Protecting Sensory Data against Sensitive Inferences

Mohammad Malekzadeh, Second Year PhD Student

Student of *Queen Mary University of London* & RA at Imperial College London

Supervisors: Prof. Andrea Cavallaro (QMUL) and Dr. Hamed Haddadi (ICL)

Research Area:

- **Machine Learning for Sensor Data Analysis**
- **Privacy-Preserving Methods for Users of Edge Devices**

Queen Mary University of London

CIS centre for intelligent sensing

Imperial College London

Problem, Motivation, Solution, Implication

- How to protect ***user's privacy*** without losing ***apps' utility***

  - More Devices/Sensors + More Cloud Services: **More Data**.
  - Less control over personal data: **Less privacy**.

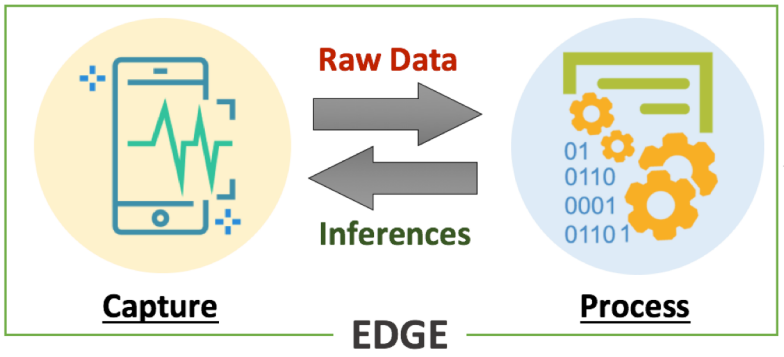
- Encoding raw data into the **feature set** at the **Edge** side,
- **Replacement / Elimination of** features corresponding to **sensitive inferences,**
- Sending a **privacy-preserving representation** of raw data the **Cloud** side to benefit from desired services.

- It enables **application-specific data release**
- It prevents **untrusted parties** to infer **sensitive information**
- It needs to establish a **utility-privacy-cost tradeoff**

Privacy-Utility Tradeoff

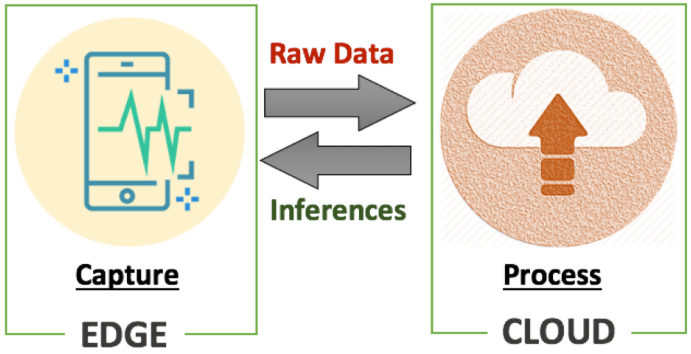
I. Edge-Based

Perfect Privacy vs. **Weak Utility**



II. Cloud-Based

Weak Privacy vs. **Perfect Utility**

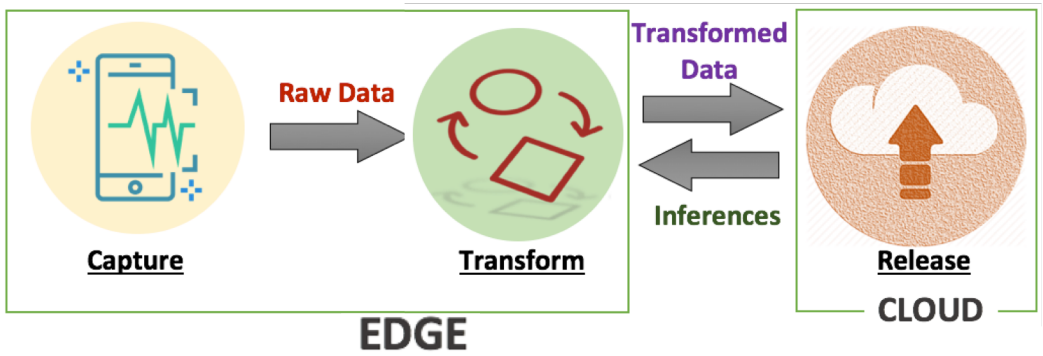


III. Hybrid

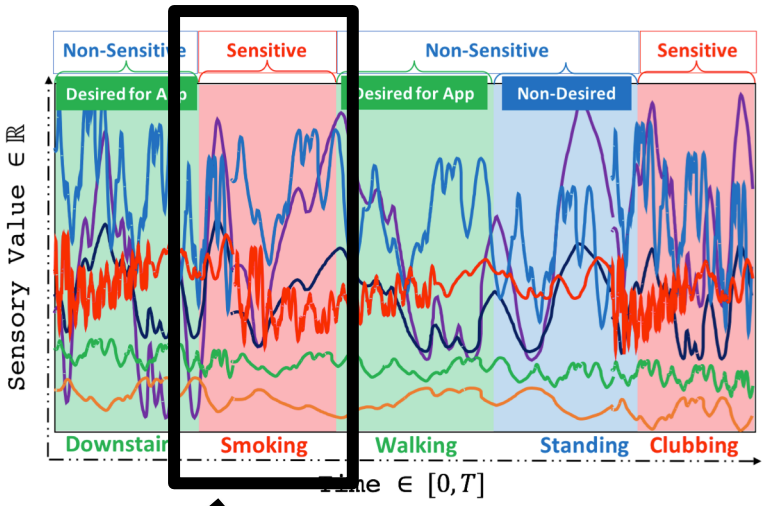
Good Privacy

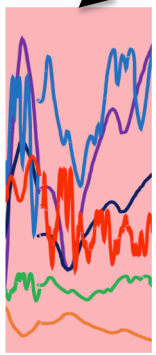
and

Good Utility

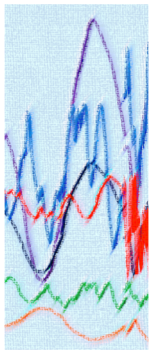


Replacement for Temporal Inferences





Replacement

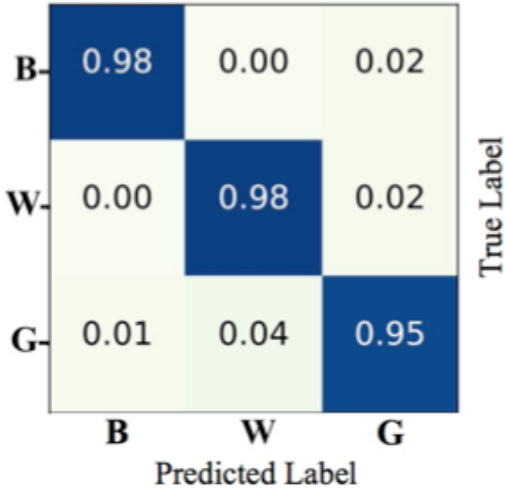


Original Data

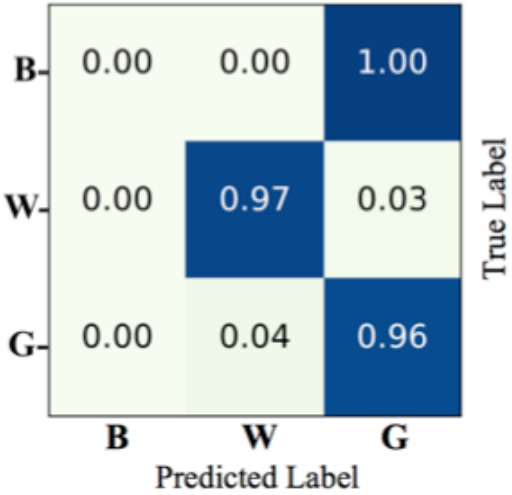
Transformed Data

Results on Skoda Dataset

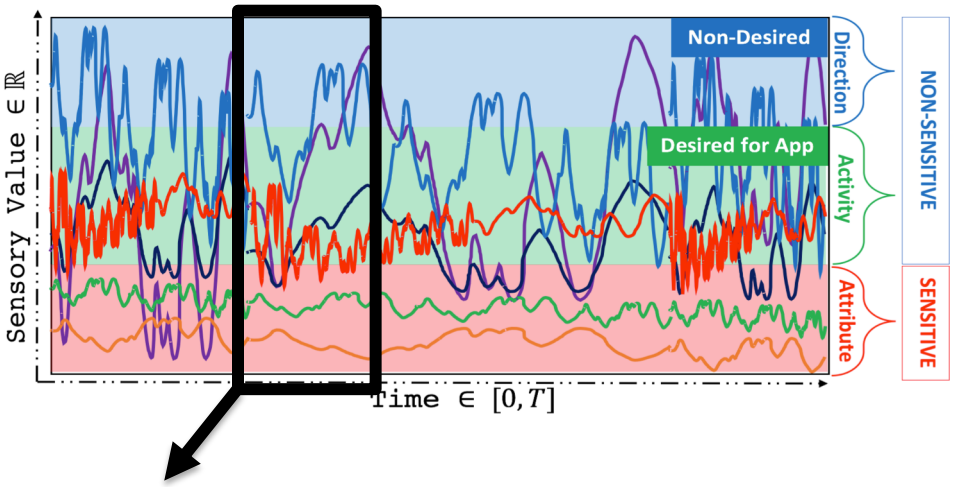
Original Data

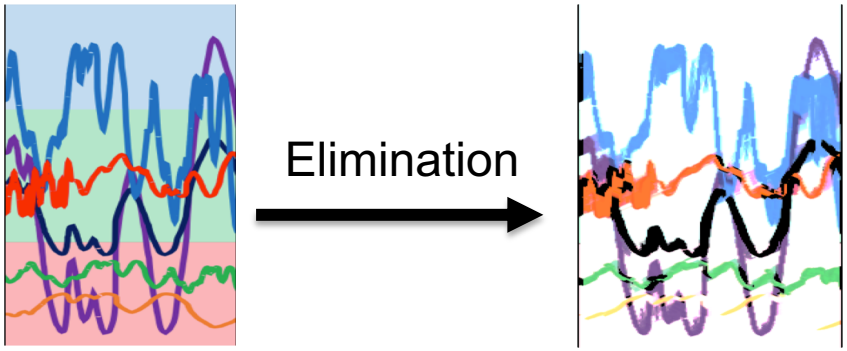


Transformed Data



Elimination for Concurrent Inferences





Original Data **Transformed Data**

	Original	Transformed
activity	95.08	93.71
gender	95.15	49.32

Results on MotionSense Dataset

	Original	Transformed
activity	94.31	90.46
gender	93.74	49.83

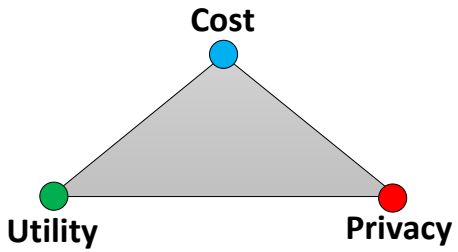
Results on MobiAct Dataset

Future Directions

- How we can provide a **statistical guarantee** (probabilistic bound) for sensitive information which can still be inferred from the transformed data?
  - Differential Privacy : **Composition Theorem?**
  - Mutual Information : **Joint Distributions?**

- Correlation among repeated measurement:
  - little by little information leakage

- The Complexity / Cost of the solution on Edge devices?



Thank You

Mohammad Malekzadeh

PhD Candidate in Computer Science

Link to the Repositories:

(Replacement): bit.ly/rep-dw18

(Elimination): bit.ly/eli-dw18

http://malekzadeh.uk/
github.com/mmalekzadeh
m.malekzadeh@qmul.ac.uk
@malekz4deh

Queen Mary University of London

CIS centre for intelligent sensing

Imperial College London