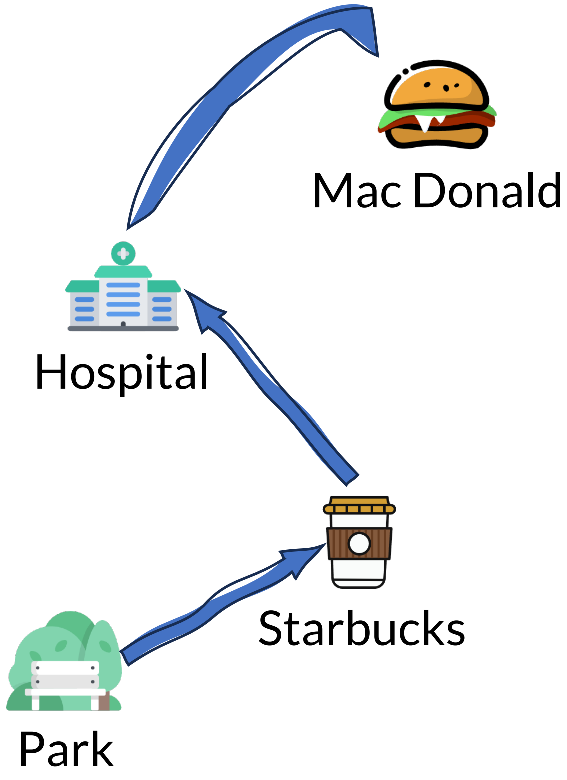
TraCS: Trajectory Collection in Continuous Space under Local Differential Privacy

Authors: Ye Zheng, Yidan Hu

Rochester Institute of Technology (RIT)

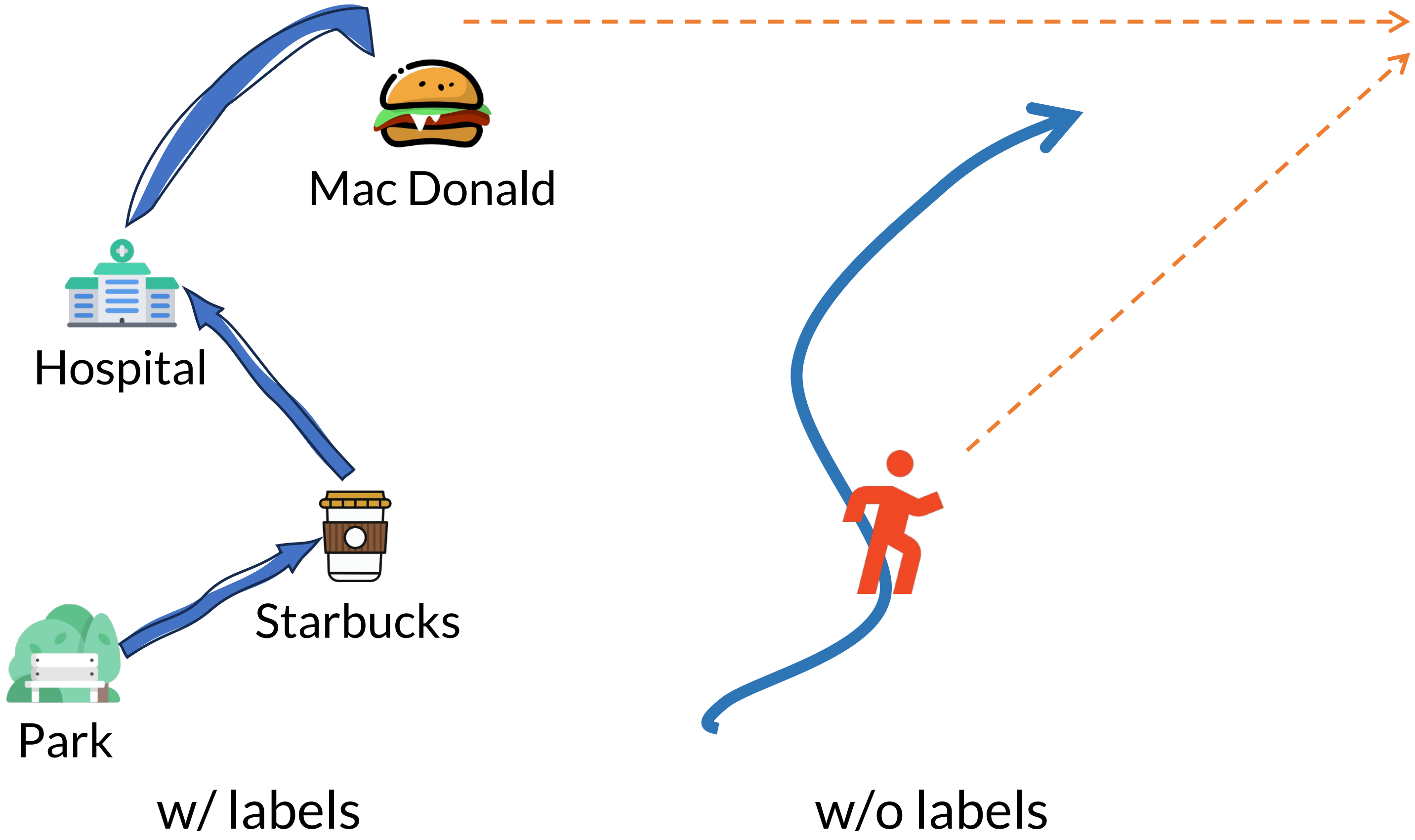
Trajectory Collection

- Private trajectories



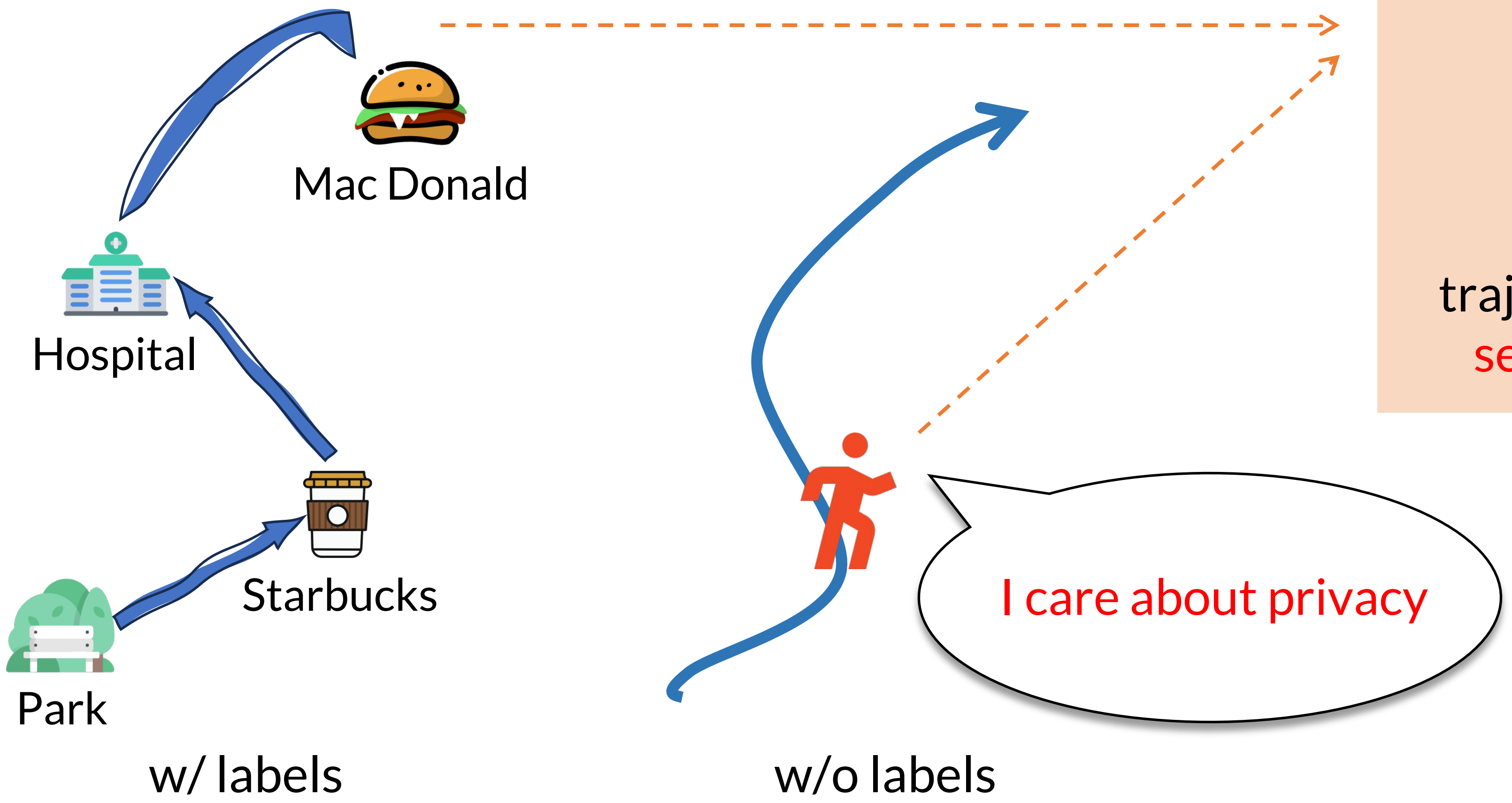
Trajectory Collection

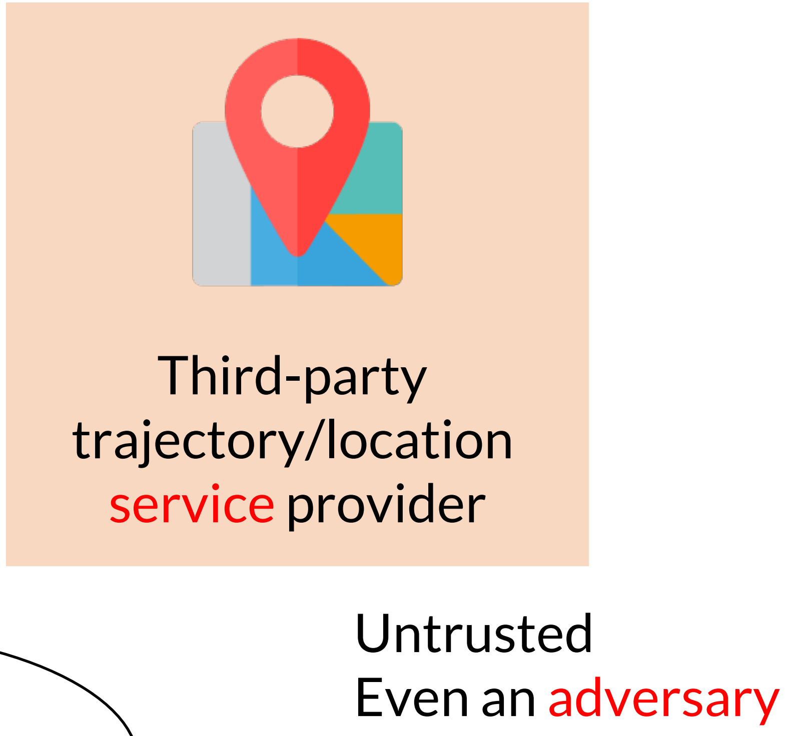
- Collecting private trajectories



Trajectory Collection

- Collecting private trajectories





LDP-fy a Trajectory

- LDP-fy: perturb a trajectory with LDP guarantee

  - cannot distinguish location τ_1 from τ_2 with confidence level e^{ε}

$$\forall \tau_1, \tau_2, \tau_* \in \mathcal{S}: \quad \frac{\Pr[\mathcal{M}(\tau_1) = \tau_*]}{\Pr[\mathcal{M}(\tau_2) = \tau_*]} \leq e^{\varepsilon}$$

distinguish level

location space

probability difference of outputting τ_*

LDP-fy a Trajectory

- LDP-fy: perturb a trajectory with LDP guarantee

  - cannot distinguish location τ_1 from τ_2 with confidence level e^{ε}

$$\forall \tau_1, \tau_2, \tau_* \in \mathcal{S}: \quad \frac{\Pr[\mathcal{M}(\tau_1) = \tau_*]}{\Pr[\mathcal{M}(\tau_2) = \tau_*]} \leq e^{\varepsilon}$$

distinguish level

location space

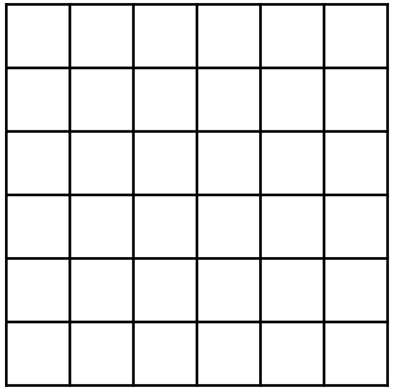
probability difference of outputting τ_*

- **We want:** $\mathcal{M}$ for continuous location space

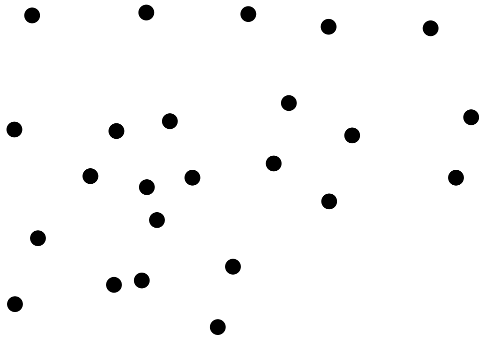
  - existing methods are all for discrete location space $\rightarrow$ not always available
  - continuous space $\supset$ discrete space $\rightarrow$ can apply to discrete space

Existing Methods

- Location space $\mathcal{S}$ is a set of cells or points



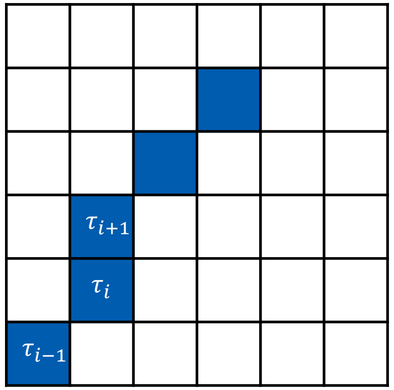
$\mathcal{S} = \{c_1, \ldots, c_n\}$



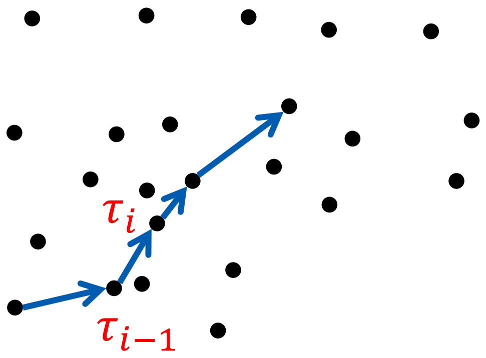
$\mathcal{S} = \{p_1, \ldots, p_n\}$

Existing Methods

- Location space $\mathcal{S}$ is a set of cells or points



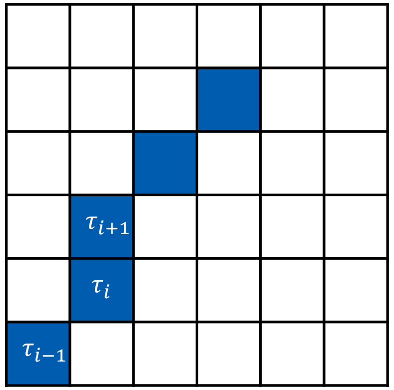
$\mathcal{S} = \{c_1, \dots, c_n\}$



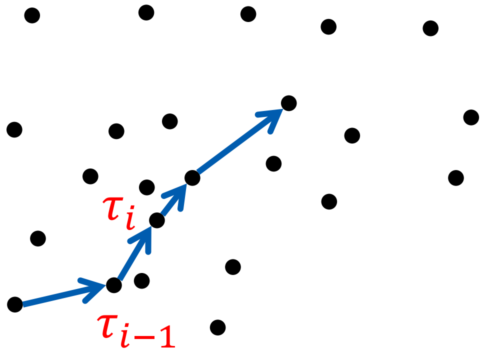
$\mathcal{S} = \{p_1, \dots, p_n\}$

Existing Methods

- Location space $\mathcal{S}$ is a set of cells or points



$\mathcal{S} = \{c_1, \ldots, c_n\}$



$\mathcal{S} = \{p_1, \ldots, p_n\}$

- Perturb each location τ using the Exponential Mechanism

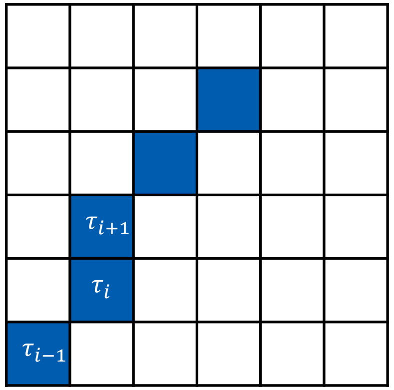
$$\Pr[\mathcal{M}_{\exp}(\tau) = \tau^*] = \frac{\exp\left(kd(\tau, \tau^*)\right)}{\sum_{\tau' \in S} \exp\left(kd(\tau, \tau')\right)}$$

distance

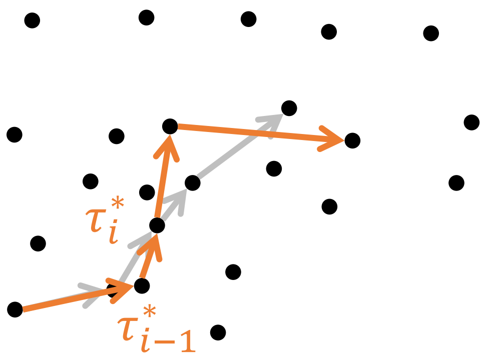
sum of distance

Existing Methods

- Location space $\mathcal{S}$ is a set of cells or points



$\mathcal{S} = \{c_1, \dots, c_n\}$



$\mathcal{S} = \{p_1, \dots, p_n\}$

Perturbed trajectory (ensuring LDP)

- Perturb each location τ using the Exponential Mechanism

$$\Pr[\mathcal{M}_{\exp}(\tau) = \tau^*] = \frac{\exp\big(kd(\tau, \tau^*)\big)}{\sum_{\tau' \in S} \exp\big(kd(\tau, \tau')\big)}$$

distance

sum of distance

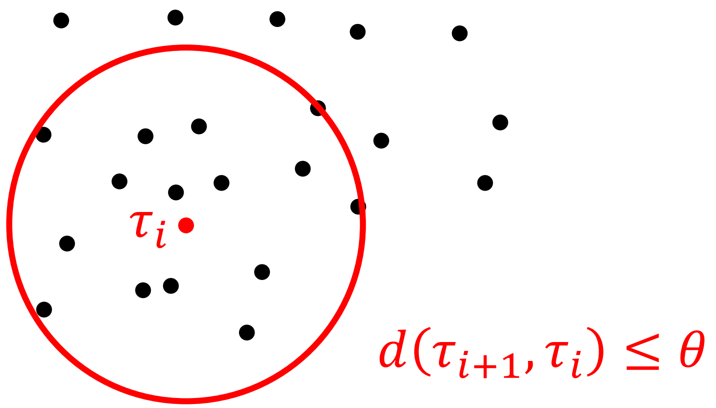
Existing Methods

- NGram [2021]: reachability constraint from public knowledge

  - e.g. distance reachability

  - the next location cannot be too far

  - $\mathcal{M}_{\exp}(\tau)$ on reduced $\mathcal{S}$



Existing Methods

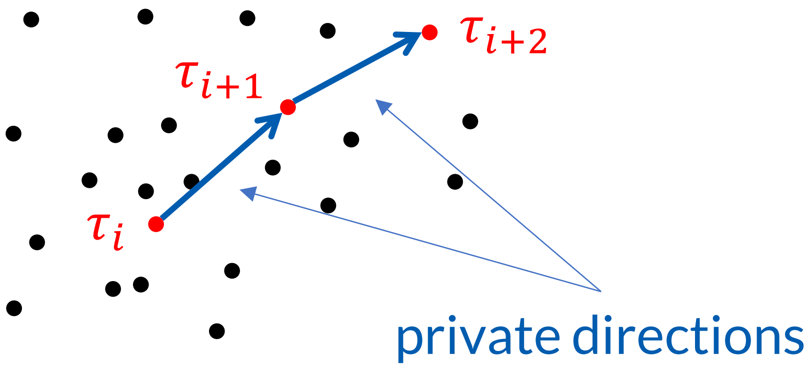
- NGram [2021]: reachability constraint from public knowledge

  - e.g. distance reachability

  - the next location cannot be too far

  - $\mathcal{M}_{\text{exp}}(\tau)$ on reduced $\mathcal{S}$

- ATP [2023]: direction perturbation



Existing Methods

- NGram [2021]: reachability constraint from public knowledge

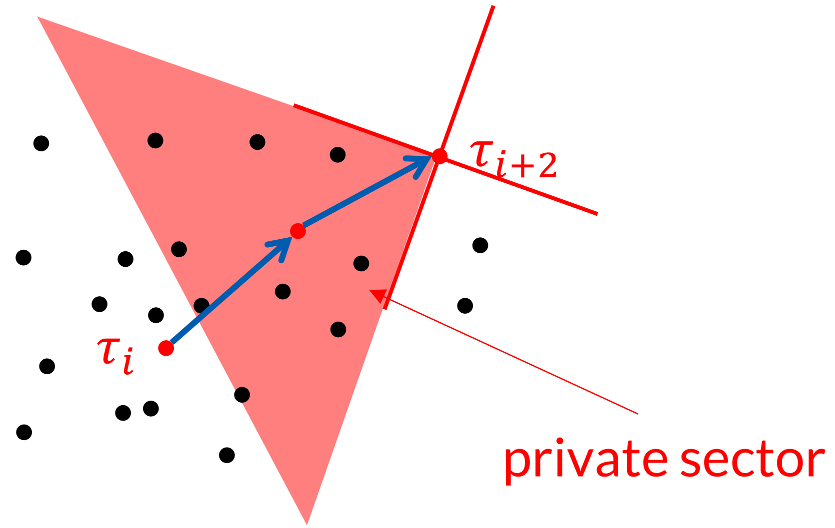
  - e.g. distance reachability

  - the next location cannot be too far

  - $\mathcal{M}_{\text{exp}}(\tau)$ on reduced $\mathcal{S}$

- ATP [2023]: direction perturbation

  - *i.* divide direction sectors, e.g. $k = 4$



Existing Methods

- NGram [2021]: reachability constraint from public knowledge

  - e.g. distance reachability

  - the next location cannot be too far

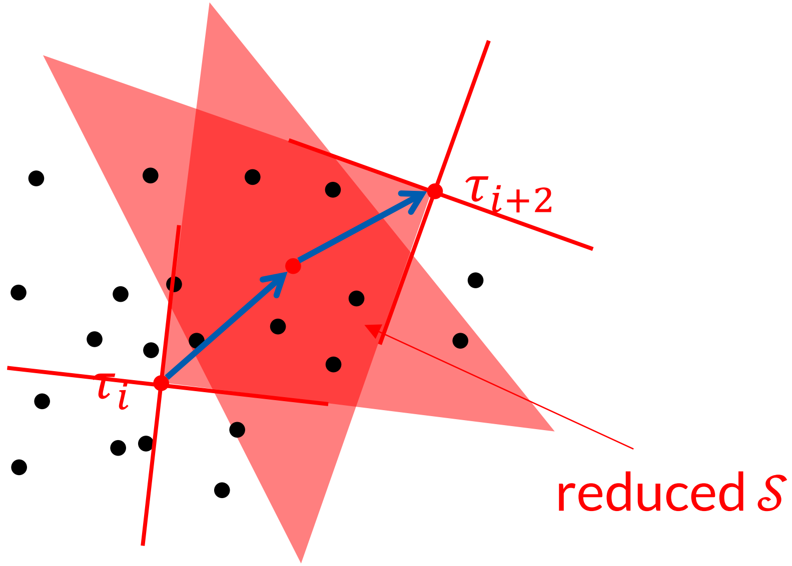
  - $\mathcal{M}_{\exp}(\tau)$ on reduced $\mathcal{S}$

- ATP [2023]: direction perturbation

  - *i.* divide direction sectors, e.g. $k = 4$

  - *ii.* perturb sector

  - *iii.* $\mathcal{M}_{\exp}(\tau)$ on reduced $\mathcal{S}$



Existing Methods: **Weaknesses**

- NGram [2

  - e.g. dist

  - the nex

  - $\mathcal{M}_{\text{exp}}(\tau$

- ATP [202

  - *i.* divide

  - *ii.* pertu

  - *iii.* $\mathcal{M}_{\text{ex}}$

- Privacy: only for discrete space

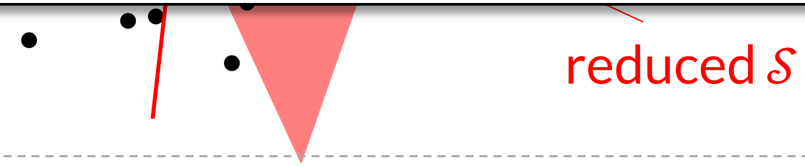
  - indistinguishability for finite # locations/cells/sectors (weak privacy)

- Utility: relies on the Exponential Mechanism

  - efficacy depends on # locations/cells/sectors

  - high time complexity (in computing and sampling)

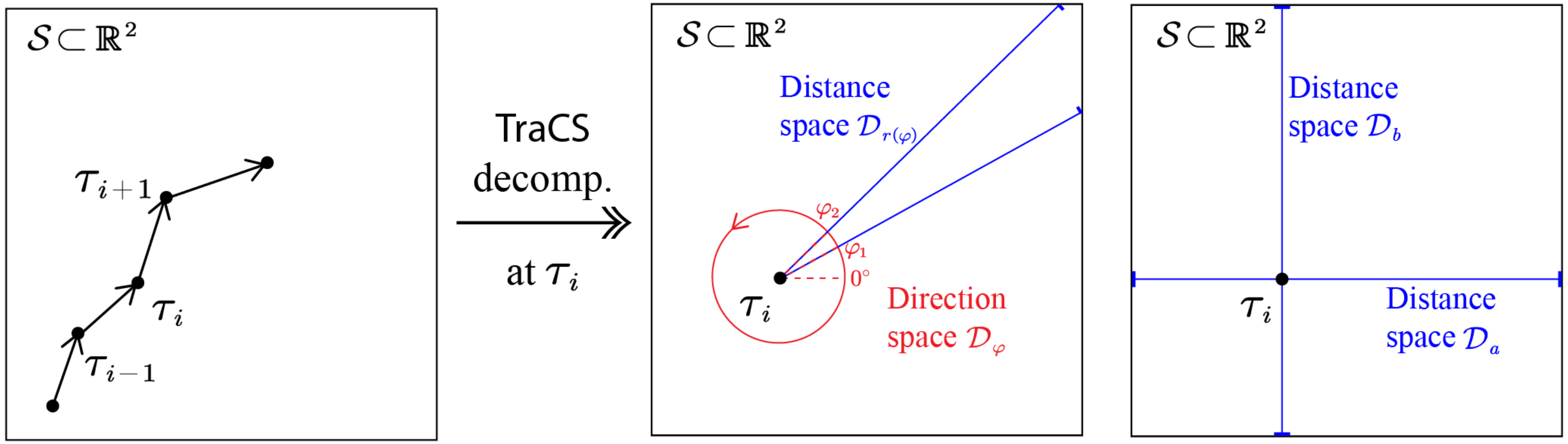
  - discrete locations not always available



This Paper: Continuous Space

- TraCS-D: direction-distance perturbation
- TraCS-C: coordinates perturbation

- **Key idea:** decomposes $\mathcal{S}$ into two subspaces

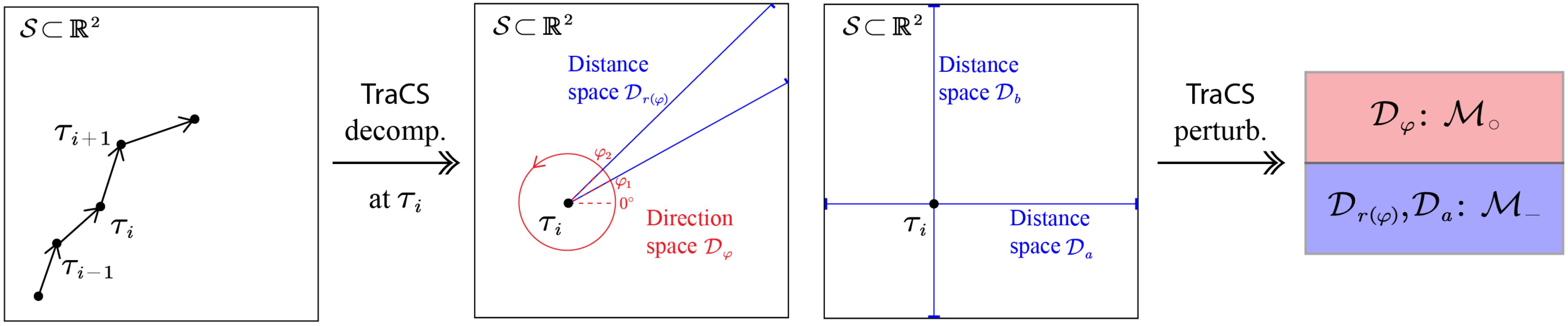


TraCS-D: $\mathcal{D}_\varphi \otimes \mathcal{D}_{r(\varphi)}$ TraCS-C: $\mathcal{D}_a \times \mathcal{D}_b$

This Paper: Continuous Space

- TraCS-D: direction-distance perturbation
- TraCS-C: coordinates perturbation

- **Key idea:** decomposes $\mathcal{S}$ into two subspaces $\rightarrow$ design $\mathcal{M}$ for each subspace



TraCS-D: $\mathcal{D}_\varphi \otimes \mathcal{D}_{r(\varphi)}$

TraCS-C: $\mathcal{D}_a \times \mathcal{D}_b$

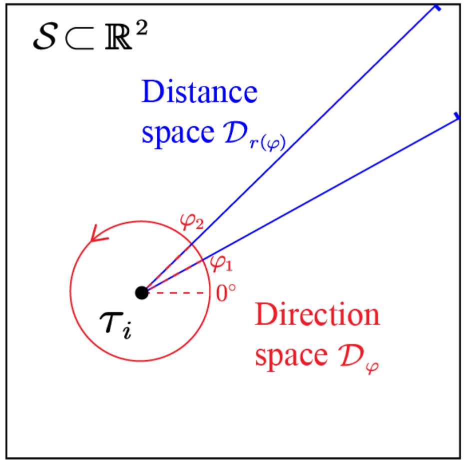
Decomposition of Continuous Space

- $\mathcal{S} = \mathcal{D}_{\varphi} \otimes \mathcal{D}_{r(\varphi)}$

2D space 1D subspaces

- Each location $\tau_{i+1} \in \mathcal{S}$ has a unique representation $(\varphi, r(\varphi))$

- Perturb φ and $r(\varphi)$ using 1D mechanisms



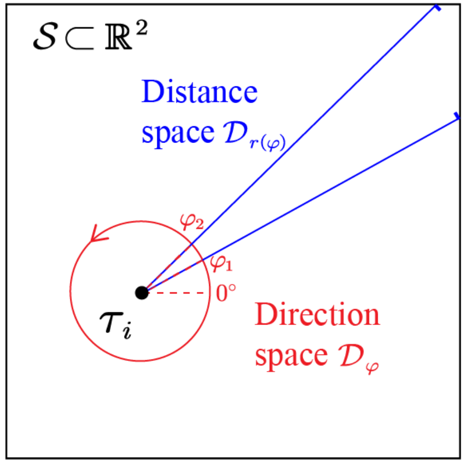
TraCS-D: $\mathcal{D}_{\varphi} \otimes \mathcal{D}_{r(\varphi)}$

Decomposition of Continuous Space

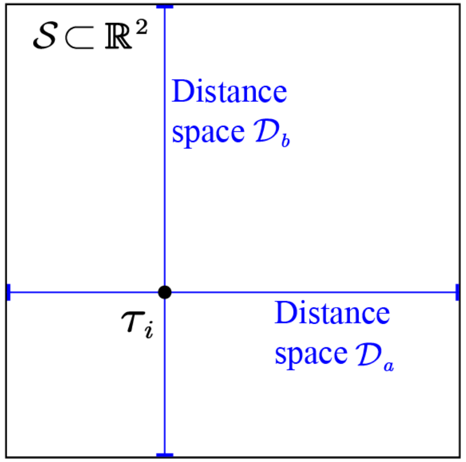
- $\mathcal{S} = \mathcal{D}_{\varphi} \otimes \mathcal{D}_{r(\varphi)}$

2D space 1D subspaces

- Each location $\tau_{i+1} \in \mathcal{S}$ has a unique representation $(\varphi, r(\varphi))$
- Perturb φ and $r(\varphi)$ using 1D mechanisms



- $\mathcal{S} = \mathcal{D}_a \times \mathcal{D}_b$
- Each location $\tau_{i+1} \in \mathcal{S}$ has a unique representation (a, b)
- Perturb a and b using 1D mechanisms



TraCS-C: $\mathcal{D}_a \times \mathcal{D}_b$

$\mathcal{M}$ for Continuous Space

- **Q:** How to design LDP mechanisms for:

  - circular space $[0,2\pi) \to [0,2\pi)$? linear space $[a_{\text{sta}}, a_{\text{end}}) \to [a_{\text{sta}}, a_{\text{end}})$?

$\mathcal{M}$ for Continuous Space

- **Q:** How to design LDP mechanisms for:

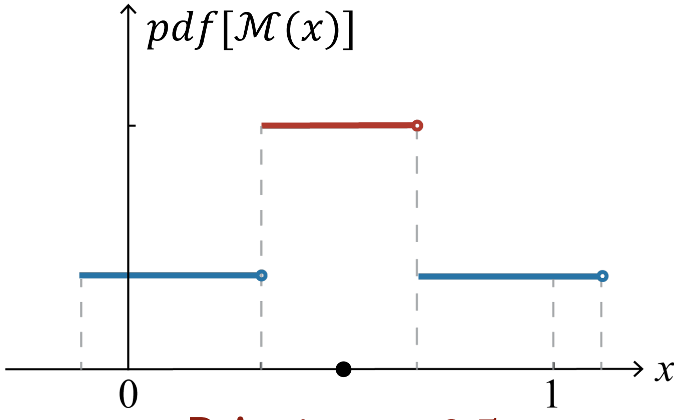
  - circular space $[0,2\pi) \rightarrow [0,2\pi)$? linear space $[a_{\mathrm{sta}}, a_{\mathrm{end}}) \rightarrow [a_{\mathrm{sta}}, a_{\mathrm{end}})$?

- **S:** Piecewise-based mechanism

  - originally designed for mean * /distribution estimation**

$$pdf[\mathcal{M}(x) = y] = \begin{cases} p_\varepsilon & \text{if } y \in [l_{x,\varepsilon}, r_{x,\varepsilon}), \\ p_\varepsilon / \exp(\varepsilon) & \text{otherwise,} \end{cases}$$

  - ensure LDP for $[0,1) \rightarrow [-C, C)$



Private $x = 0.5$

* Collecting and Analyzing Multidimensional Data with Local Differential Privacy, ICDE 2019

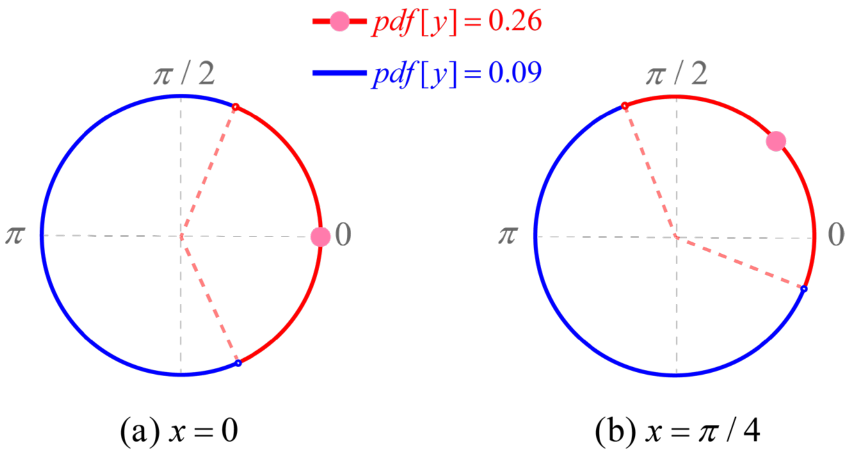
** Estimating Numerical Distributions under Local Differential Privacy, SIGMOD 2020

$\mathcal{M}_\circ$ and $\mathcal{M}_-$

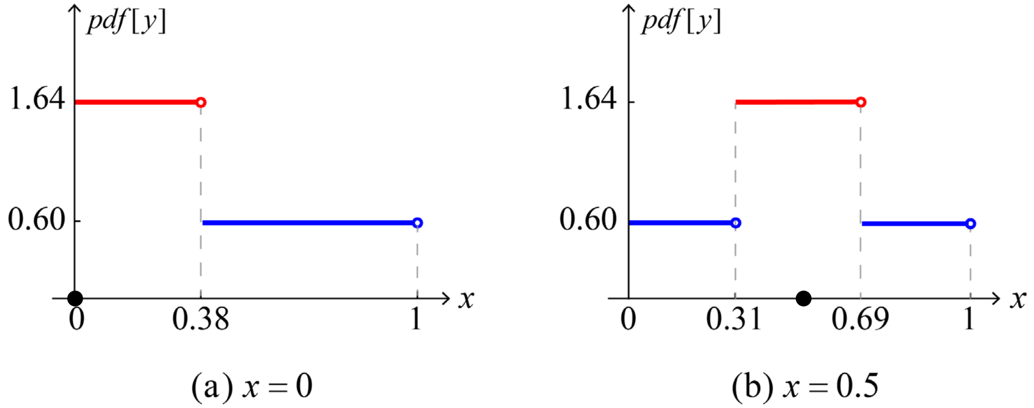
- Design piecewise-based mechanisms $\mathcal{M}_\circ$ and $\mathcal{M}_-$

$[0,2\pi) \rightarrow [0,2\pi)$ $[0,1) \rightarrow [0,1)$

- Examples of $\mathcal{M}_\circ$ and $\mathcal{M}_-$



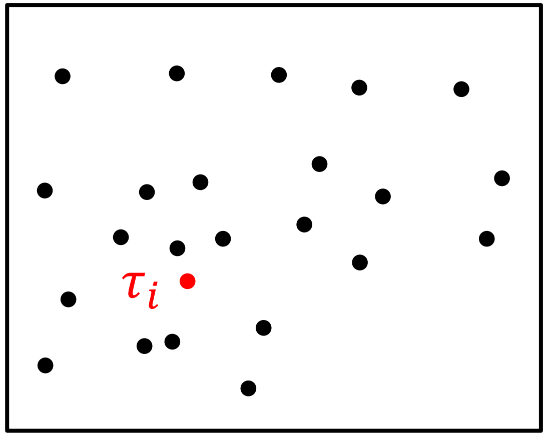
$\mathcal{M}_\circ(x; \varepsilon = 1)$



$\mathcal{M}_-(x; \varepsilon = 1)$

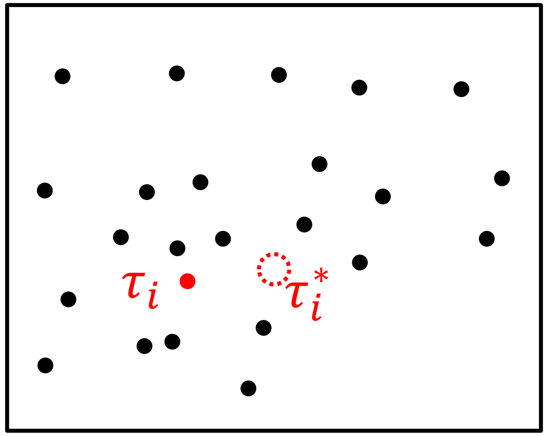
Rounding to Discrete Spaces

- $\mathcal{S}$ = [longtitude] × [latitude]



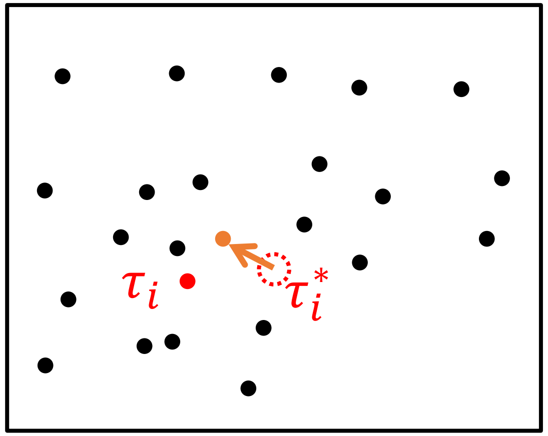
Rounding to Discrete Spaces

- $\mathcal{S}$ = [longtitude] × [latitude] → apply TraCS to $\mathcal{S}$



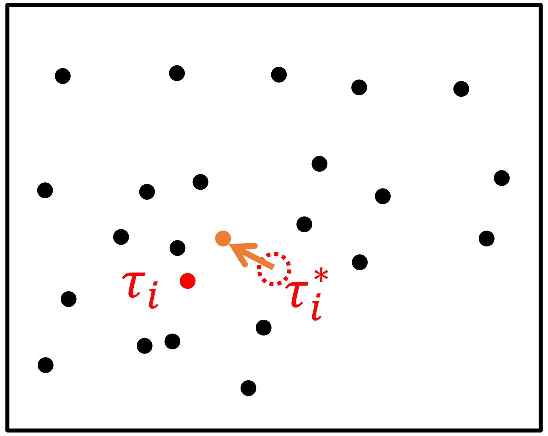
Rounding to Discrete Spaces

- $\mathcal{S}$ = [longtitude] × [latitude] → apply TraCS to $\mathcal{S}$ → round to discrete locations



Rounding to Discrete Spaces

- $\mathcal{S}$ = [longtitude] × [latitude] → apply TraCS to $\mathcal{S}$ → round to discrete locations



- Advantages:

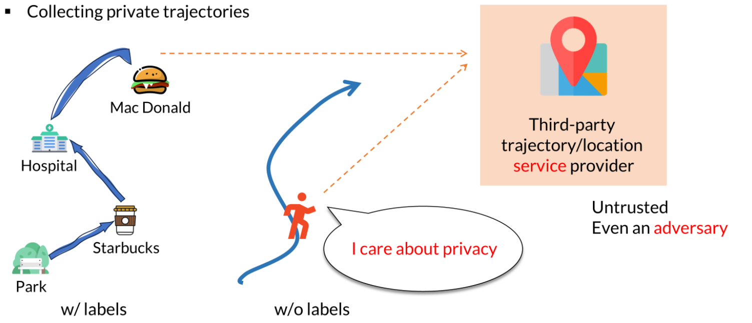
  - not affected by # locations/cells

  - efficient sampling ($\Theta(1)$ time complexity; EM: $\Theta(m)$)

TraCS: Trajectory Collection in Continuous Space under Local Differential Privacy

Trajectory Collection

- Collecting private trajectories



LDP-fy a Trajectory

- LDP-fy: perturb a trajectory with LDP guarantee
  - cannot distinguish location τ_1 from τ_2 with confidence level e^ε

$$\forall \tau_1, \tau_2, \tau_* \in \mathcal{S}: \frac{\Pr[\mathcal{M}(\tau_1) = \tau_*]}{\Pr[\mathcal{M}(\tau_2) = \tau_*]} \le e^\varepsilon$$

distinguish level

location space

probability difference of outputting τ_*

- **We want:** $\mathcal{M}$ for continuous location space
  - existing methods are all for discrete location space → not always available
  - continuous space ⊃ discrete space → can apply to discrete space

Existing Methods: **Weaknesses**

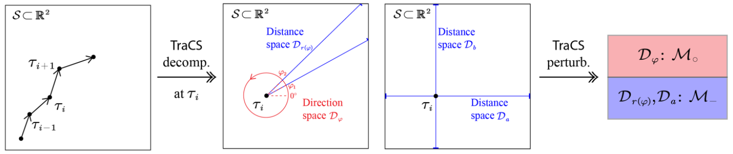
- NGram [
  - e.g. dist
  - the nex
  - $\mathcal{M}_{\exp}(\tau$
- ATP [202
  - *i.* divide
  - *ii.* pertu
  - *iii.* $\mathcal{M}_{\exp}$

- Privacy: only for discrete space
  - indistinguishability for finite # locations/cells/sectors (weak privacy)
- Utility: relies on the Exponential Mechanism
  - efficacy depends on # locations/cells/sectors
  - high time complexity (in computing and sampling)
  - discrete locations not always available

reduced $\mathcal{S}$

This Paper: Continuous Space

- TraCS-D: direction-distance perturbation
- TraCS-C: coordinates perturbation
- **Key idea:** decomposes $\mathcal{S}$ into two subspaces → design $\mathcal{M}$ for each subspace



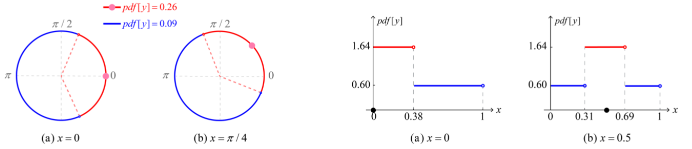
TraCS-D: $\mathcal{D}_\varphi \otimes \mathcal{D}_{r(\varphi)}$ TraCS-C: $\mathcal{D}_a \times \mathcal{D}_b$

$\mathcal{M}_\circ$ and $\mathcal{M}_-$

- Design piecewise-based mechanisms $\mathcal{M}_\circ$ and $\mathcal{M}_-$

$[0,2\pi) \to [0,2\pi)$ $[0,1) \to [0,1)$

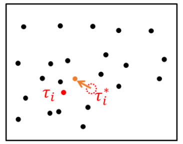
- Examples of $\mathcal{M}_\circ$ and $\mathcal{M}_-$



$\mathcal{M}_\circ(x; \varepsilon = 1)$ $\mathcal{M}_-(x; \varepsilon = 1)$

Rounding to Discrete Spaces

- $\mathcal{S}$ = [longtitude] × [latitude] → apply TraCS to $\mathcal{S}$ → round to discrete locations



- Advantages:
  - not affected by # locations/cells
  - efficient sampling ($\Theta(1)$ time complexity; EM: $\Theta(m)$)

Thank you!