Kernels for Below-Upper-Bound Parameterizations of the Hitting Set and Directed Dominating Set Problems

Gregory Gutin

Royal Holloway, University of London
Joint work with Mark Jones and Anders Yeo

WorKer 2011, Vienna
TCS **412** (2011), 5744–5751

Outline

1. Results on Parameterizations of Hitting Set Problem
2. Hypergraph Terminology and Notation
3. Results on HitSet($m-k$,k)
4. Results on HitSet($n-k$,k) and HitSet($n-k$,$k+d$)
5. Open Problem

Outline

1. Results on Parameterizations of Hitting Set Problem
2. Hypergraph Terminology and Notation
3. Results on HitSet($m-k$,k)
4. Results on HitSet($n-k$,k) and HitSet($n-k$,$k+d$)
5. Open Problem

Generic Hitting Set Problem

HitSet*(p,κ)*
Instance: A set V, a collection $\mathcal{F}$ of subsets of V.
Parameter: κ.
Question: Does $(V,\mathcal{F})$ have a hitting set S of size at most p? (A subset S of V is called a hitting set *if $S \cap F \neq \emptyset$ for each $F \in \mathcal{F}$.)*

In what follows, n stands for the size of V and m for the size of $\mathcal{F}$.

Applications

- software testing, Jones and Harrold (2003)
- computer networks, Kuhn, Rickenbach, Wattenhofer, Welzl and Zollinger (2005)
- bioinformatics, Ruchkys and Song (2002)
- medicine, Vazquez (2009)
- medicine, Mellor, Prieto, Mathieson and Moscato (2010) [they use an Abu-Khzam-like kernelization]

Known Results

$s = \max\{|F| : F \in \mathcal{F}\}$

- $\textsc{HitSet}(p,p)$ is W[2]-complete (Paz and Moran, 1981)
- For $\textsc{HitSet}(p,p+s)$:
  - Kernel of size $\leq s^p$ (see Downey and Fellows, 1999)
  - Kernel of size $O(sp^s s!)$ (Flum and Grohe, 2006)
  - Kernel of order $\leq (2s-1)p^{s-1}+p$ (Abu-Khzam, 2010)
- Dom, Lokshtanov and Saurabh, 2009:
  - $\textsc{HitSet}(p,p+s)$, $\textsc{HitSet}(p,p+m)$ and $\textsc{HitSet}(p,p+n)$ have no poly kernels unless coNP$\subseteq$NP/poly
  - $\textsc{HitSet}(p,p+m)$ and $\textsc{HitSet}(p,p+n)$ are fpt

Our Results

- $\textsc{HitSet}(m-k,k)$ has a kernel with $\leq k4^k$ vertices and sets and it has no poly kernel unless coNP$\subseteq$NP/poly
- $\textsc{HitSet}(n-k,k)$ is W[1]-complete
- $\textsc{HitSet}(n-k,k+d)$ has a poly kernel, where d is the degeneracy [defined later] of $(V,\mathcal{F})$
- Linear Kernel for Directed Nonblocker (I'll not speak of it)

Outline

1. Results on Parameterizations of Hitting Set Problem
2. Hypergraph Terminology and Notation
3. Results on HitSet($m-k$,k)
4. Results on HitSet($n-k$,k) and HitSet($n-k$,$k+d$)
5. Open Problem

Edges Sharing Common Vertices

For a hypergraph $H = (V, \mathcal{F})$ and a vertex $v \in V$:

- $\mathcal{F}[v]$ is the set of edges containing v.
- The *degree* of v is $d(v) = |\mathcal{F}[v]|$.
- For a subset T of vertices, $\mathcal{F}[T] = \bigcup_{v \in T} \mathcal{F}[v]$.

Deletions

For a hypergraph $H = (V, \mathcal{F})$, a vertex v, an edge e and a set $X \subset V$:

- $H - e$: delete e and all isolated vertices.
- $H - v$: delete v and v from any edge.
- $H \ominus X$: delete all edges hit by X and all isolated vertices.

Hypergraph Degeneracy

- A hypergraph $H=(V,\mathcal{F})$ is *d-degenerate* if, for all $X\subset V$, the subhypergraph $H\ominus X$ contains a vertex of degree at most d.
- The *degeneracy* $\deg(H)$ of a hypergraph H is the smallest d for which H is d-degenerate.
- $\deg(H)$ can be calculated in linear time:
  - Set $d:=0$.
  - while H nonempty choose a vertex v of minimum degree and set $d:=\max\{d,d(v)\}$ and $H:=H\ominus\{v\}$.

Outline

1. Results on Parameterizations of Hitting Set Problem
2. Hypergraph Terminology and Notation
3. **Results on HitSet($m-k$,k)**
4. Results on HitSet($n-k$,k) and HitSet($n-k$,$k+d$)
5. Open Problem

Three Reduction Rules

Red. Rule 1 If there exist distinct $e, e' \in \mathcal{F}$ such that $e \subseteq e'$, set $H := H - e'$ and $k := k - 1$.

Red. Rule 2 If there exist $u, v \in V$ such that $u \neq v$ and $\mathcal{F}[u] \subseteq \mathcal{F}[v]$, set $H := H - u$.

Red. Rule 3 If there exist $v \in V$, $e \in \mathcal{F}$ such that $\mathcal{F}[v] = \{e\}$ and $e = \{v\}$, then delete v and e.

Lemma

Let $(H = (V, \mathcal{F}), k)$ be a hypergraph reduced by Rules 1, 2 and 3 and $\mathcal{F} \neq \emptyset$. Then for all $v \in V$, $d(v) \geq 2$, and for all $e \in \mathcal{F}$, $|e| \geq 2$.

Mini-hitting Set

Definition

A *mini-hitting set* is a set $S_{\mathrm{MINI}} \subseteq V$ such that $|S_{\mathrm{MINI}}| \leq k$ and $|\mathcal{F}[S_{\mathrm{MINI}}]| \geq |S_{\mathrm{MINI}}| + k$.

Lemma (Mini-hit Lemma)

A reduced hypergraph $H = (V, \mathcal{F})$ has a hitting set of size at most $m-k$ iff it has a mini-hitting set. Moreover,

1. *Given a mini-hitting set S_{MINI}, we can construct a hitting set S with $|S| \leq m-k$ s.t. $S_{\mathrm{MINI}} \subseteq S$ in polynomial time.*
2. *Given a hitting set S with $|S| \leq m-k$, we can construct a mini-hitting set S_{MINI} s.t. $S_{\mathrm{MINI}} \subseteq S$ in polynomial time.*

Greedy Algorithm

Start with $S^* = \emptyset$. While $|\mathcal{F}[S^*]| < |S^*| + k$ and there exists $v \in V$ with $|\mathcal{F}[v] \backslash \mathcal{F}[S^*]| > 1$, do the following: Pick a vertex $v \in V$ s.t. $|\mathcal{F}[v] \backslash \mathcal{F}[S^*]|$ is as large as possible, and add v to S^*.

Let $\mathcal{C} = \mathcal{F}[S^*]$.

Lemma

Suppose S^* is not a mini-hitting set. Then we have the following:

1. $|\mathcal{C}| < 2k$.
2. *For all* $v \in V$, $|\mathcal{C}[v]| \geq 1$.
3. *For all* $v \in V$, $d(v) \leq k$.

Kernel

Red. Rule 4: For any $\mathcal{C}' \subseteq \mathcal{C}$, let $V(\mathcal{C}') = \{v \in V : \mathcal{C}[v] = \mathcal{C}'\}$. If $|V(\mathcal{C}')| > k$, pick a vertex $v \in V(\mathcal{C}')$ and set $H := H - v$.

Theorem

$\textsc{HitSet}(m-k,k)$ *has a kernel with at most* $k4^k$ *vertices and at most* $k4^k$ *edges.*

Proof: Let (H, k) be an instance irreducible by the four reduction rules and let $H = (V, \mathcal{F})$. The number of possible subsets $\mathcal{C}' \subseteq \mathcal{C}$ is $2^{|\mathcal{C}|} < 2^{2k}$. Therefore by Rule 4 $n = |V| < k2^{2k} = k4^k$.
To bound $m = |\mathcal{F}|$ recall that $d(v) \leq k$ for all $v \in V$, and $|e| \geq 2$ for all $e \in \mathcal{F}$. It follows that $|\mathcal{F}| \leq k|V|/2 < k^2 2^{2k-1}$. This can be improved.

No Poly Kernel

Theorem

HitSet($m-k$,k) does not have a polynomial kernel, unless coNP $\subseteq$ NP/poly.

Proof: Reduction from HitSet(p,$m+p$).

Outline

1. Results on Parameterizations of Hitting Set Problem
2. Hypergraph Terminology and Notation
3. Results on HitSet($m-k$,k)
4. **Results on HitSet($n-k$,k) and HitSet($n-k$,$k+d$)**
5. Open Problem

W[1]-hardness

Theorem

HitSet($n - k$,k) is $W[1]$-complete.

Proof: Reduction from (Graph) Independence Set.

If we let the parameter be $k + \max_{e \in \mathcal{F}} |e|$, the problem is still $W[1]$-hard.

HitSet($n-k$,$k+d$)

$d = \deg(H)$.

Lemma

The chromatic number of a d-degenerate hypergraph is at most $d+1$.

Red. Rule 5: If there exist $v \in V$, $e \in \mathcal{F}$ such that $e = \{v\}$, then replace $H = (V, \mathcal{F})$ by $H \ominus \{v\}$. Keep k the same.

Kernel for HitSet($n-k$,$k+d$)

Theorem

HitSet($n-k$,$k+d$) admits a kernel with less than $(d+1)k$ vertices and $d(d+1)k$ edges.

Proof (for vertices): Using Rule 5 as long as possible, we reduce $H=(V,\mathcal{F})$ to a d-degenerate hypergraph with no edge of cardinality 1. By the previous lemma, $\chi(H)\leq d+1$. Thus, there is an independent set S s.t. $|S|\geq |V|/(d+1)$. But T is a hitting set of H iff $V\setminus T$ is an independent set. Thus, if $|V|/(d+1)\geq k$, the answer to HitSet($n-k$,$k+d$) is Yes. Otherwise, $|V|<(d+1)k$.

Kernel for HitSet($n-k$,$k+d$) cont'd

Proof (for edges): To prove that $|\mathcal{F}| < d(d+1)k$, choose a vertex v of minimum degree and observe that $d(v) \le d$. Now delete v from V and $\mathcal{F}[v]$ from $\mathcal{F}$, and choose a vertex v of minimum degree again, and observe that $d(v) \le d$. Continuing this procedure we will delete all edges in $\mathcal{F}$ and thus $|\mathcal{F}| \le d|V| < d(d+1)k$.

Outline

1. Results on Parameterizations of Hitting Set Problem
2. Hypergraph Terminology and Notation
3. Results on HitSet($m-k$,k)
4. Results on HitSet($n-k$,k) and HitSet($n-k$,$k+d$)
5. **Open Problem**

Independent Set Parameterized Above Tight Lower Bound

For a hypergraph H, let $\alpha(H)$ be the maximum size of an independent set of H. If H is d-degenerate then $\chi(H) \leq d+1$ and $\alpha(H) \geq n/(d+1)$.
Open Problem: What is the complexity of deciding whether for a d-degenerate hypergraph H we have $\alpha(H) \geq n/(d+1)+\kappa$, where κ is the parameter?

This problem is open even for graphs. However, it's easy (a linear kernel) if H is a graph and $d=\Delta(H)$: by Brooks' Theorem, $\chi(H) \leq d$ unless one of the connectivity components of H is K_{d+1} or $d=2$ and one of the connectivity components of H is an odd cycle.