ACM multimedia
Seattle, USA Oct 12-16 2020

ACM Multimedia 2020

A tutorial on Deep Learning for Privacy in Multimedia

Part 3:
Deep Learning for Privacy and Utility Preserving Sensor Data Transformations

Mohammad Malekzadeh

PhD Student in CS at QMUL

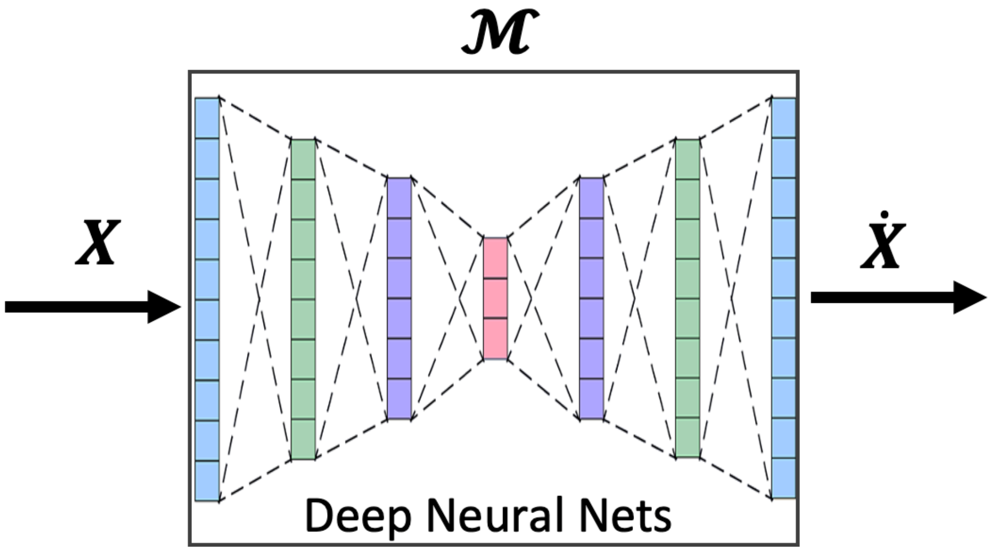
m.malekzadeh@qmul.ac.uk

CIS centre for intelligent sensing

Queen Mary University of London

Outline

1. Motivations (Mobile & Wearable Sensors)
2. Problem Definition (User’s Privacy & Data Utility)
3. How to Protect Users’ Sensitive
   - I. Activities
   - II. Attributes
   - III. Activities & Attributes
4. Conclusion and Open Questions
5. Q & A (Sharing the Code Examples)



1. Motivation

Mobile and Wearables

1970

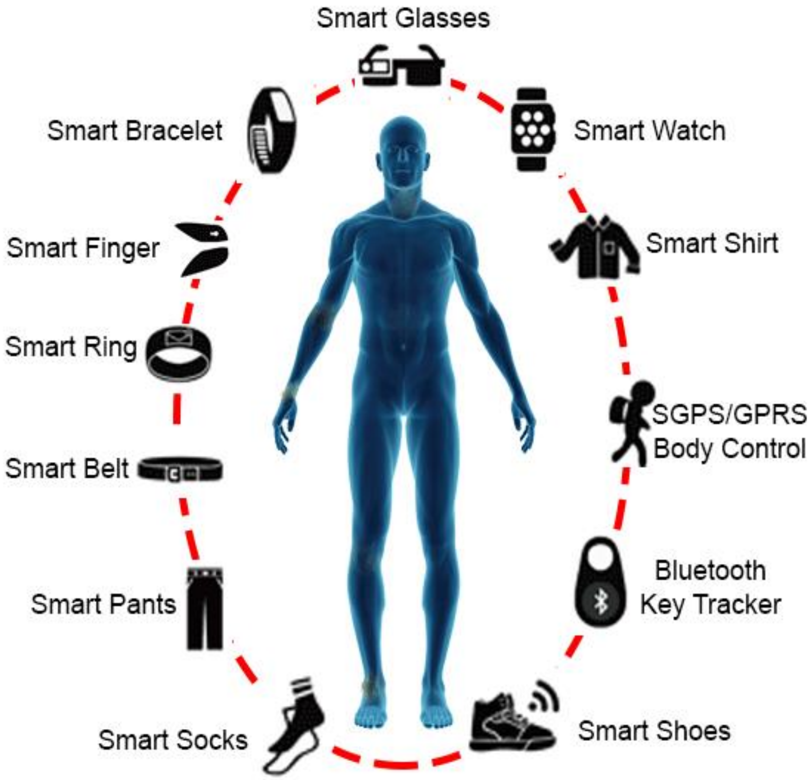
- BTS Location
- Microphone

2020

- GPS Location
- Microphone
- Accelerometer
- Gyroscope
- Magnetometer
- Barometer
- Thermometer
- Proximity
- Ambient Light
- Heart Rate
- …

Motion Sensors { Microphone, Accelerometer, Gyroscope

Several types of wearable technology*

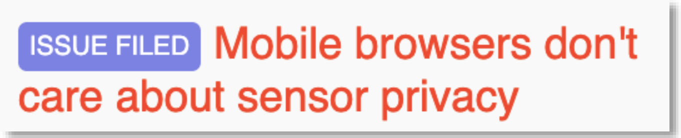


*Rodrigues, J. J., et. al. (2018). Enabling technologies for the internet of health things. *IEEE Access*, *6*, 13129-13141.

Applications and Threats

- Applications:
  - Health and Wellness,
  - Patient and Elderly Monitoring,
  - Gaming and VR, etc.

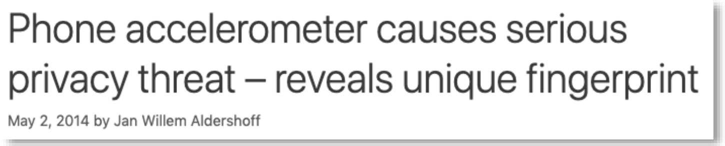
- Privacy Threats:
  - Revealing sensitive **activities:**
  - Leaking sensitive **attributes**
  - The **re-identification** of the user
  - Pin Code Inference, Targeted advertising, etc.



ISSUE FILED Mobile browsers don't care about sensor privacy

ANDY GREENBERG SECURITY 08.14.14 06:30 AM

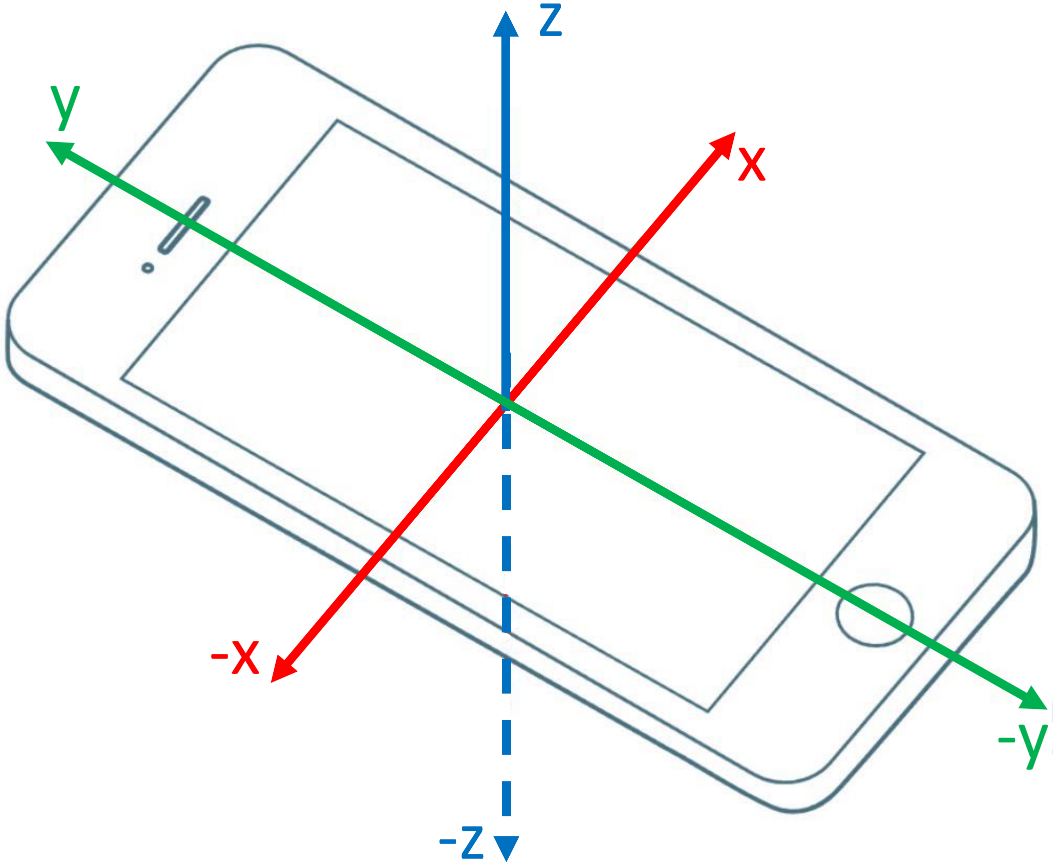
THE GYROSCOPES IN YOUR PHONE COULD LET APPS EAVESDROP ON CONVERSATIONS

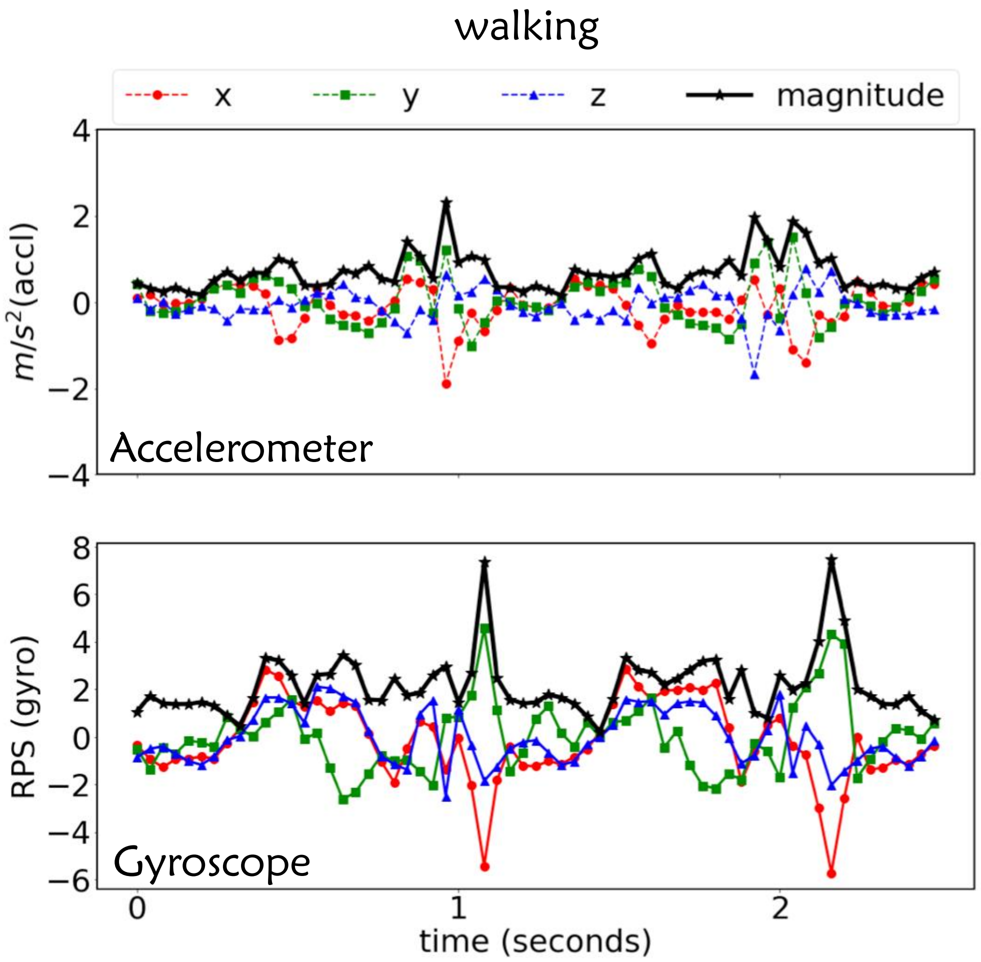


Phone accelerometer causes serious privacy threat – reveals unique fingerprint

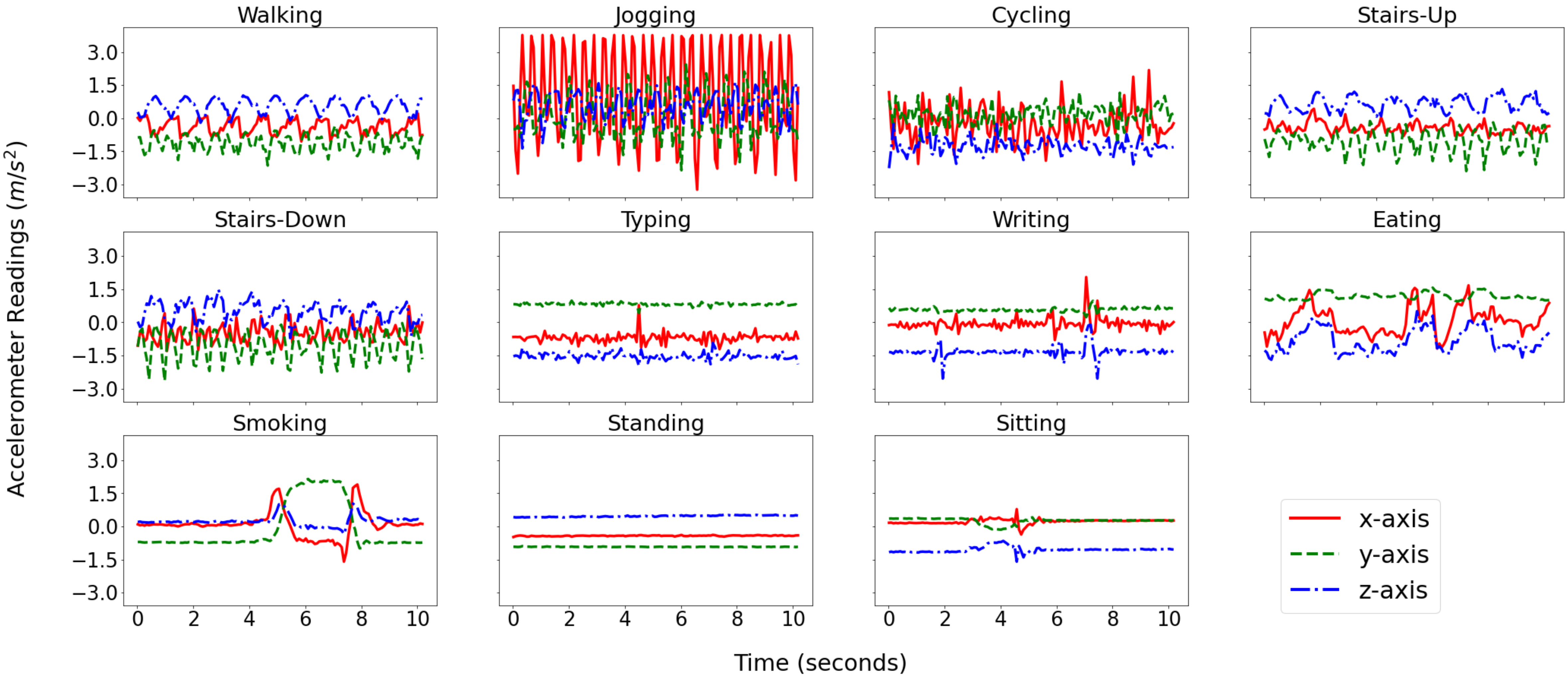
May 2, 2014 by Jan Willem Aldershoff

Motion Sensors





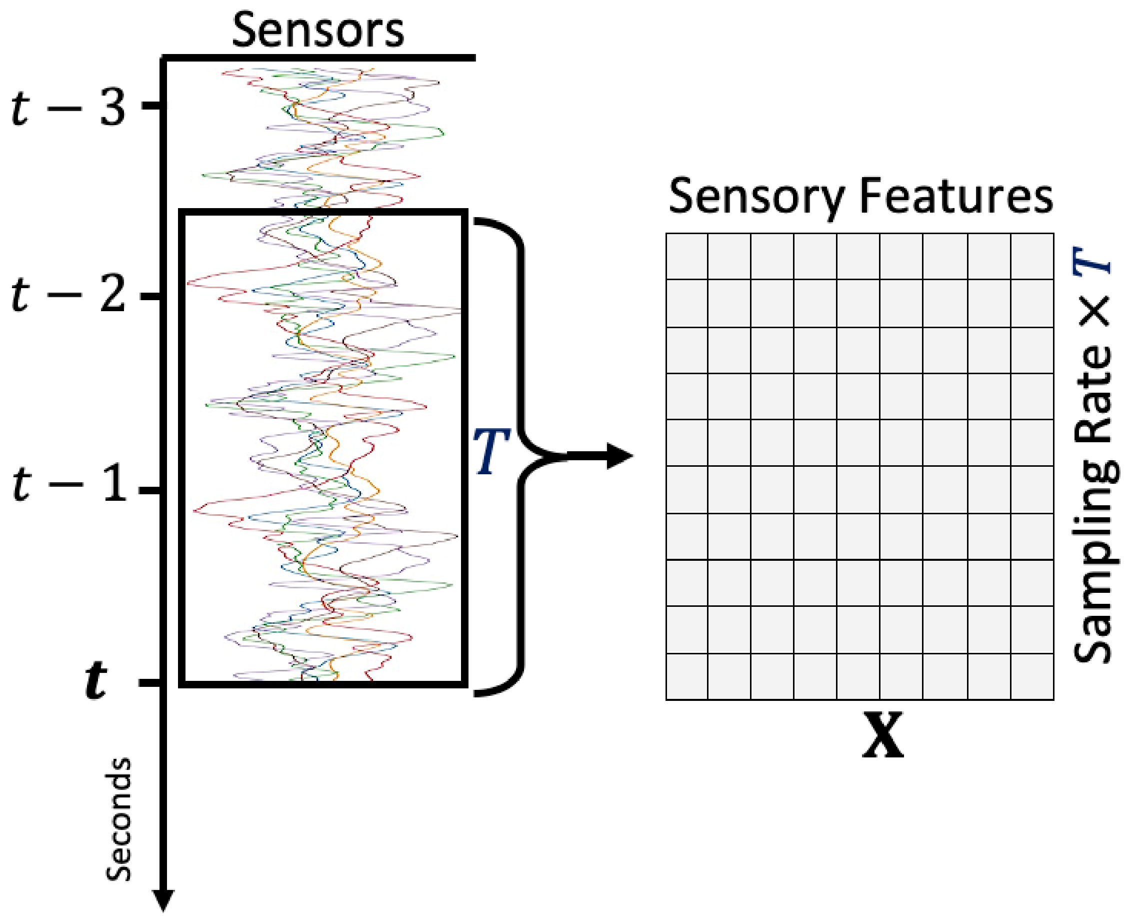
Activity Recognition



Data of a **smartwatch** worn on the right wrist of the user.

2. Problem Definition

Window-Based Classification



Example:

- 3 sensors
- 2 seconds’ time-window
- Stride length : ½ seconds
- Sampling rate: 50 Hz

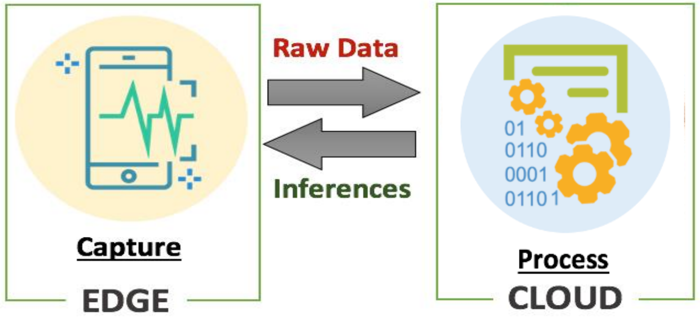
↓

Dimensions of X --> 100 x 9

Three Approaches to Classification

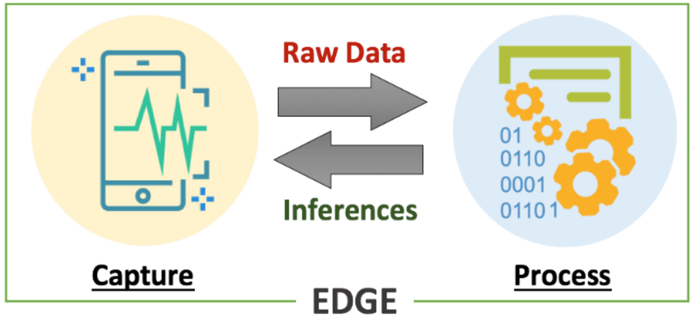
I. Cloud-Based

Weak Privacy but Perfect Utility



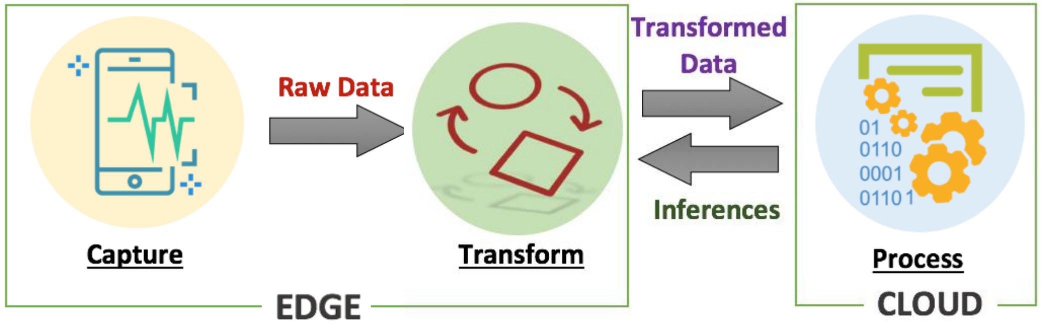
II. Edge-Based

Perfect Privacy but Weak Utility



III. Hybrid (edge and cloud)

Good Privacy and Good Utility



Privacy-Preserving Mechanisms

- Filtering
  - To avoid releasing the original data if it includes sensitive information.

- Noise Addition
  - independent or correlated noise to the original data.

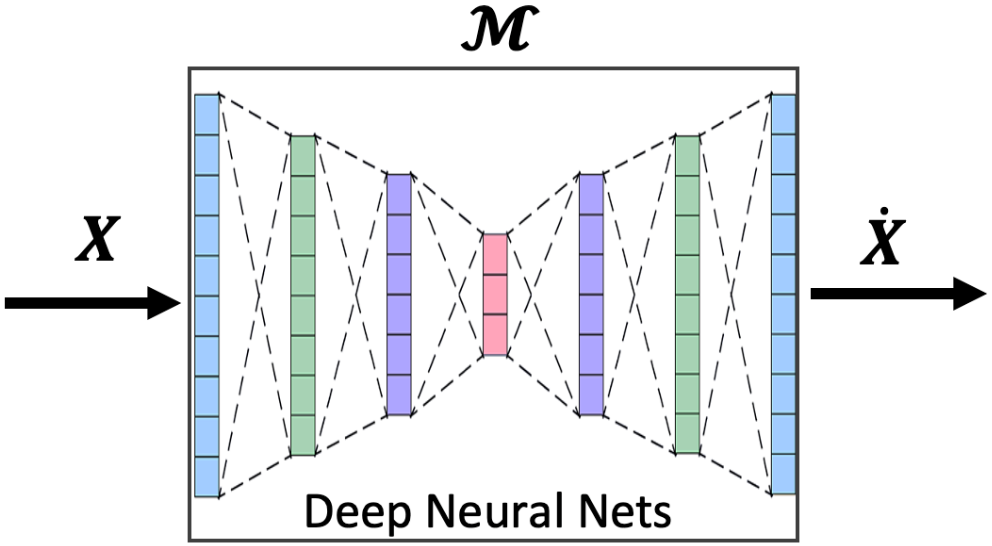
- Transformation ←
  - To generate a transformed version of the original data that:
    - is still informative about the required task.

    and

    - is invariant to the user’s sensitive attribute.

Utility and Privacy Preserving Data Transformation

- Considering:
  - $\boldsymbol{X}$**:** the user's data
  - $\boldsymbol{\mathcal{M}}$**:** the desired transformation mechanism



- The aim is to release $\dot{\boldsymbol{X}} = \boldsymbol{\mathcal{M}}(\boldsymbol{X})$ such that:

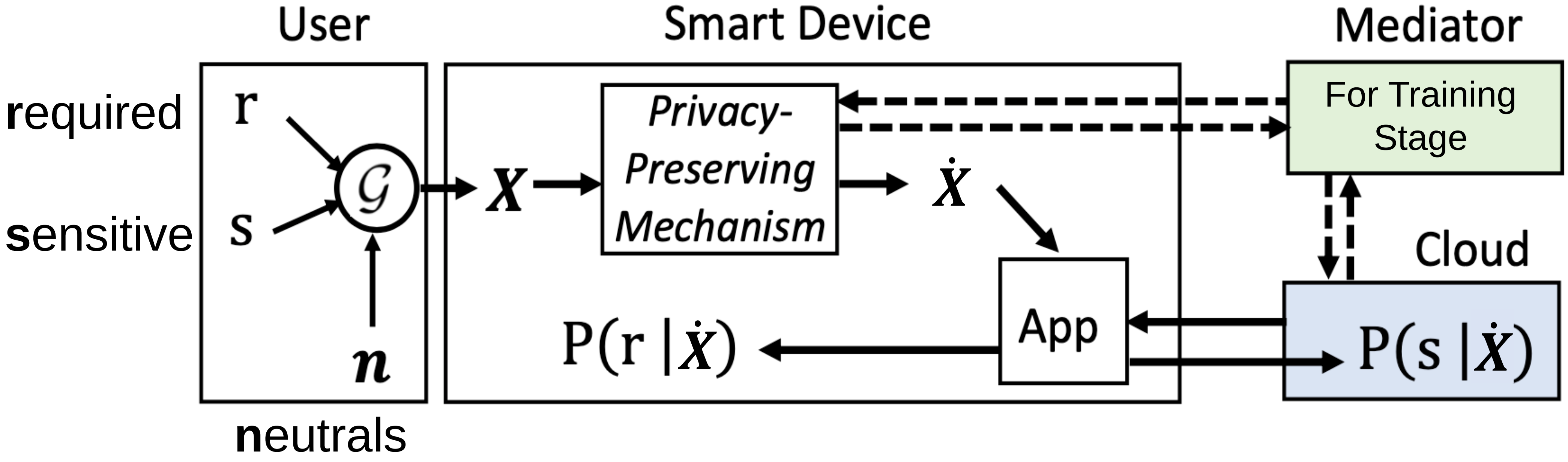
  - *Required* data, $\mathbf{r}$, can be inferred from $\dot{\boldsymbol{X}}$,

    as accurate as possible to what one could have inferred if we would have released $\boldsymbol{X}$.

  - No information about the *sensitive* data, $\mathbf{s}$, can be inferred from $\dot{\boldsymbol{X}}$,

    ideally, one cannot have a better guess than the random guess on the possible values

The Motivated Setting



Metrics

Utility: $\operatorname{argmax}_r P(r \mid \dot{X}) = \text{True } r\,?$

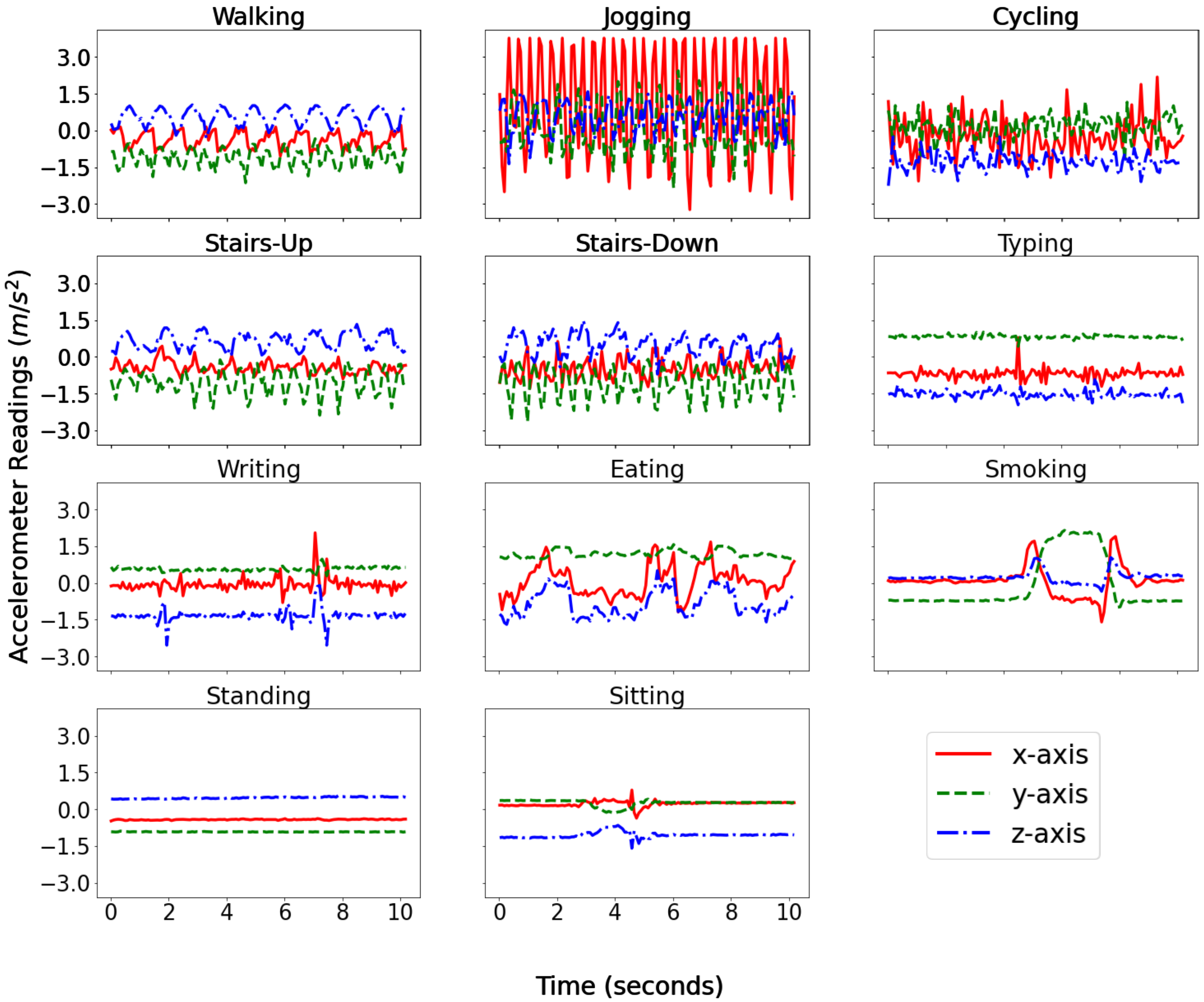
Privacy: $| P(s \mid \dot{X}) - \text{Random Guess on } s\,|\,?$

3.I. How to Protect User’s Sensitive **Activities**

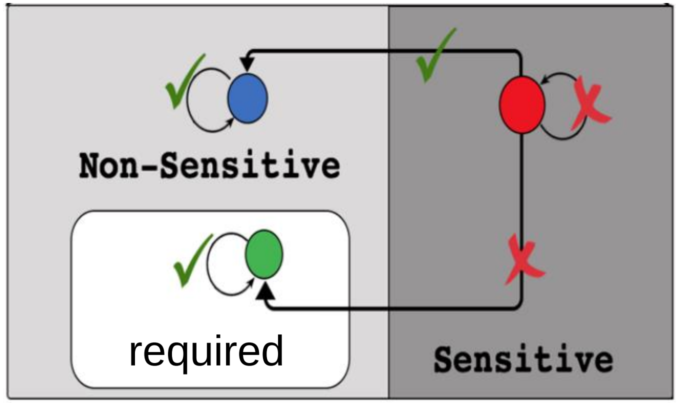
Three types of activities

As an Example:
a step counter application

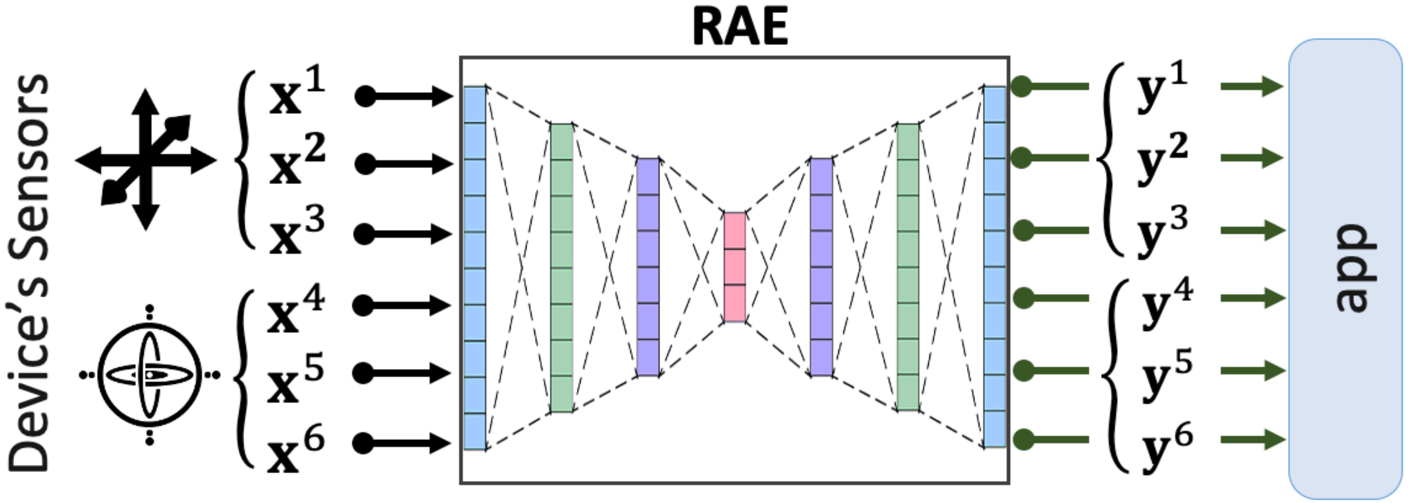
- **Required**: activities which the user gains utility from sharing with the app.
- **Sensitive**: activities which the user wish to keep private and should not be revealed to the app.
- **Neutral:** activities that are not sensitive to the user that these activities can be recognized by the server and it is also not useful in gaining utility from the server.



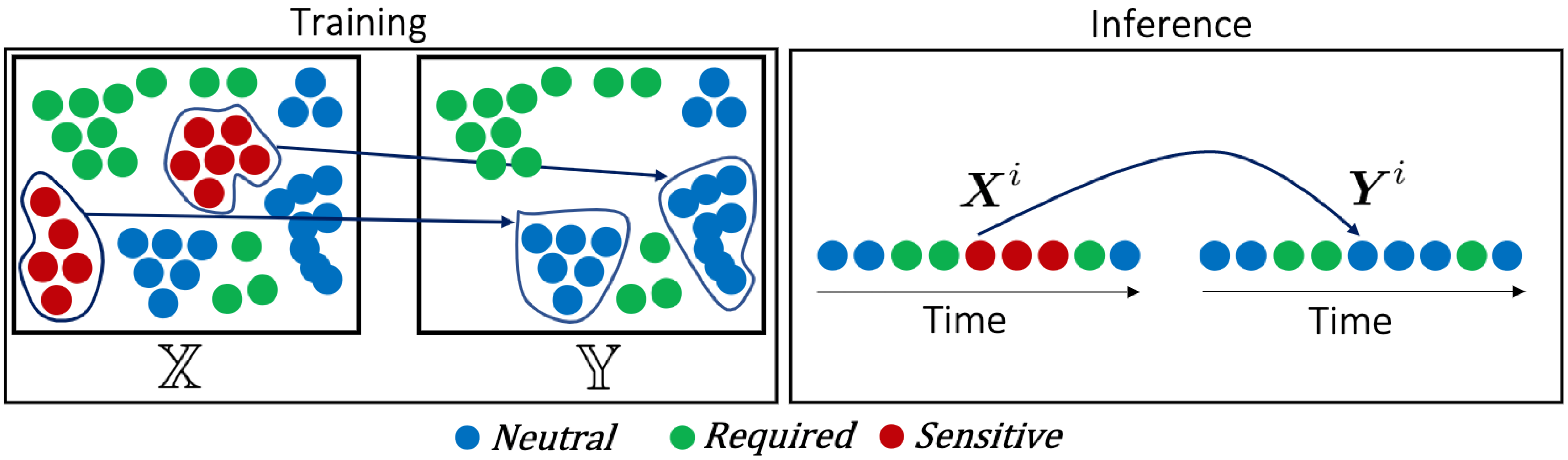
Replacement Approach



$$Y = \mathcal{M}(\boldsymbol{X}) = \begin{cases} \boldsymbol{Z} & \text{if } \boldsymbol{X} \text{ contain sensitive data patterns,} \\ \boldsymbol{X} & \text{otherwise,} \end{cases}$$



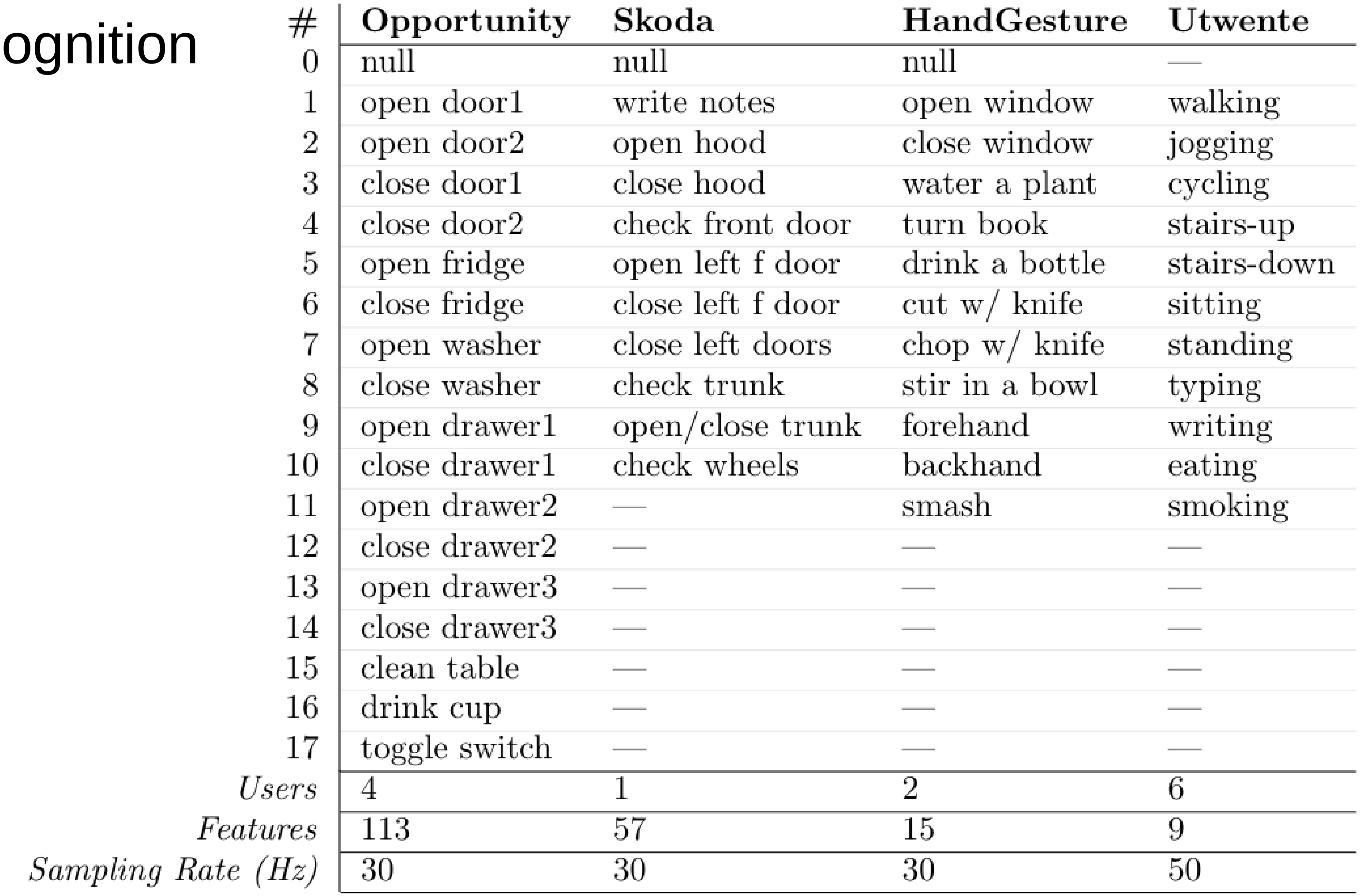
Pairing Datasets for Training



$$\theta^* = \arg\min_{\theta} \mathcal{L}\Big(\mathcal{M}(\mathbb{X};\theta), \mathbb{Y}\Big)$$

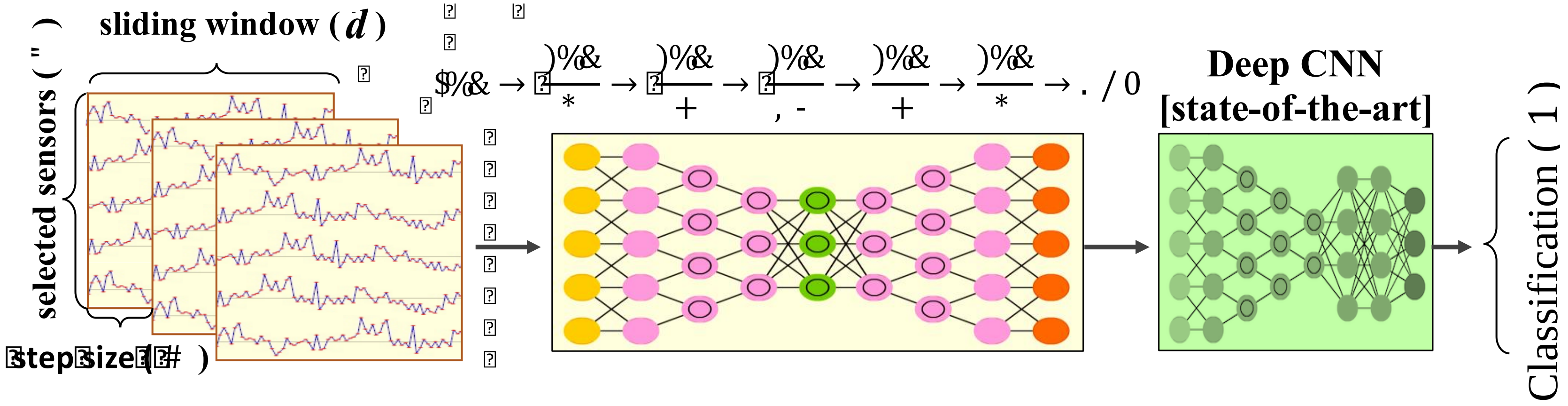
Datasets

- Activity Recognition



#	Opportunity	Skoda	HandGesture	Utwente
0	null	null	null	—
1	open door1	write notes	open window	walking
2	open door2	open hood	close window	jogging
3	close door1	close hood	water a plant	cycling
4	close door2	check front door	turn book	stairs-up
5	open fridge	open left f door	drink a bottle	stairs-down
6	close fridge	close left f door	cut w/ knife	sitting
7	open washer	close left doors	chop w/ knife	standing
8	close washer	check trunk	stir in a bowl	typing
9	open drawer1	open/close trunk	forehand	writing
10	close drawer1	check wheels	backhand	eating
11	open drawer2	—	smash	smoking
12	close drawer2	—	—	—
13	open drawer3	—	—	—
14	close drawer3	—	—	—
15	clean table	—	—	—
16	drink cup	—	—	—
17	toggle switch	—	—	—
Users	4	1	2	6
Features	113	57	15	9
Sampling Rate (Hz)	30	30	30	50

Evaluation Setting



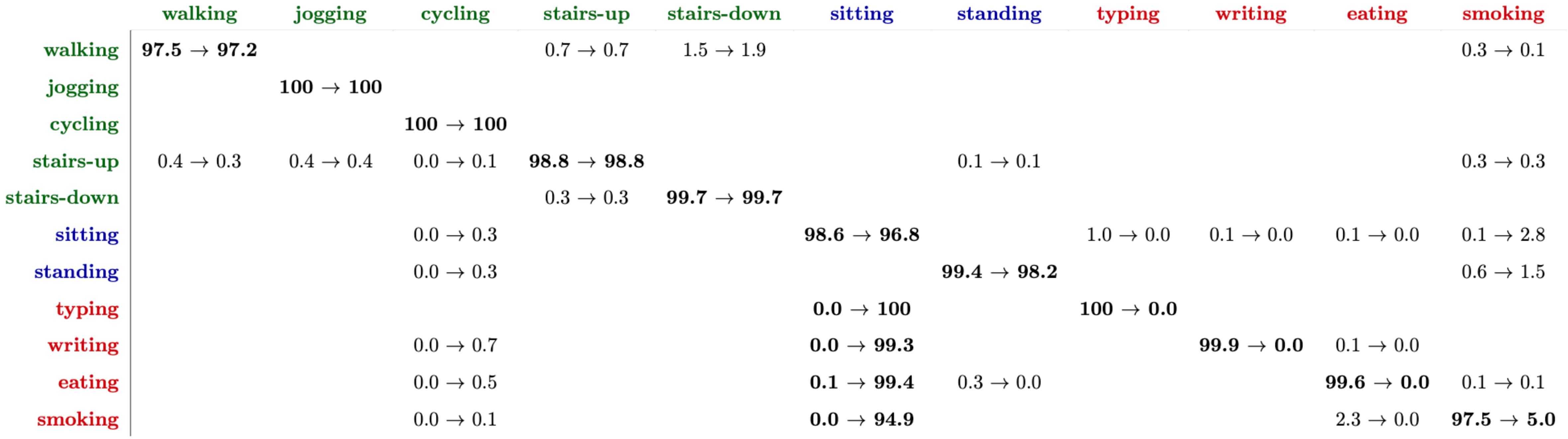
- RAE : A 7-layers Deep Autoencoder
- Activity Recognizer: A Deep Convolutional Autoencoder
  - One of the state-of-the-art for activity recognition using sensor data*

*J. Yang, et. al., “Deep convolutional neural networks on multichannel time series for human activity recognition.” in *IJCAI*, 2015, pp. 3995–4001.

Experimental Result

Classifier's Accuracy

on original data $\boldsymbol{X}$ → on transformed data $\dot{\boldsymbol{X}}$



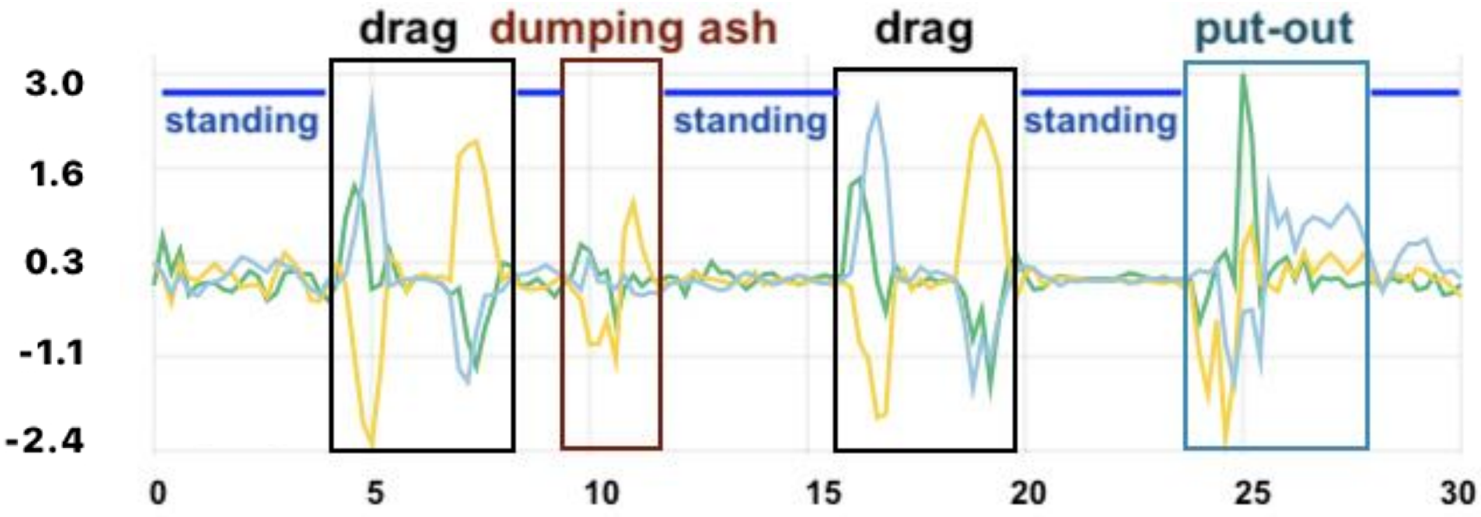
	walking	jogging	cycling	stairs-up	stairs-down	sitting	standing	typing	writing	eating	smoking
walking	**97.5 → 97.2**			0.7 → 0.7	1.5 → 1.9						0.3 → 0.1
jogging		**100 → 100**									
cycling			**100 → 100**								
stairs-up	0.4 → 0.3	0.4 → 0.4	0.0 → 0.1	**98.8 → 98.8**			0.1 → 0.1				0.3 → 0.3
stairs-down				0.3 → 0.3	**99.7 → 99.7**						
sitting			0.0 → 0.3			**98.6 → 96.8**		1.0 → 0.0	0.1 → 0.0	0.1 → 0.0	0.1 → 2.8
standing			0.0 → 0.3				**99.4 → 98.2**				0.6 → 1.5
typing						**0.0 → 100**		**100 → 0.0**			
writing			0.0 → 0.7			**0.0 → 99.3**			**99.9 → 0.0**	0.1 → 0.0	
eating			0.0 → 0.5			**0.1 → 99.4**	0.3 → 0.0			**99.6 → 0.0**	0.1 → 0.1
smoking			0.0 → 0.1			**0.0 → 94.9**				2.3 → 0.0	**97.5 → 5.0**

UTwente Dataset: Complex Human Activities Dataset*

* https://www.utwente.nl/en/eemcs/ps/research/dataset/

Smoking Sub-Activities

Accelerometer Data



time

Experimental Result

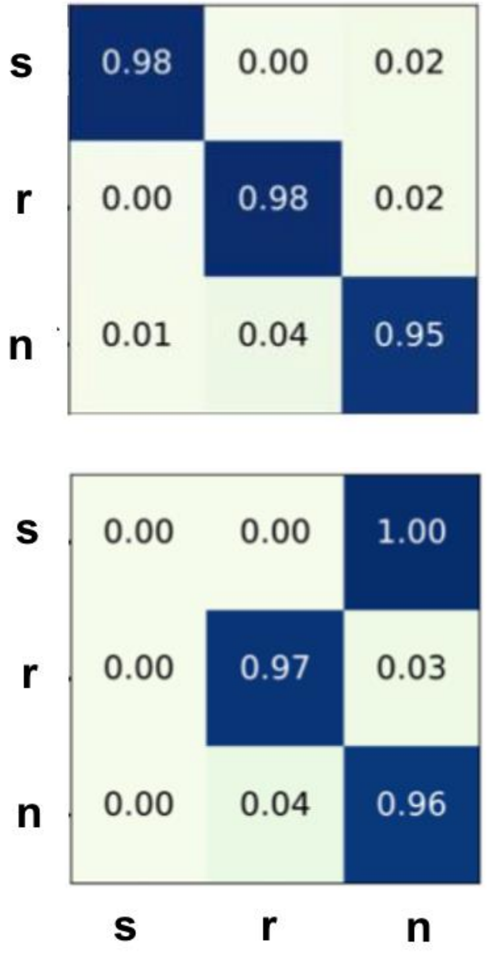
Classifier’s accuracy on the

- $\boldsymbol{X}$: original data
- $\dot{\boldsymbol{X}}$: transformed data

#	*Set of activities*	$\boldsymbol{X}$	$\dot{\boldsymbol{X}}$
1	$\boldsymbol{r} = \{2,3,5,6,7,9\}$	96.5	93.2
	$\boldsymbol{s} = \{4,8,10\}$	97.9	0.0
	$\boldsymbol{n} = \{0,1\}$	93.9	94.8
2	$\boldsymbol{r} = \{4,8,9,10\}$	97.9	96.3
	$\boldsymbol{s} = \{1,5,6,7\}$	96.2	0.0
	$\boldsymbol{n} = \{0,2,3\}$	94.3	93.4

Skoda Dataset*

Confusion Matrix



	s	r	n
s	0.98	0.00	0.02
r	0.00	0.98	0.02
n	0.01	0.04	0.95

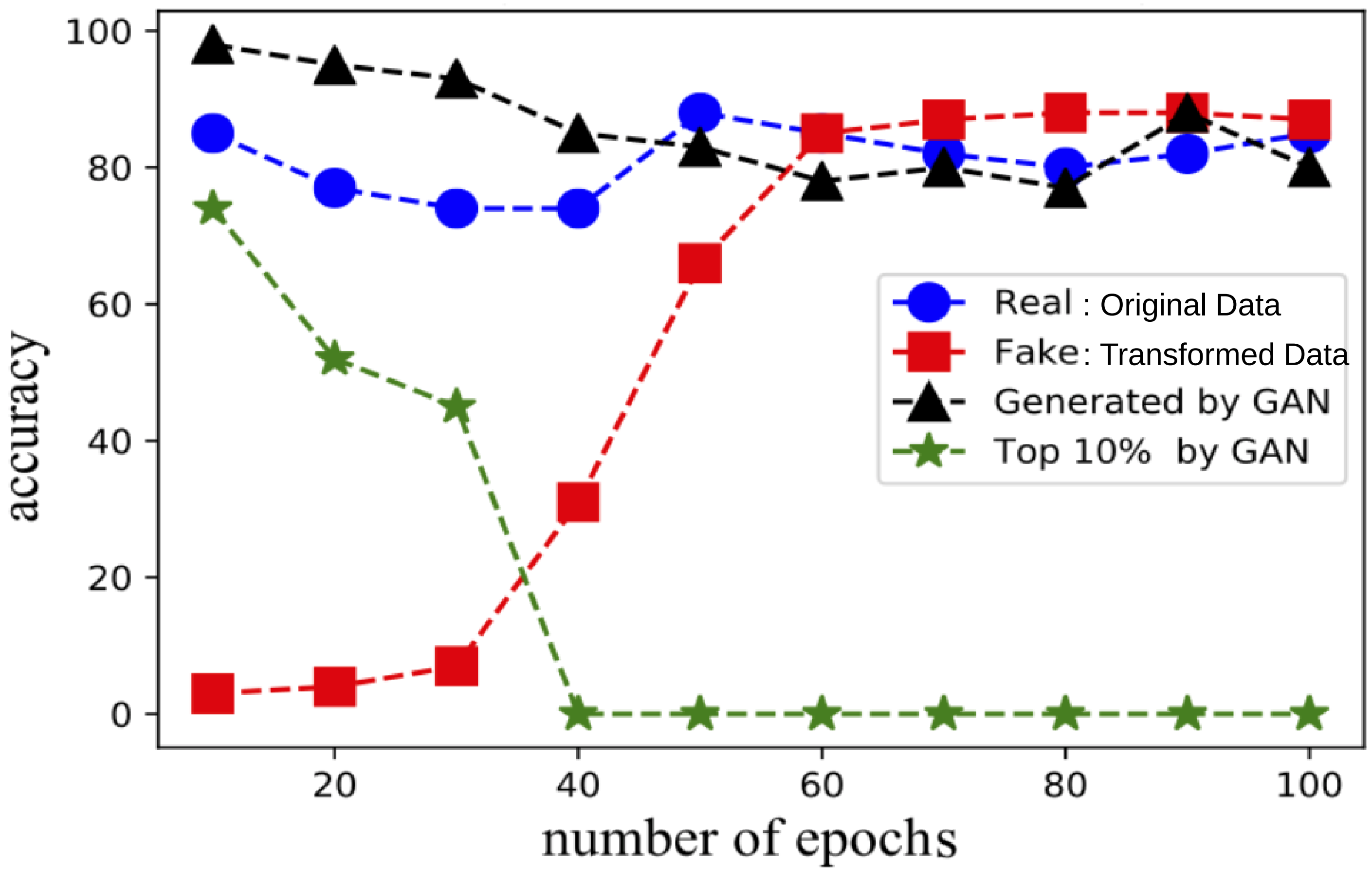
	s	r	n
s	0.00	0.00	1.00
r	0.00	0.97	0.03
n	0.00	0.04	0.96

* http://har-dataset.org/doku.php?id=wiki:dataset

A Potential Attack

- Using a Deep Convolutional Generative Adversarial Net. (DC-GAN)

Assuming that the adversary have access to a dataset of the target user



Code

https://github.com/mmalekzadeh/replacement-autoencoder

3.ii How to Protect User's Sensitive **Attributes**

Sensor Data Anonymization

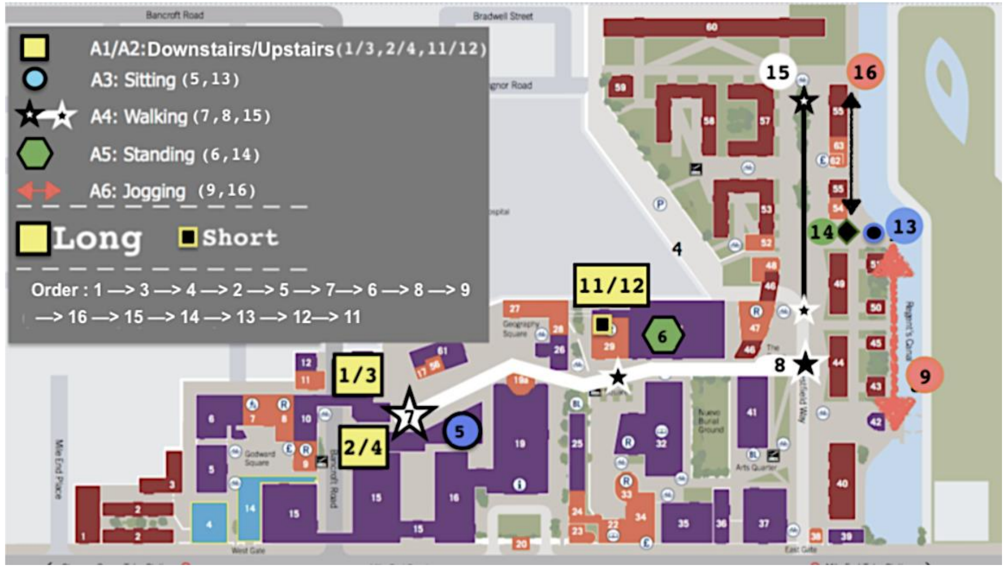
- Privacy is not only about sensitive activates.

- Information that might be discovered from non-sensitive activities:
  - gender
  - race
  - weight

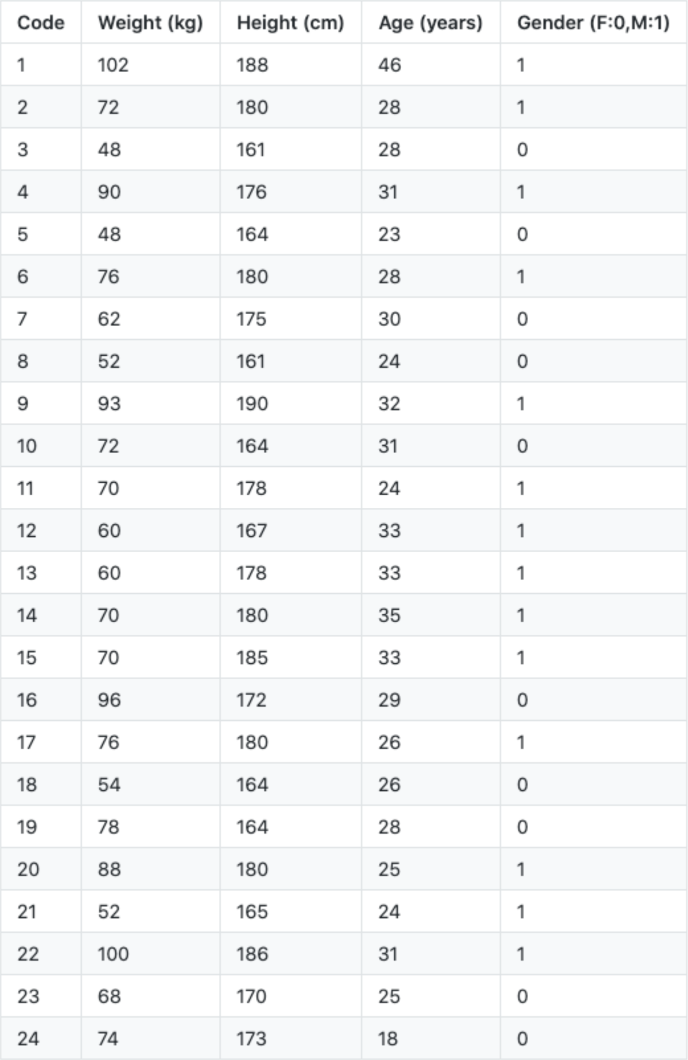
- Re-identification
  - to figure out whether the observed data belongs to a specific person or not,
  - for example, by taking advantage of some data collected through other channels.

An experiment

- **iPhone** 6 in the **front pocket** of participants' **trousers.**
- **2 Sensors:**
  Device Acceleration (Accelerometer)
  Device Rotation (Gyroscope)
- **24 Subjects:**
  Gender, Age, Weight, Height
- **6 Activities:**
  Walking, Jogging, Downstairs, Upstairs
  Sat, Stand-up

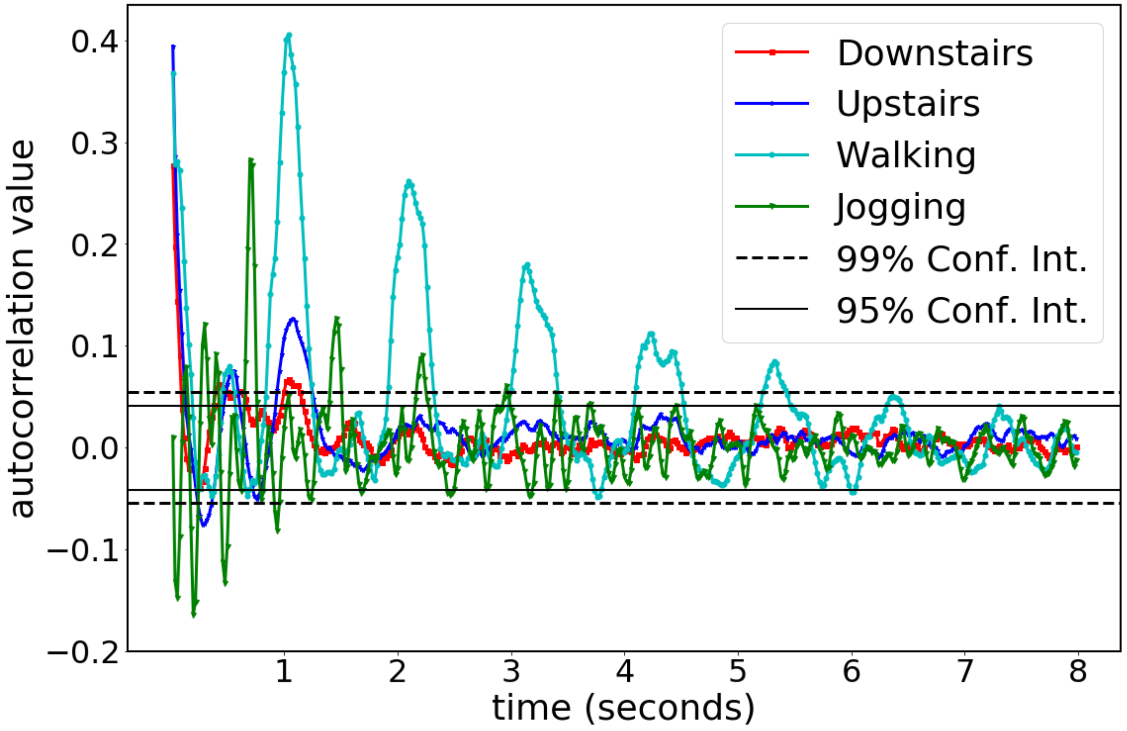


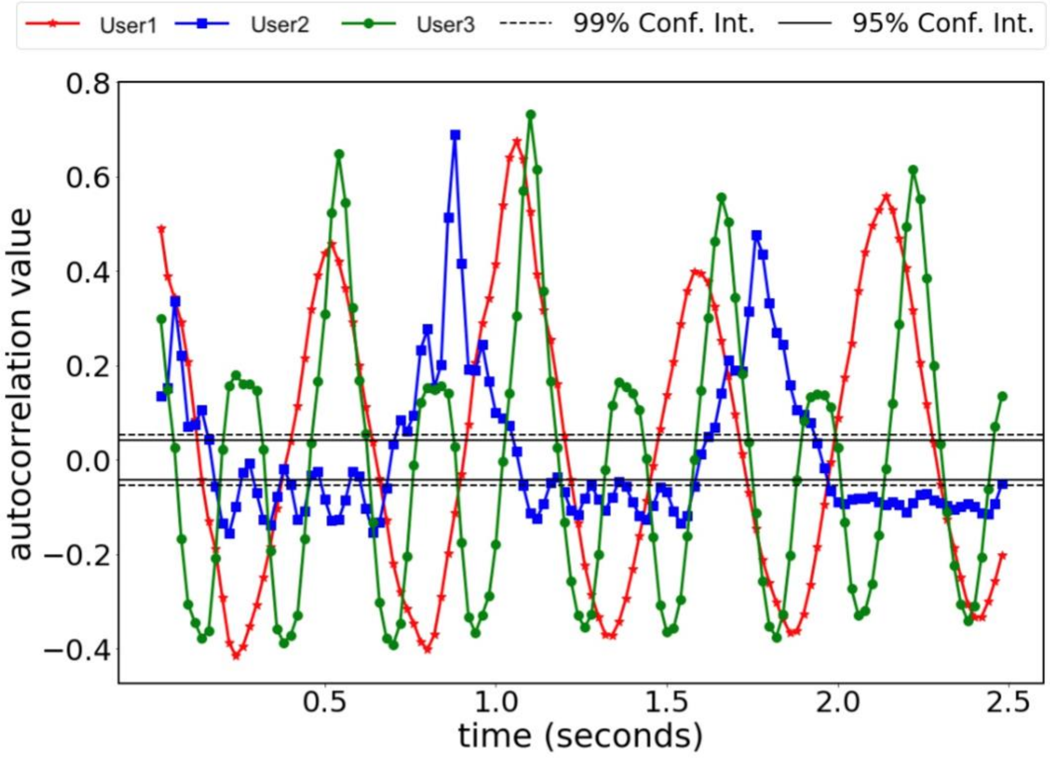
The campus of Queen Mary University of London



Code	Weight (kg)	Height (cm)	Age (years)	Gender (F:0,M:1)
1	102	188	46	1
2	72	180	28	1
3	48	161	28	0
4	90	176	31	1
5	48	164	23	0
6	76	180	28	1
7	62	175	30	0
8	52	161	24	0
9	93	190	32	1
10	72	164	31	0
11	70	178	24	1
12	60	167	33	1
13	60	178	33	1
14	70	180	35	1
15	70	185	33	1
16	96	172	29	0
17	76	180	26	1
18	54	164	26	0
19	78	164	28	0
20	88	180	25	1
21	52	165	24	1
22	100	186	31	1
23	68	170	25	0
24	74	173	18	0

Correlation

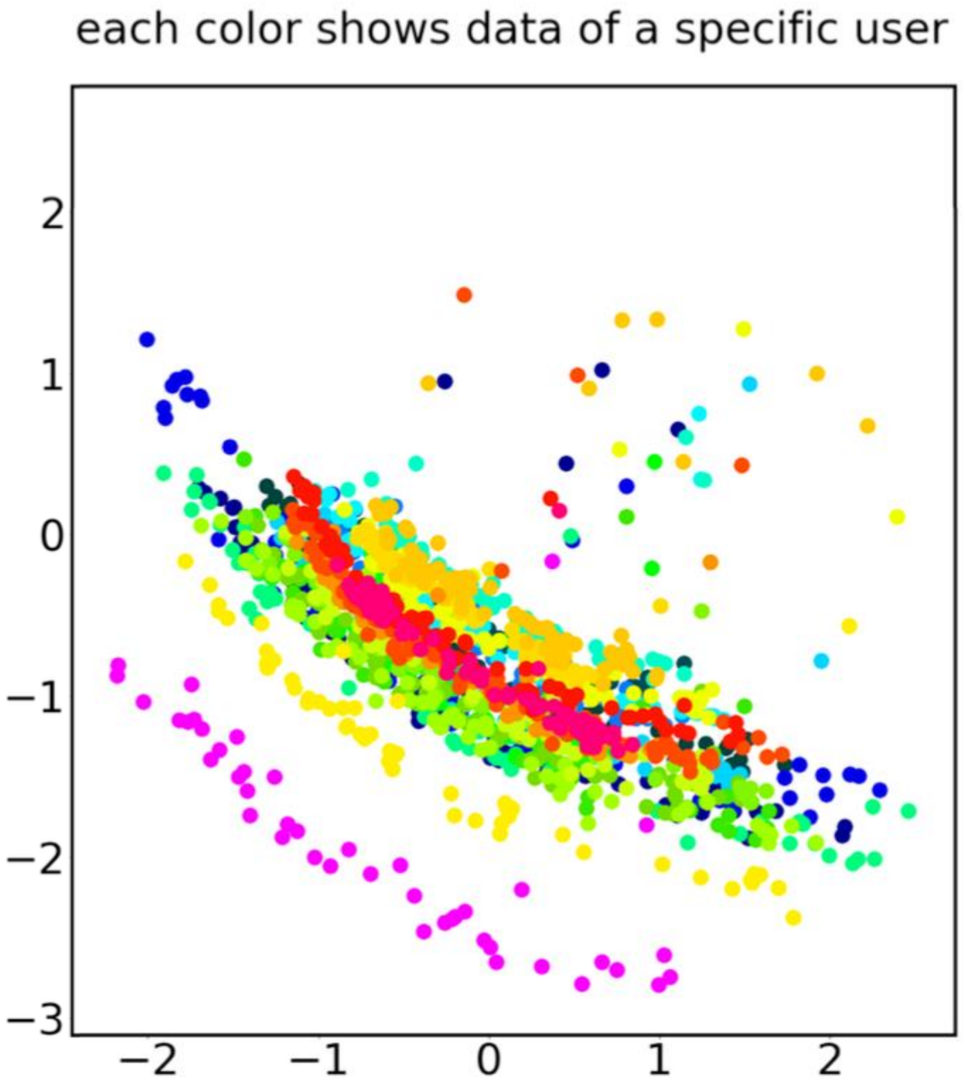




walking

User-Specific Info.

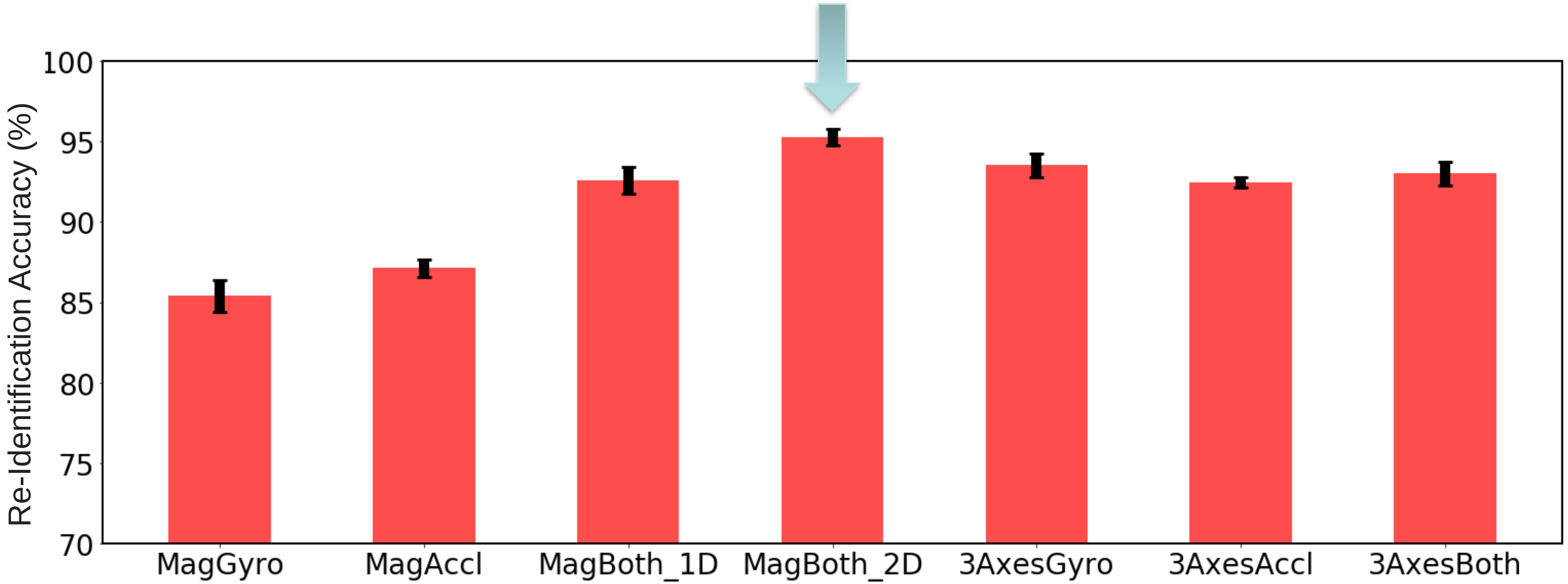
- 2D visualization Using t-SNE*
  - Jogging Activity
  - 2.5 seconds Time Window
  - 24 Users



*Maaten, L. V. D., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of machine learning research*, *9*(Nov), 2579-2605.

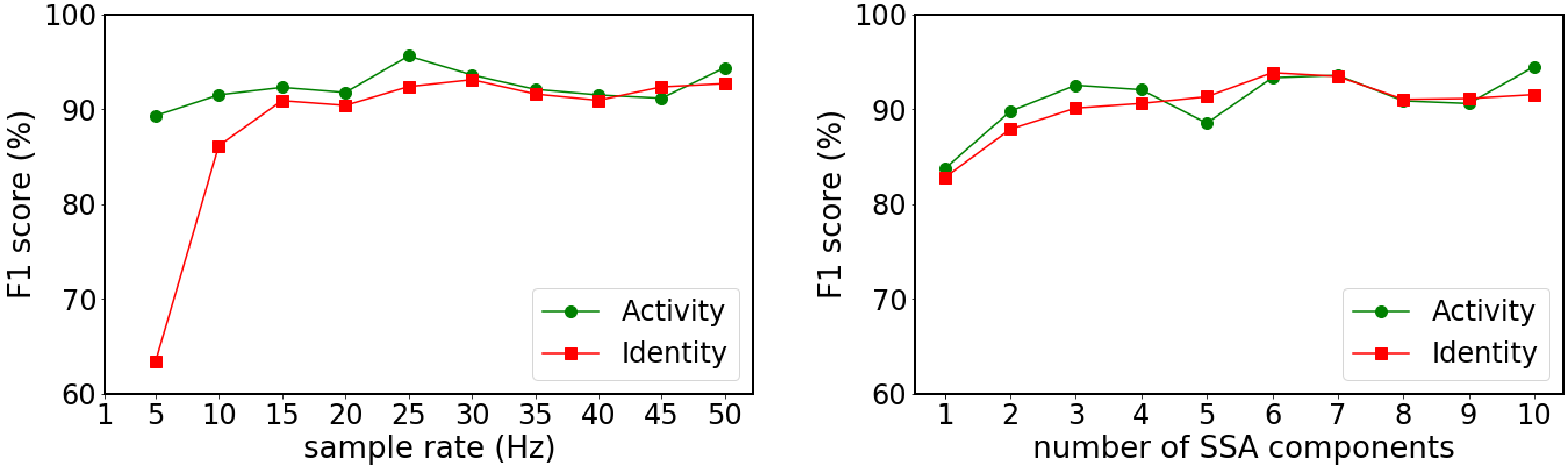
Single and Multivariate Data

Processing **magnitude** values for **both** sensors using **2D** convolutional filters



$$\text{magnitude} = \sqrt{x^2 + y^2 + z^2}$$

The Effect of Reducing the Granularity



A window of 2.5 sec. sensor data is processed by a ConvNet* to recognize users' activity and to re-identify the user

*J. Yang, et. al., "Deep convolutional neural networks on multichannel time series for human activity recognition." in *IJCAI*, 2015, pp. 3995–4001.

Informational Privacy

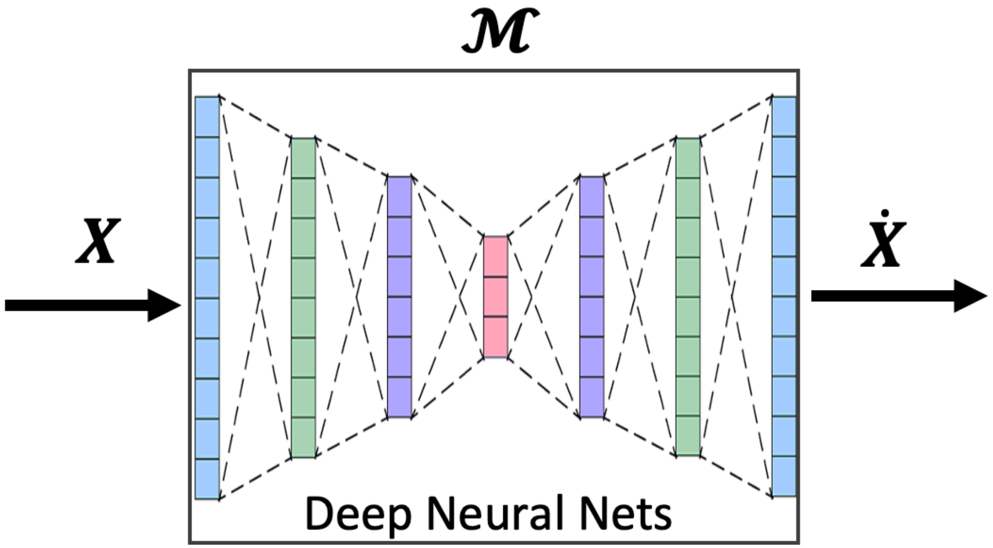
$$\inf_{\boldsymbol{\mathcal{M}}: \boldsymbol{X} \to \dot{\boldsymbol{X}}} \mathrm{I}(\mathbf{s}; \dot{\boldsymbol{X}}) \text{ subject to } \mathrm{I}(\boldsymbol{r}, \boldsymbol{X}) - \mathrm{I}(\boldsymbol{r}, \dot{\boldsymbol{X}}) \leq \delta$$

$\mathbf{r}$: required data

$\mathbf{s}$: sensitive data

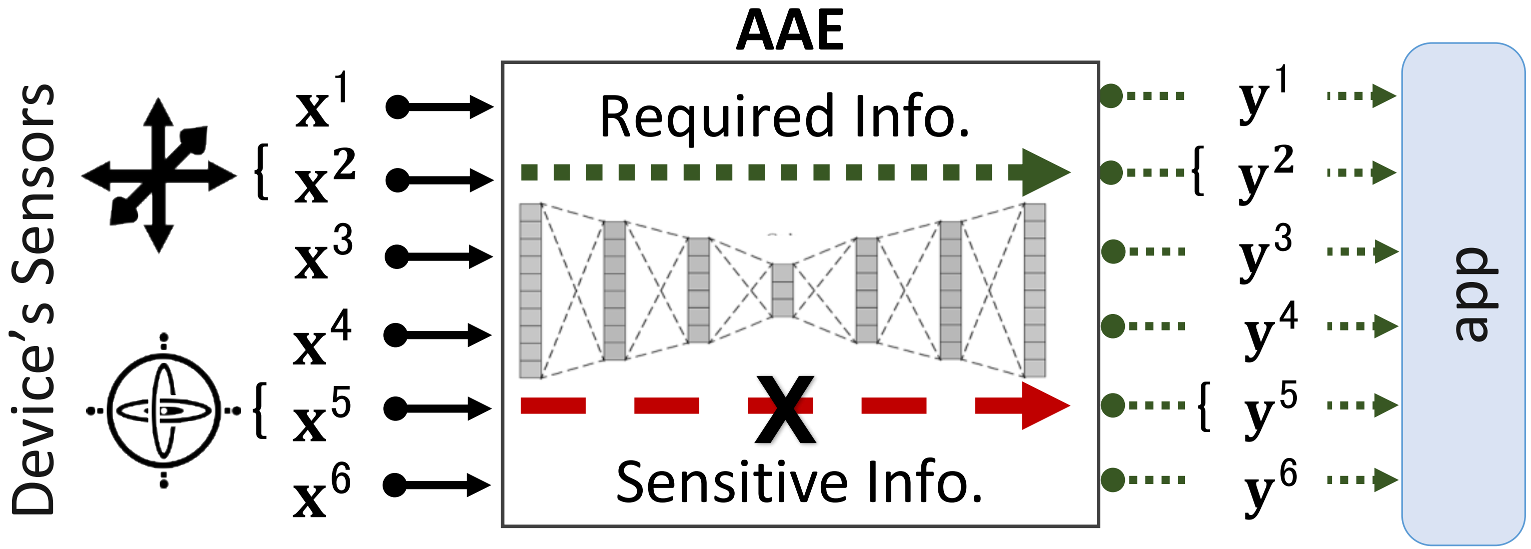
I : mutual information

δ : allowed distortion

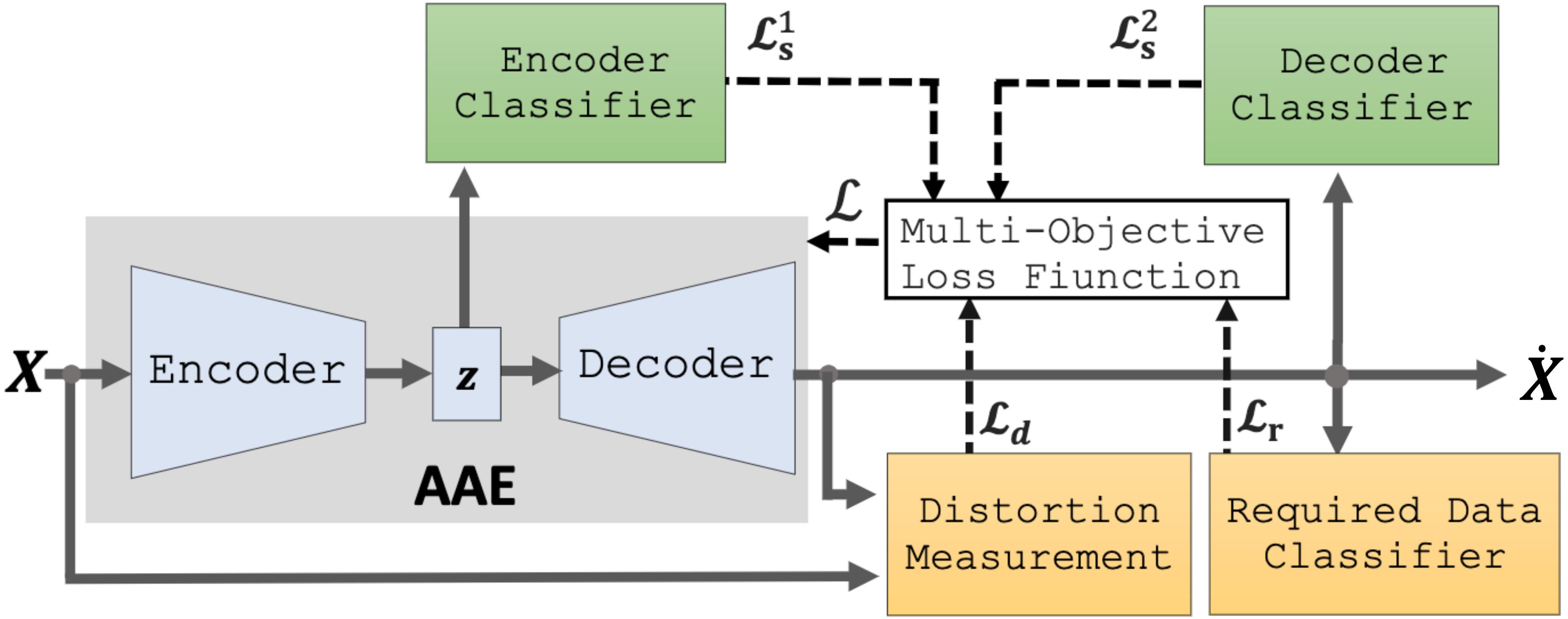


Anonymizing Approach

$$\theta^* = \underset{\theta\in\Theta}{\arg\min}\, \beta_s \mathrm{I}\Big(\mathbf{s};\hat{\mathcal{M}}(\mathbf{X};\theta)\Big) - \beta_r \mathrm{I}\Big(\mathbf{r};\hat{\mathcal{M}}(\mathbf{X};\theta)\Big) + \beta_d \mathcal{D}\Big(\mathbf{X},\hat{\mathcal{M}}(\mathbf{X};\theta)\Big)$$



Multi-Objective Loss Function



$$\mathcal{L} = \beta_s \underbrace{\mathcal{L}_s}_{\text{customized}} + \beta_r \underbrace{\mathcal{L}_r}_{\text{Categorical Cross Entropy}} + \beta_d \underbrace{\mathcal{L}_d}_{\text{MSE}}$$

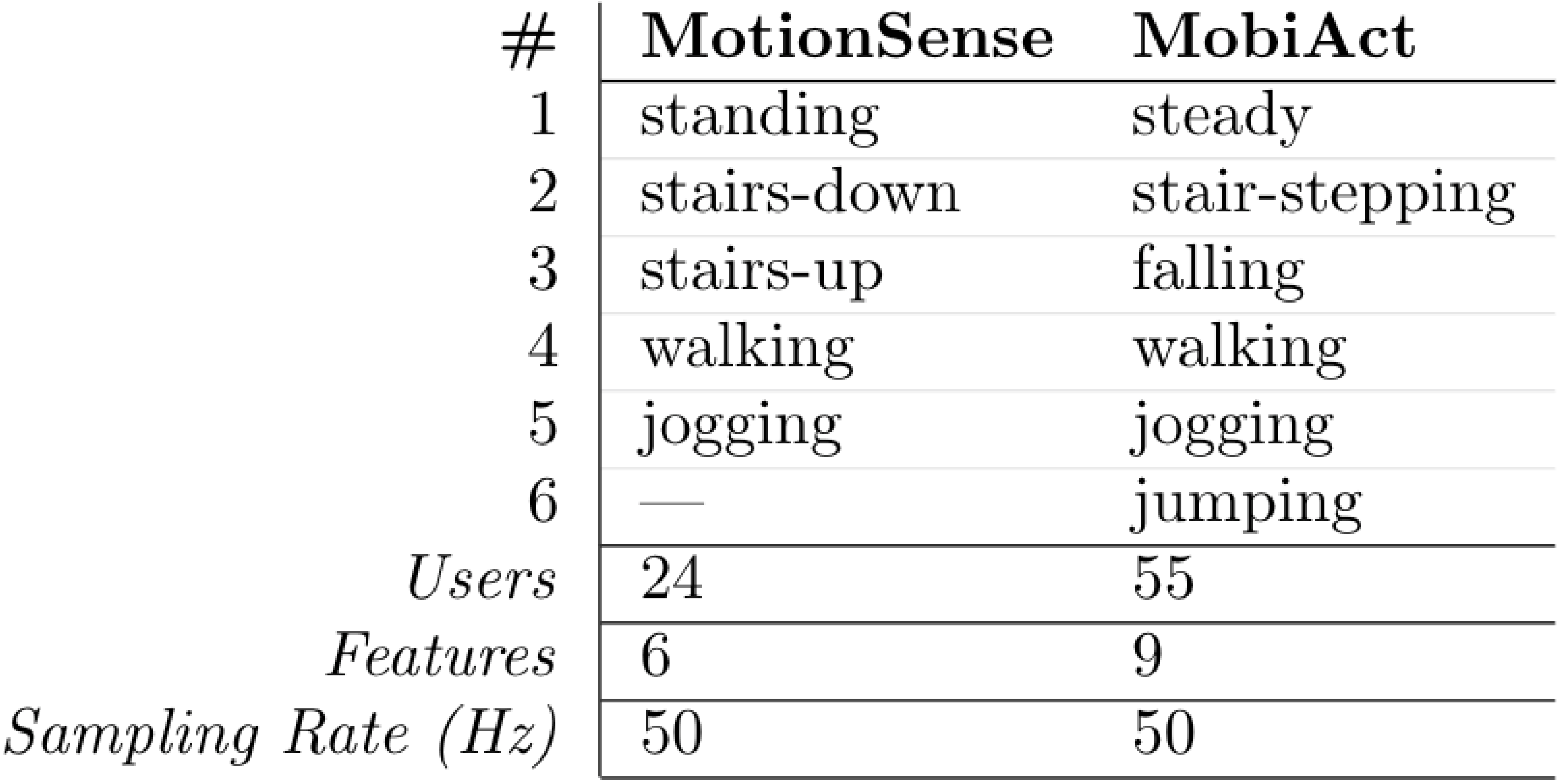
Intuition

$$\mathcal{L}_s = -\left(\mathbf{s} \cdot \log(\mathbf{1}^S - \hat{\mathbf{s}}) + \log\left(1 - \max(\hat{\mathbf{s}})\right)\right)$$

Predicted $\hat{\mathbf{s}}$	1 - $\hat{\mathbf{s}}$	$\mathcal{L}_s$
[.20 .20 .20 .20 **.20**]	[.80 .80 .80 .80 **.80**]	1.12
[.12 .12 .12 .12 **.50**]	[.88 .88 .88 .88 **.50**]	1.23
[.01 .09 .10 .30 **.50**]	[.99 .91 .90 .70 **.50**]	1.26
[.01 .01 .09 .09 **.80**]	[.99 .99 .91 .91 **.20**]	1.82
[.01 .01 .01 .01 **.96**]	[.99 .99 .99 .99 **.04**]	3.04

e.g. True $\mathbf{s} = \mathbf{5}$

Datasets



#	MotionSense	MobiAct
1	standing	steady
2	stairs-down	stair-stepping
3	stairs-up	falling
4	walking	walking
5	jogging	jogging
6	—	jumping
Users	24	55
Features	6	9
Sampling Rate (Hz)	50	50

Experimental Results

	Raw Data	Our AAE Transformation	Censoring Representations[1]	Singular Spectrum Analysis	Resampling by FFT
	Activity Recognition Accuracy (%) : Using ConvNets				
Utility	92.5 ± 2.0	**92.9 ± 0.3**	~ 91.5 ± 0.9	~ 87.4 ± 0.9	88 ± 1.8
	Re-Identification of Users Accuracy (%) : Using ConvNets				
Privacy	96.2	**7.0**	15.9	16.1	13.5
	Data Similarity Rank : Using Dynamic Time Warping				
Fidelity	0	**6.6**	10.7	9.5	9.3

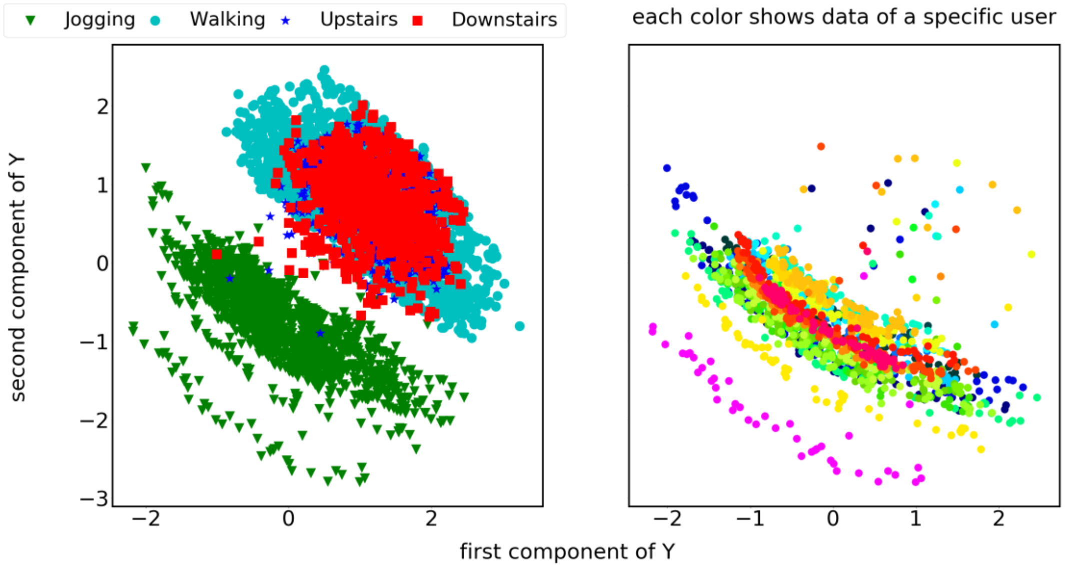
Motion Sense Dataset

CIS centre for intelligent sensing

Queen Mary University of London

Visualization: t-SNE

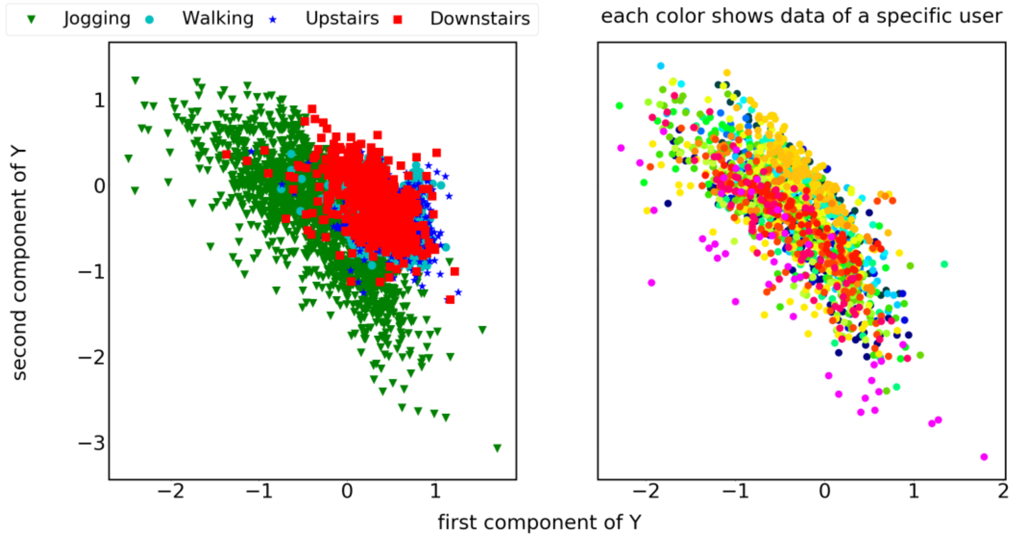
Raw



Activities

Users

Transformed



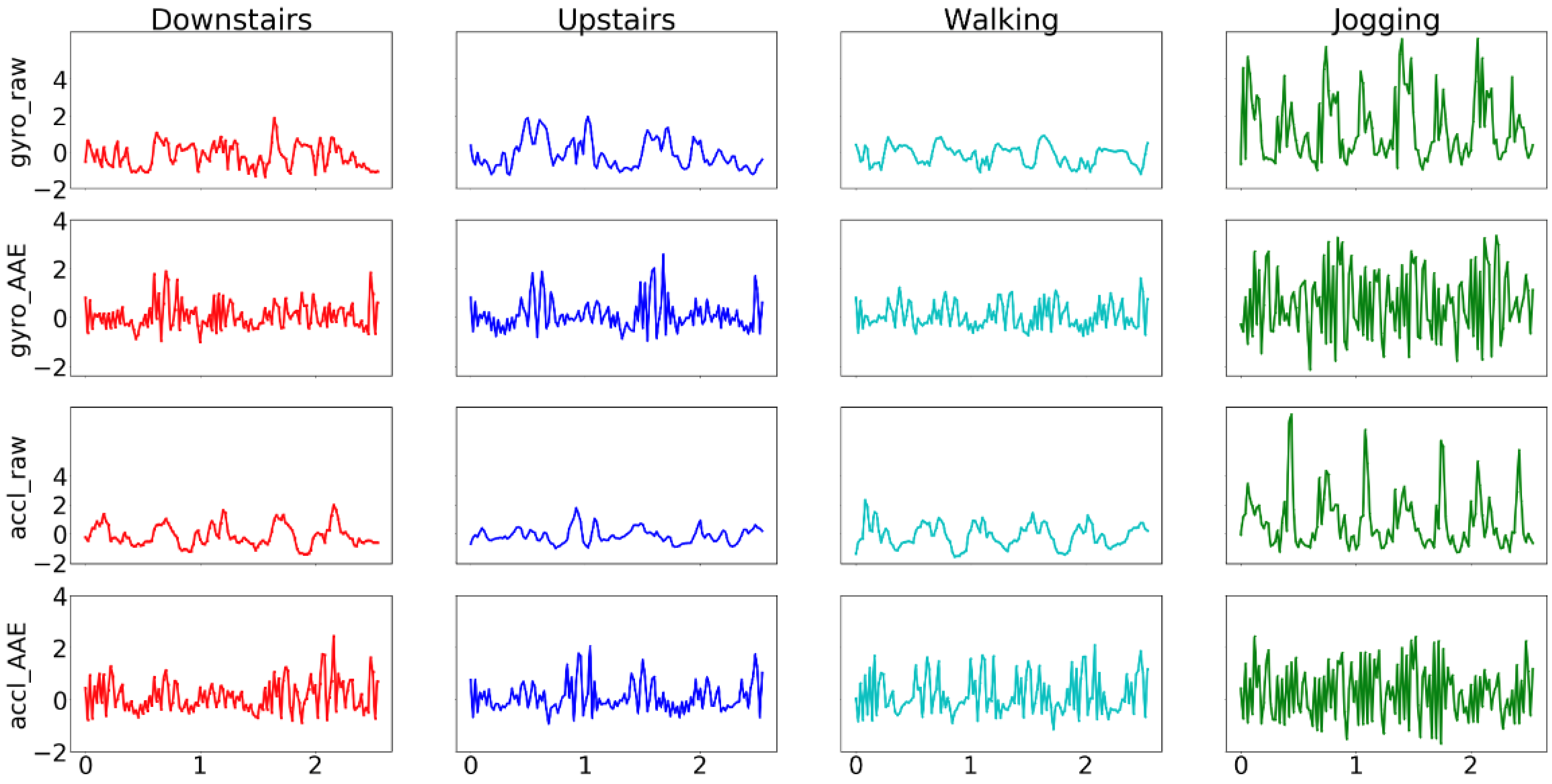
Activities

Users

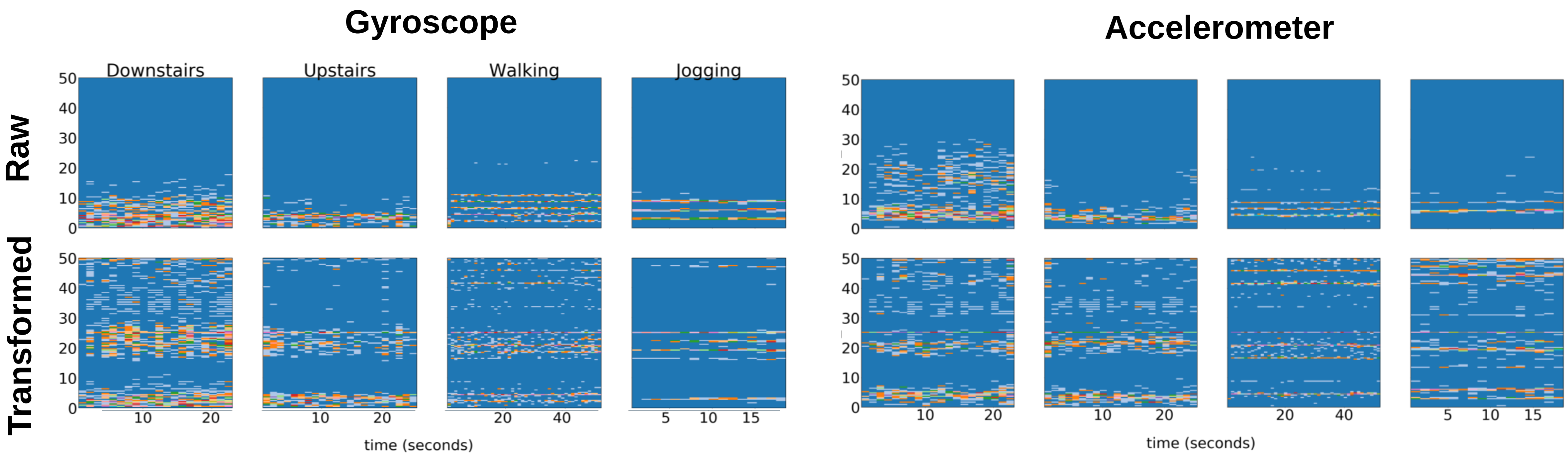
Time Domain

Gyroscope

Accelerometer



Spectrogram



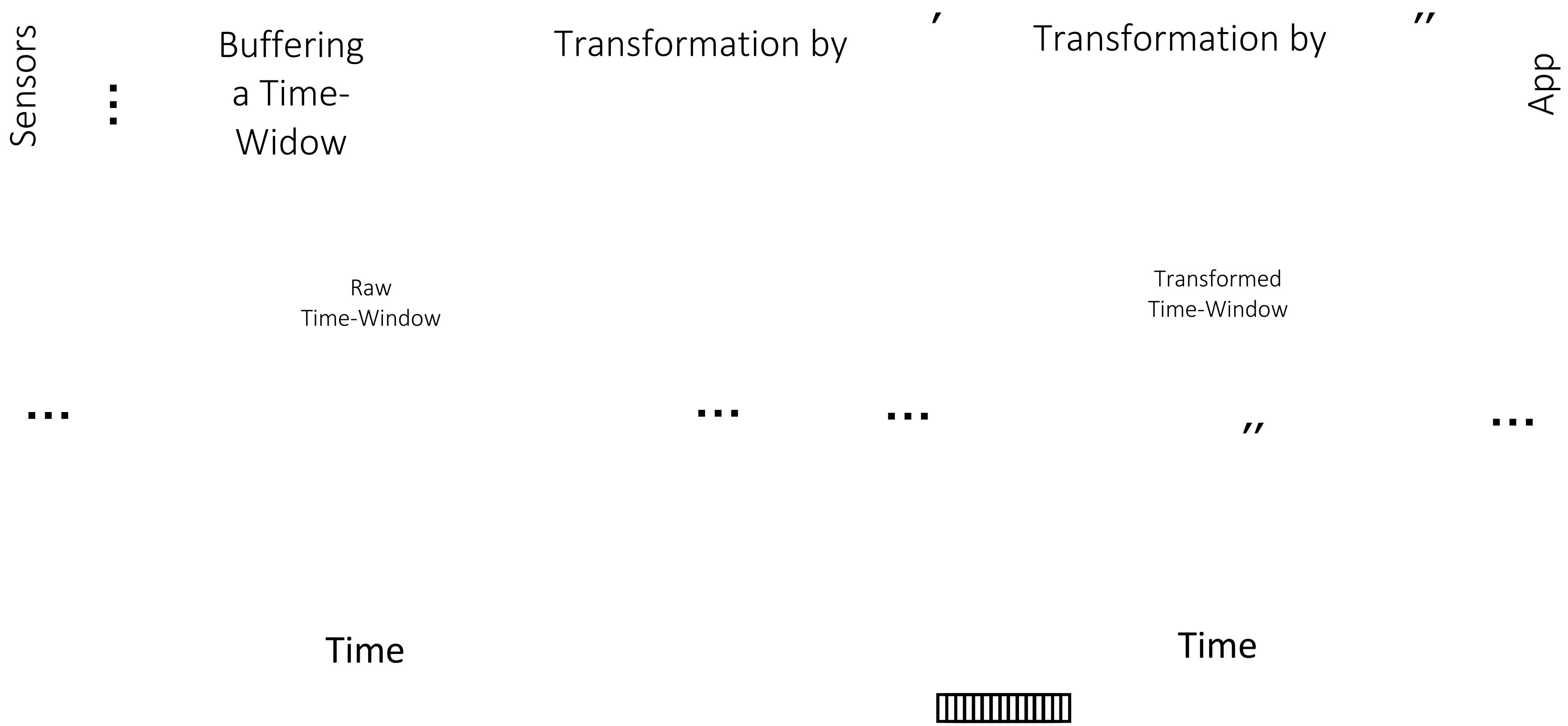
New **periodic components** are introduced in the data and some of the original ones are **obscured**.

Code

https://github.com/mmalekzadeh/motion-sense

3.iii How to Protect User’s Sensitive **Activities** and **Attributes**

A Compound Solution



Sensors

⋮

Buffering a Time-Widow

Transformation by ′

Transformation by ″

App

Raw Time-Window

Transformed Time-Window

…

…

…

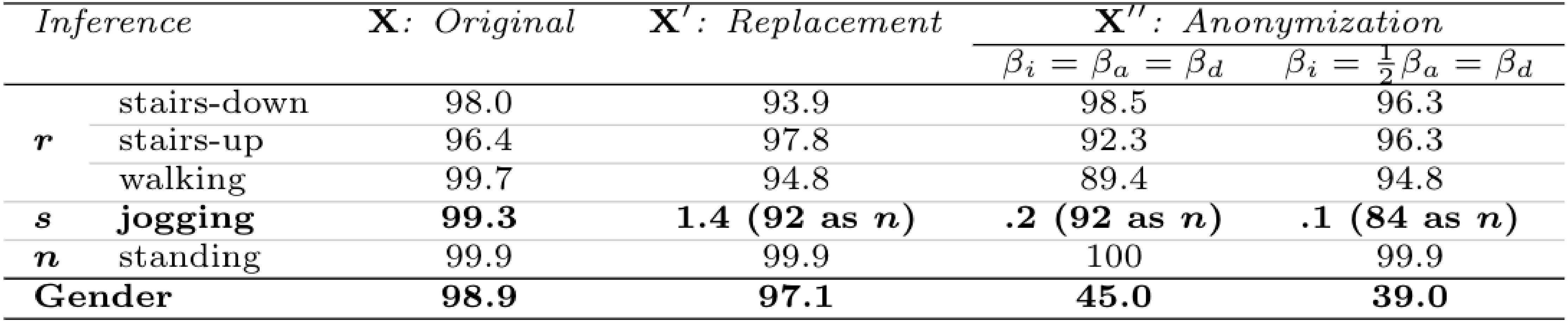
″

…

Time

Time

Experimental Results

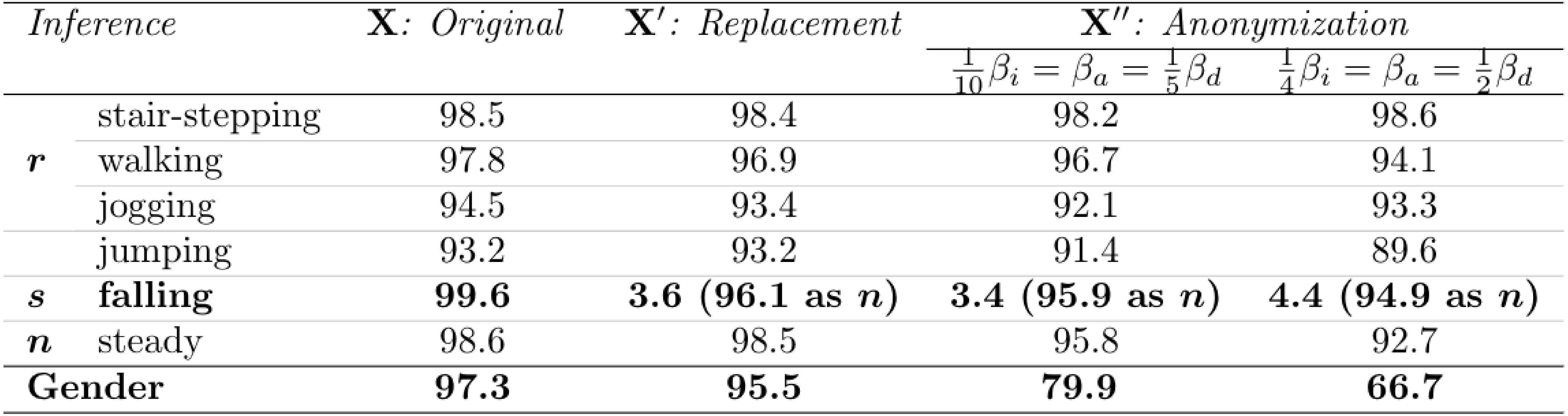


Inference		$\mathbf{X}$: *Original*	$\mathbf{X}'$: *Replacement*	$\mathbf{X}''$: *Anonymization*	
				$\beta_i = \beta_a = \beta_d$	$\beta_i = \frac{1}{2}\beta_a = \beta_d$
r	stairs-down	98.0	93.9	98.5	96.3
	stairs-up	96.4	97.8	92.3	96.3
	walking	99.7	94.8	89.4	94.8
s	**jogging**	**99.3**	**1.4 (92 as n)**	**.2 (92 as n)**	**.1 (84 as n)**
n	standing	99.9	99.9	100	99.9
Gender		**98.9**	**97.1**	**45.0**	**39.0**

Motion Sense Dataset*

* https://github.com/mmalekzadeh/motion-sense

Experimental Results



Inference		$\mathbf{X}$: *Original*	$\mathbf{X}'$: *Replacement*	$\mathbf{X}''$: *Anonymization*	
				$\frac{1}{10}\beta_i = \beta_a = \frac{1}{5}\beta_d$	$\frac{1}{4}\beta_i = \beta_a = \frac{1}{2}\beta_d$
$\boldsymbol{r}$	stair-stepping	98.5	98.4	98.2	98.6
	walking	97.8	96.9	96.7	94.1
	jogging	94.5	93.4	92.1	93.3
	jumping	93.2	93.2	91.4	89.6
$\boldsymbol{s}$	**falling**	**99.6**	**3.6 (96.1 as $\boldsymbol{n}$)**	**3.4 (95.9 as $\boldsymbol{n}$)**	**4.4 (94.9 as $\boldsymbol{n}$)**
$\boldsymbol{n}$	steady	98.6	98.5	95.8	92.7
Gender		**97.3**	**95.5**	**79.9**	**66.7**

Mobi Act Dataset*

* Vavoulas, George, et al. "The MobiAct Dataset: Recognition of Activities of Daily Living using Smartphones." ICT4AgeingWell. 2016.

Code

https://github.com/mmalekzadeh/motion-sense

4. Conclusion and Open Questions

Summary

- Data generated by motion sensors is informative
  - user’s activities
  - user’s attributes
- We can train deep autoencoders for data transformation
  - release required data
  - remove sensitive data
- The trained model can generalize
  - model can be used by a for unseen user during training
  - Model can be used for on-device data transformation or for offline dataset publishing

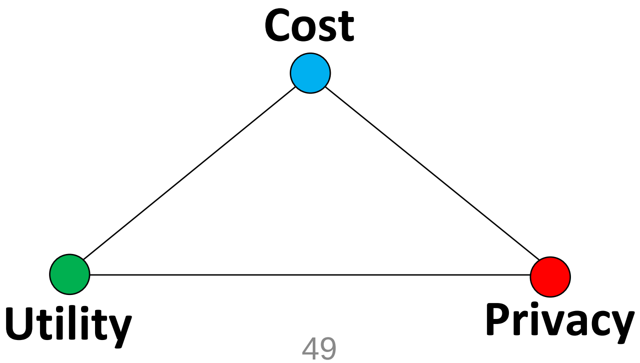
Open Questions for Future Directions

1. Probabilistic and/or mathematical bound on the privacy and utility guarantees.
   - *Differential Privacy* is not a suitable metric for continual sharing of multi-dimensional data.
   - *Information Privacy* needs a complete knowledge of the data distribution.

2. Correlation among consecutive data release:
   - An approach to account and track of the privacy loss occurred
   - A Bayesian approach might be useful

3. Datasets including more fine-grained activities and more users.

4. The cost and complexity of such solutions for running them on the edge devices?



Q & A

Thanks for your attention

Resources:

- https://github.com/mmalekzadeh/motion-sense
- https://github.com/mmalekzadeh/dana
- https://github.com/mmalekzadeh/replacement-autoencoder
- https://github.com/mmalekzadeh/privacy-preserving-bandits