Is Geo-Indistinguishability What You Are Looking for?

Simon Oya, Carmela Troncoso, Fernando Pérez-González

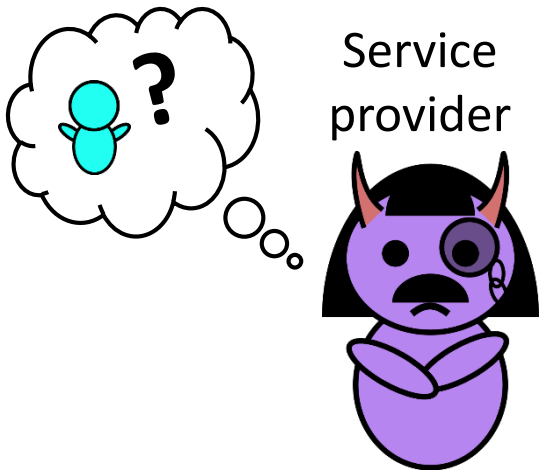
Universidad de Vigo
GPSC
Signal Processing in Communications Group

atlanTTic
research center for Telecommunication Technologies

institute imdea software

Motivation. Obfuscation-Based Location Privacy.

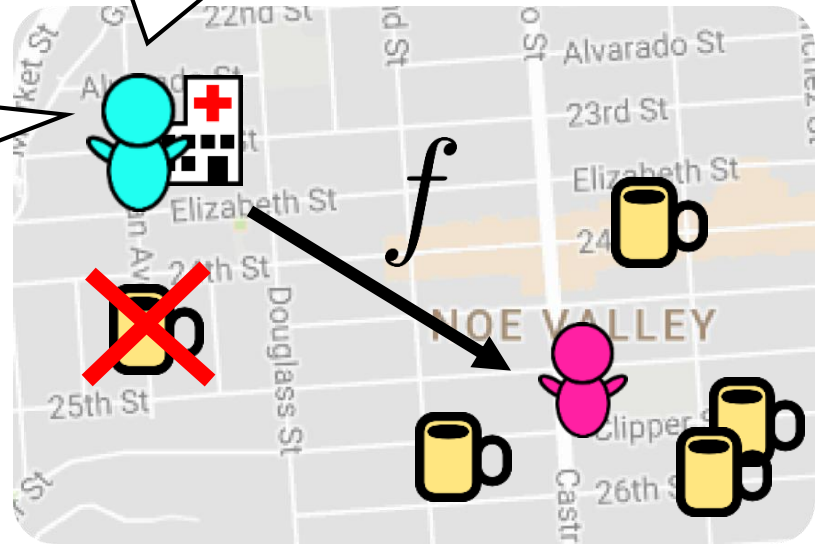
- Location information is sensitive.
- Solution: obfuscation mechanisms $f(\cdot|\cdot)$



I’m at the fake location , closest ?

Here you go!

I want to use location services without disclosing my location



- We get some privacy.
- We lose some quality of service.
- There are many metrics to assess the privacy of $f(\cdot|\cdot)$
- A popular notion is **<u>geo-indistinguishability</u>**.

<u>In this work</u>

We study the privacy implications of geo-indistinguishability, revealing some of its issues.

Geo-Indistinguishability [1]

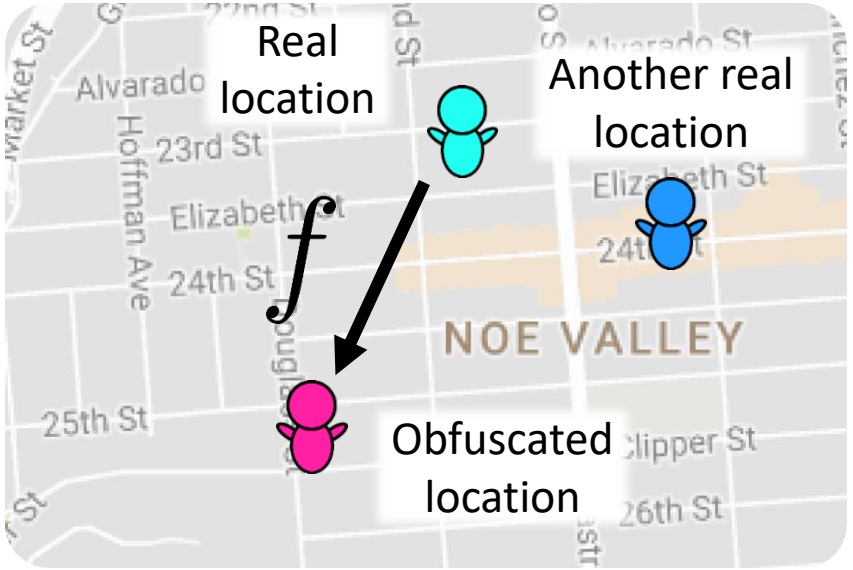
- GeoInd means ensuring that  and  are “indistinguishable” given  .
- Mathematically: $\forall$ 

$$f(\text{pink}|\text{cyan}) \leq e^{\epsilon \cdot d(\text{cyan},\text{blue})} \cdot f(\text{pink}|\text{blue})$$

Obfuscation mechanism

Privacy parameter

Distance metric (e.g., Euclidean)



ϵ ↑ Less privacy

ϵ ↓ More privacy

$d(\text{cyan},\text{blue})$ ↑ Less privacy (easier to distinguish)

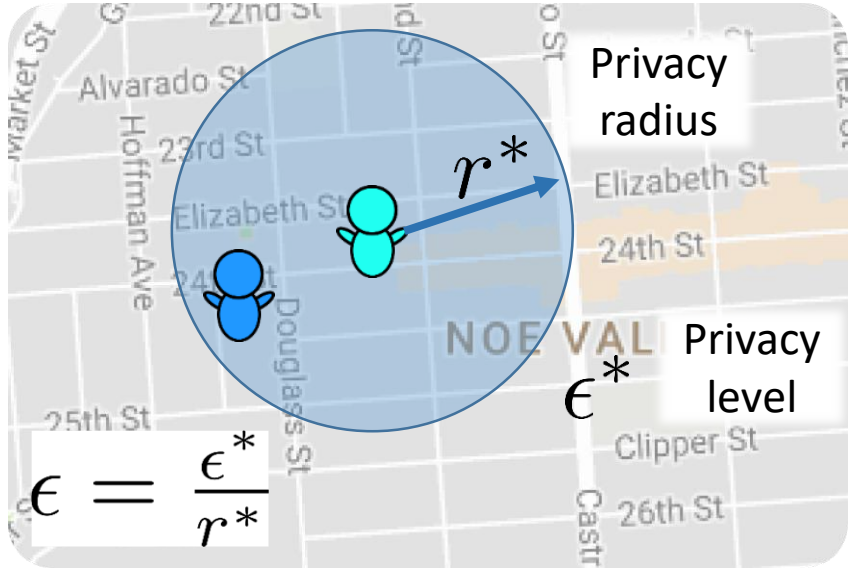
$d(\text{cyan},\text{blue})$ ↓ More privacy (harder to distinguish)

[1] Andrés, Miguel E., et al. "Geo-indistinguishability: Differential privacy for location-based systems." *CCS’13*.

Choosing the GeoInd Privacy Parameter

- How do we choose ϵ ?
- Typical approach:

$$f(\text{👶}|\text{👶}) \leq e^{\epsilon \cdot d(\text{👶},\text{👶})} \cdot f(\text{👶}|\text{👶})$$



- How do we choose ϵ^*?
  - From log(1.4) to log(10).
  - Normally, log(2).
- Example: $\left.\begin{array}{r} r^* = 0.5\text{km} \\ \epsilon^* = \log(2) \end{array}\right\} \epsilon \approx 0.60\text{km}^{-1}$
- Inside the region, we get:

$$f(\text{👶}|\text{👶}) \leq 2 \cdot f(\text{👶}|\text{👶})$$

Hard to interpret

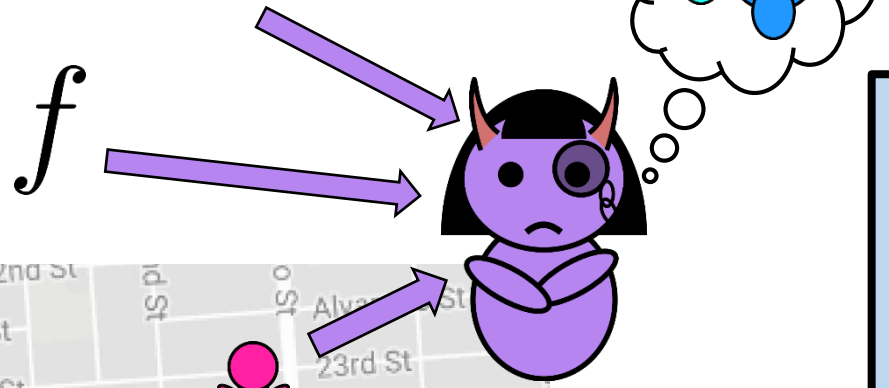
$$d(\text{👶},\text{👶}) \leq r^* \Rightarrow f(\text{👶}|\text{👶}) \leq e^{\epsilon^*} \cdot f(\text{👶}|\text{👶})$$

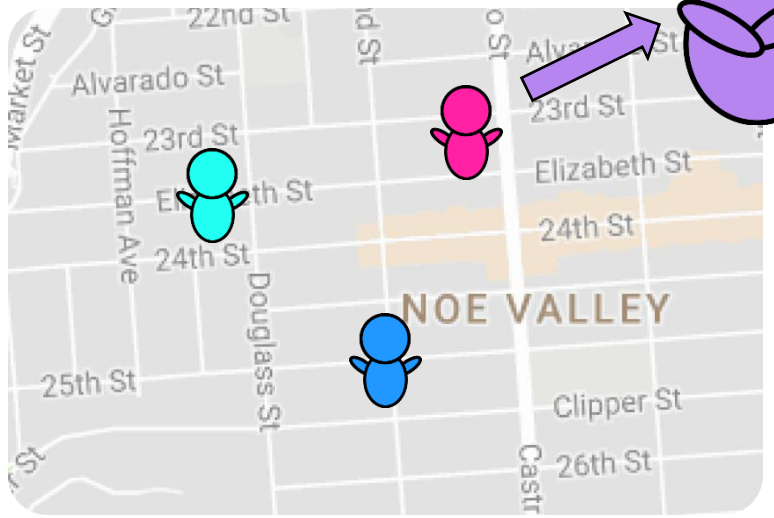
GeoInd as an Adversary Error

- Decision Adversary:

Assume $f(\text{🩷}|\text{🩵}) \leq f(\text{🩷}|\text{💙})$, so the adv. decides 💙.

$$\pi(\text{🩵}) = \pi(\text{💙}) = 0.5$$



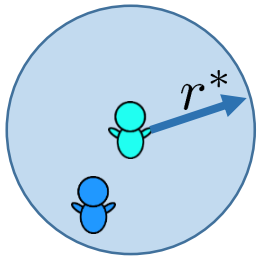


$$p_e(\text{🩵},\text{💙},\text{🩷}) = \frac{f(\text{🩷}|\text{🩵})}{f(\text{🩷}|\text{🩵}) + f(\text{🩷}|\text{💙})}$$

f gives GeoInd if and only if, $\forall$🩵,💙,🩷:

$$p_e(\text{🩵},\text{💙},\text{🩷}) \geq p_e^* = \frac{1}{1+e^{\epsilon \cdot d(\text{🩵},\text{💙})}}$$

- Previous example:

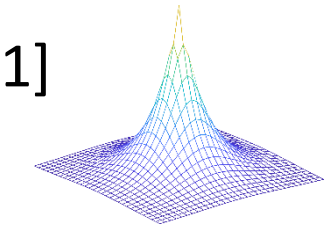


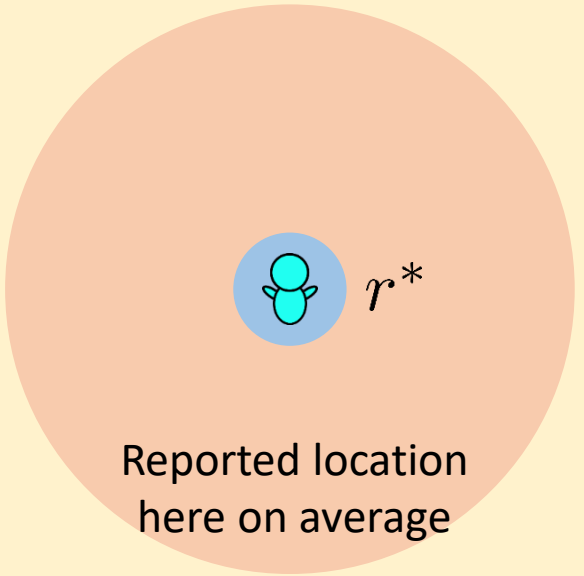
$$f(\text{🩷}|\text{🩵}) \leq 2 \cdot f(\text{🩷}|\text{💙}) \Rightarrow p_e \geq 0.33$$

Easier to interpret

GeoInd in Numbers

- Two GeoInd mechanisms: Laplace [1] and Laplace with remapping [2].
- Example.
  - Privacy goal: $p_e^* = 0.4$ for locations in r^*
  - Laplace: $\bar{r} \approx 5r^* \quad r_{95} \approx 12r^*$



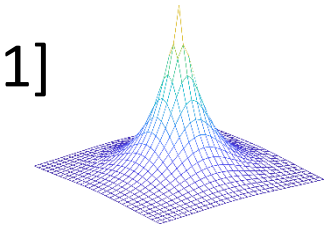


[1] Andrés, Miguel E., et al. "Geo-indistinguishability: Differential privacy for location-based systems." *CCS'13*.

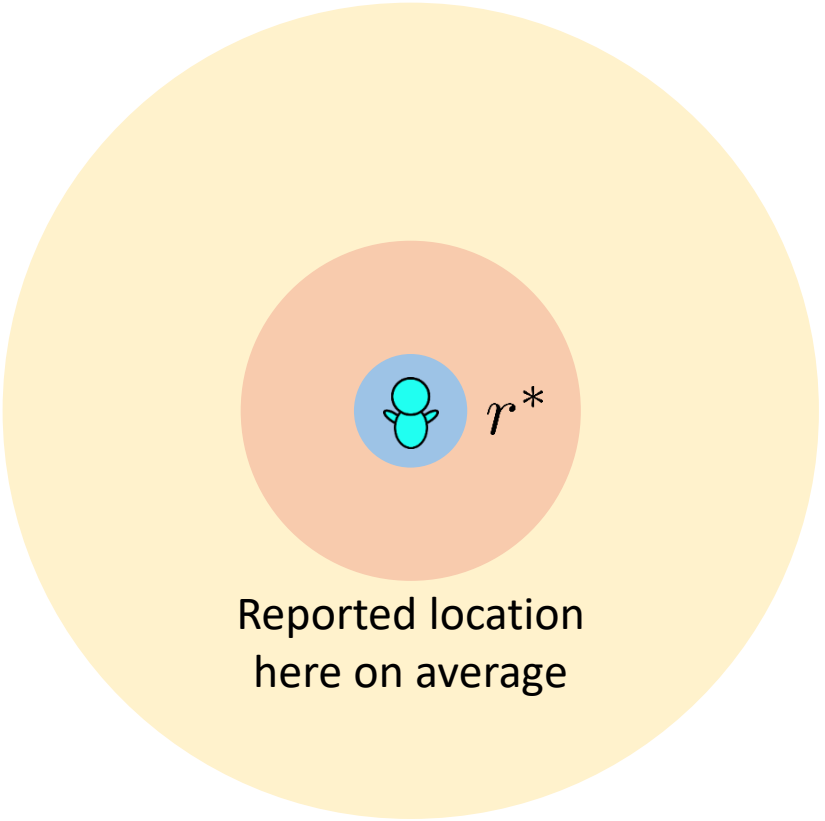
[2] Chatzikokolakis, Konstantinos, Ehab ElSalamouny, and Catuscia Palamidessi. "Efficient Utility Improvement for Location Privacy." *PoPETS'17.* 308-328.

GeoInd in Numbers

- Two GeoInd mechanisms: Laplace [1] and Laplace with remapping [2].



- Example.
  - Privacy goal: $p_e^* = 0.4$ for locations in r^*
  - Laplace: $\bar{r} \approx 5r^* \quad r_{95} \approx 12r^*$
  - Laplace + RM: $\bar{r} \approx 3r^* \quad r_{95} \approx 10r^*$
    (Gowalla dataset)

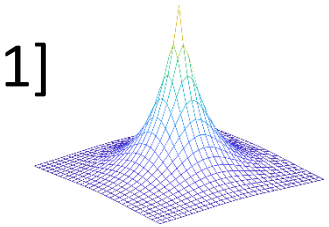


[1] Andrés, Miguel E., et al. "Geo-indistinguishability: Differential privacy for location-based systems." *CCS'13*.

[2] Chatzikokolakis, Konstantinos, Ehab ElSalamouny, and Catuscia Palamidessi. "Efficient Utility Improvement for Location Privacy." *PoPETS'17.* 308-328.

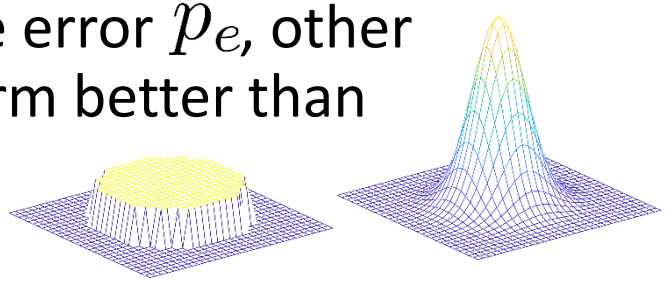
GeoInd in Numbers

- Two GeoInd mechanisms: Laplace [1] and Laplace with remapping [2].

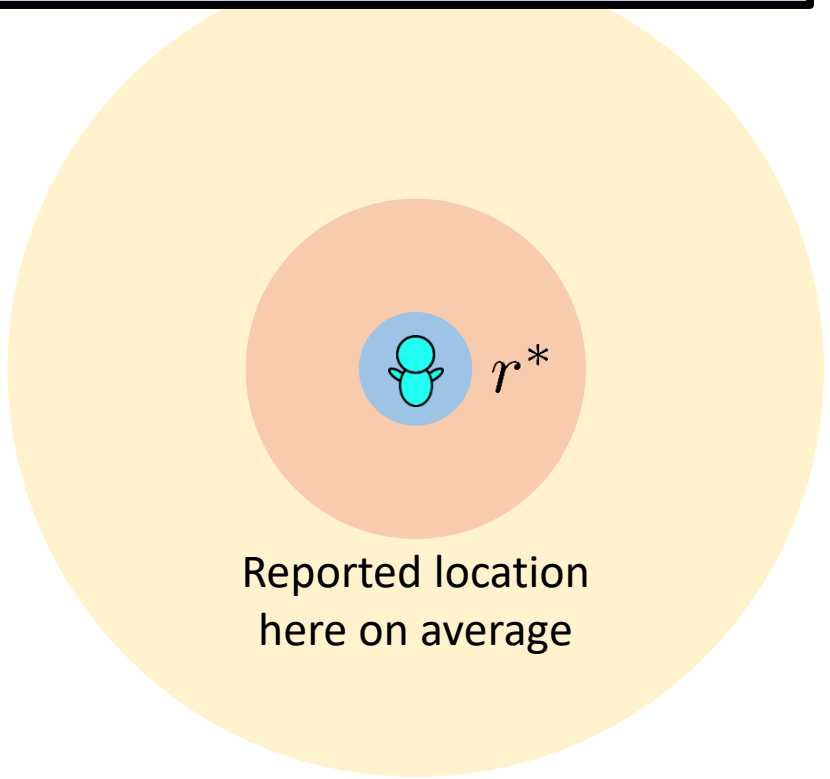


- Example.
  - Privacy goal: $p_e^* = 0.4$ for locations in r^*
  - Laplace: $\bar{r} \approx 5r^* \quad r_{95} \approx 12r^*$
  - Laplace + RM: $\bar{r} \approx 3r^* \quad r_{95} \approx 10r^*$
    (Gowalla dataset)

- In terms of average error $\bar{p}_e$, other mechanisms perform better than Laplace.



The price we pay is too high for the privacy we get!!
Bad privacy-utility trade-off



[1] Andrés, Miguel E., et al. "Geo-indistinguishability: Differential privacy for location-based systems." *CCS'13*.

[2] Chatzikokolakis, Konstantinos, Ehab ElSalamouny, and Catuscia Palamidessi. "Efficient Utility Improvement for Location Privacy." *PoPETS'17.* 308-328.

Where is the problem?

- GeoInd comes from differential privacy.
- Differential Privacy scenarios: **low sensitivity queries.**
  - It is possible to achieve $f(\text{🚺}|\text{🚺}) \leq e^{\epsilon^*} \cdot f(\text{🚺}|\text{🚹})$ with high privacy $\begin{cases} \epsilon^* = 0.1 \\ \epsilon^* = 0.01 \end{cases}$
- User-centric Location Privacy: **high sensitivity queries**!

$$\left.\begin{array}{c} \epsilon^* = 0.01 \\ d(\text{🚺},\text{🚹}) = 100\text{m} \end{array}\right\} \bar{r} = 20\,000\text{m}$$

<u>Solutions?</u>

- Re-design location queries to have low sensitivity [1].
- Use bandwidth as a resource to improve utility [1].
- Use less ambitious privacy metrics…

[1] Andrés, Miguel E., et al. "Geo-indistinguishability: Differential privacy for location-based systems." *CCS'13*.

Conclusions

- Evaluate privacy and quality loss **numerically**.
- GeoInd as an adversary error can help in this regard.
- Understand what GeoInd means:
  - If you want average protection, use something else!
  - If you really want GeoInd, re-design queries, use bandwidth as a resource, etc.

Thank you!!

simonoya@gts.uvigo.es