Geometric clustering: fixed-parameter tractability and lower bounds with respect to the dimension

Sergio Cabello
University of Ljubljana
Slovenia

Panos Giannopoulos
Tel Aviv University
Christian Knauer
FU Berlin
Dániel Marx
Budapest University T& E
Günter Rote
FU Berlin
Motivation

Computational geometry parameterized by the dimension

Parameterized Complexity  \hspace{1cm} \hspace{1cm} Computational Geometry
$k$-center problem

- $k$-center optimization problem
  - *Input*: a set of $n$ points $S$ in $\mathbb{R}^d$
  - *Task*: find the smallest $k$ congruent balls that cover $S$

- $k$-center decision problem
  - *Input*: a set of $n$ points $S$ in $\mathbb{R}^d$
  - *Question*: can $S$ be covered with $k$ unit balls?
\textit{k-center problem}

- \textit{k-center optimization problem}
  \begin{itemize}
  \item \textit{Input}: a set of \( n \) points \( S \) in \( \mathbb{R}^d \)
  \item \textit{Task}: find the smallest \( k \) congruent balls that cover \( S \)
  \end{itemize}

- \textit{k-center decision problem}
  \begin{itemize}
  \item \textit{Input}: a set of \( n \) points \( S \) in \( \mathbb{R}^d \)
  \item \textit{Question}: can \( S \) be covered with \( k \) unit balls?
  \end{itemize}

- most discussion about decision problem
- we will consider \( L_2 \) and \( L_\infty \) metrics
- \( d \) is not a constant
$k$-center problem in $L_2$

- $k = 1$
  - linear programming in $d + 1$ dimensions
  - solvable in $O(f(d)n) = O(3^{d^2}n)$ time
- $k = 2$
  - easily solvable in $O(n^{2d+2})$ time using arrangements
  - NP-hard [Megiddo 90]
\textit{k-center problem in } L_2 \\

\begin{itemize}
  \item \textit{k = 1}
    \begin{itemize}
      \item linear programming in \(d + 1\) dimensions
      \item solvable in \(O(f(d)n) = O(3^{d^2}n)\) time
    \end{itemize}
  \item \textit{k = 2}
    \begin{itemize}
      \item easily solvable in \(O(n^{2d+2})\) time using arrangements
      \item NP-hard
      \item W[1]-hard [Megiddo 90]
    \end{itemize}
\end{itemize}
New results: 2-center problem in $L_2$

Theorem

2-center problem parameterized by the dimension is $W[1]$-hard.

- if there is an algorithm solving 2-center in $O(f(d)n^c)$ time
  - we can find $k$-cliques in graphs in $O(g(k)n^{c'})$ time
  - we can solve 3-SAT in $O(2^{o(n)})$ time
  - some hierarchy collapses
New results: 2-center problem in $L_2$

Theorem
2-center problem parameterized by the dimension is $W[1]$-hard.

- if there is an algorithm solving 2-center in $O(f(d)n^c)$ time
  - we can find $k$-cliques in graphs in $O(g(k)n^{c'})$ time
  - we can solve 3-SAT in $O(2^{o(n)})$ time
  - some hierarchy collapses
- an algorithm solving 2-center in $O(f(d)n^{100})$ time is unlikely
- an algorithm solving 2-center in $O(f(d)n^{o(d)})$ time is unlikely
$k$-center problem in $L_\infty$

- $k = 1$
  - trivial $O(dn)$ time
- $k = 2$
  - solvable in $O(dn^2)$ time \[\text{[Megiddo 90]}\]
- $k = 3$
  - solvable in $O(n^{\lfloor d/3 \rfloor} \log n)$ time \[\text{[Assa, Katz 99]}\]
  - NP-hard \[\text{[Megiddo 90]}\]
$k$-center problem in $L_\infty$

- $k = 1$
  - trivial $O(dn)$ time

- $k = 2$
  - solvable in $O(dn^2)$ time  \[ \text{[Megiddo 90]} \]

- $k = 3$
  - solvable in $O(n^{\lfloor d/3 \rfloor} \log n)$ time \[ \text{[Assa, Katz 99]} \]
  - NP-hard \[ \text{[Megiddo 90]} \]
  - $O(6^d \cdot dn \log(dn))$ time
\section*{$k$-center problem in $L_\infty$}

\begin{itemize}
\item $k = 1$
  \begin{itemize}
  \item trivial $O(dn)$ time
  \end{itemize}
\item $k = 2$
  \begin{itemize}
  \item solvable in $O(dn^2)$ time \hfill [Megiddo 90]
  \end{itemize}
\item $k = 3$
  \begin{itemize}
  \item solvable in $O(n^{\lfloor d/3 \rfloor} \log n)$ time \hfill [Assa, Katz 99]
  \item NP-hard \hfill [Megiddo 90]
  \item $O(6^d \cdot dn \log(dn))$ time
  \end{itemize}
\item $k = 4$
  \begin{itemize}
  \item W[1]-hard
  \item an algorithm solving 4-center in $O(f(d)n^{100})$ is unlikely
  \end{itemize}
\end{itemize}
What is new?

Finer classification of $k$-center for unbounded dimension

- $L_2$
  - easy for $k = 1$
  - $W[1]$-hard for $k = 2$

- $L_\infty$
  - easy for $k = 1, 2$
  - NP-hard, but fixed-parameter tractable for $k = 3$
  - $W[1]$-hard for $k = 4$

Other related work:

- $k$-center problem parameterized by $k$ is $W[1]$-hard for $d \geq 2$ [Marx 05]
What is new?

Finer classification of $k$-center for unbounded dimension

- $L_2$
  - easy for $k = 1$
  - $W[1]$-hard for $k = 2$

- $L_\infty$
  - easy for $k = 1, 2$
  - NP-hard, but fixed-parameter tractable for $k = 3$
  - $W[1]$-hard for $k = 4$

Other related work:

- $k$-center problem parameterized by $k$ is $W[1]$-hard for $d \geq 2$
  [Marx 05]
Outline

- Introduction
- What is new?
- Ideas
  - Solving 3-center in $L_\infty$
  - $W[1]$-hardness of 2-center in $L_2$
- Conclusions
Solving 3-center in $L_\infty$ – Frame

- consider decision problem
  
  Input: a set of $n$ points $S$ in $\mathbb{R}^d$

  Question: can $S$ be covered with 3 unit cubes?

- the points are denoted 1, 2, \ldots, $n$
  
  - $u$ a generic point

- the cubes are denoted $A, B, C$
  
  - $X$ a generic cube
Solving 3-center in $L_\infty$ – Frame

- consider decision problem
  
  *Input*: a set of $n$ points $S$ in $\mathbb{R}^d$
  
  *Question*: can $S$ be covered with 3 unit cubes?

- the points are denoted 1, 2, …, $n$
  - $u$ a generic point

- the cubes are denoted $A, B, C$
  - $X$ a generic cube

- decision $\rightarrow$ optimization
  - easy using [Frederickson, Johnson ’84]
Solving 3-center in $L_\infty$ – General Idea

- cube $X$ covers point $u$ iff $\pi_j(u) \in \pi_j(X)$ for each coordinate projection $\pi_j$

- classify possible solutions according to certain patterns
  - for each pattern
    - reduce the problem to 2-SAT
Solving 3-center in $L_{\infty}$ – Patterns

- the pattern of 3 cubes $A, B, C$ is
  
  $$(L_1, M_1, R_1), (L_2, M_2, R_2), \ldots, (L_d, M_d, R_d),$$

  where
  - $(L_j, M_j, R_j)$ a permutation of $(A, B, C)$
  - $\pi_j(L_j)$ left of $\pi_j(M_j)$ left of $\pi_j(R_j)$

- example with pattern $(A, B, C), (B, C, A)$ in $d = 2$
Solving 3-center in $L_\infty$ – Patterns

- the pattern of 3 cubes $A, B, C$ is

$$(L_1, M_1, R_1), (L_2, M_2, R_2), \ldots, (L_d, M_d, R_d),$$

where

- $(L_j, M_j, R_j)$ a permutation of $(A, B, C)$
- $\pi_j(L_j)$ left of $\pi_j(M_j)$ left of $\pi_j(R_j)$

- there are $6^d$ possible patterns
- each pattern explored independently
- each pattern, one 2-SAT problem
Solving 3-center in $L_\infty$ – A pattern

- consider a pattern $(L_1, M_1, R_1), (L_2, M_2, R_2), \ldots, (L_d, M_d, R_d)$
- we can fix the position of $\pi_j(L_j)$ using $l_j$
- idem for $\pi_j(R_j)$ using $r_j$
- the position of $\pi_j(M_j)$ is unclear
- Boolean variable $y_{\chi u} \equiv$ point $u$ covered by cube $X$
Solving 3-center in $L_\infty$ – A pattern

- consider a pattern $(L_1, M_1, R_1), (L_2, M_2, R_2), \ldots, (L_d, M_d, R_d)$
- we can fix the position of $\pi_j(L_j)$ using $l_j$
- idem for $\pi_j(R_j)$ using $r_j$
- the position of $\pi_j(M_j)$ is unclear
- Boolean variable $y_{Xu} \equiv$ point $u$ covered by cube $X$
Solving 3-center in $L_\infty$ – SAT

- each point $u$ is covered
  \[ y_Au \lor y_Bu \lor y_Cu \] for each point $u$

- incompatible pairs; for each dimension
  \[ \neg y_{L_j u} \] for each point $u$ with $\pi_j(u) > l_j + 1$
  \[ \neg y_{R_j u} \] for each point $u$ with $\pi_j(u) < r_j - 1$
  \[ \neg y_{M_j u} \lor \neg y_{M_j v} \] for each points $u, v$ with $|\pi_j(u) - \pi_j(v)| > 1$
Solving 3-center in $L_\infty – \text{SAT}$

- each point $u$ is covered
  \[ y_{Au} \lor y_{Bu} \lor y_{Cu} \]  for each point $u$

- incompatible pairs; for each dimension
  \[ \neg y_{L_j u} \]  for each point $u$ with $\pi_j(u) > l_j + 1$
  \[ \neg y_{R_j u} \]  for each point $u$ with $\pi_j(u) < r_j - 1$

- for each points $u, v$ with $|\pi_j(u) - \pi_j(v)| > 1$

- there are 3 cubes covering with the given pattern iff all clauses satisfiable simultaneously

- 3-SAT instance with $O(dn^2)$ clauses
Solving 3-center in $L_\infty$ – 2-SAT

$y_{Au} \lor y_{Bu} \lor y_{Cu} \quad \forall$ points $u$

$\neg y_{L_j u} \quad \forall j, \forall$ points $u$ with $\pi_j(u) > l_j + 1$

$\neg y_{R_j u} \quad \forall j, \forall$ points $u$ with $\pi_j(u) < r_j - 1$

$\neg y_{M_j u} \lor \neg y_{M_j v} \quad \forall j, \forall$ points $u, v$ with $|\pi_j(u) - \pi_j(v)| > 1$
Solving 3-center in $L_\infty - 2$-SAT

\[ y_{Au} \lor y_{Bu} \lor y_{Cu} \quad \forall \text{ points } u \]
\[ \neg y_{L_j u} \quad \forall j, \forall \text{ points } u \text{ with } \pi_j(u) > l_j + 1 \]
\[ \neg y_{R_j u} \quad \forall j, \forall \text{ points } u \text{ with } \pi_j(u) < r_j - 1 \]
\[ \neg y_{M_j u} \lor \neg y_{M_j v} \quad \forall j, \forall \text{ points } u, v \text{ with } |\pi_j(u) - \pi_j(v)| > 1 \]

For each point $u$ either

- $y_{Au} \lor y_{Bu} \lor y_{Cu}$ reducible to 2-SAT clause, or
- point $u$ always covered
Solving 3-center in $L_\infty$ – SAT

$$y_{Au} \vee y_{Bu} \vee y_{Cu} \forall \text{ points } u$$

$$\neg y_{L_j u} \forall j, \forall \text{ points } u \text{ with } \pi_j(u) > l_j + 1$$

$$\neg y_{R_j u} \forall j, \forall \text{ points } u \text{ with } \pi_j(u) < r_j - 1$$

$$\neg y_{M_j u} \vee \neg y_{M_j v} \forall j, \forall \text{ points } u, v \text{ with } |\pi_j(u) - \pi_j(v)| > 1$$

- deciding for a pattern $\rightarrow$ 2-SAT with $O(dn^2)$ clauses
- deciding for a pattern takes $O(dn^2)$ time
- can be reduced to $O(dn)$ time per pattern

- $O(dn \log n + 6^d dn)$ time for decision 3-center

- $j$-th coordinate
  $L_j$  $R_j$

- $l_j$  $r_j$
Outline

- Introduction
- What is new?
- Ideas
  - Solving 3-center in $L_\infty$
  - $W[1]$-hardness of 2-center in $L_2$
- Conclusions
Hardness 2-center in $L_2$ – Idea

- consider the decision 2-center
- assumption: we cannot find $k$-cliques in $O(f(k)n^c)$
- polynomial-time reduction from clique to 2-center

\[
\text{Clique} \quad \text{reduction} \quad \text{2-center}
\]

\[
(G,k) \quad S = S(G,k) \subset \mathbb{R}g(k)
\]

where $G$ has $k$-clique iff $S$ can be 2-covered

- if 2-center solvable in $O(f(d)n^c)$ time
  \[\Rightarrow (G, k) \text{ solvable in } O(f(g(k))n^{c'}) \text{ time}\]
Hardness 2-center in $L_2$ – Point set

- $k$ orthogonal planes $E_1, \ldots E_k$ and one $Z$ axis
- point set in $\mathbb{R}^{2k+1}$
- in $Z$ 2 points with $z = 2$ and $z = -2$

- in each $E_i$ a point set like
- choose appropriate radius
- bijection $k$-tuples of $V(G)$ and 2-coverings of $S$
- add extra points killing $k$-tuples with non-adjacent vertices
Conclusions

Finer classification of $k$-center problem for unbounded dimension

- $L_2$
  - easy for $k = 1$
  - $W[1]$-hard for $k = 2$
    - reduction from parameterized-clique
    - lots of symmetry

- $L_\infty$
  - easy for $k = 1, 2$
  - fixed parameter tractable for $k = 3$
    - reduction to 2-SAT
    - simple
  - $W[1]$-hard for $k = 4$