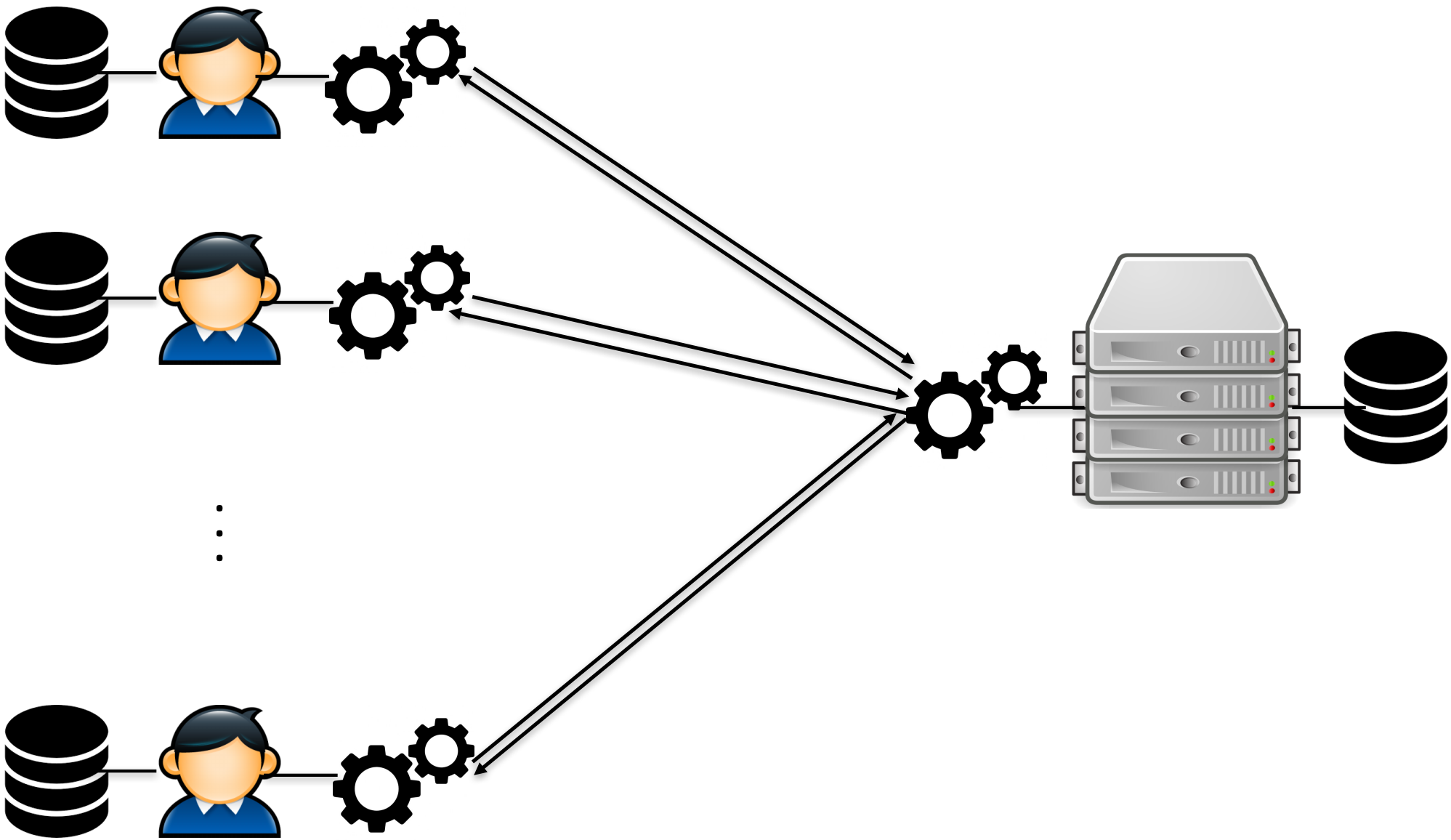
Building and Measuring Privacy-Preserving Mobility Analytics

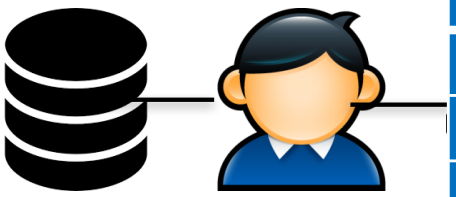
Emiliano De Cristofaro

University College London (UCL)

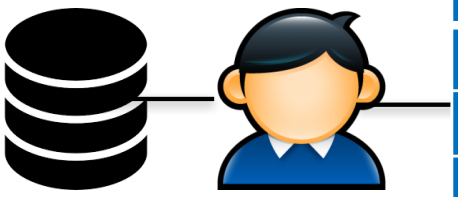
https://emilianodc.com



Example: Recommender System

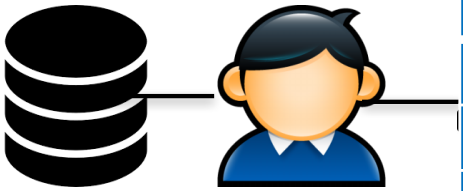


	Dr Who	Sherlock	Earth
Dr Who	1	-	-
Sherlock	1	1	-
Earth	0	0	0



	Dr Who	Sherlock	Earth
Dr Who	1	-	-
Sherlock	1	1	-
Earth	1	1	1

	Dr Who	Sherlock	Earth
Dr Who	3	-	-
Sherlock	2	2	-
Earth	1	1	1



	Dr Who	Sherlock	Earth
Dr Who	1	-	-
Sherlock	0	0	-
Earth	0	0	0

Privacy in Recommender System?

Collaboration with BBC R&D

BBC iPlayer

- Web platform with 500-1000 programs, free in the UK
- No account required (just promise you have TV licence)
- No tracking, no ads

Still useful to collect statistics, offer personalized recommendations to users

- Which programs are successful?
- Increase viewership

Privacy-Preserving Aggregation

Goal: aggregator collects matrix, s.t.

- Can only learn aggregate counts (e.g., 237 users have watched both Dr Who and Sherlock)
- Not who has watched what

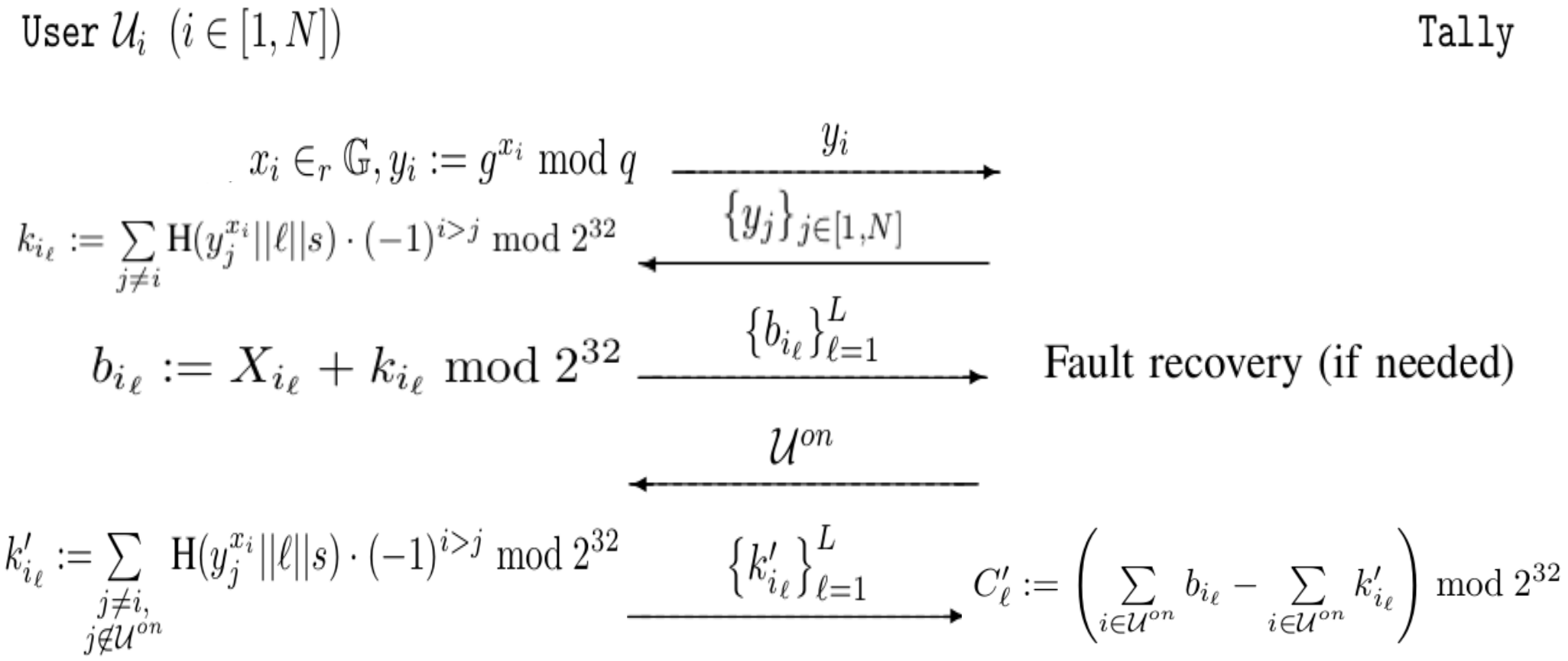
Use additively homomorphic encryption

- $Enc_{PK}(a)*Enc_{PK}(b) = Enc_{PK}(a+b)$

Keys summing up to zero

Users U_1, U_2, …, U_N, each has k_1, k_2, …, k_N s.t.

$k_1 + k_2 + \ldots + k_N = 0$



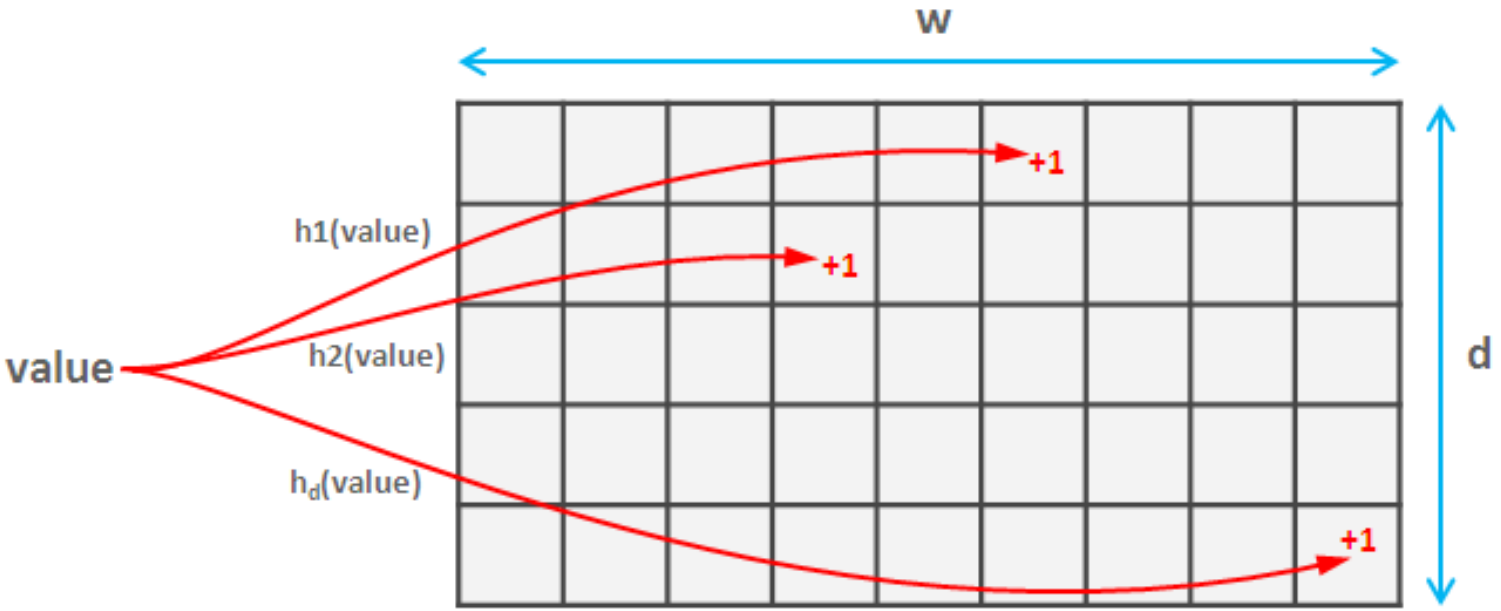
Is this efficient?

Count-Min Sketch

Estimate an item's frequency in a stream

Mapping a stream of values (of length T) to a matrix of size O(logT)

The sum of two sketches results in the sketch of the union of the two data streams



More details in the paper [1]

Security

In the honest-but-curious model under the CDH assumption

Prototype implementation:

Tally as a Node.js web server

Client-side run in the browser or as a mobile cross-platform application (Apache Cordova)

Collecting statistics in the Tor network

Median → different data structure, differential privacy

[1] Luca Melis, George Danezis, Emiliano De Cristofaro. Efficient Private Statistics with Succinct Sketches. In NDSS 2016

More Applications

Infectious disease modeling via Google queries

- Need how many people searched for certain keywords, from certain areas, not who
- Ongoing collaboration with London School of Hygiene and Tropical Medicine

Mobility Analytics

- Next…

Mobility Analytics

Use location/movement data to improve urban and transportation planning

- London studies, Waze, and many more

Raises privacy concerns…

- Infer life-style, political/religious inclinations
- Anonymization of location traces is ineffective

How about using only aggregate statistics?

- How many people at location X at time t? (Not who)

Our Work

1. Mobility analytics using aggregate locations? [2]

Is it useful? What tasks can we perform?

Real-world deployability?

2. How much privacy do aggregates leak? [3]

How can we quantify it?

How do we test it?

[2] Apostolos Pyrgelis, Gordon Ross, Emiliano De Cristofaro. Privacy-Friendly Mobility Analytics using Aggregate Location Data. In ACM SIGSPATIAL 2016

[3] Apostolos Pyrgelis, Carmela Troncoso, Emiliano De Cristofaro. What Does The Crowd Say About You? Evaluating Aggregation-based Location Privacy. In PETS 2017

Analytics on Aggregate Locations

Experiment with mobility datasets (TFL, SFC)

Tasks:

1. Forecasting traffic volumes in regions of interest (ROIs)
2. Detecting mobility anomalies
3. Improving traffic volume predictions in the presence of anomalies

Empirical evaluation of complexity, energy, etc.

TFL Data

Logs of anonymized oyster card trips including Underground (LUL), National Rail (NR), Overground (LRC), Docklands Light Railway (DLR)

Monday, March 1 to Sunday, March 28, 2010

60 million trips as performed by 4 million unique users, over 582 stations

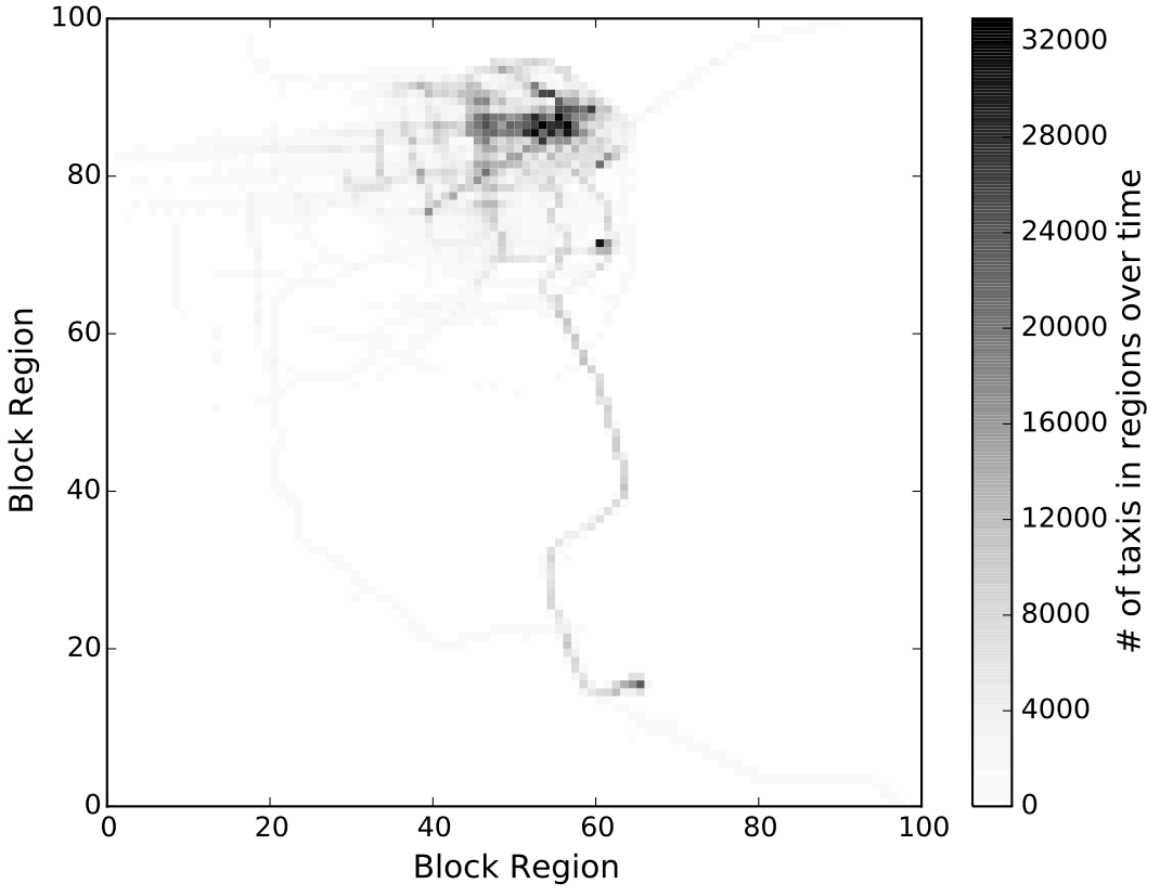
San Francisco Cabs (SFC)

Mobility traces of 536 cabs in SF (May 19 to June 8, 2008)

11 million GPS coords

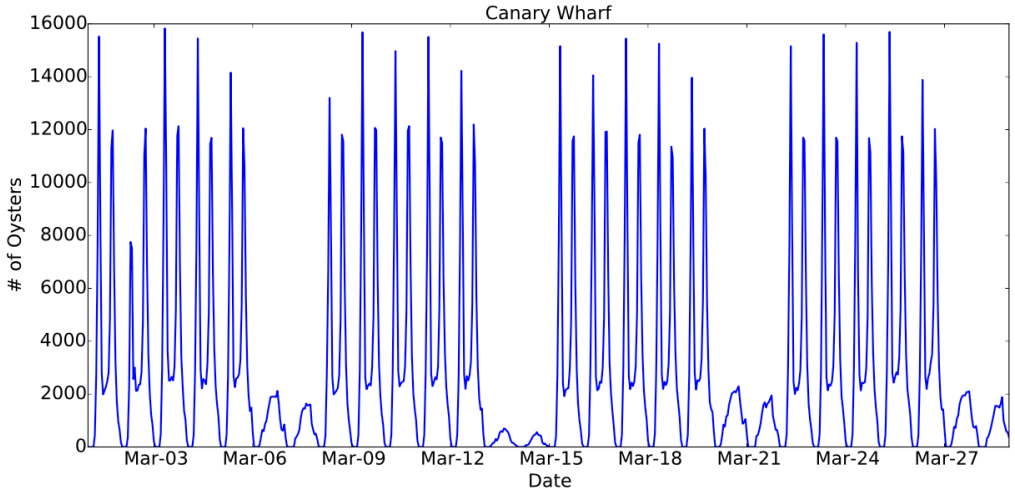
San Francisco grid of 100 x 100 regions

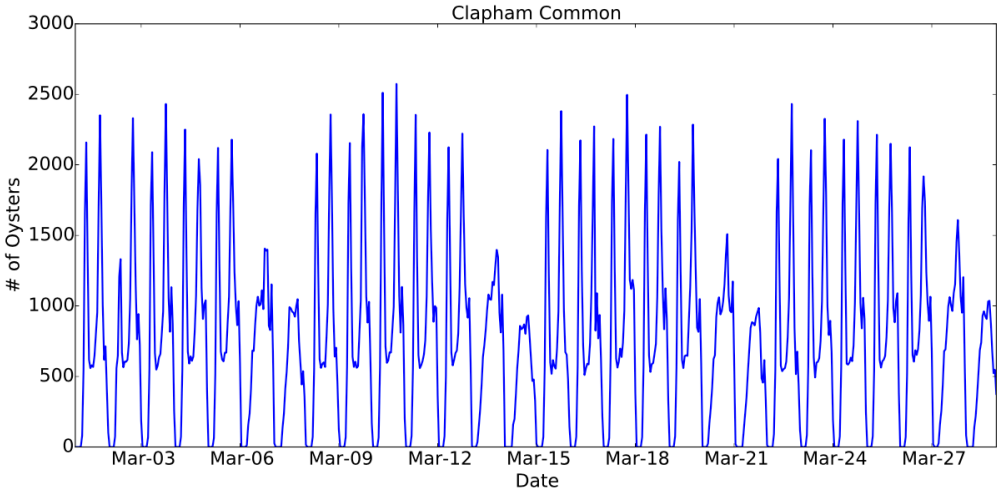
0.19 × 0.14 sq mi



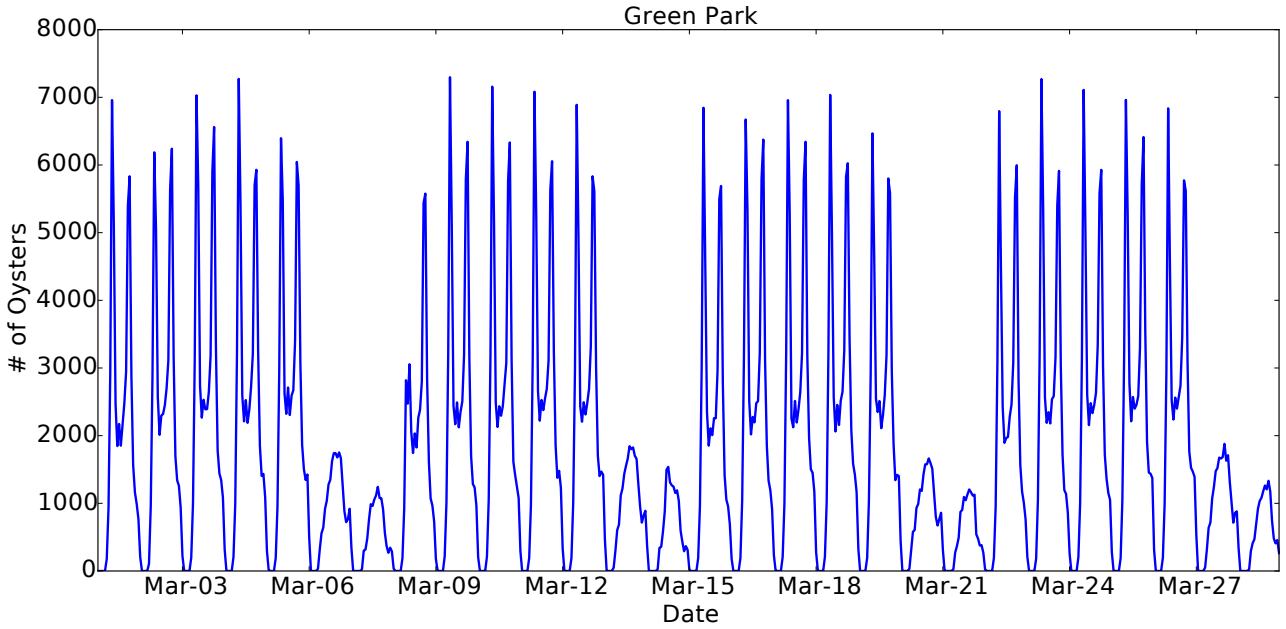
TFL Time-Series

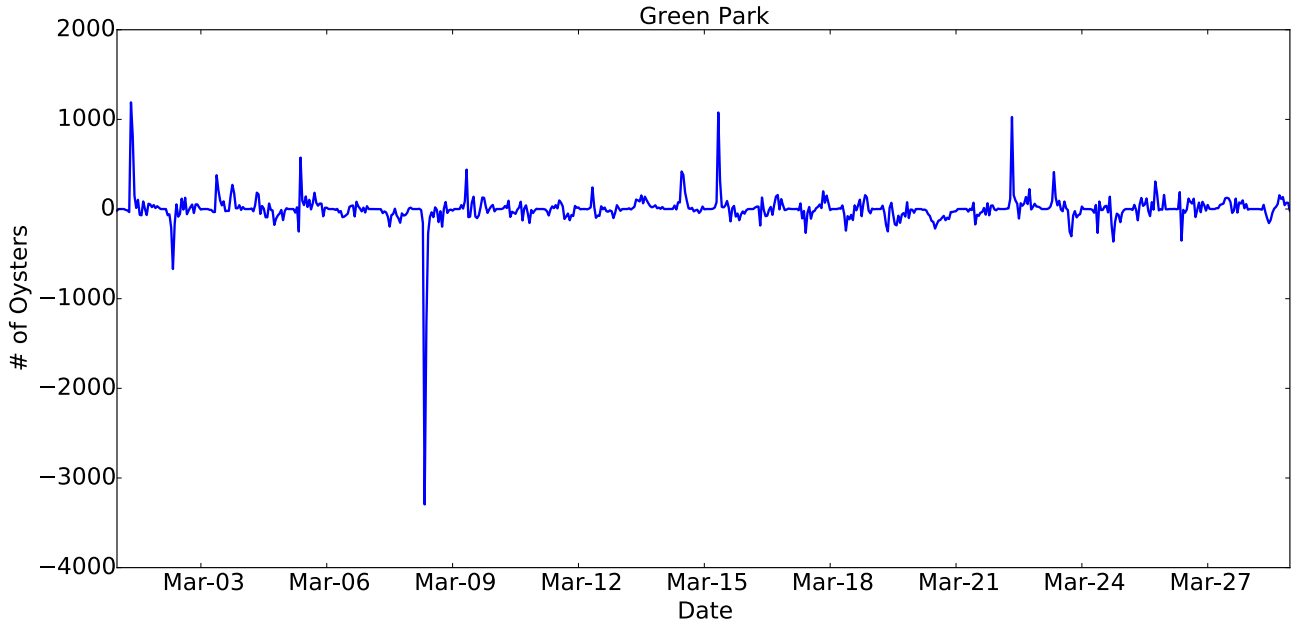
We build hourly time series of stations, counting # of users tapping-in/out at each station



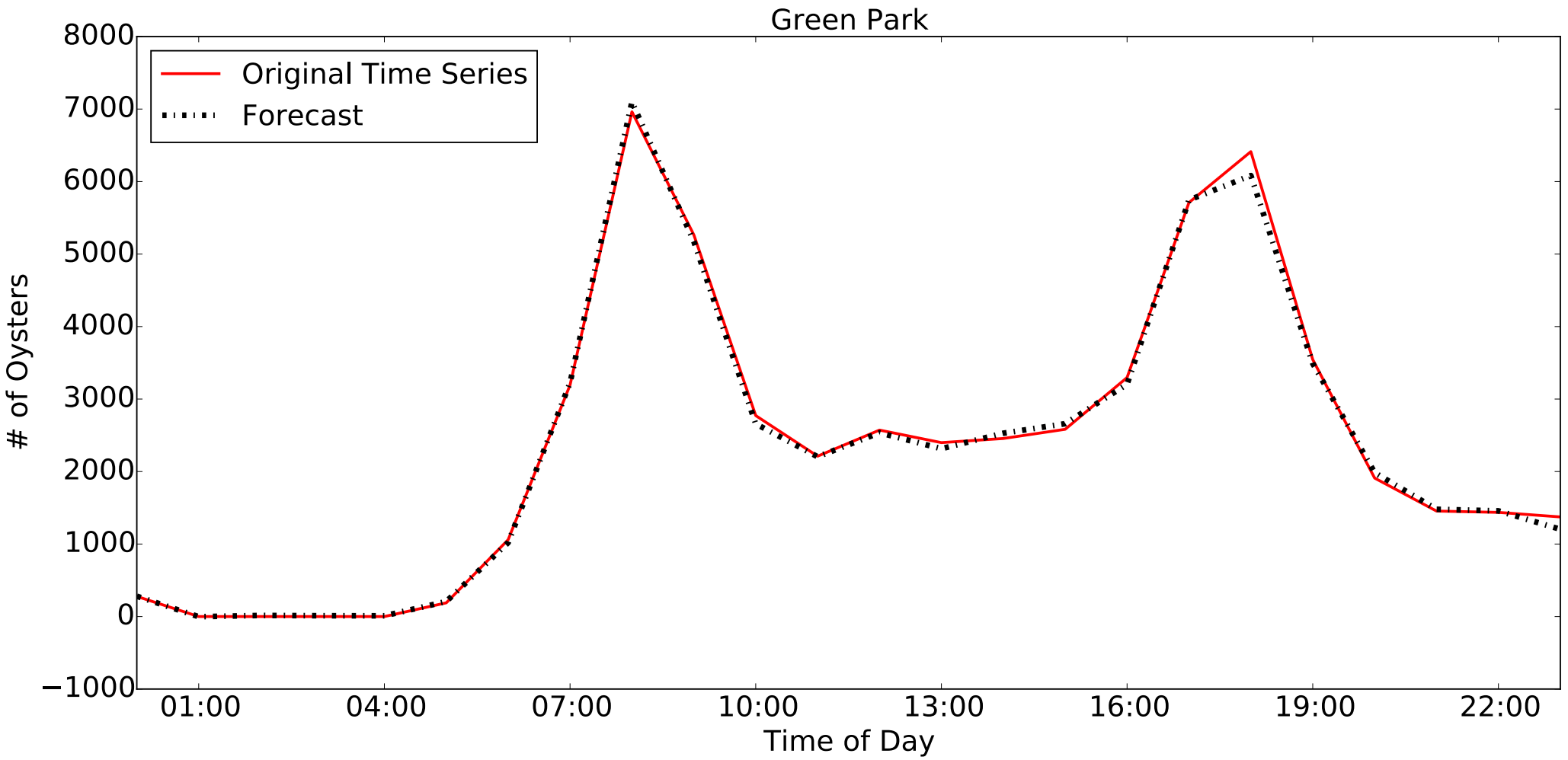


Removing Seasonality





Forecasting Traffic Volumes



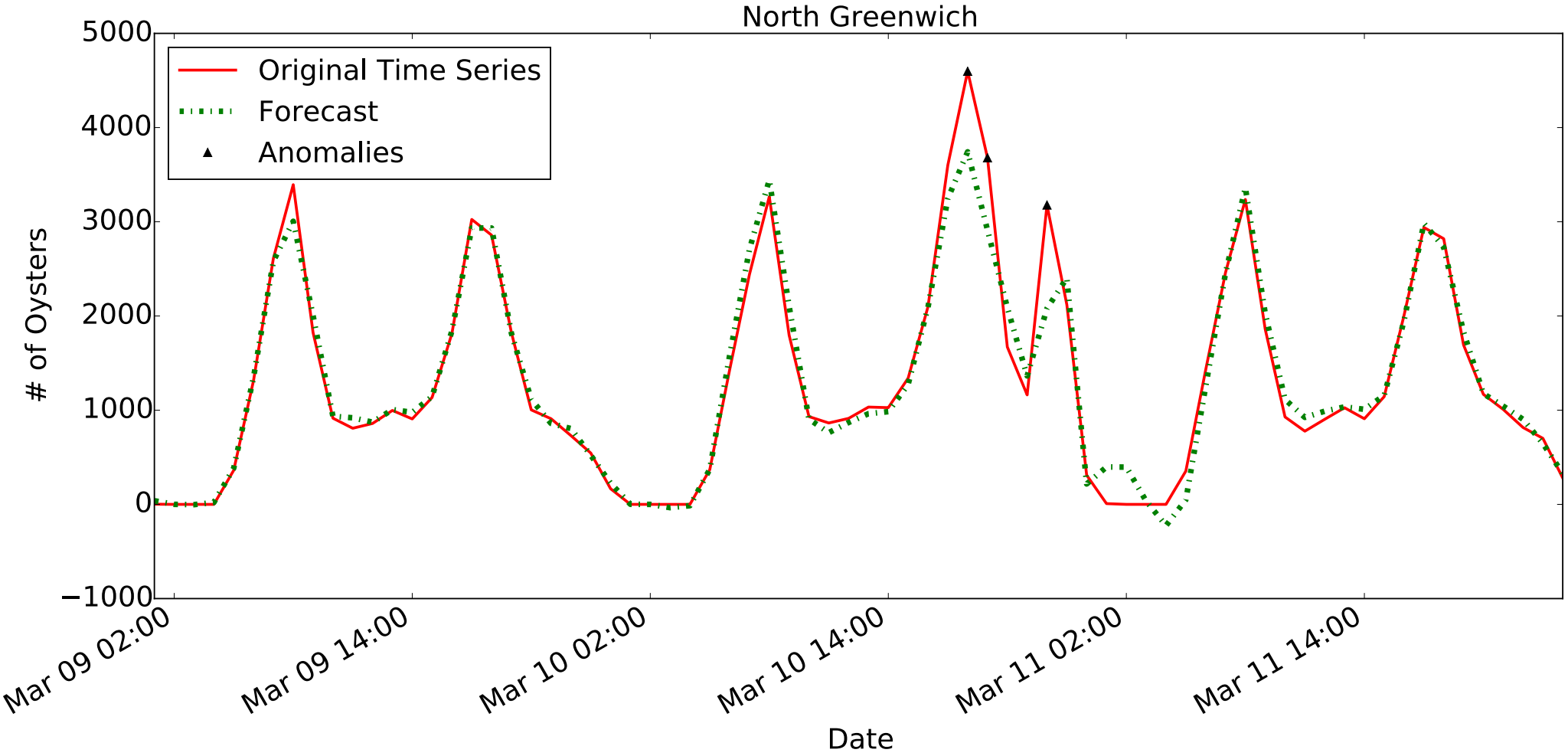
Detecting Anomalies

Train the ARMA model with 1 week data, test it against the rest of weeks

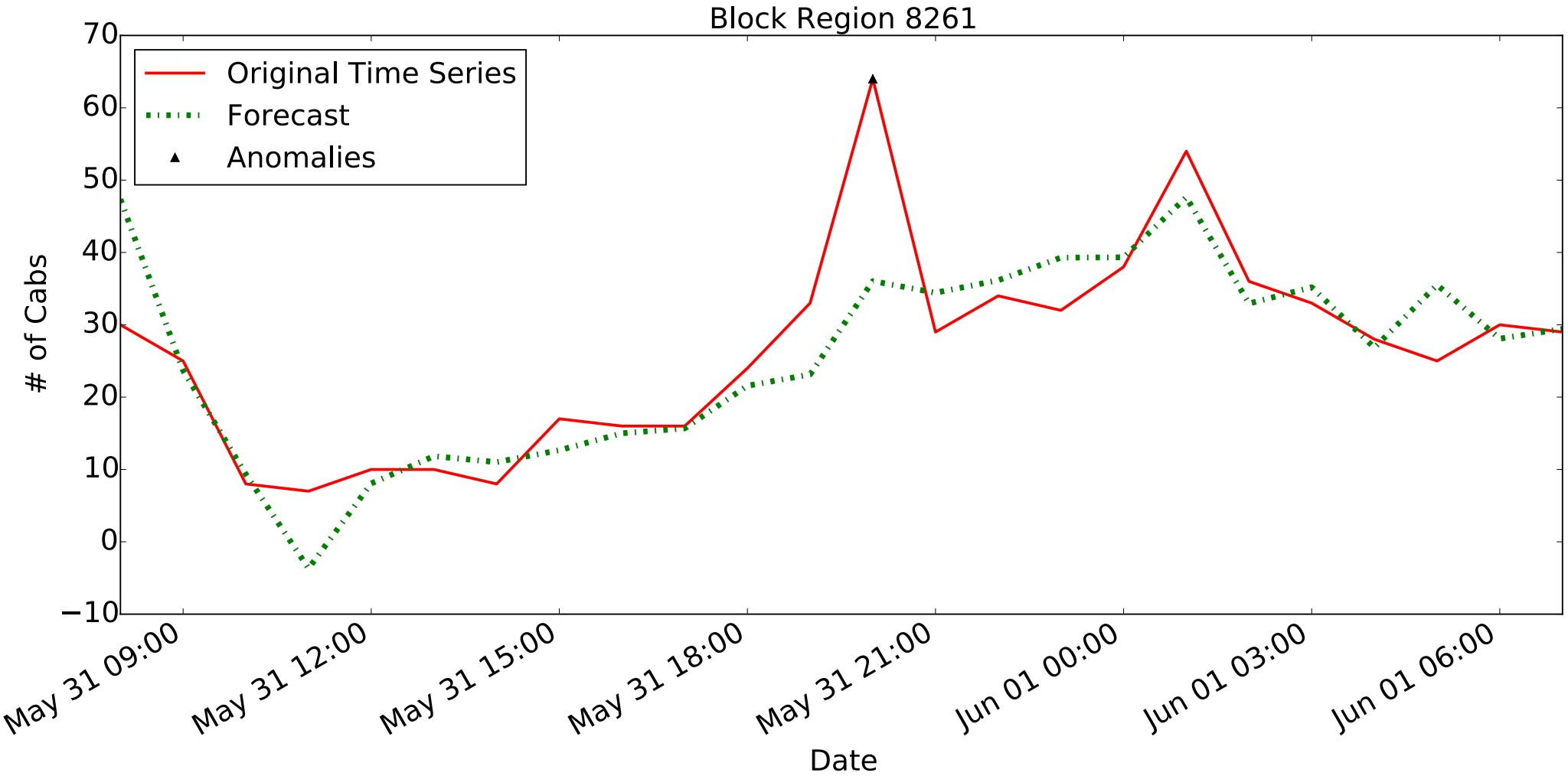
- Top 100 TFL stations: 896 anomalies
- Top 100 SFC blocks: 366 anomalies

No ground truth though…

Detecting Anomalies



Detecting Anomalies



Improving Prediction During Anomaly

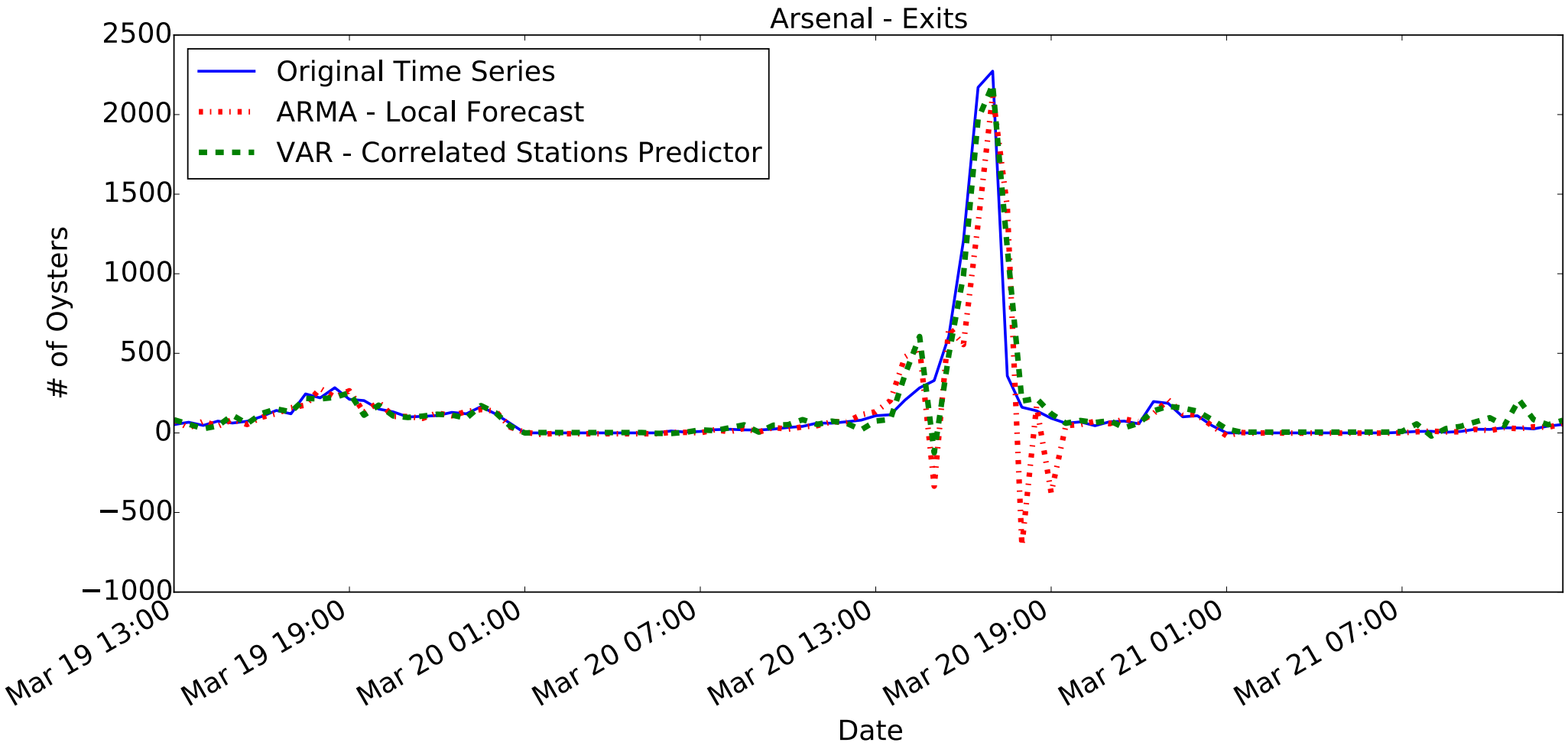
Improve predictions in the presence of an anomaly?

- Discover *correlated* ROIs by sliding their time series
- Use a VAR model to capture linear interdependencies between time series

Experiments

- Experiment with 10% of anomalies of TFL and SFC
- Train a VAR model w/ information from 10 correlated ROIs
- Compare against a baseline: ARMASEAS model trained on local data

Improving Prediction During Anomaly



Mobility & Privacy

Aggregation often considered as a privacy defense [NDSI'12, CCS'15, NDSS'16]

But do users lose privacy from the aggregates?

Differential Privacy (DP) to the rescue?

Add noise to the statistics to bound the privacy leakage

(Input or output perturbation)

The problem with DP…

Does it really tell us about the privacy loss?

Epsilon gives a theoretical upper-bound (indistinguishability)

How do we tune it? What does it mean in practice?

Other location privacy work

Geo-indistinguishability

Great, but only focus on a single user accessing location-based service (LBS)

Not sure how to apply it to aggregation

EPFL quantification framework

Great, but again, focus on single user accessing LBS

[Specifically, evaluate efficacy of obfuscation techniques]

Roadmap

Framework to reason about privacy from aggregates

- Adversary has a prior about the user, runs inference attacks to improve her prior using the aggregates
- Quantify the improvement (“privacy loss”)

Experimental evaluation on raw aggregates

- Using datasets obtained from Transport for London (TFL) and San Francisco Cabs (SFC)

Experimental evaluation on DP techniques

- Both input and output perturbation

Adversarial Goals

Profiling

- Infer the probability of a user being at a ROI at a certain time
- Performance: JS-divergence from ground truth
- Privacy loss: normalized improvement of JS with vs without aggregates

Localization

- Predict where a user is going to be at a certain time
- Performance: (1-F1 score), privacy loss: same as above

Timeline of the attack

- Observation period: used to build a prior
- Inference period: launch the attacks

Prior Knowledge

In real-life, some information about a user may be available to the adversary through social networks, data leaks, location traces released by providers, personal knowledge

In our work, we explore possible approaches…

Probabilistic priors: model knowledge of a user profile (probability distribution)

Assignment priors: knowledge of a user location (binary)

Probabilistic Priors

FREQ_ROI

Frequent locations over time

ROI_DAY

Most frequent ROIs for each instance of a day

ROI_DAY_WEEK

Most frequent ROIs for each instance of a week

TIME_DAY

Most frequent time instances of a day, reporting ROIs

TIME_DAY_WEEK

Most frequent time instances of a week, reporting ROIs

Assignment Priors

POP

Popular ROIs (above a certain threshold)

CEIL

All prior ROIs

LAST_WEEK/DAY/HOUR

Last week's, day's, or hour's ROIs

Inference Strategies

Bayesian Updating

Posterior probability of a user being at a certain ROI at a certain time given the prior and the aggregate

Max-ROI

Greedy strategy aiming at maximizing the total probability for each ROI by assigning most probable users to each location

Max-User

Greedy strategy aiming at maximizing each user's probability over the ROIs by assigning them to their most likely locations

Roadmap

Framework to reason about privacy from aggregates

- Adversary has a prior about the user, runs inference attacks to improve her prior using the aggregates
- Quantify the improvement (“privacy loss”)

Experimental evaluation on raw aggregates

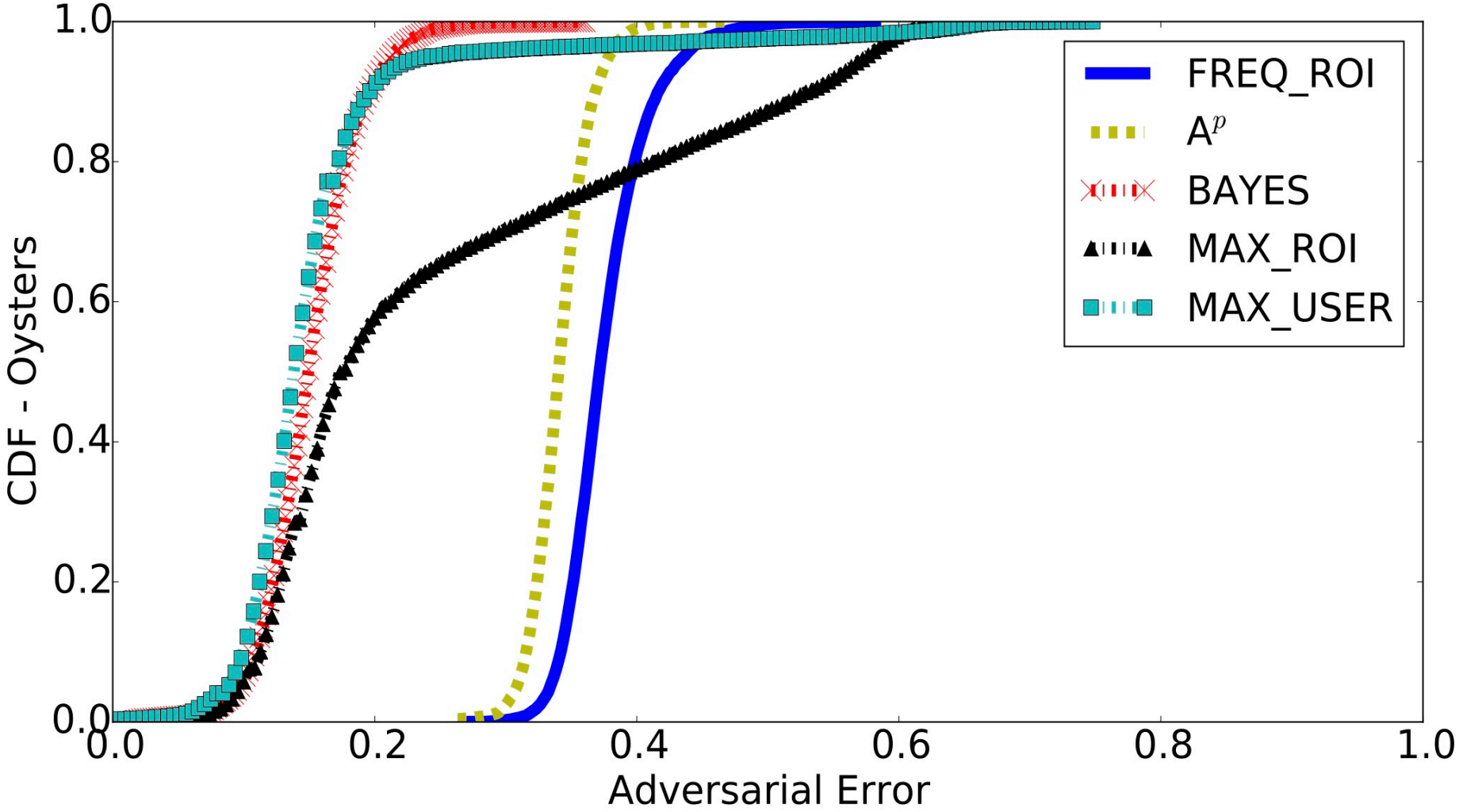
- Using datasets obtained from Transport for London (TFL) and San Francisco Cabs (SFC)

Experimental evaluation on DP techniques

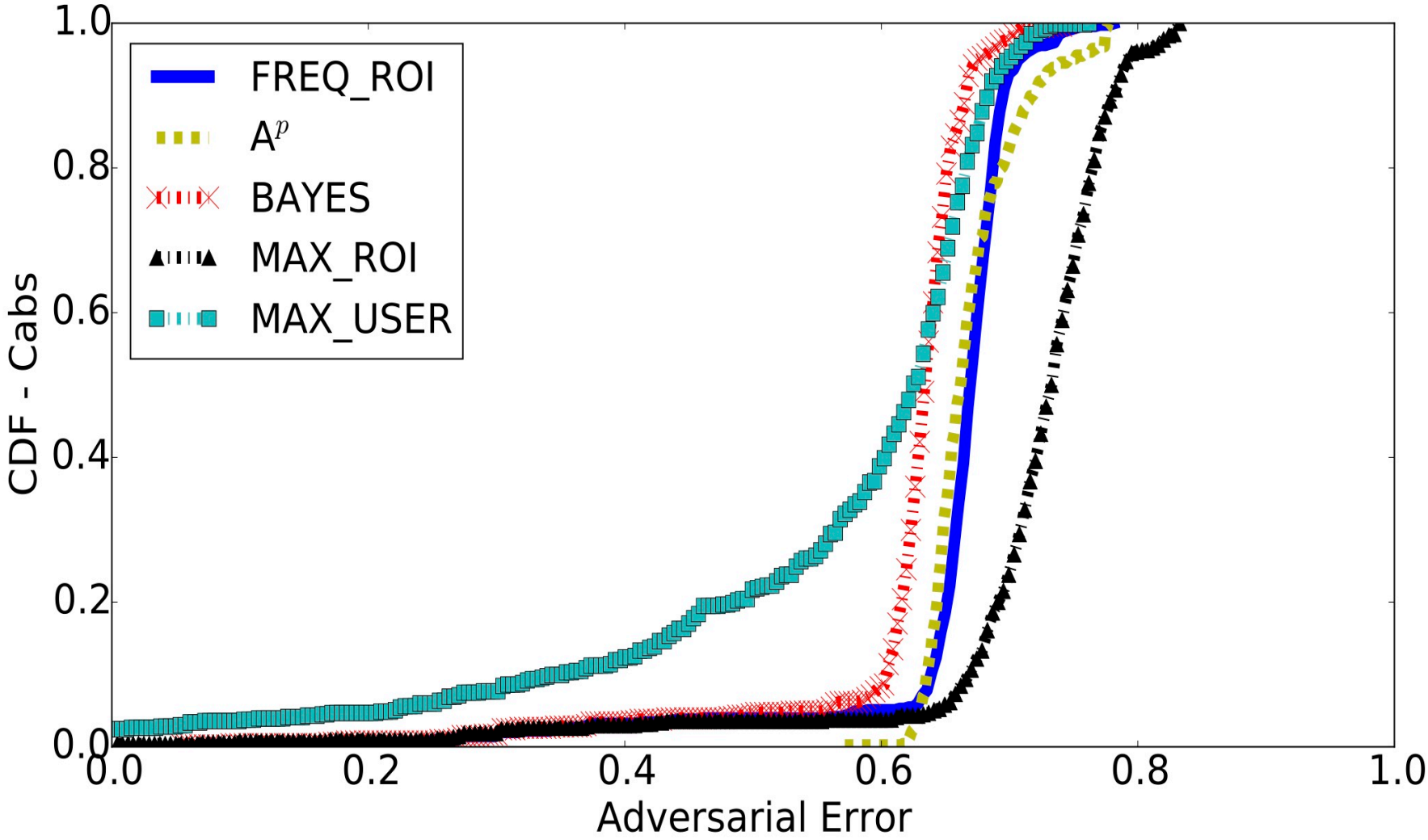
- Both input and output perturbation

Evaluation of Raw Aggregates

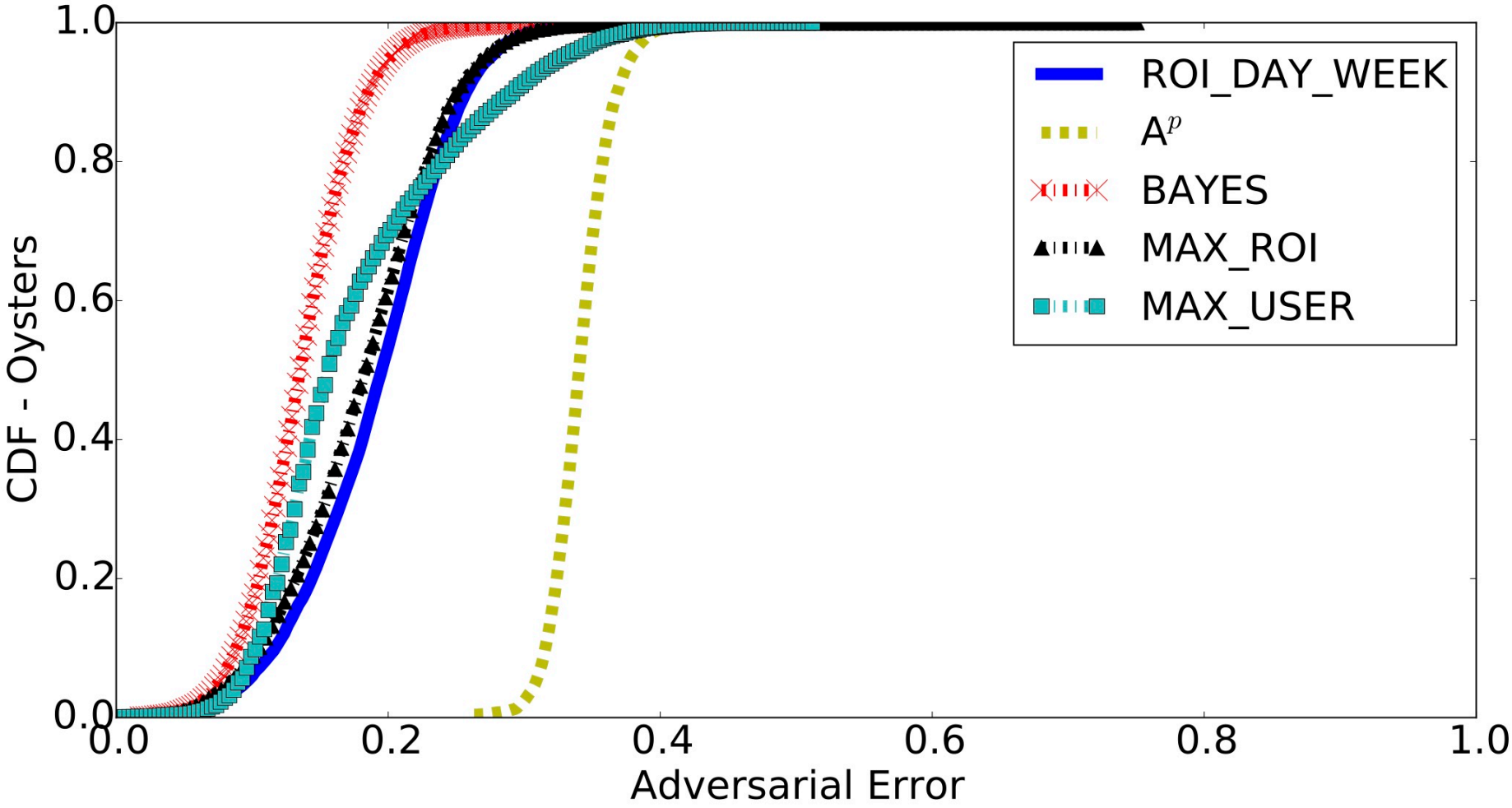
Profiling @ TFL - FREQ_ROI



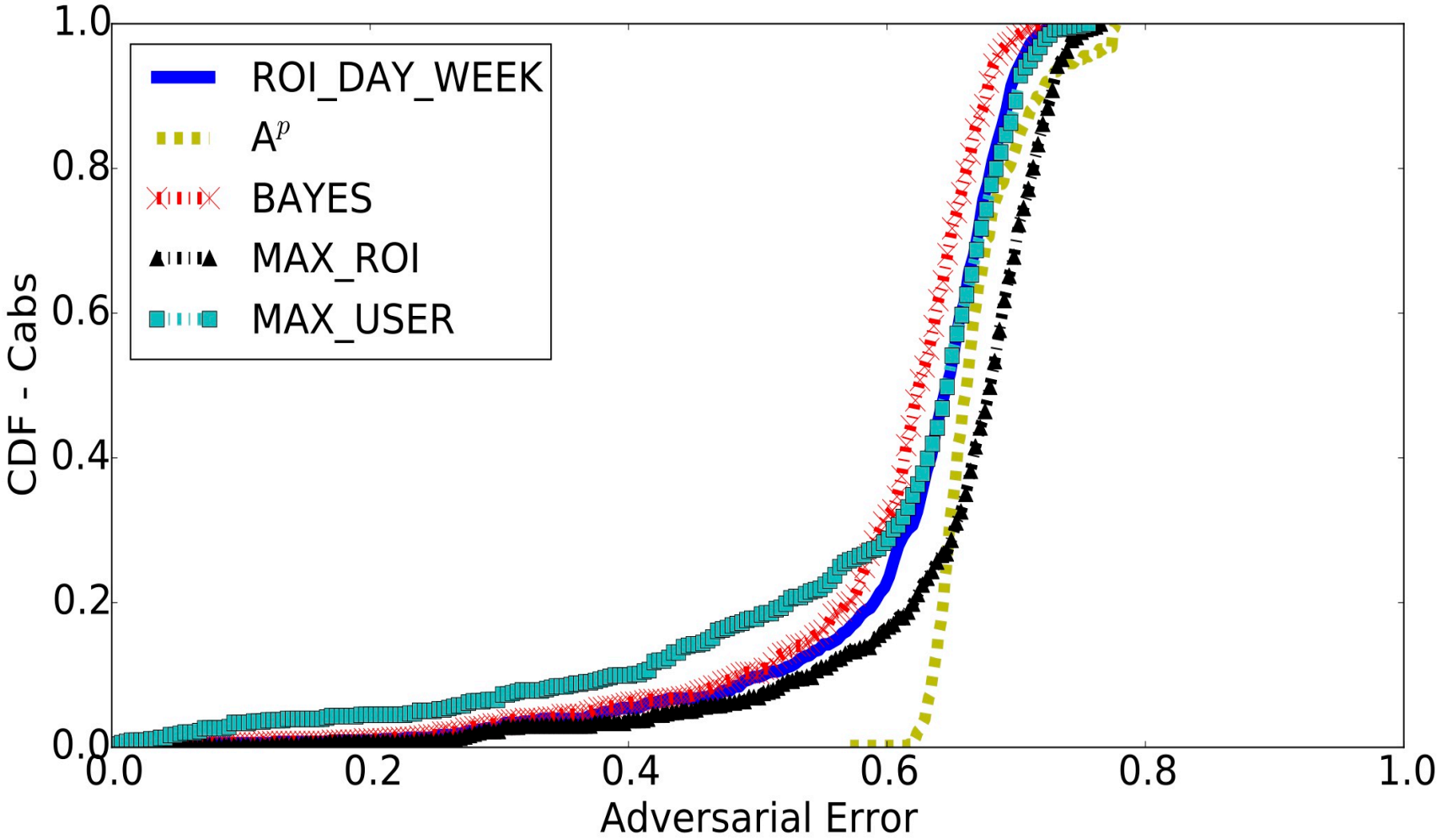
Profiling @ SFC - FREQ_ROI



Profiling @ TFL - ROI_DAY_WEEK



Profiling @ SFC - ROI_DAY_WEEK



Profiling

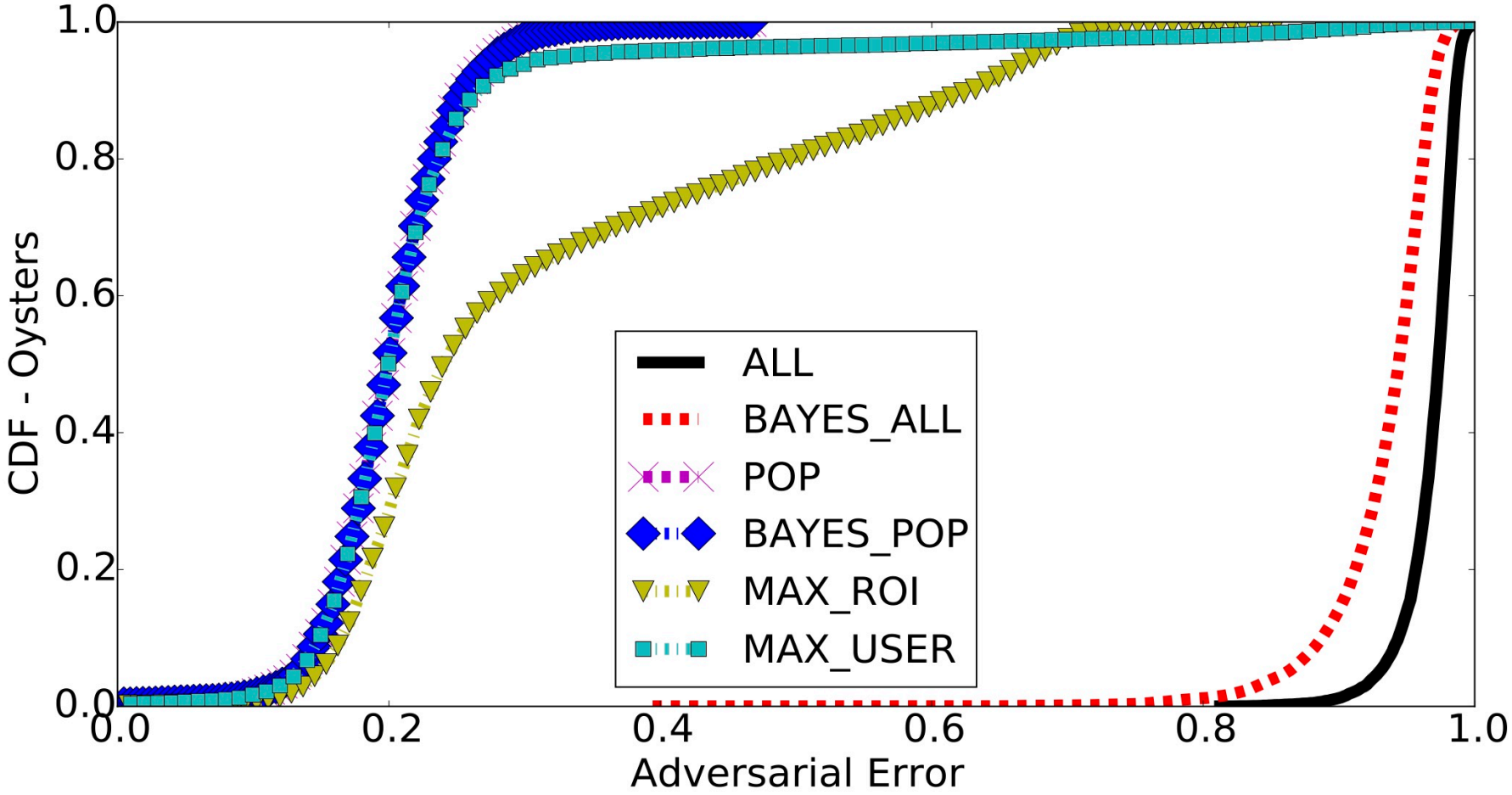
Overall aggregates do help the adversary to profile users

- Actual degree of privacy loss depends on the prior
- Assignment ones yield smaller privacy leakages, as they are already quite informative for the adversary compared to probabilistic ones

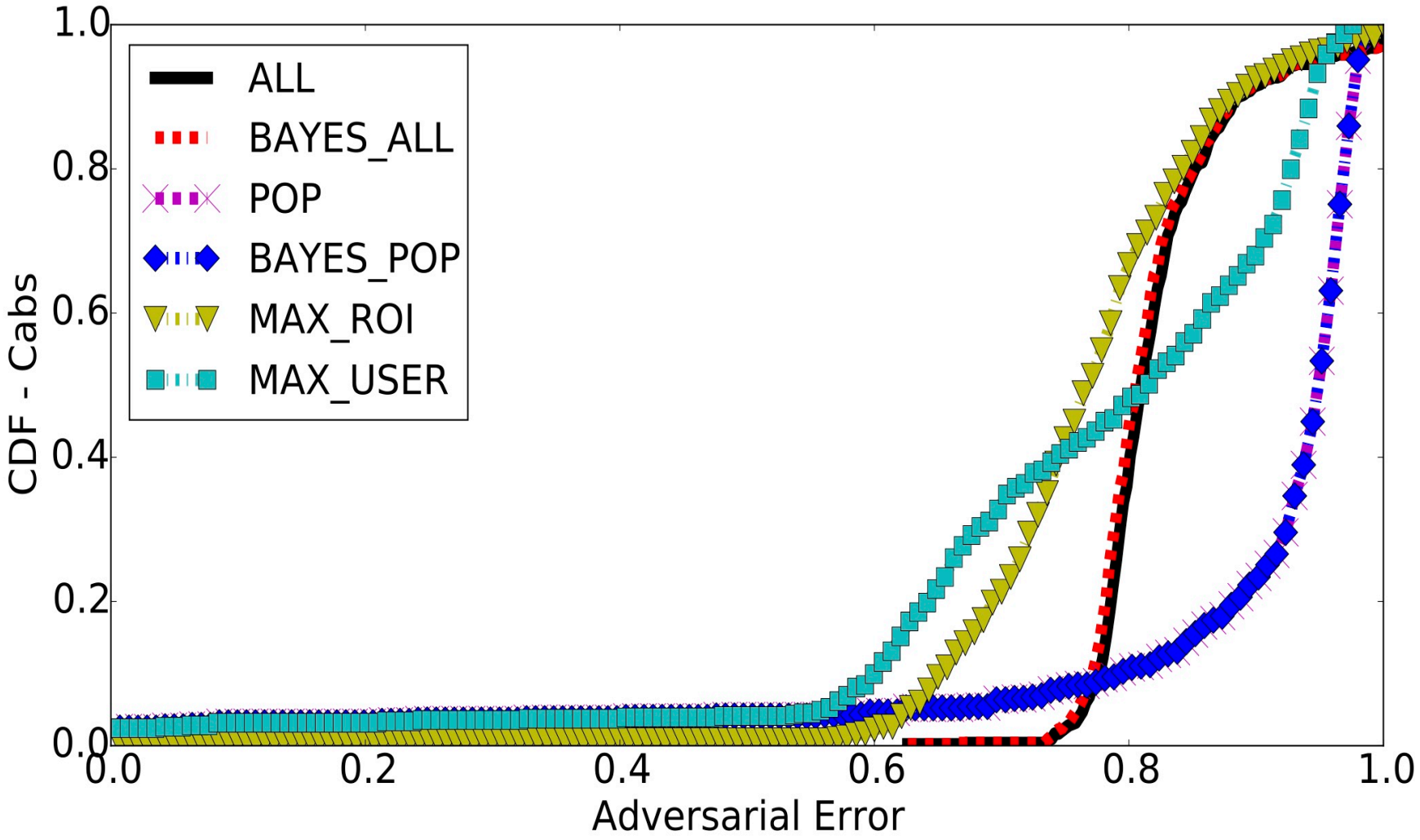
TFL vs SFC

- Inferring mobility profiles of commuters significantly easier than cabs, as cabs' patterns are not as regular

Localization @ TFL - FREQ_ROI



Localization @ SFC - FREQ_ROI



Localization

Aggregates enable our adversary to localize users

Assignment priors are (once again) more revealing and yield insignificant privacy leakage

TFL vs SFC: Commuters are best localized via their most popular ROIs whereas cabs via their last hour's ROIs

Roadmap

Framework to reason about privacy from aggregates

- Adversary has a prior about the user, runs inference attacks to improve her prior using the aggregates
- Quantify the improvement (“privacy loss”)

Experimental evaluation on raw aggregates

- Using datasets obtained from Transport for London (TFL) and San Francisco Cabs (SFC)

Experimental evaluation on DP techniques

- Both input and output perturbation

Evaluation of DP Techniques

How much "additional privacy" does DP provide?

Vis-à-vis epsilon and utility, of course

Privacy

We define **Privacy Gain (PL)** as the normalized reduction of the adversarial error compared to raw aggregates

Utility

Mean Relative Error (MRE)

Output Perturbation

Simple Counter Mechanism (SCM)

Straightforward extension of the Laplace mechanism

Only guarantees event-level privacy, i.e., protects whether or not a user was in a ROI at one specific time slot, but can be extended to provide stronger guarantees

Fourier Perturbation Algorithm (FPA)

Improves privacy/utility trade-off by reducing the amount of noise, specifically, performing the noise addition in the compressed domain

Input Perturbation

Randomized Response

Users report to be in a ROI with some probability p, or report the truth with probability 1–p

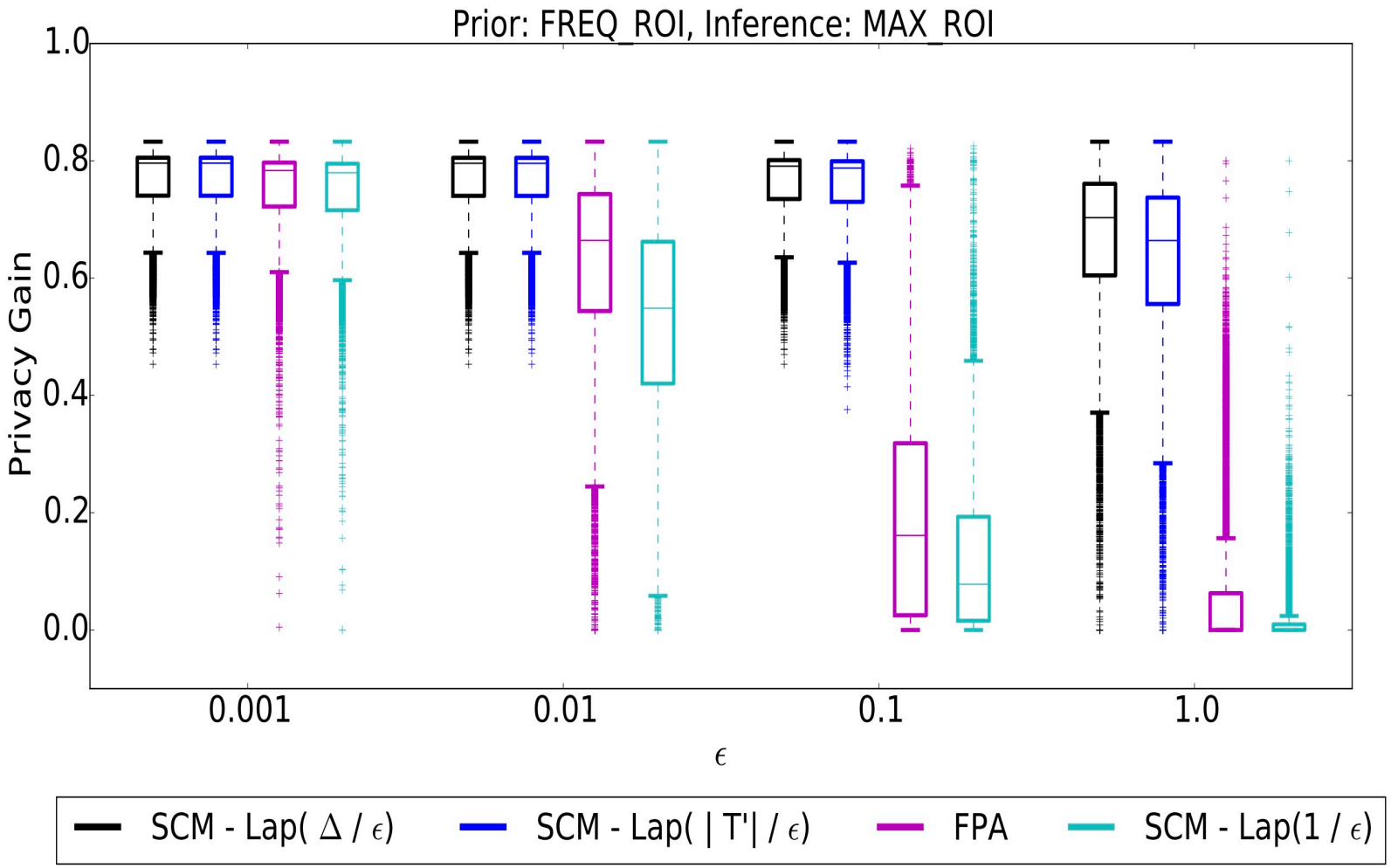
For aggregate location time-series, state-of-the-art is SpotMe

Output Perturbation – Utility

ϵ	0.001	0.01	0.1	1.0
SCM - Lap($\|S\| \cdot \|T'\|/\epsilon$)	739.9	743.2	735.8	709.4
SCM - Lap(Δ/ϵ)	720.1	605.1	168.9	16.7
SCM - Lap($\|T'\|/\epsilon$)	719.8	549.6	123.5	12.8
FPA	117.1	11.7	1.3	0.3
SCM - Lap($1/\epsilon$)	74.4	7.8	0.9	0.1

Table 3. TFL: MRE (Utility) of output perturbation mechanisms.

Output Perturbation – Privacy

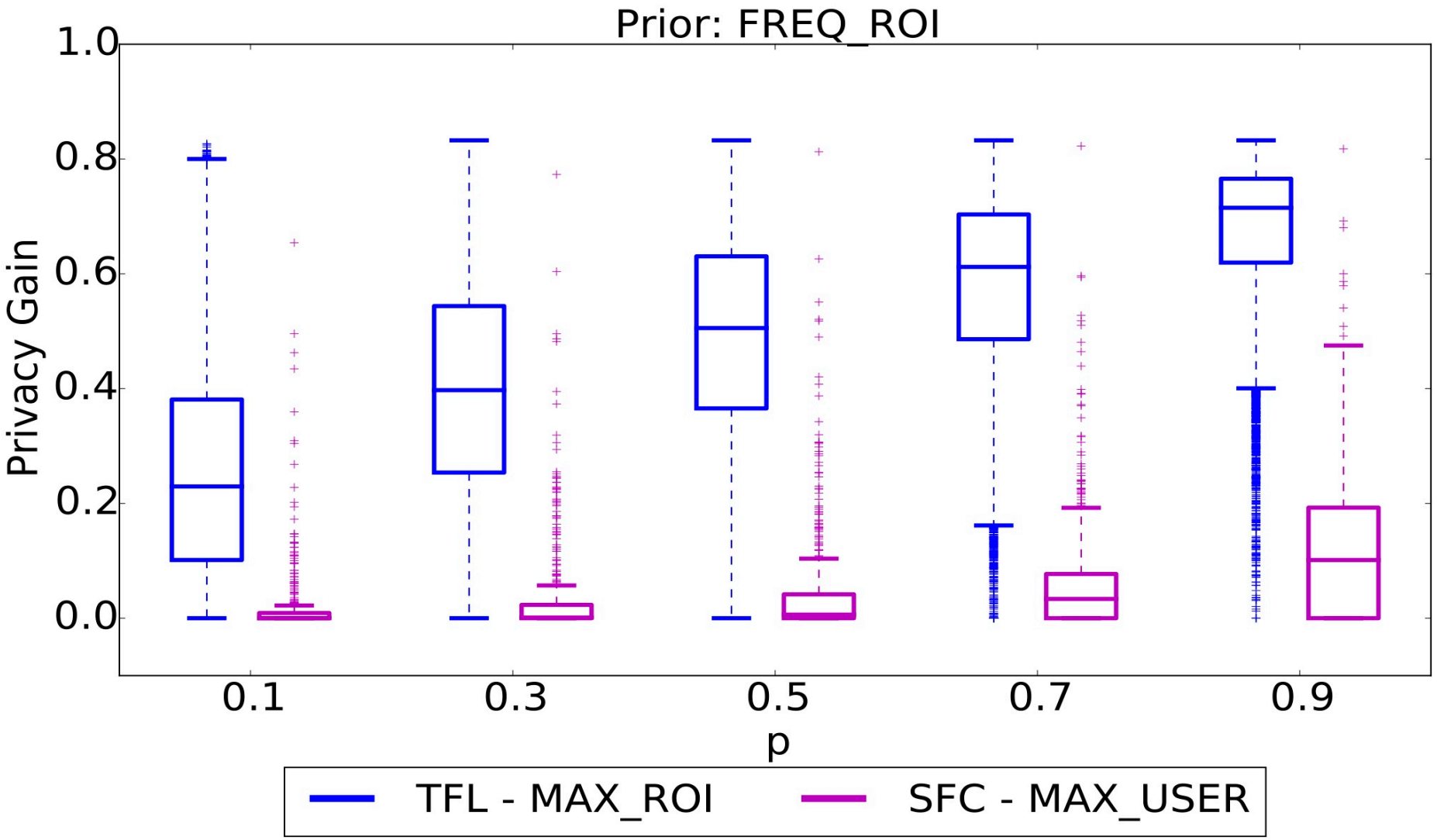


Input Perturbation – Utility

p	0.1	0.3	0.5	0.7	0.9
TFL - MRE	2.1	3.9	6.1	9.3	17.6
SFC - MRE	0.4	0.7	1.1	1.6	2.9

Table 5. SpotMe [36]: MRE (Utility) for increasing values of p, on TFL and SFC datasets.

Input Perturbation



Discussion

- Location aggregates enable an adversary with some prior knowledge to profile and localize users
- DP mechanisms should be carefully evaluated before application as they enhance users' privacy when the noise introduced destroys the utility of the time-series
- Membership inference attacks possible?
- Novel defense mechanisms for analytics on location time-series needed?

Thank you!