Radboud University

Privacy friendly location-based services

Beatrise Bertule (s1105119), Julianne Polman (s1173079), Sarah Rahman (s1144280)

Agenda

1. Location Data Sensitivity
2. Location-based Service Architecture Assumptions
3. PET Solutions:
    3.1. Obfuscation-based Solutions
    3.2 Cryptography-based Solutions
    3.3. Policy-based Solutions
4. Conclusion

Agenda

1. Location Data Sensitivity
2. Location-based Service Architecture Assumptions
3. PET Solutions:
    - 3.1. Obfuscation-based Solutions
    - 3.2 Cryptography-based Solutions
    - 3.3. Policy-based Solutions
4. Conclusion

Location Based Services in Daily Life

Location Based Services (LBS)

- Use our location to provide services
- Requires continuous data collection

- Navigation (maps, routes)
- Ride sharing (Uber, Bolt)
- Nearby search

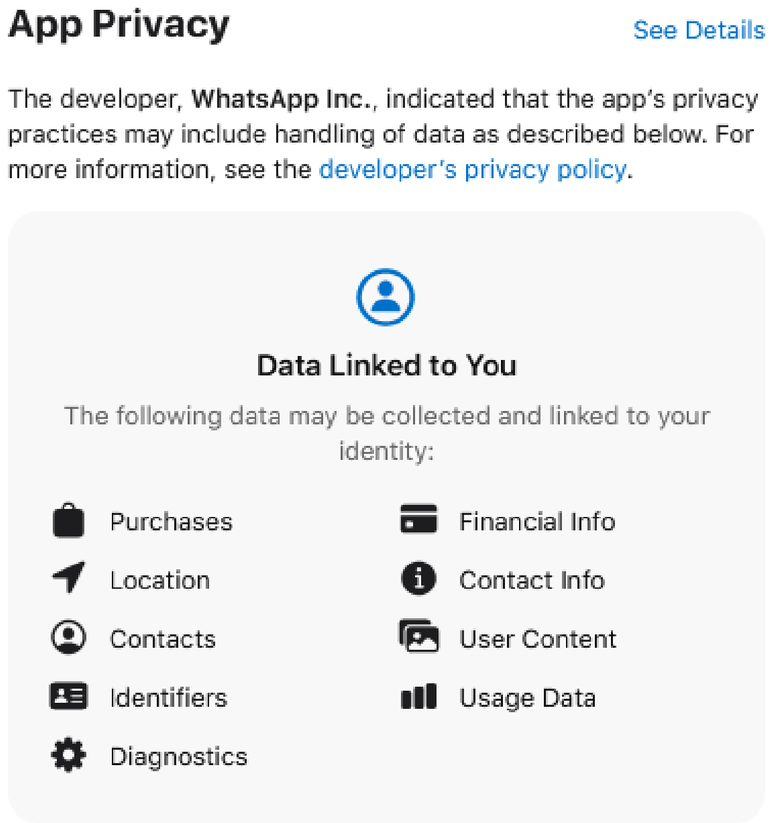
Privacy Communication in Apps

- Privacy policies
- App store privacy labels
- “We protect your privacy”
- “Used to improve services”

(Krämer, 2024)

Is anonymous data really anonymous?

- A few location points can identify you
- Patterns reveal identity
- Real world evidences exist



New York Times Investigation (2019)

- Anonymous location data tracked real individuals
- Home and workplace easily identified
- Sensitive location revealed

Even Governments Use Location Data

How the FBI can conduct mass surveillance – even without AI

Nick Robins-Early

Anthropic fought against the government's misuse of its technology, but authorities are buying Americans' data, enabling them to surveil citizens at scale

Governments purchase location data from brokers

No direct collection required

Can bypass traditional legal safeguards

Small Actions, Big Risks

Strava reveals location of French warship after naval officer logs run on aircraft carrier deck

News By Alex Blake published March 23, 2026

When fitness becomes a security threat

When you purchase through links on our site, we may earn an affiliate commission. Here's how it works.

Everyday apps can expose sensitive locations

Eg: fitness tracking data leak

Even normal usage can create risks

How many location points do you think are needed to identify a person?

Research Finding

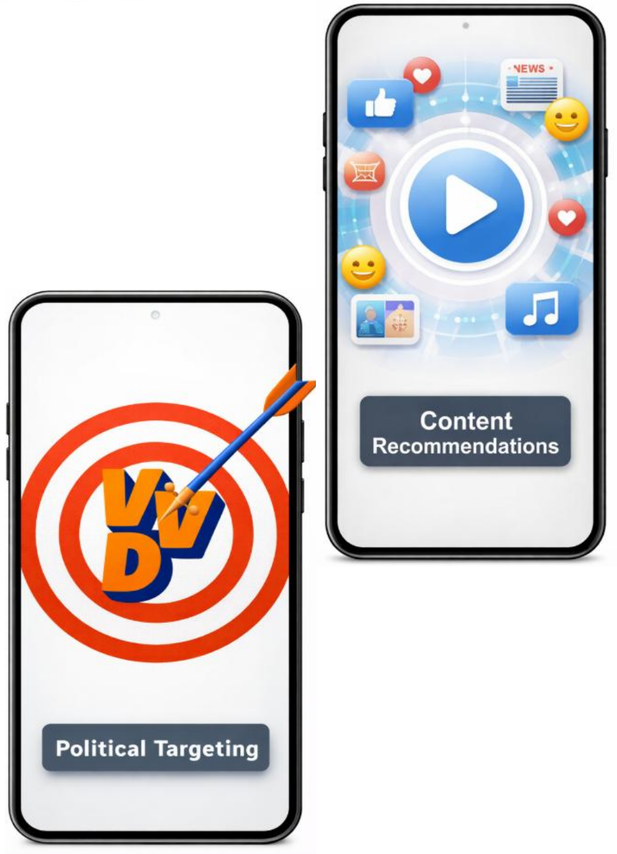
Only 4 location points can identify 95% of people

Why should we care?

- Tracking and surveillance
- Behavioral profiling
- Loss of Privacy
- Influence & manipulation

How This Affects Us

- Personalized advertising
- Content recommendation
- Political targeting



Agenda

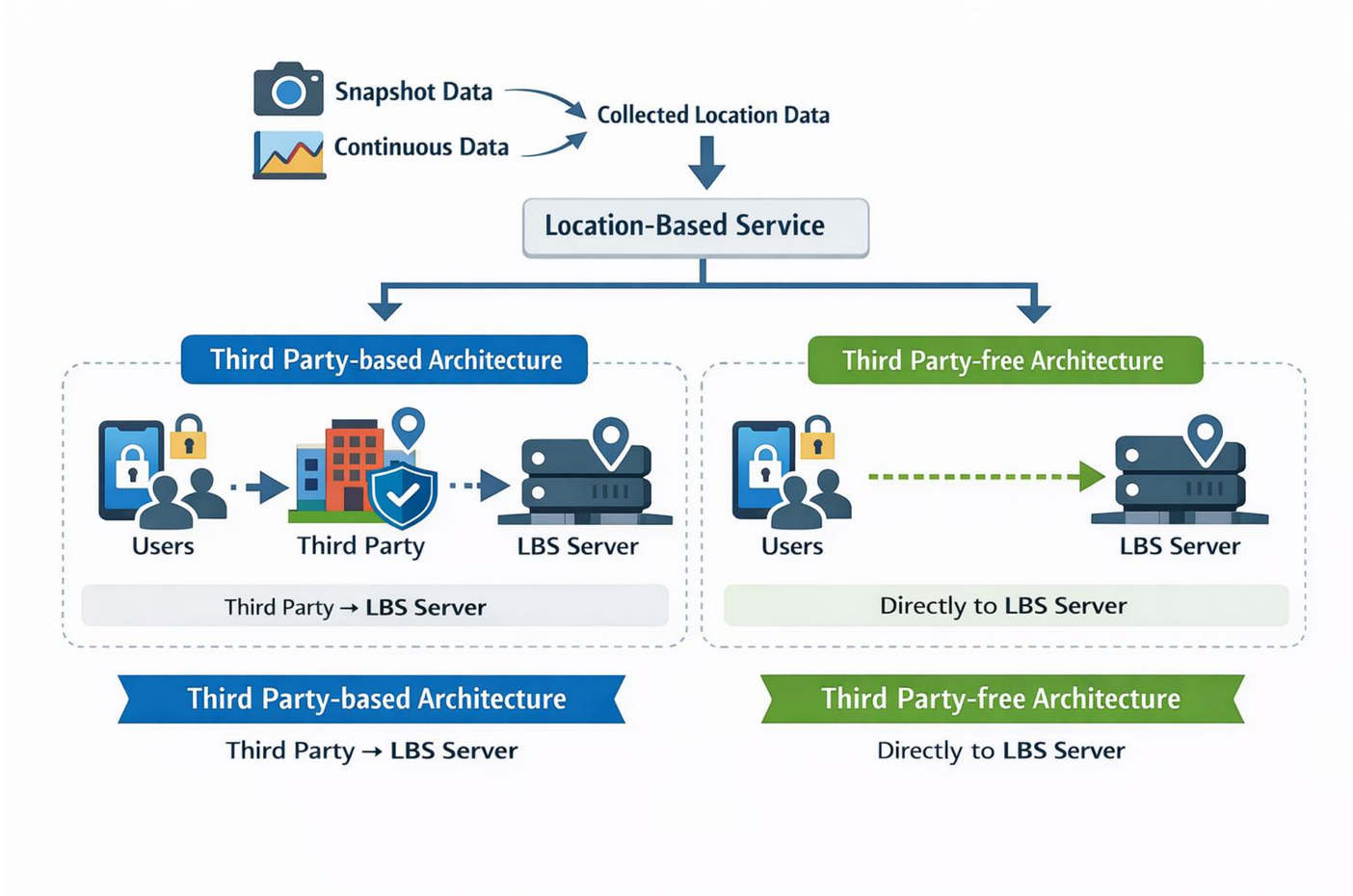
1. Location Data Sensitivity
2. Location-based Service Architecture Assumptions
3. PET Solutions:
    - 3.1. Obfuscation-based Solutions
    - 3.2 Cryptography-based Solutions
    - 3.3. Policy-based Solutions
4. Conclusion

The Data

- Snapshot data
  - Single point
- Continuous data
  - Multiple points

- Why would we want this distinction?
- Some privacy preserving solutions work for one but not for the other

Architectures

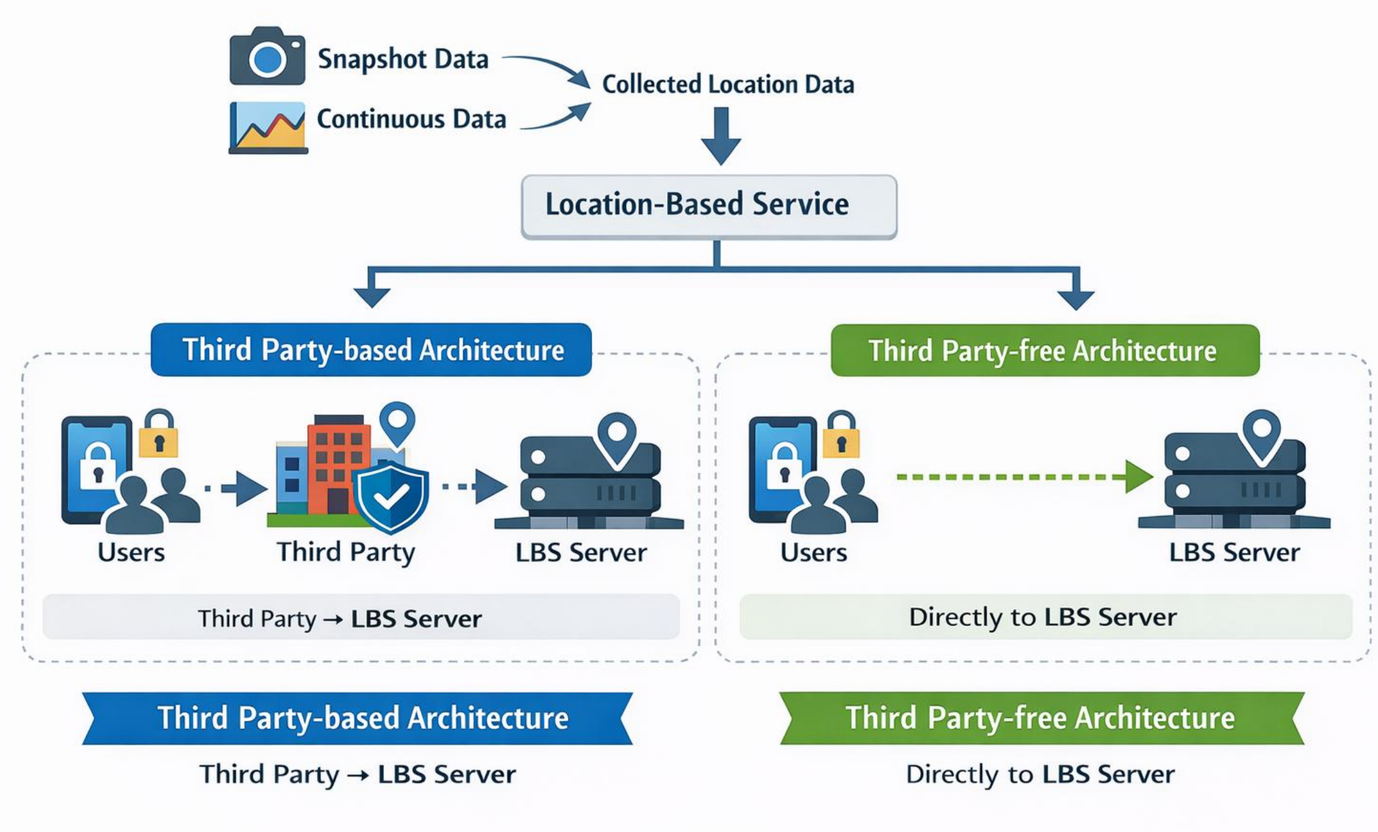


Privacy Preserving Solutions

- **Obfuscation-based Solution**
  - Can we reduce the accuracy?
  - K-anonymity, Cloaking, Dummy Locations, Differential Privacy
- **Cryptography-based Solution**
  - Can we make it so that no one sees the data?
  - Space Transformation
- **Policy-based Solution**
  - Do we really need this information?
  - GDPR

Privacy Preserving Solutions

- Obfuscation-based
- Cryptography-based
- Policy-based



Agenda

1. Location Data Sensitivity
2. Location-based Service Architecture Assumptions
3. PET Solutions:
    3.1. Obfuscation-based Solutions:
        - Spatial and Temporal Cloaking
        - Dummy Locations
        - Differential Privacy
    3.2 Cryptography-based Solutions
    3.3. Policy-based Solutions
4. Conclusion

Obfuscation–based Solution

Spatial and Temporal Cloaking

Obfuscation

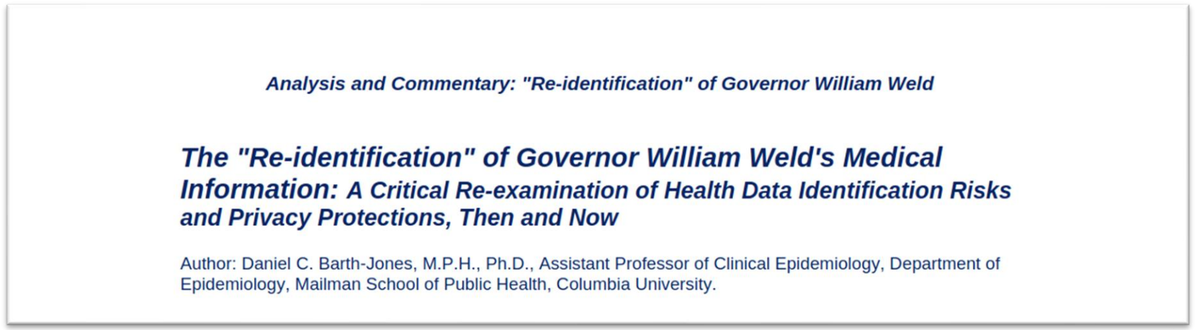
Spatial and Temporal Cloaking

- Based on concept of **k-anonymity**: a user is **indistinguishable** from at least **k – 1** others
- Prevents **re-identification** attacks through **linking** with external information
- Applied to location data:
  - User's position is generalized
  - Shared with at least k users

k-anonymity

Re-Identification Example

- William Weld – a governor of Massachusetts (1991–1997)
- Link Medical Record and Voter List datasets together
  → **re-identify** a William Weld among the medical records

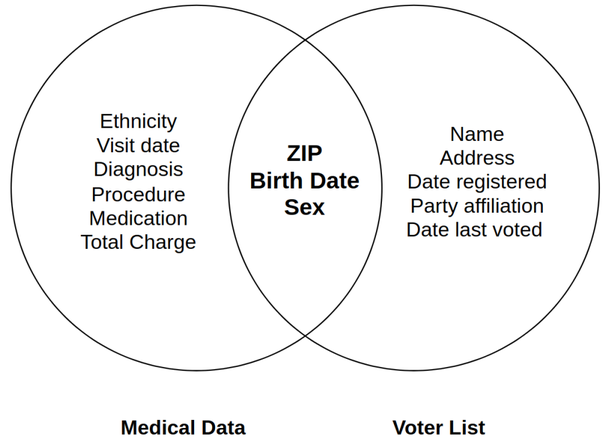


Analysis and Commentary: "Re-identification" of Governor William Weld

***The "Re-identification" of Governor William Weld's Medical Information:** A Critical Re-examination of Health Data Identification Risks and Privacy Protections, Then and Now*

Author: Daniel C. Barth-Jones, M.P.H., Ph.D., Assistant Professor of Clinical Epidemiology, Department of Epidemiology, Mailman School of Public Health, Columbia University.

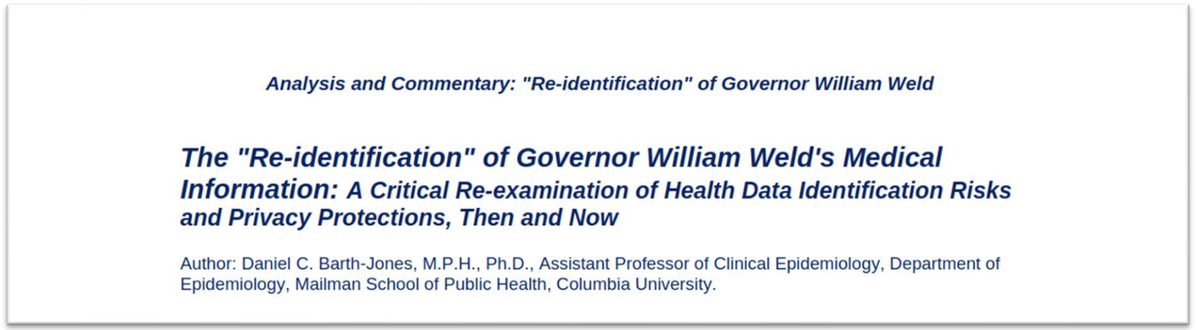
https://fpf.org/wp-content/uploads/2025/05/DBJ_Weld_Re-Identification.pdf
https://biographytree.com/biography/william-weld-former-governor-of-massachusetts/



k-anonymity

Re-Identification Example

- William Weld – a governor of Massachusetts (1991–1997)
- Link Medical Record and Voter List datasets together
  → **re-identify** a William Weld among the medical records
- Link based on **quasi-identifiers**

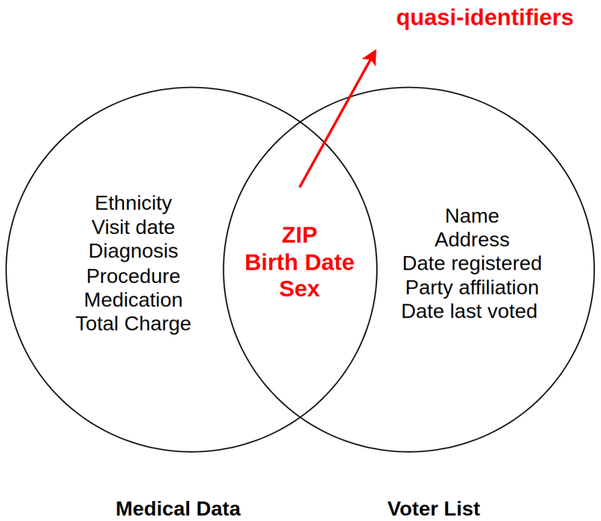


Analysis and Commentary: "Re-identification" of Governor William Weld

The "Re-identification" of Governor William Weld's Medical Information: A Critical Re-examination of Health Data Identification Risks and Privacy Protections, Then and Now

Author: Daniel C. Barth-Jones, M.P.H., Ph.D., Assistant Professor of Clinical Epidemiology, Department of Epidemiology, Mailman School of Public Health, Columbia University.

https://fpf.org/wp-content/uploads/2025/05/DBJ_Weld_Re-Identification.pdf



k-anonymity

k-anonymous Dataset Example

- Quasi-Identifiers = {Race, Birth, Gender, ZIP}
- Entries t1 and t2 are **k = 2 indistinguishable** from each other

	Race	Birth	Gender	ZIP	Problem
t1	Black	1965	m	0214*	short breath
t2	Black	1965	m	0214*	chest pain
t3	Black	1965	f	0213*	hypertension
t4	Black	1965	f	0213*	hypertension
t5	Black	1964	f	0213*	obesity
t6	Black	1964	f	0213*	chest pain
t7	White	1964	m	0213*	chest pain
t8	White	1964	m	0213*	obesity
t9	White	1964	m	0213*	short breath
t10	White	1967	m	0213*	chest pain
t11	White	1967	m	0213*	chest pain

https://dataprivacylab.org/projects/kanonymity/kanonymity.pdf

Obfuscation–based Solution

Spatial and Temporal Cloaking

Apply k-anonymity to Location Data

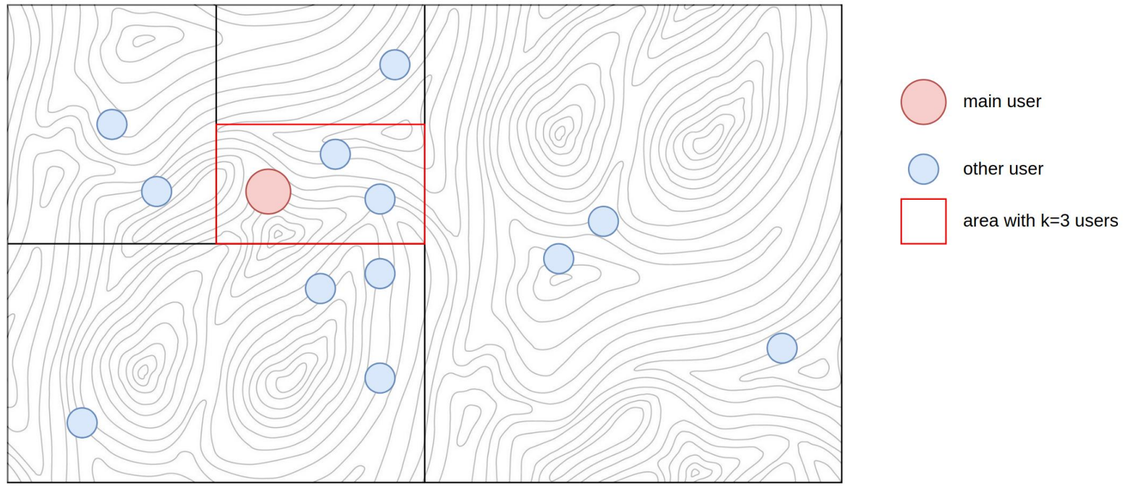
Spatial Cloaking

- Extend k-anonymity to location data
- Exact user location → generalized **spatial** or **temporal** region
- At least **k users** share same spatial/temporal region

Apply k-anonymity to Location Data

Spatial Cloaking

- Reduce **location** precision
- Find an area that contains at least **k users**
- Output a region instead of exact point



Apply k-anonymity to Location Data

Temporal Cloaking

- Reduce **time** precision
- Find a time interval that contains at least **k users** queries
- Output a **time interval** instead of exact time value

Temporal and Spatial Cloaking

Limitations

- In **dense user areas**, cloaking algorithms perform well
  → Sufficient nearby users allow effective anonymization while preserving data utility

- In **sparse user areas**, precision must be reduced significantly
  → Cloaking enlarges the region too much, making the data less useful

- **Key challenge:**
  Balancing privacy (via spatial/temporal cloaking) with **data usability**

Agenda

1. Location Data Sensitivity
2. Location-based Service Architecture Assumptions
3. PET Solutions:
    3.1. Obfuscation-based Solutions:
        - Spatial and Temporal Cloaking
        - Dummy Locations
        - Differential Privacy
    3.2 Cryptography-based Solutions
    3.3. Policy-based Solutions
4. Conclusion

Obfuscation-based Solution

Dummy Location

Dummy Locations

- Hide real location among multiple locations
- Real + fake locations sent together
- Attacker cannot distinguish

Real life situation

Why this works

- Creates uncertainty
- Increases location entropy
- Harder to identify real position

Problem: Naive dummy locations fail

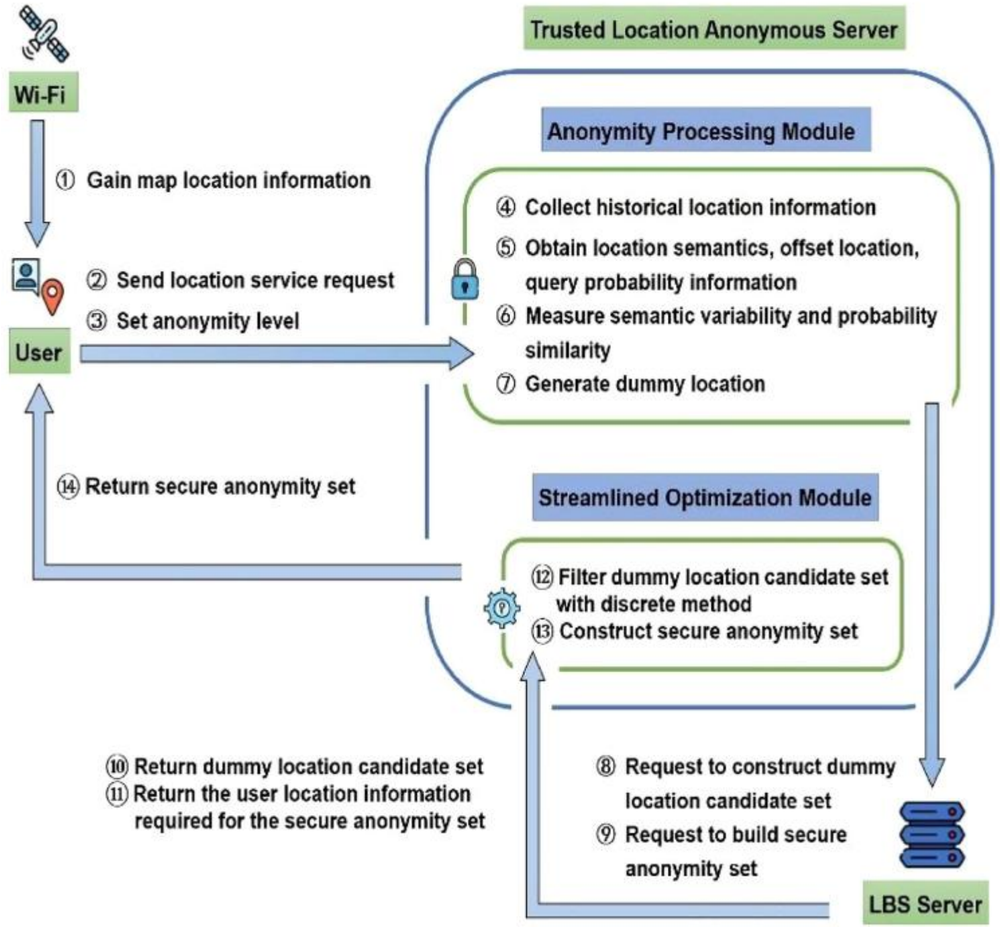
- Background knowledge attacks
- Location distribution attacks
- Probability filtering
- Semantic similarity attacks

Dummy location interference privacy protections(DLIP): A smarter Privacy Approach

How DLIP Works

- Semantic filtering (WordNet)
- Offset location
- Query probability matching
- Spatial dispersion
- K- anonymity set

System Architecture



Source: Zhang & Li, 2022 (DLIP paper)

Practical Problems

- Requires trusted middleman
- Higher computation & communication cost
- Increased battery and data usage
- Still vulnerable to advanced attacks

BREAK

Agenda

1. Location Data Sensitivity
2. Location-based Service Architecture Assumptions
3. PET Solutions:
    3.1. Obfuscation-based Solutions:
        - Spatial and Temporal Cloaking
        - Dummy Locations
        - Differential Privacy
    3.2 Cryptography-based Solutions
    3.3. Policy-based Solutions
4. Conclusion

Obfuscation–based Solution

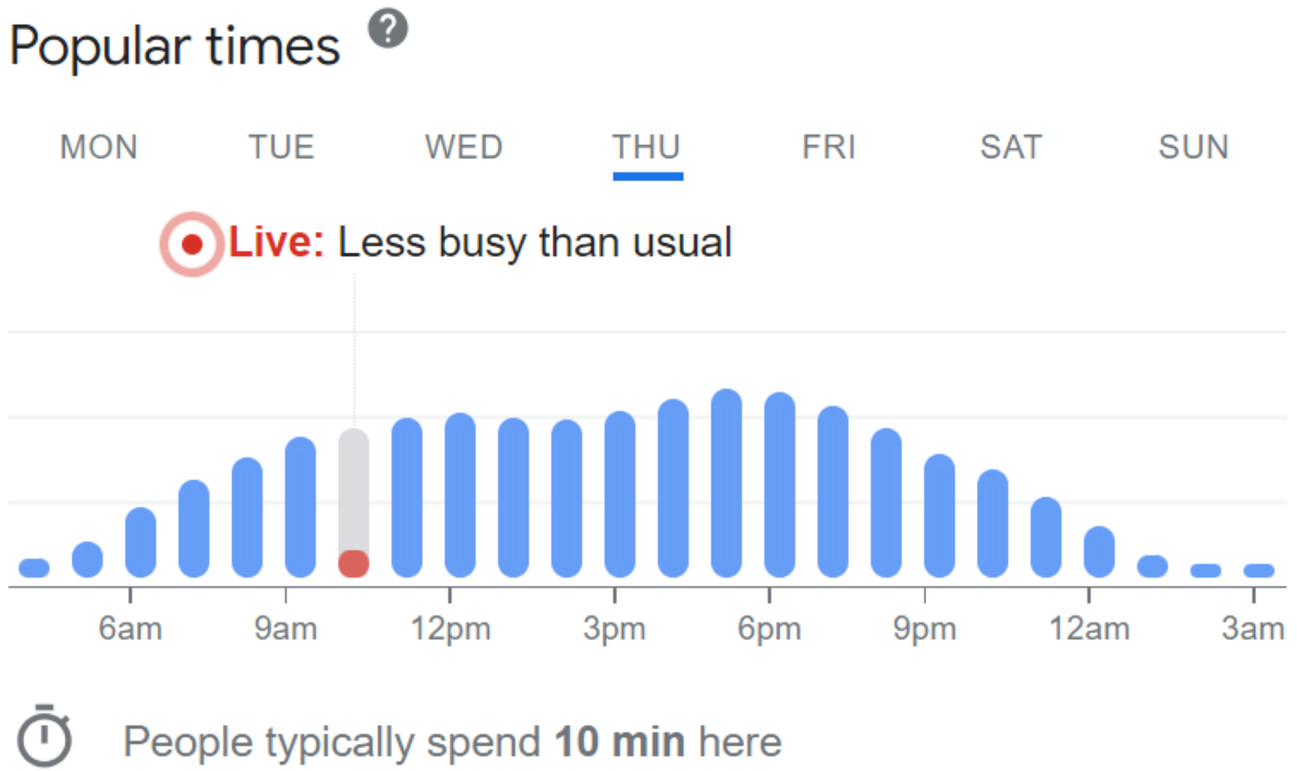
Differential Privacy

Differential Privacy (DP)

No significant difference in the query if you are in the dataset or not.

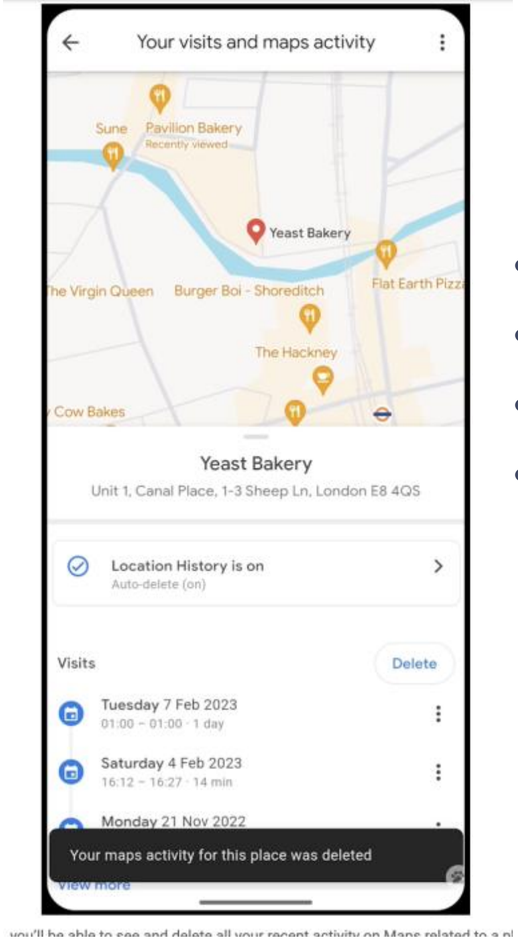
Add noise to ‘obfuscate’ after the result

Google Maps Popular Times



https://www.google.com/url?sa=t&source=web&rct=j&url=https%3A%2F%2Fwww.sterlingsky.ca%2Fpopular-times-wait-times-visit-duration%2F&ved=0CBYQjRxqFwoTCMCzkuCT4JMDFQAAAAAdAAAAABAI&opi=89978449

DP in Practice



- Collection from google timeline
- Data will be anonymized
- Combine it all into one large dataset
- Add noise on top of the dataset

https://helios-i.mashable.com/imagery/articles/01JIuJtgY2OSivnxvzHFrij/hero-image.fill.size_1248x702.v1717611852.png

DP Definition

ε – Differential privacy: $\Pr(M(x) \in S) \leq e^{\varepsilon}\Pr(M(x') \in S)$
(ε, δ) – Differential privacy: $\Pr(M(x) \in S) \leq e^{\varepsilon}\Pr(M(x') \in S) + \delta$

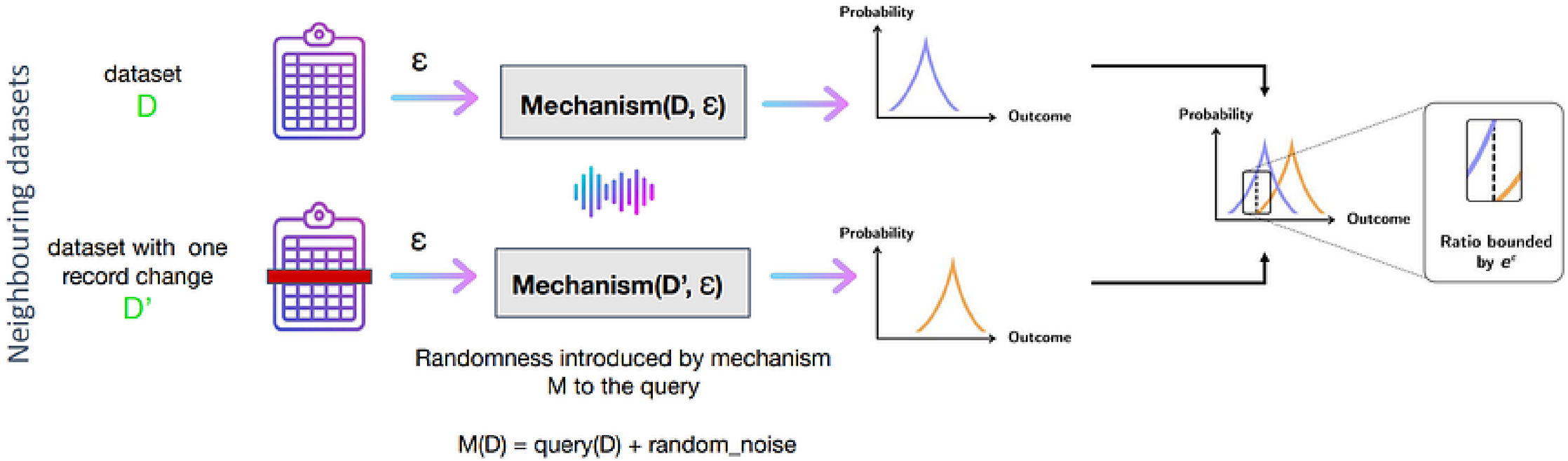
- M → query
- x & x′ → neighboring datasets
- ε → noise
- δ → probability of failure

DP Noise

- Laplace noise
  - Laplace Distribution
- Gaussian noise
  - Gaussian Distribution (Normal distribution)

Laplace Noise

ε – Differential privacy: $\Pr(M(x) \in S \leq e^{\varepsilon}\Pr(M(x') \in S)$

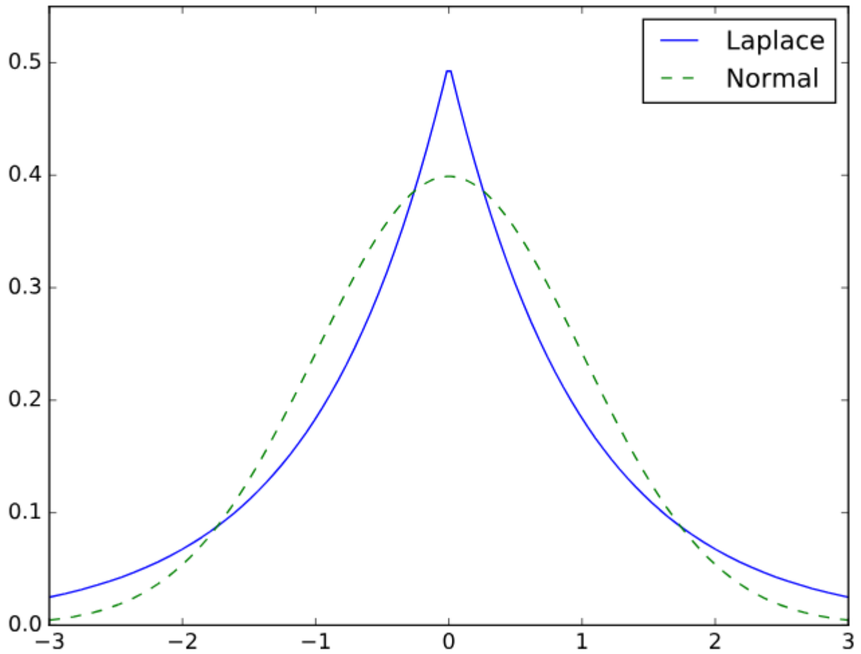


https://medium.com/dsaid-govtech/protecting-your-data-privacy-with-differential-privacy-an-introduction-abee1d7fcb63

Gaussian Noise

(ε, δ) – Differential privacy: $\Pr(M(x) \in S \leq e^{\varepsilon}\Pr(M(x') \in S) + \delta$

- Difference in delta
- Easier to compute



https://www.johndcook.com/normal_laplace.svg

Noise Comparison

- https://lpanavas.github.io/mechanism-comparison/

Differential Privacy – Conclude

- For location data ‘geo-indistinguishability’
  - Close locations should provide similar outputs
  - Instead of databases we use two locations
- Works well for queries
- Can work on both continuous and snapshot data
- Continuous data is harder since it needs more noise
- Can work both third-party & third-party free

Agenda

1. Location Data Sensitivity
2. Location-based Service Architecture Assumptions
3. PET Solutions:
    - 3.1. Obfuscation-based Solutions
    - 3.2 Cryptography-based Solutions:
        - Space Transformation
    - 3.3. Policy-based Solutions
4. Conclusion

Cryptography–based Solution

Space Transformation

Cryptography

Space Transformation

Query the location-based service without revealing the user's exact coordinates

Core concept:

- Transform the coordinate space
- Convert location coordinates from **2D** space to **1D** space

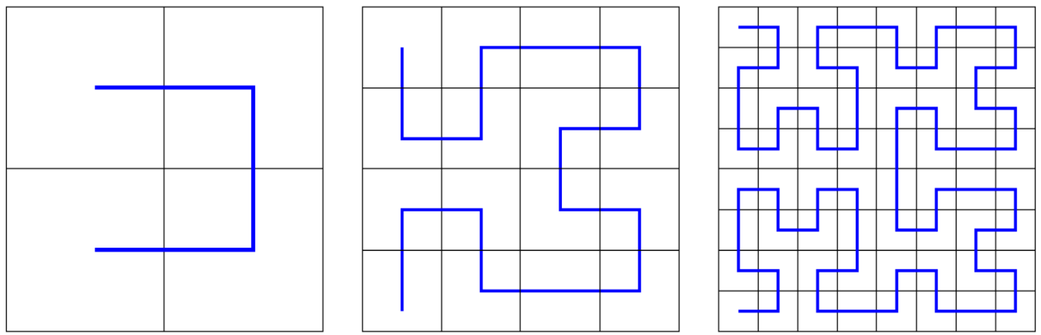
To do so, use:

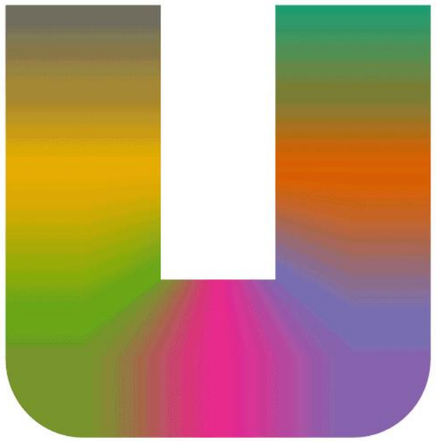
- Space-filling curves such as **Hilbert Curves**
- Cryptographic **hash** to make the transformation **irreversible**

Space Transformation

Hilbert Curves

- Type of **space-filling curve**
- Maps a point 2D → 1D
- Preserves **spatial locality**
- Nearby points → similar values



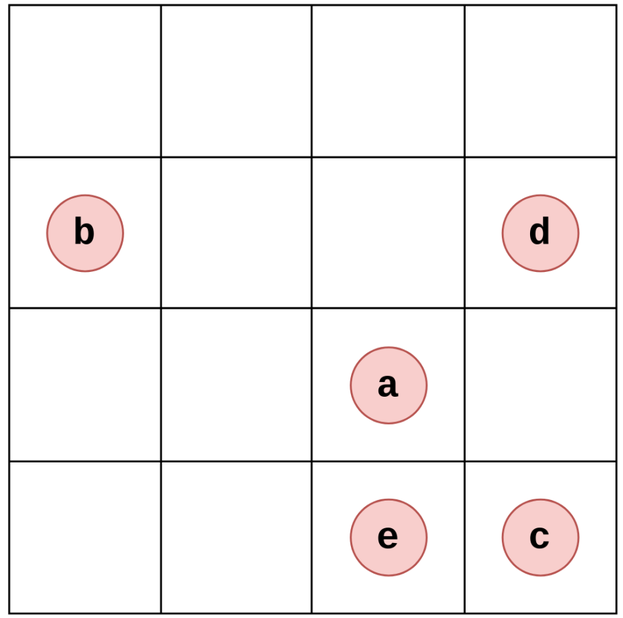


https://en.wikipedia.org/wiki/Hilbert_curve

Space Transformation

Hilbert Curves Example

- **Original** 2D grid space with 5 points
- Each point represents a location coordinate
- No transformation applied yet

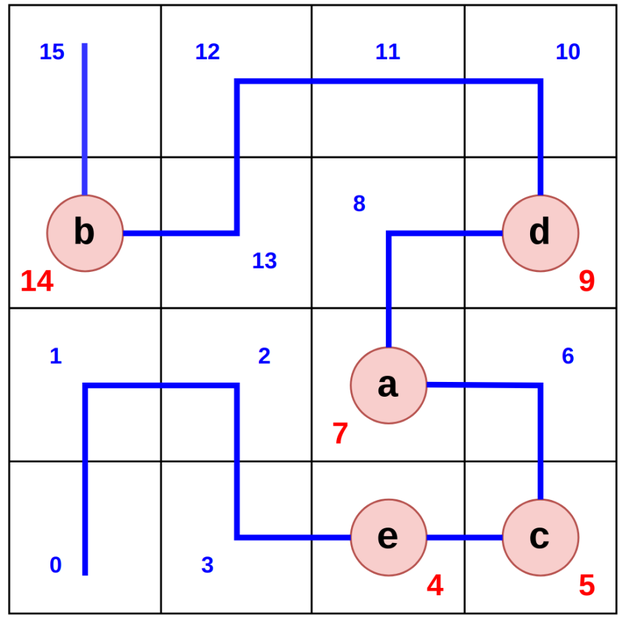


Space Transformation

Hilbert Curves Example

- Apply **2nd order Hilbert Curve for 2D space**
  onto the original grid space
- The Hilbert curve maps
  each **2D coordinate** to a unique **1D Hilbert value**

- Observe:
  - points that are **geographically close**
    (e.g., points *a* and *d*) have **similar** Hilbert values
  - points that are **far apart**
    (e.g., points *a* and *b*) have **very different** Hilbert values



Cryptography

Space Transformation

Space transformation consists of two phases:

- **Offline phase** (service side)
- **Online phase** (user query)

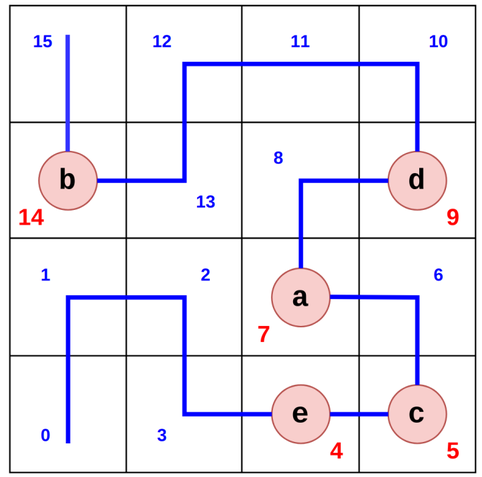
Space Transformation

Offline Phase

Server side:

- Convert object location coordinates: 2D coordinates → **Hilbert values**
- Apply a **cryptographic hash** function
- Build an **encoded lookup table** (ELT) for each object o:

$$ELT = < \Phi(H(o.x, o.y)), \varepsilon_k(o.t), \textit{prev}, \textit{next} >$$



Space Transformation

Online Phase

User query:

- Convert user location: 2D coordinates → **Hilbert values**
- Apply cryptographic hash function
- Send transformed query to the location-based service

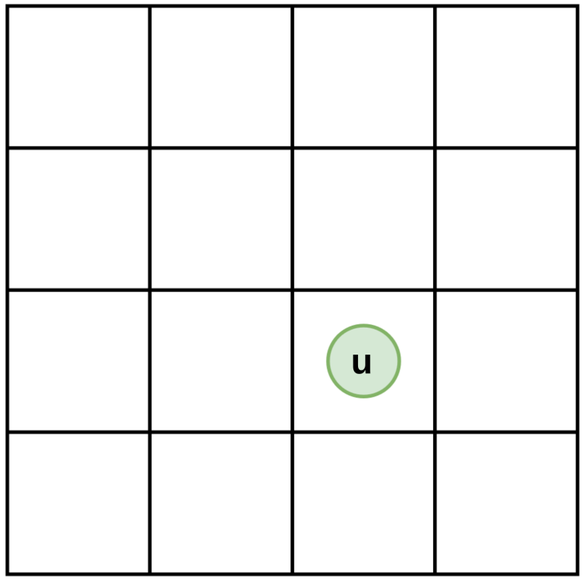
Space Transformation

User Query Process

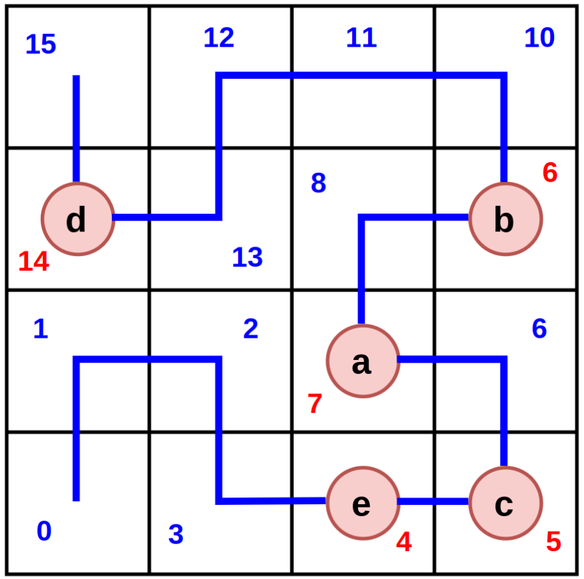
Server-side process:

- Receives the **hashed Hilbert-transformed** user location
- Identifies the **matching** entry in the encoded lookup table
- Explores adjacent entries using stored pointers
- Returns the ***k* nearest neighbors (KNN)**

Space Transformation
Initial State

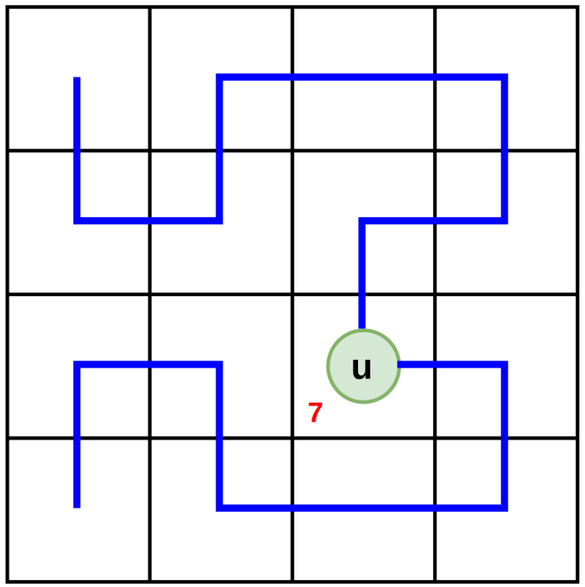


Original Space with a user

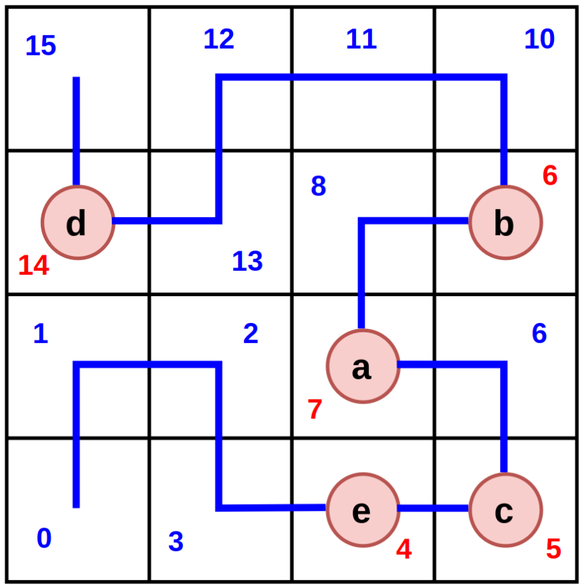


ELT table with encoded objects

Space Transformation
Transform Original Space

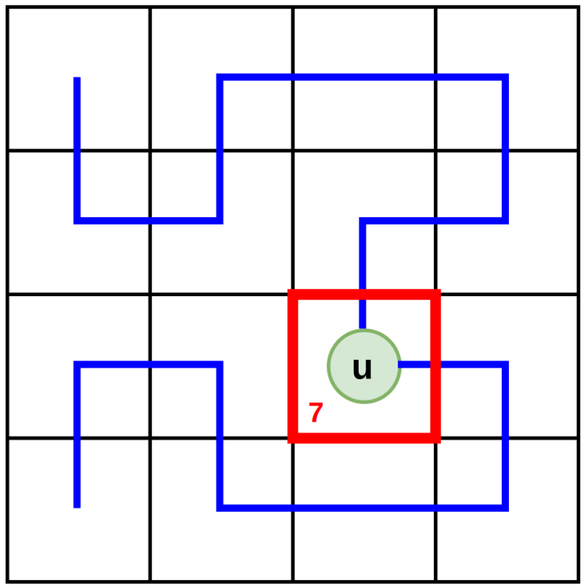


Original Space with a user

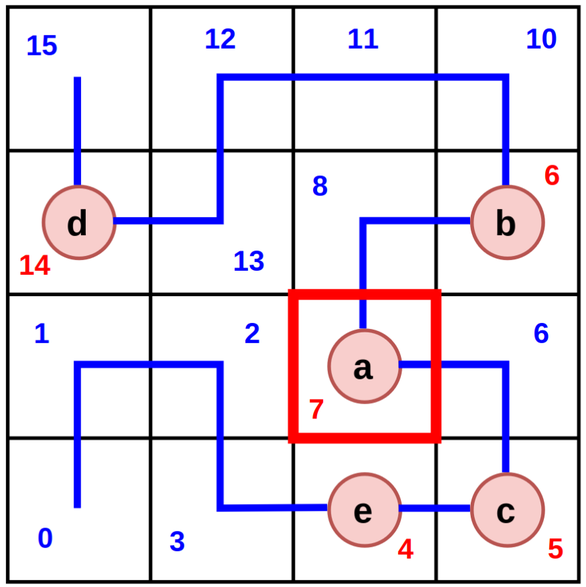


ELT table with encoded objects

Space Transformation
Match the Hash

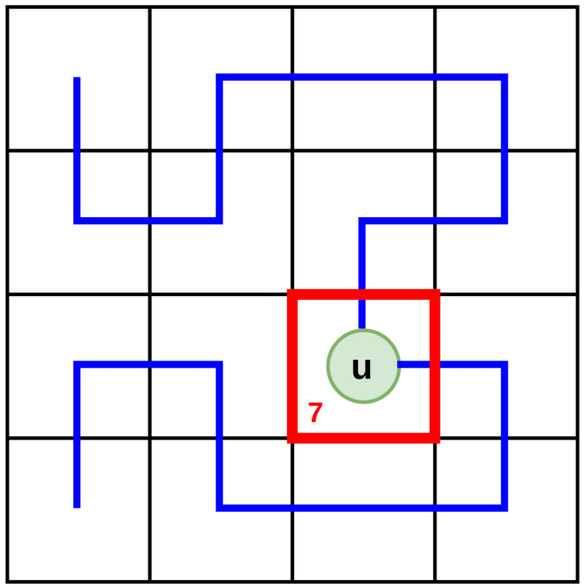


Transformed Space of a user

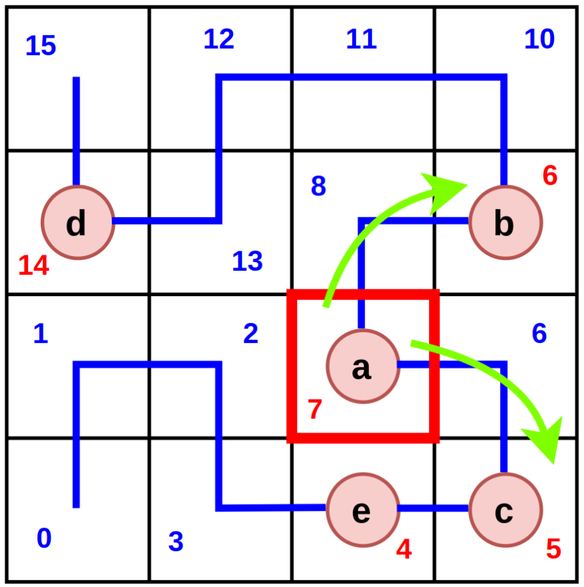


ELT table with encoded objects

Space Transformation
Traverse Neighbour Objects (KNN)

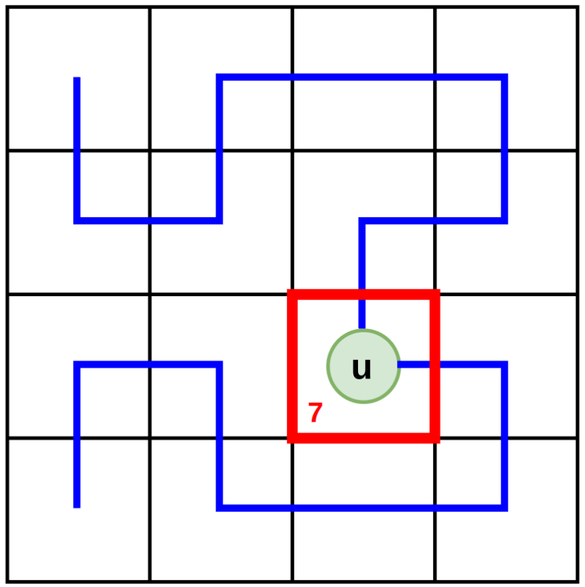


Transformed Space of a user

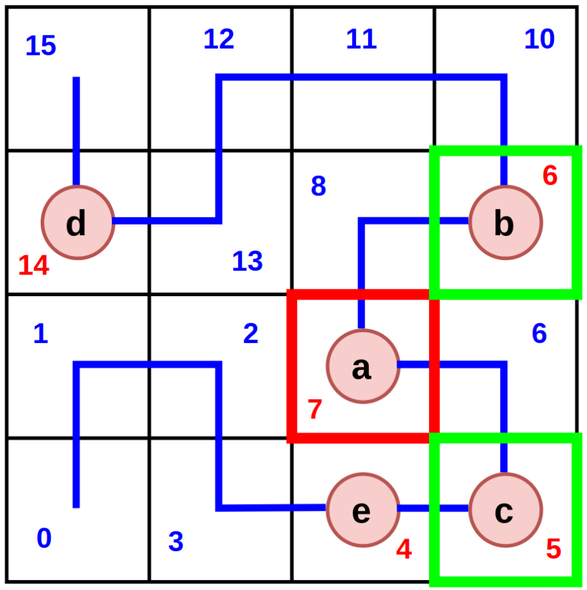


ELT table with encoded objects

Space Transformation
Traverse Neighbour Objects (KNN)

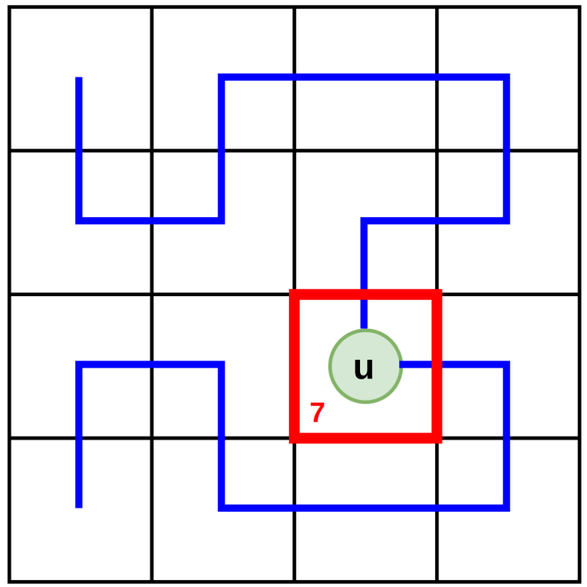


Transformed Space of a user

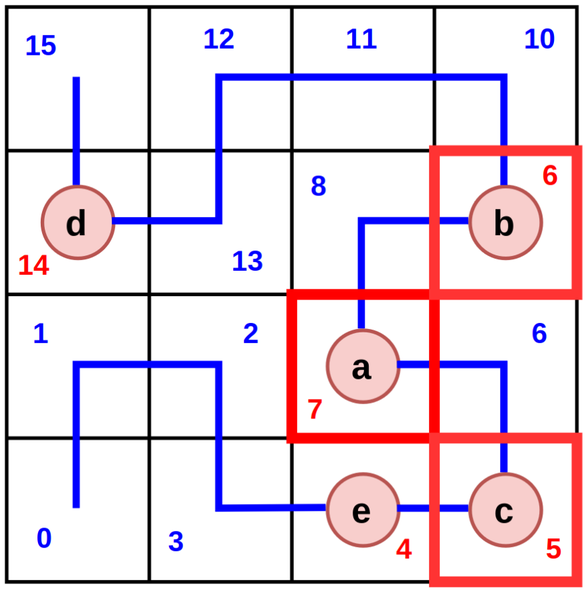


ELT table with encoded objects

Space Transformation
Retrieve Neighbour Objects



Transformed Space of a user



ELT table with encoded objects

k=2

Space Transformation

Limitations

- Dependence on a third party
  → Required for generating and managing encoding/lookup tables
- High computational cost
  → Transformations are resource-intensive and slow to compute
- Reliance on public key infrastructure (PKI)
  → Requires secure key management and added system complexity

Agenda

1. Location Data Sensitivity
2. Location-based Service Architecture Assumptions
3. PET Solutions:
    3.1. Obfuscation-based Solutions
    3.2 Cryptography-based Solutions
    3.3. Policy-based Solutions:
        - GDPR
4. Conclusion

Policy-based Solution

GDPR

What GDPR Does

- Article 4(1): Personal Data Definition
- Identified or identifiable natural person
- Includes indirect identification

Location Data = Personal Data

- Location data enables identification
- Acts as a quasi-identifier
- Enables inference of behavior

Yola Georgiadou (2019)

GDPR Principles

- Article 5: Data Principles
- Data minimization
- Purpose limitation
- Storage limitation

- Article 6: Legal Basis for Processing

- Article 7: Consent

- Article 25: Privacy by Design

GDPR (from Law to System Design)

Three Core Requirements

- Notice → Transparency about data usage
- Consent → User permission & control
- Control → User rights over data

Key Challenge

- No clear technical guidelines
- Complex for developers
- Trade-off: usability vs privacy

Important Insight

- Location data is highly sensitive – Reveals behavior, routines, identity

Mehrnaz Ataei (2018)

Location Data in Different Legal Contexts

- GDPR: strict data protection framework
- Other systems vary in protection
- Example: US (sector-based laws)
- Location data used in investigations
- Sensitive implications (e.g., clinic visits)

GDPR in Practice

Dutch SA imposes a fine of 290 million euro on Uber because of transfers of drivers' data to the US

26 August 2024

France Netherlands Belgium Austria Croatia Czech Republic Denmark Finland Sweden Estonia Spain Germany Greece Hungary Ireland Italy Malta Poland Portugal Romania Slovakia Norway

Background information

- Date of final decision: 22 July 2024
- Cross-border case
  One-Stop-Shop Procedure: the decision was taken by national supervisory authorities following the One-Stop-Shop cooperation procedure (OSS).
- Netherlands
- and CSAs: All SAs, except for Bulgaria, Cyprus, Iceland, Latvia, Liechtenstein, Luxembourg and Slovenia.
- Legal Reference(s): Article 44 (General principle for transfers), Article 46 (Transfers by way of appropriate safeguards), Article 49 (Derogations for specific situations)
- Decision: Administrative fine
- Key words: Administrative fine, International transfer, Third party access to personal data

Latest news

EDPB conference on cross-regulatory cooperation: what we learned

24 March 2026 EDPB

EDPB and EDPS support strengthening EU's cybersecurity and easing compliance while protecting individuals' personal data

19 March 2026 EDPB EDPS

- €290 million fine by Dutch DPA
- Data transfer outside EU (to US)
- Included driver location data
- Insufficient data protection safeguards

https://www.edpb.europa.eu/news/news/2024/dutch-sa-imposes-fine-290-million-euro-uber-because-transfers-drivers-data-us_en

GDPR is Necessary but It is Not Sufficient

Technical Gap

- GDPR ≠ technical protection
- Technology evolves rapidly (AI, data analytics)
- Regulation takes time to adapt
- Gap between law and real-world capability

Agenda

1. Location Data Sensitivity
2. Location-based Service Architecture Assumptions
3. PET Solutions:
    3.1. Obfuscation-based Solutions
    3.2 Cryptography-based Solutions:
    3.3. Policy-based Solutions
4. Conclusion

Recap

- We went over obfuscation, cryptography and policy-based solutions
- For obfuscation we have k-anonymity, cloaking, dummy locations & differential privacy
- Cryptography we have space transformation
- Policy, we have GDPR
- We saw that everyone of them has their own use case and limitations. Often, we see combinations.

Our Opinion

- Sensitivity of location data often underestimated
- Need more laws and enforcements, for example in the app stores.

References

1. Gruteser, M., & Grunwald, D. (2003) – https://doi.org/10.1145/1066116.1189037
2. Sweeney, L. (2002) – https://doi.org/10.1142/S0218488502001648
3. de Montjoye, Y.-A., Hidalgo, C. A., Verleysen, M., & Blondel, V. D. (2013) – https://www.nature.com/articles/srep01376
4. Georgiadou, Y., de By, R. A., & Kounadi, O. (2019) – https://doi.org/10.3390/ijgi8030157
5. Zhang, A., & Li, X. (2022) – https://doi.org/10.1177/15501329221125111
6. Zhang, S., Li, M., Liang, W., Sandor, V. K. A., & Li, X. (2022) – https://doi.org/10.3390/s22166141
7. Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006) – https://doi.org/10.29012/jpc.v7i3.405
8. Dong, J., Roth, A., & Su, W.J. (2019) – https://doi.org/10.48550/arXiv.1905.02383

THANK YOU!