Towards a Fixed Parameter Tractability of Geometric Hitting Set Problem for Axis-Parallel Squares Intersecting a Given Straight Line

Daniel KHACHAY\textsuperscript{1,2,*}, Michael KHACHAY\textsuperscript{1,2,3} and Maria POBERIY\textsuperscript{1}

\textsuperscript{1}Krasovskiy Institute of Mathematics and Mechanics, 16 S. Kovalevskoy str., Ekaterinburg 620990, Russia
\textsuperscript{2}Ural Federal University, 19 Mira str., Ekaterinburg 620002, Russia
\textsuperscript{3}Omsk State Technical University, 11 Mira ave., Omsk 644050, Russia

*Corresponding author

Keywords: Hitting set problem, Fixed parameter tractability, Dynamic programming.

Abstract. The Hitting Set Problem (HSP) is the well-known extremal problem adopting interest of researchers in the fields of statistical learning theory, combinatorial optimization, and computational geometry for decades. It is known, that the problem is NP-hard in its general case and remains intractable even in very specific geometric settings, e.g., for axis-parallel rectangles intersecting a given straight line. Recently, for the special case, where all the rectangles are unit squares, a polynomial but very time-consuming exact algorithm was proposed. We improve this algorithm to decrease its complexity bound more than 100 degrees of magnitude. Also, we extend it on the more general case and prove that the geometric HSP for axis-parallel (not necessarily unit) squares intersected by a line is polynomially solvable for any fixed range of the squares to hit. Hence, this geometric setting of the HSP belongs to the class of Fixed Parameter Tractable (FPT) problems.

Introduction and Related Work

We consider a geometric setting of the well-known Hitting Set Problem (HSP), which is actively investigated by researchers in machine learning, combinatorial optimization, and computational geometry.

On the one hand, to the best of our knowledge, the HSP appears to be the first intractable combinatorial problem, whose approximability was substantially improved (see, e.g. [11]) based on statistical learning techniques, especially the well-known Vapnik and Chervonenkis theory [15]. Studying the randomized approximation algorithms for the HSP and related problems sourcing from seminal papers [1] and [6] stems to the new drastically developing field in computational geometry.

On the other hand, concepts of the hitting set and the ensemble classifier taking its decisions by some kind of voting are very close. In particular, the HSP and its dual Set Cover problem are closely related to well-known boosting learning procedures [14], especially to the Minimum Committee problem that of finding the ensemble majority voting classifier of the minimum VC-dimension (see, e.g. [9] and [8, 10]).

Finally, investigation of the efficient polynomial time exact and approximate algorithms for the Hitting Set and Set Cover problems seems to be very relevant in modern engineering, for instance in development of highly reliable wireless networks (see, e.g. [13]).

The geometric HSP for axis-parallel rectangles is the well-studied problem. It is known [5] that the problem is NP-hard even in the case of unit squares. In papers [7] and [2] first polynomial time approximation schemes were proposed for unit and arbitrary axis-parallel squares, respectively. Chepoi and Felsner [3] proposed 6-approximation polynomial time algorithm for the case of axis-parallel rectangles intersected by an axis-monotone Lagrangian curve. Recently, Correa et al. [4] showed that this problem is NP-hard even in the case of a straight line and proposed a 4-approximation algorithm for this particular case.

In this paper, we extend another recent result obtained by Mudgal and Pandit [10]. They proposed the first polynomial time algorithm constructing a minimal hitting set for a collection of axis-parallel
unit squares intersected by a straight line. That result seems to be very promising since the most settings of the HSP (even very special) are proven to be NP-hard. Unfortunately, their algorithm has a very high (but polynomial) bound of time consumption. In Section 3, we improve the algorithm for this setting of the HSP to decrease its complexity bound more than 100 degrees of magnitude. Further, in Section 4, we extend the algorithm to the case of square collections of varying side lengths and show that this setting belongs to a class of Fixed Parameter Tractable (FPT) problems.

**Problem Statement**

We consider the Hitting Set Problem in the following geometric setting. In the Euclidean plane, the finite collection \( S = \{Q_1, \ldots, Q_n\} \) of axis-parallel (closed) squares intersecting the given straight line \( d \) is given. For the collection \( S \), it is required to find a hitting set \( P \) of a minimum size, i.e.

\[
\arg \min \{|P|: P \subset \mathbb{R}^2, P \cap Q_j \neq \emptyset, \ j = 1, \ldots, n \}.
\]

Figure 1. Value \( K \) does not exceed the number of rectangular cells induced by the lines defining borders of \( Q_1, \ldots, Q_n \).

Without loss of generality we assume that the line \( d \) is defined by the equation \( kx + y = 0 \) for some \( k \geq 1 \).

The collection \( S \) induces a partition of the plane onto mutually disjunctive regions \( \theta_1, \ldots, \theta_K \) such that, any points \( p_1 \) and \( p_2 \) belong to the same \( \theta_k \) if and only if

\[
\left( \forall Q_j \in S \right) \left( (p_1 \in Q_j) \iff (p_2 \in Q_j) \right).
\]

Since each optimal hitting set contains at most one point \( p_k \) from any region \( \theta_k \), the initial continuous problem is polynomially equivalent to the corresponding combinatorial one of finding a minimal hitting set among subsets of the finite set \( P = \{p_1, \ldots, p_K\}, p_k \in \theta_k \cup \cup_{l \neq k} \theta_l \). Indeed, for any collection of \( n \) axis-parallel squares (and even rectangles), the corresponding set \( P \) contains at most \( O(n^2) \) elements (see Fig. 1) and can be constructed in polynomial time.

**Improved Algorithm for Unit Squares**

We start with the similar (but not the same) notation to that was introduced in [12]. First, we partition the plane by straight lines \( l_0, \ldots, l_{r+2} \) orthogonal to \( d \) with distance of \( \sqrt{2}/2 \) between each neighboring lines such that for each square \( Q_j \in S \) its center \( c_j \) is located between \( l_1 \) and \( l_{r+1} \) (hereinafter all tights are broken arbitrarily). For any \( i = 0, \ldots, r + 1 \), denote by \( R_i \) the stripe located between \( l_i \) and \( l_{i+1} \). Next, we introduce the notation \( S_i = \{Q_j: Q_j \cap R_i \neq \emptyset\}, S_i^{in} = \{Q_j \in S_i: c_j \in R_i\} \), and \( S_i^{out} = S_i \setminus S_i^{in} \). By construction, \( S_i^{out} \subset S_{i-1} \cup S_{i+1}^{in} \).
Figure 2. Any unit square $Q_j \in S_i^{in}$ is hit by one of the centers A and B of $\sqrt{2}/2$–squares.

As in [12], we find an optimal hitting set recursively, by the dynamic programming procedure presented in Algorithm 1. Indeed, for any $i \in 1, \ldots, r$, denote $P_i = P \cap R_i$. Let, for $U \subset P_{i-1}$ and $V \subset P_i$, $T(i, U, V)$ be the size of a smallest hitting set $P$ for $\bigcup_{l \geq i} S_l$ such that $P \cap P_{i-1} = U$ and $P \cap P_i = V$. Similarly to [12], we express $T(i, U, V)$ in terms of $T(i + 1, U', V')$ but for a substantially smaller subsets $U'$ and $V'$.

The following Theorem summarizes the properties of Algorithm 1.

**Theorem 1.** For $q = 6$, Algorithm 1 finds an optimal hitting set for the collection $S$ in time of $O(n^{37})$.

**Proof.** We start with the following simple fact. By construction, for any $i \in 1, \ldots, r$ and any $j \in S_i^{in}$, $Q_j \cap \{A, B\} \neq \emptyset$ (see Fig. 2 for details). As a consequence, for any optimal hitting set $P$ and any $i \in 1, \ldots, r$, $|P_i| \leq 6$, where $P_i = P \cap R_i$. Indeed, assume by contradiction that, for some $i, |P_i| > 6$. Since $S_i \subset S_{i-1} \cup S_i^{in} \cup S_{i+1}^{in}$ and $P_i \cap Q_j = \emptyset$ for any $Q_j \notin S_i$, we can substitute $P_i$ by an appropriate 6-point subset $P_i'$ such that $P \cup P_i' \setminus P_i$ remains a hitting set for $S$ and $|P'| < |P|$. The contradiction obtained with optimality of $P$ finalizes our argument. Hence, Algorithm 1 finds an optimal hitting set for the given collection $S$.

Let us obtain an upper bound for its running time. Obviously, the loop 5-9 having $r - 1 = O(n)$ iterations is the most time consuming part of Algorithm 1. In each iteration, $O(|P_{i-1}|^6) \times O(|P_i|^6) = O(n^{24})$ subproblems each having time complexity of $O(n^{12})$ should be solved. Therefore, the overall running time is $O(n^{37})$.

---

**Algorithm 1 Parameterized exact DP based algorithm**

*Input:* a collection $S = \{Q_1, \ldots, Q_n\}$ of axis-parallel unit squares intersecting a given straight line $d$.

*Output:* an upper bound $q$ of the size of subsets to search for.

1. Construct a set $\mathcal{P}$ induced by the collection $S$, let $\mathcal{P}_i = \mathcal{P} \cap R_i$.
2. for all $U \subset \mathcal{P}_{i-1}$ and $V \subset \mathcal{P}_i$, s.t. $|U|, |V| \leq q$ do
3. define $\mathcal{W}_i = \{W \subset \mathcal{P}_i : |W| \leq q, U \cup V \cup W \cap S_i \neq \emptyset\}$ and
4. $T(i, U, V) = \bigg\{ \min\{U \cup V \cup W : W \in \mathcal{W}_i\}, \quad$ if $\mathcal{W}_i \neq \emptyset$,
  $\infty, \quad$ otherwise.
5. end for
6. for all $U \subset \mathcal{P}_{i-1}$ and $V \subset \mathcal{P}_i$, s.t. $|U|, |V| \leq q$ do
7. define $\mathcal{W}_i = \{W \subset \mathcal{P}_i : |W| \leq q, U \cup V \cup W \cap S_i \neq \emptyset\}$ and
8. $T(i, U, V) = \bigg\{ |U| + \min\{T(i+1, V, W) : W \in \mathcal{W}_i\}, \quad$ if $\mathcal{W}_i \neq \emptyset$,
  $\infty, \quad$ otherwise.
9. end for
10. Output $\arg\min\{T(1, U, V) : U \subset \mathcal{P}_0, V \subset \mathcal{P}_1, |U|, |V| \leq q\}$. 

---
Hitting Squares Whose Sizes Differ

Obviously, due to the scaling, the result of Section 3 remains valid in the case of equal squares of any side-length. In this section we extend this result to the more general case. Let $a$ and $b$ be the minimum and the maximum values of side-lengths of the given squares. W.o.l.g, assume that $a = 1$.

**Case of $k = 1$**

We proceed with the following observation. For $k = 1$, as in Section 3, any square $Q$ of size at least 1, whose center belongs to some stripe $R'$ of width $\sqrt{2}/2$ orthogonal to the line $d$, is hit by the points $A$ and $B$ (like in Fig. 2). Therefore, in this case, we can adapt Algorithm 1 to take into account the squares whose side-lengths are greater than 1. Indeed, as above, consider stripes $R_i$ of width $b\sqrt{2}/2$ consisting all the squares. Then, partition each of them onto $\lfloor b \rfloor$ substripes of width $\sqrt{2}/2$ (see Fig. 3) and use all other notation introduced in Section 3 as is. The following assertion is valid.

**Theorem 2.** Let the given collection $S$ consists of squares with side-lengths from $[1, b]$. Algorithm 1 with $q = 6\lfloor b \rfloor$ finds an optimal hitting set for this collection in time of $O(n^{6q+1})$. For the sake of brevity, we skip the proof of Theorem 2. As a corollary, we obtain that the hitting set problem in this case is a Fixed Parameter Tractable (FPT) problem with parameter $b$.

![Figure 3. Partition of the plane for $b = 4$.](image)

**What if $k > 1$**

Unfortunately, in this case, situation like presented in Fig. 2 takes place only for unit squares. Indeed, consider a stripe of width $\sqrt{2}/2$; to any $k > 1$ assign the largest $b = b(k)$, for which any square $Q$ with center $c$ in this stripe and intersecting the line $d$ is hit by the set $\{A, B\}$ (as in Fig. 2). It can be proven that

$$\inf\{b(k) : k \geq 1\} = 1.$$  

But, to proceed with our extension, for squares of a larger size we can adopt the centers of the neighboring $\sqrt{2}/2$ squares. In particular, we proved that, for any $k \geq 1$, the set $\{A, B, A', B'\}$ (see Fig. 4) hits any square $Q$ intersecting the line $d$ with the center in a stripe of width $\sqrt{2}/2$ with side-length at most $\approx 4.43$ As a corollary, we obtain

**Theorem 3.** For any $k \geq 1$, for $q = 8$, Algorithm 1 finds a minimal hitting set of any given square collection $S$ with side-lengths of $[1, 4]$ in time $O(n^{49})$.

**Conclusion**

In the paper, the improvement of the exact polynomial time hitting set construction algorithm for axis-parallel squares intersecting the given straight line introduced in [12] is proposed. Our
modification has better upper time complexity bound by 100 orders of magnitude. In addition, we propose an extension of this algorithm to the case of non-unit squares and proved that this problem is fixed parameter tractable with respect to the maximum of their side-length. As for the future work, it would be interesting to establish the complexity status of the considered problem in the case, where this parameter is unbounded.

Acknowledgement
This research was supported by Russian Foundation for Basic Research, grant no. 16-07-00266 and Russian Science Foundation, grant no. 14-11-00085.

References