_M|$ and k by analyzing the effectiveness of a simple greedy algorithm for the problem.

Greedy algorithms for SET SPLITTING seem to do quite well and it is indeed possible to prove that there is a polynomial time algorithm that splits at least half of the sets. For our graph problem this is the equivalent of proving that it is always possible to two-color V_X such that at least half of V_M has a colorful neighborhood.

Lemma 3 *It is always possible to find a partitioning* B, W *of* V_X *such that at least half of the vertices in* V_M *has a colorful neighborhood.*

Proof. For a subset $V'_X \subseteq V_X$ we define $M(V'_X) = \{v_m \mid v_m \in V_M, N(v_m) \subseteq V'_X\}$.

We now induct on the size of V'_X.

Base case: If $|V'_X| = 1$, $M(V'_X) = \emptyset$. Thus the statement is trivially true.

Inductive Hypothesis: We assume that for all sets $V'_X \subseteq V_X$ of size n_0 we can find a partitioning B', W' of V'_X such that at least half of the vertices in $M(V'_X)$ has a colorful neighborhood.

Inductive Step: Assume any set $V''_X \subseteq V_X$ where $|V''_X| = n_0 + 1$. Let $v_x \in V''_X$ be an arbitrary vertex in V''_X, and let $M' = M(V''_X \setminus v_x)$. By the inductive hypothesis we can find a partitioning B', W' such that half of the vertices in M' has a colorful neighborhood. Since $\delta(V_M) > 1$, every vertex in $M(V''_X) \setminus M'$ has at least one neighbor $B' \cup W'$. We can assume without loss of generality that half of the vertices of $M(X'') \setminus M'$ have a neighbor in B'. Hence the partitioning $B', W' \cup \{v_x\}$ ensures that at least half of the vertices in $M(V''_X)$ has a colorful neighborhood.

$\square$

The following corollary follows directly from the above lemma. It is easy to design a greedy algorithm that mimic the inductive procedure in the proof and produces the necessary partitioning.

Corollary 2 *All instances where* $k \leq \frac{|V_M|}{2}$ *are trivially Yes-instances.*

Theorem 1 *k-BCN can be solved in time* $\mathcal{O}^*(4^k)$ *and has a linear kernel where* $|V_X| \leq |V_M| < 2k$.

Proof. By Corollary 2 we have that for a nontrivial instance $(G, k), k > \frac{|V_M|}{2}$. By Corollary 1 we have that $|V_X| \leq |V_M|$. Thus we have the inequality $|V_X| \leq |V_M| < 2k$. We can try every possible coloring of the elements of V_X in time $\mathcal{O}^*(4^k)$. $\square$

The following corollary then follows by a transformation of the kernel back to k-Set Splitting.

Corollary 3 k-SET SPLITTING *can be solved in time* $\mathcal{O}^*(4^k)$ *and has a linear kernel of* $2k$ *sets and* $2k$ *elements.*

4 An application to Max Cut

In this section we mention that our result also applies to the more known MAX CUT, which can be encoded using SET SPLITTING.

> MAX CUT
> INSTANCE: A graph $G = (V, E)$, and a positive integer k
> PARAMETER: k
> QUESTION: Is there a partitioning of V into two sets V',V'' such that the number of edges between V' and V'' is at least k?

Let the set of elements $X = V$ and for every edge $vu \in E$ create a set $\{v, u\}$. A splitting of a set vu now correspond to placing u and v in different partitions in Max Cut. The results on Set Splitting now apply to Max Cut.

Observation 1 *k-MAX CUT has a linear kernel of $2k$ vertices and $2k$ edges.*

Using the best known exact algorithm for this problem, a $\mathcal{O}^*(2^{|E|/4})$ algorithm by Fedin and Kulikov [FK02] we get a running time of $\mathcal{O}^*(2^{k/2})$ which is equivalent to Prieto's algorithm in [P04] where she used the *Method of Extremal Structure*,another well known FPT technique, to reach a kernel of k vertices and $2k$ edges. Earlier Mahajan, Raman[MR99] has used yet another technique to reach the same number of edges.

5 Conclusion

We have shown that k-SET SPLITTING has a linear kernel and we have improved the current best algorithm of $\mathcal{O}^*(8^k)$ to $\mathcal{O}^*(4^k)$. [DFRS04] also used crown decomposition, but on another graph model. This shows that when using crown decompositions the selection of the right graph model is crucial.

Since k-SET SPLITTING also encodes k-MAX CUT we have a linear kernel for this problem also. We have reached a good kernel, equalling the best for this problem using a different approach.

Having achieved a linear kernel for Set Splitting we believe that it is now possible to improve the running time further using results from exact algorithms or a bounded search tree technique.

References

[AE97] G. Andersson and L. Engebretsen. Better approximation algorithms for set splitting and Not-All-Equal-Sat, in *Information Processing Letters*, Vol. 65, pp.305-311,1998.

[CFJ04] B. Chor, M. Fellows, and D. Juedes. Linear Kernels in Linear Time, or How to Save k Colors in $O(n^2)$ steps. Proceedings of WG2004, Springer-Verlag, *Lecture Notes in Computer Science* (2004)

[DFR03] F. Dehne, M. Fellows, and F. Rosamond. An FPT Algorithm for Set Splitting, in *Proceedings WG2004 - 30th Workshop on Graph Theoretic Concepts in Computer science*, LNCS. Springer Verlag, 2004.

[DFRS04] F. Dehne, M. Fellows, and F. Rosamond, P. Shaw. Greedy Localization, Iterative Compression and Modeled Crown Reductions: New FPT Techniques and Improved Algorithms for Max Set Splitting and Vertex Cover. To appear in the proceedings of IWPEC04.

[FK02] S. Fedin and A. Kulikov. A $2^{|E|/4}$-time Algorithm for MAX-CUT. Zapiski nauchnyh seminarov POMI, No.293, pages 129-138,(2002). English translation to appear in Journal of Mathematical Sciences.

[FHRST04] M.Fellows, P.Heggernes, F.Rosamond, C. Sloper, J.A.Telle, Finding k disjoint triangles in an arbitrary graph, *To appear WG2004*

[GJ79] M. Garey and D. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness.*W.H. Freeman, San Francisco, 1979.

[MR99] M. Mahajan, V. Raman. Parameterizing above guaranteed values: MaxSat and MaxCut. Journal of Algorithms Volume 31, Issue 2(May 1999) pages 335-354.

[P04] E. Prieto. The Method of Extremal Structure on the k-Maximum Cut Problem. Manuscript, to appear.

[PS04] E.Prieto and C. Sloper. Reducing to Independent Set Structure — the Case of k-INTERNAL SPANNING TREE. Manuscript

[Pe94] E. Petrank. The hardness of approximation: Gap location, *Computational Complexity*, Vol. 4, pp. 133- 157, 1994.

[W03] G. Woeginger. Exact Algorithms for NP-Hard Problems. A Survey, in *Proceedings of 5th International Workshop on Combinatorial Optimization-Eureka, You Shrink!* Papers dictated to Jack Edmonds, M. Junger, G. Reinelt, and G. Rinaldi (Festschrift Eds.) LNCS Vol. 2570, pp. 184-207. Springer-Verlag, 2003.

[ZL01] H. Zhang and C.X.Ling. An improved learning algorithm for augmented naive Bayes, *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (PAKDD),LNCS, Vol. 2035, pp.581-586. Springer Verlag, 2001.