A privacy mechanism with minimal noise for location-based services

By:

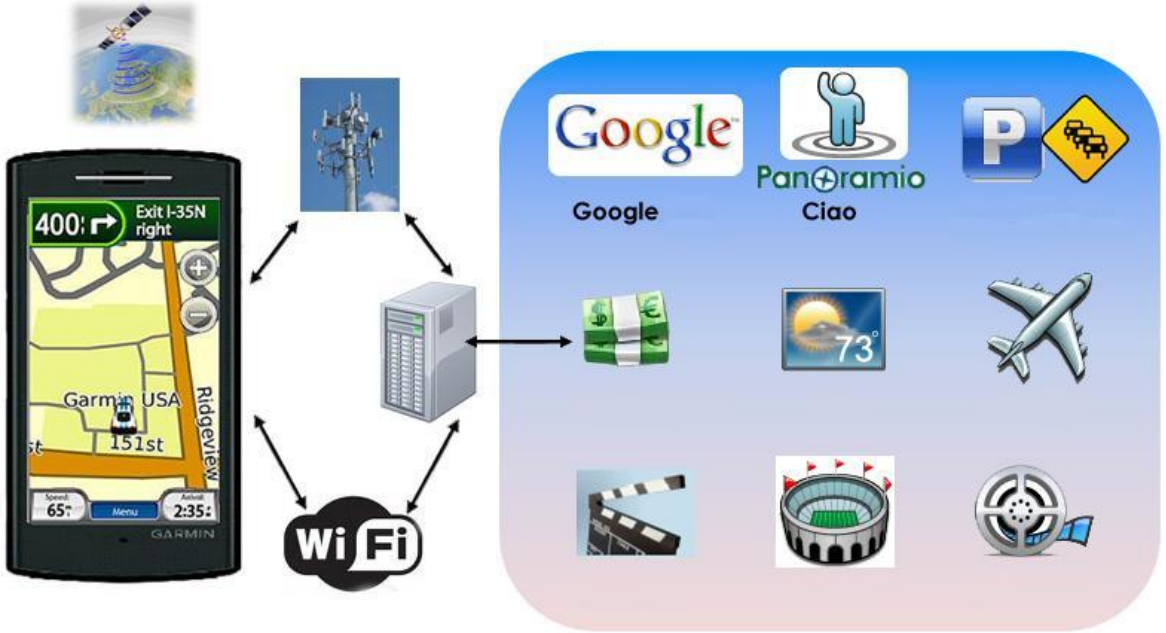
Ehab ElSalamouny

Joint work with

Sébastien Gambs

Motivation

- A user requires to access a service based on his location.



Motivation

My location is **p**

Where are the nearby restaurants ?

Service provider

Motivation

My location is **p**

Where are the nearby restaurants ?

Service provider

Motivation

My location is **p**

Where are the nearby restaurants ?

Service provider

p

Objectives:

- Protect the user's location privacy: (Part 1)

  Inhibit any adversary from precisely identifying the user's location.

- Maximise the utility: (Part 2)

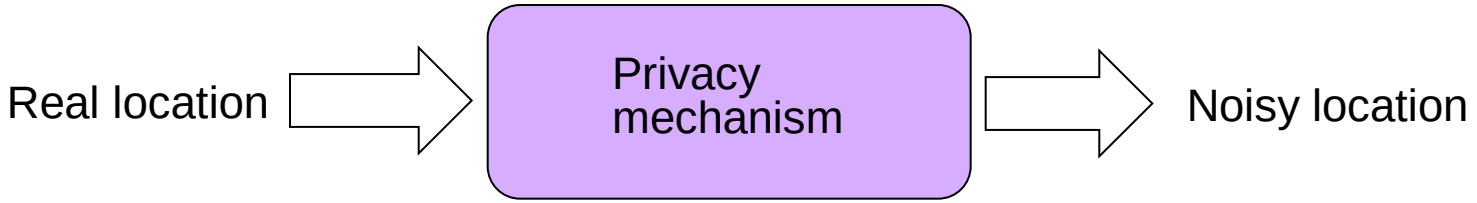
  Provide the user with a useful service relevant to his real location.

Outline

- Location privacy mechanisms.
- (D,ε)-location privacy.
- Comparison to ε-differential privacy and geo-indistinguishability.
- Asymmetric mechanisms.
- Symmetric mechanisms: privacy and utility.
- Circular noise functions.
- Optimal symmetric mechanisms.
- The stepping noise function.

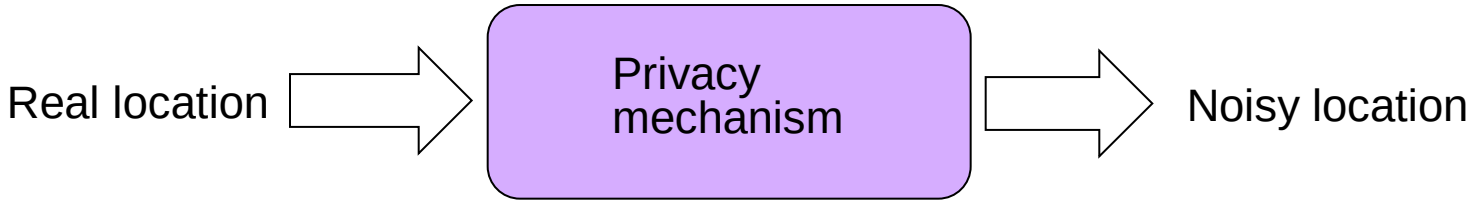
Location privacy mechanisms

- An intuitive approach to protect the location privacy is to report a noisy location instead of the real one.
- A location privacy mechanism is therefore seen a probabilistic function from the location space to itself.



Location privacy mechanisms

- An intuitive approach to protect the location privacy is to report a noisy location instead of the real one.
- A location privacy mechanism is therefore seen a probabilistic function from the location space to itself.



- Informally: A mechanism provides location privacy if any adversary is unable to distinguish (from the output) between two nearby locations $\mathbf{p_1}$,$\mathbf{p_2}$.

(D, ε)-Location privacy

- Definition:

  For a distance $D > 0$, and $\varepsilon > 0$, a mechanism K satisfies (D, ε)-location privacy if it holds for all locations $\mathbf{p_1}, \mathbf{p_2}$ where $d(\mathbf{p_1}, \mathbf{p_2}) \leq D$, and all regions S that

$$\frac{P(K(p_1) \in S)}{P(K(p_2) \in S)} \leq e^{\varepsilon}$$

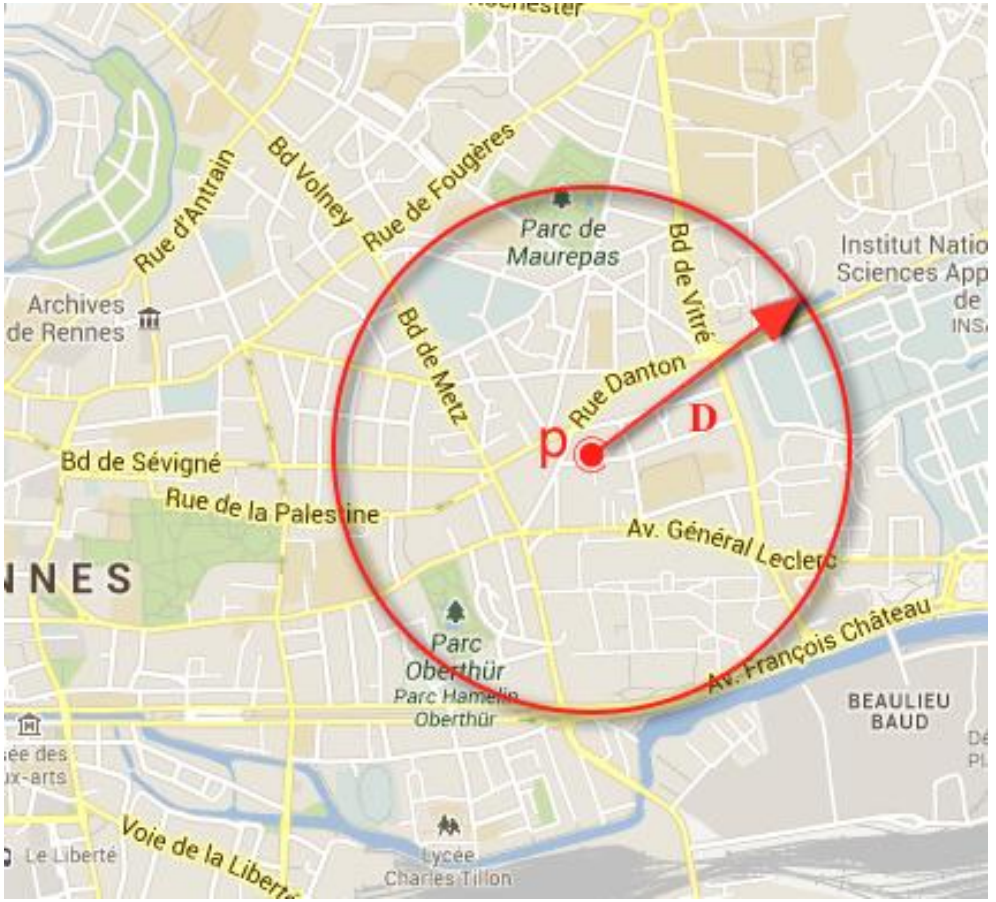
(D, ε)-Location privacy

$$\frac{P(K(p_1) \in S)}{P(K(p_2) \in S)} \leq e^{\varepsilon}$$

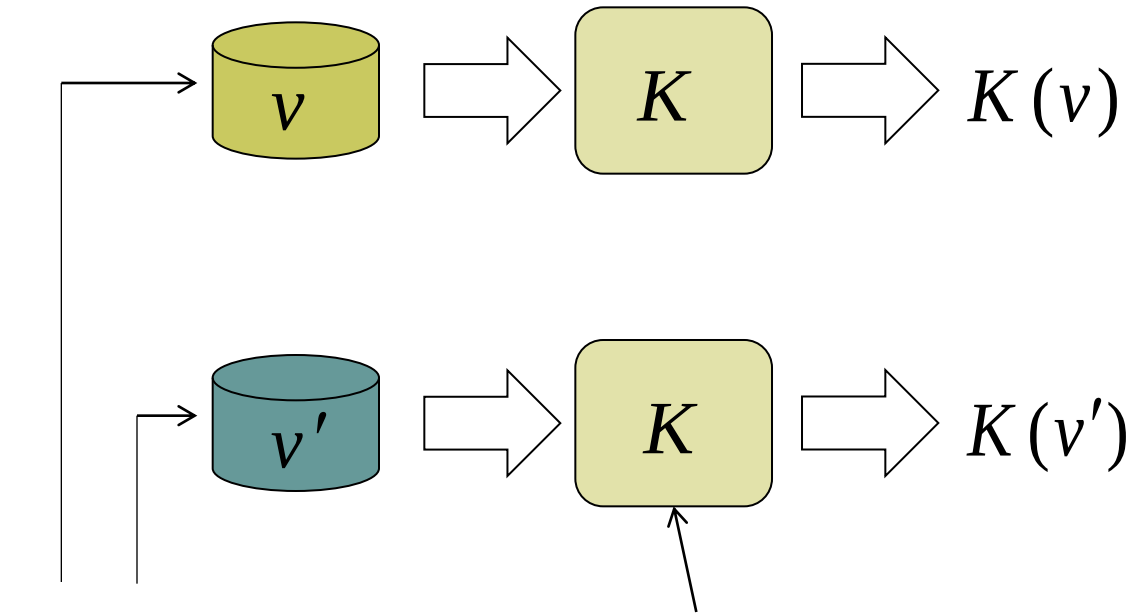
$$\forall S, p_1, p_2 : \quad d(p_1, p_2) \leq D$$

(D, ε)-Location privacy

- User's real location **p** is therefore indistinguishable from other nearby points.



Comparison to ε-Differential privacy [Dwork, 2006]



$$\frac{P(K(v) \in S)}{P(K(v') \in S)} \leq e^{\varepsilon}$$

$$\forall S, v, v' : \quad v \approx v'$$

Comparison to ε-geo-indistinguishability [Andres et al, 2012]

- A mechanism K satisfies ε-geo-indistinguishability if

$$\frac{P(K(p_1) \in S)}{P(K(p_2) \in S)} \leq e^{\varepsilon . r}$$

$$\forall S, r, p_1, p_2 : \quad d(p_1, p_2) \leq r$$

- Here, the indistinguishability level (εr) is enforced to decrease with $d(\mathbf{p}_1, \mathbf{p}_2)$.
- Thus ε-geo-indistinguishability implies (D,ε)-location privacy.

Asymmetric mechanisms: Implementing (D,ε)-LP

- Every point **p** in the space is associated with a probability density function $\mathbf{F}_p$ over the output space.
- For every **p**, the pdf $\mathbf{F}_p$ is continuous almost everywhere.

- Theorem: A mechanism K satisfies (D,ε)-LP iff for all points $\mathbf{p}_1, \mathbf{p}_2, \mathbf{u}$ such that $\mathbf{F}_{p1}$ and $\mathbf{F}_{p2}$ are continuous at **u**, and $d(\mathbf{p}_1, \mathbf{p}_2) \leq D$, it holds

$$F_{p1}(u) \leq e^{\varepsilon} F_{p2}(u).$$

Symmetric mechanisms: Implementing (D,ε)-LP

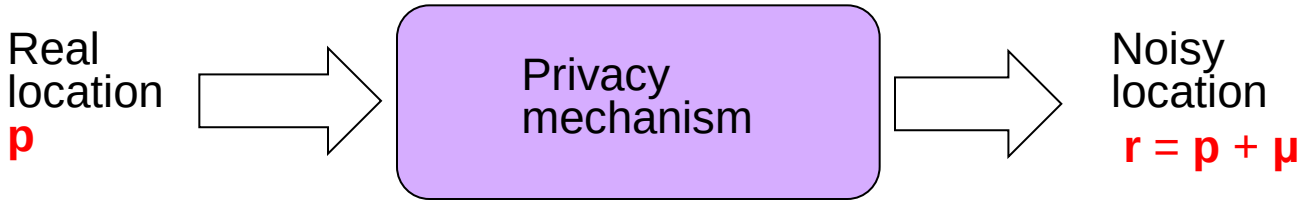
- A random point **u** is sampled according to a single pdf **F**, called a noise function.
- Given the real location is **p,** the mechanism's output is then **p** + **vec(u)**.

→ The *position vector* of **u** (noise vector)

- Theorem: A symmetric mechanism K satisfies (D,ε)-LP iff for all points **u,v** at which the noise function **F** is continuous at **u,v**, and d(**u**,**v**)≤ D, it holds

$$F(u) \leq e^{\varepsilon} F(v).$$

Symmetric mechanisms: Loss functions and disutility



- The output **r** incurs larger loss as the distance $d(p,r)$ increases.
- The loss incurred by **r**, given the real location **p**, is modelled by a loss function L of the distance $d(p,r)$.
- The loss L is assumed to be increasing with respect to $d(p,r)$.

Symmetric mechanisms: Loss functions and disutility

- We consider polynomial losses. Examples are:

$$L(d) = d, \quad L(d) = d^2, \quad L(d) = d + d^2, \quad \text{etc.}$$

Symmetric mechanisms: Loss functions and disutility

- We consider polynomial losses. Examples are:

  $L(d) = d$, $L(d) = d^2$, $L(d) = d + d^2$, etc.

- The disutility of a mechanism K_F (with a noise function F) is defined as the expected value of the loss function L.

$$\Phi(F, L) = \iint_{E^2} F(u) L(\|u\|) dA$$

Symmetric mechanisms: Loss functions and disutility

- We consider polynomial losses. Examples are:

  $L(d) = d$, $L(d) = d^2$, $L(d) = d + d^2$, etc.

- The disutility of a mechanism K_F (with a noise function F) is defined as the expected value of the loss function L.

$$\Phi(F, L) = \iint_{E^2} F(u) L(\|u\|) dA$$

Symmetric mechanisms with circular noise functions

- Definition:

  A noise function **F** is said to be circular if

$$F(u) = F(v) \quad \forall u, v : d(o, u) = d(o, v)$$

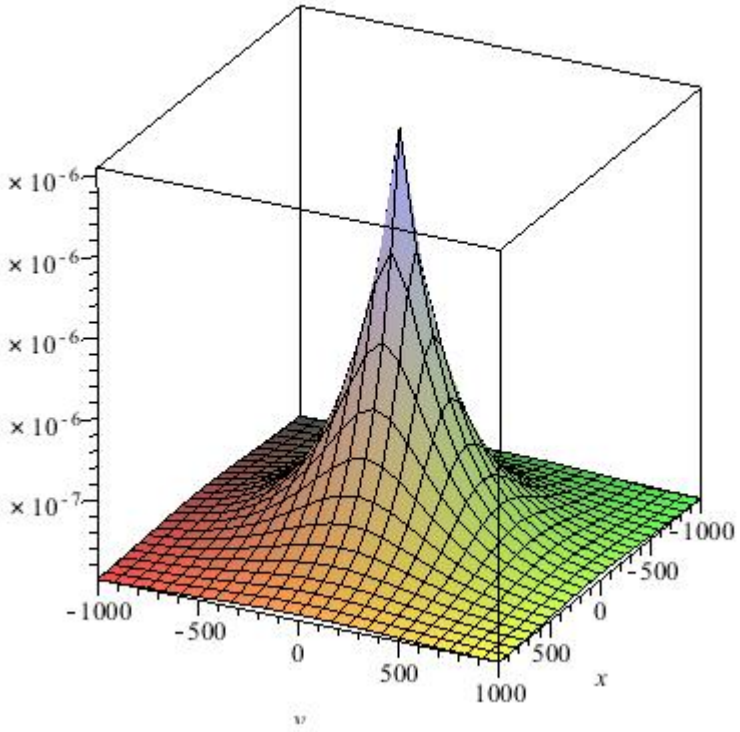
  A reference point (origin)

- That is the probability density of a noise vector depends only on its magnitude.
- Sampling from a circular noise function is simple and efficient.

Symmetric mechanisms with circular noise functions

- An example for circular noise functions satisfying (D,ε)-LP is the 2-dimensional Laplace pdf.

$$Lab_{D,\varepsilon}(u) = \left(\frac{\varepsilon}{D}\right)^2 \left(\frac{1}{2\pi}\right) e^{-\frac{\varepsilon}{D}\|u\|}$$



Symmetric mechanisms with circular noise functions

- Theorem: (circular noise functions suffice)

  Consider $D>0, \varepsilon>0$, and a loss function L. For every noise function F satisfying (D,ε)-LP, there is a circular one F^c such that

  1. F^c satisfies (D,ε)-LP,
  2. F and F^c have the same expected loss.

- Thus, using circular noise functions does not incur loss of privacy nor utility.

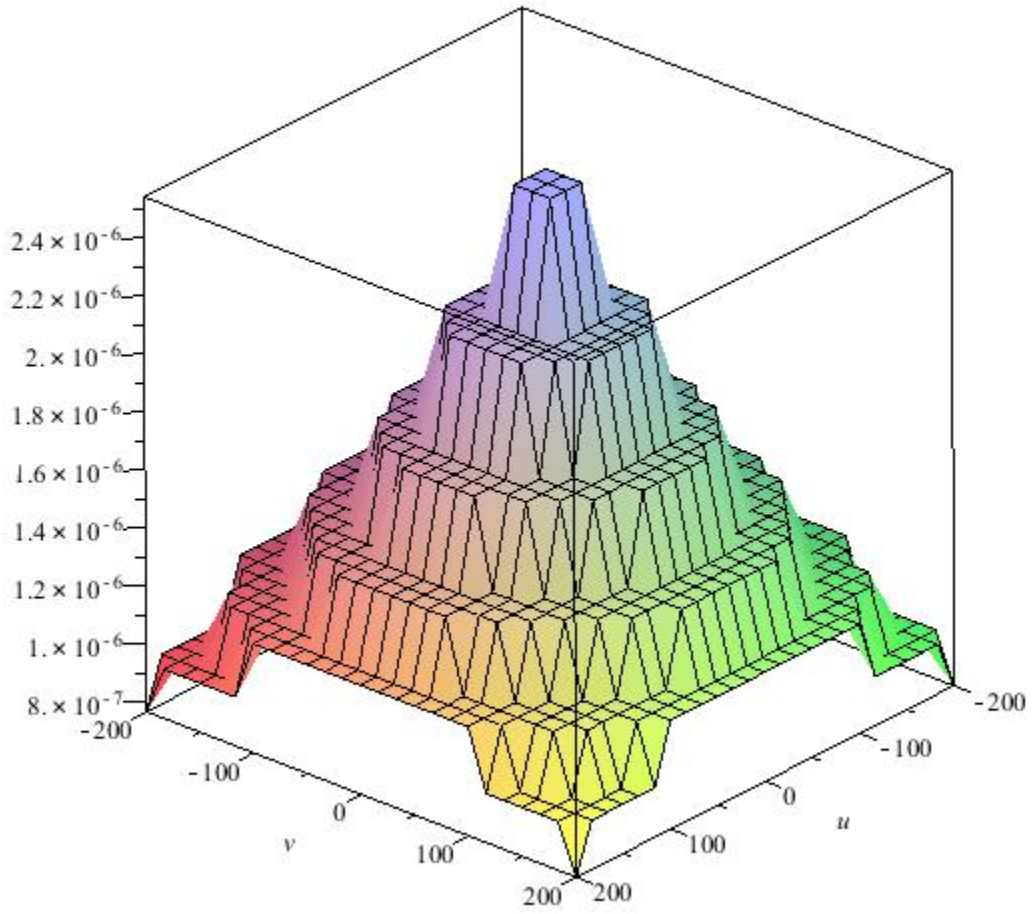
Optimal symmetric mechanisms

- Informally: A symmetric mechanism is optimal if its noise function minimises the expected loss function while satisfying (D, ε)-LP.

Difficulties:

- The optimal mechanism (noise function) depends on the privacy parameters and the loss function.
- Linear programming techniques are not appropriate:
  - Inputs and outputs (locations) are in a continuous domain.
  - The noise functions can be discontinuous.
- The calculus of variations for instance is not appropriate since the noise function is allowed to have restricted discontinuities.

The stepping noise function



The stepping noise function

- Conjecture:

  For any $D>0$, $\varepsilon>0$, and a polynomial increasing loss function L, there is s: $0 \leq s < D$ such that the stepping noise function with parameters D, s, ε is optimal for D, ε, L.

- The parameter s depends on the chosen privacy parameters D, ε, and the loss function L.

- The proof is under construction .. Coming soon !

Questions ?