ON THE PRIVACY OF URBAN DATA APPLICATIONS

Dimitrios Tomaras [1], Vana Kalogeraki [1]

[1]Department of Informatics, Athens University of Economics and Business, Greece

1. Introduction

- Real-time urban data
  - Provide city-scale information (traffic, people activities, etc.)
  - People are capable of sharing systematically their activities in the urban space
  - Storing and processing data not sufficient for crowd apps
- Exploit urban data for human-centered applications
  - Mapping
  - Finding nearby businesses
  - Alerting
  - Real-time interactions with friends and business associates
  - etc.
- Address urban data privacy implications and their extent: A fundamental challenge while developing crowdsourcing apps and toward achieving smart city sustainability

2. Research Issues

- Analyze urban data to gain valuable insights
  - Users do not realize that continuously sharing their location and trajectory data with online systems may end up revealing a great amount of information in terms of their behavior, mobility patterns and social relationships
- Exploit urban data insights for preserving user privacy
  - Privacy preferences are subjective by nature: only a small percentage of the users of these systems realize the serious privacy implications that may arise and their extent
  - Develop a set of user tunable privacy techniques that exploit the mobility analysis insights in order to satisfy different levels of users privacy in urban data applications

3. Urban Data Analysis

- Smart Cities provide tremendous opportunities
  - For monitoring the city at city scale
  - For improving the quality of life of their citizens
- Urban Data: Location and Trajectory Data
  - Large amount of data generated by location-based social apps
  - Different sampling rates, volumes, scales, veracity
  - Provide precise modeling of human mobility in the urban environment
- People are willing to share their data with a cloud environment for:
  - Tracking family members and friends (e.g. friends trying to find each other in busy places such as shopping centers or parks)
  - Get rewards (citizens earning free parking time in smart cities)
  - Receive recommendations about places of interest etc.

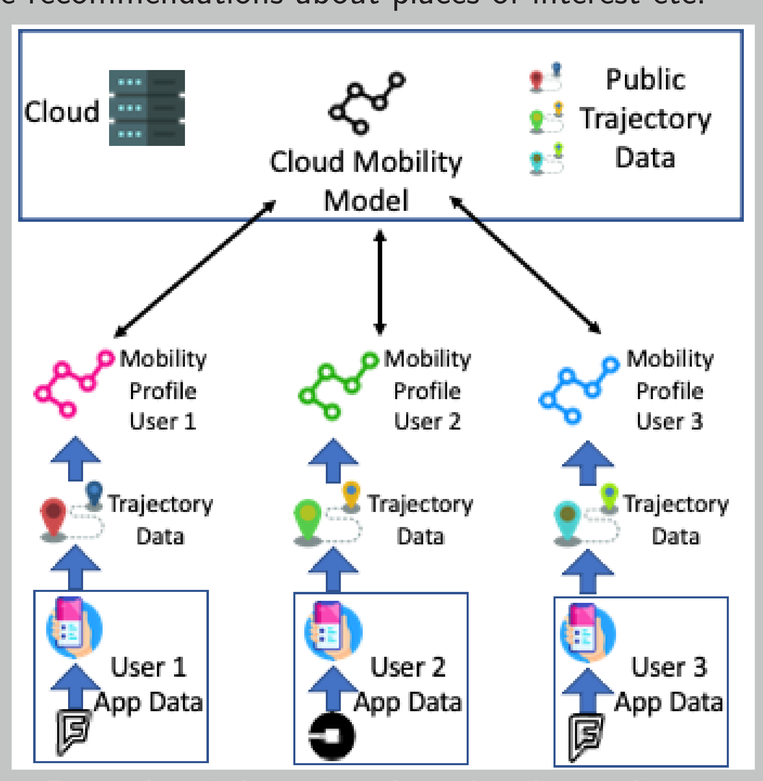
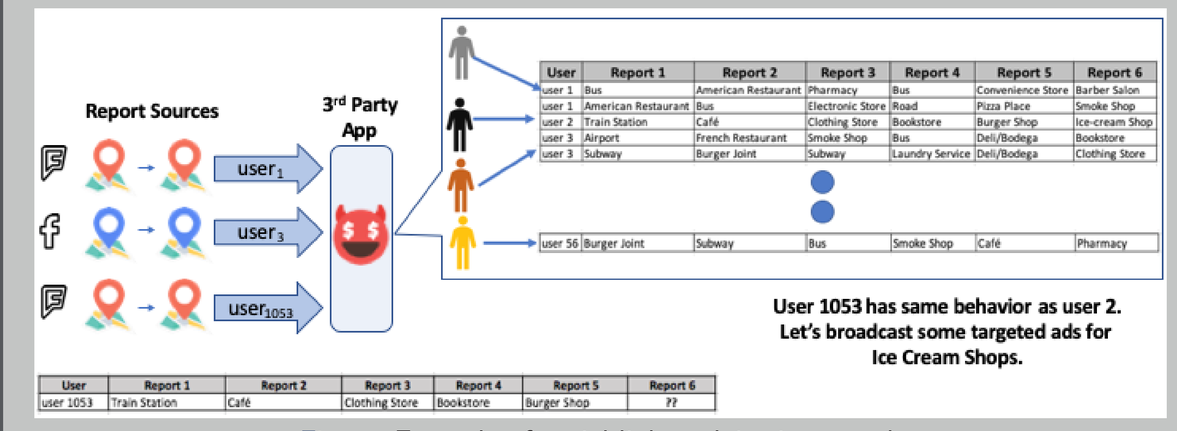


Figure: A typical overview of an urban data application

4. Understanding Privacy Implications

- Urban Data encapsulate users mobility patterns. An adversary can:
  - Extract social trajectory-based data to identify social ties among the users of these systems
  - Target *individual users* for marketing campaigns
  - Monitor user movements to compromise one's personal safety
  - Act on behalf of a third company: eg. an insurance company can evaluate the health status of the users and appropriately adjust their insurance rates



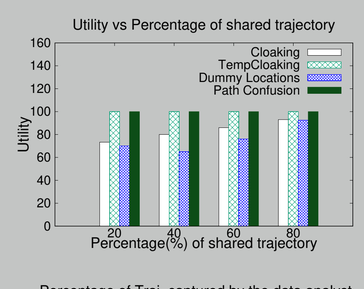
User	Report 1	Report 2	Report 3	Report 4	Report 5	Report 6
user 1	Bus	American Restaurant	Pharmacy	Bus	Convenience Store	Barber Salon
user 1	American Restaurant	Bus	Electronic Store	Road	Pizza Place	Smoke Shop
user 2	Train Station	Café	Clothing Store	Bookstore	Burger Shop	Ice-cream Shop
user 3	Airport	French Restaurant	Smoke Shop	Bus	Deli/Bodega	Bookstore
user 3	Subway	Burger Joint	Subway	Laundry Service	Deli/Bodega	Clothing Store
user 56	Burger Joint	Subway	Bus	Smoke Shop	Café	Pharmacy

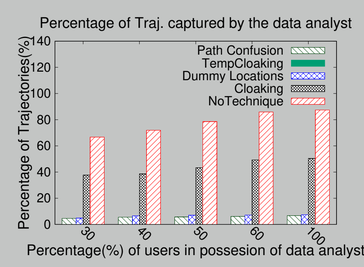
User	Report 1	Report 2	Report 3	Report 4	Report 5	Report 6
user 1053	Train Station	Café	Clothing Store	Bookstore	Burger Shop	??

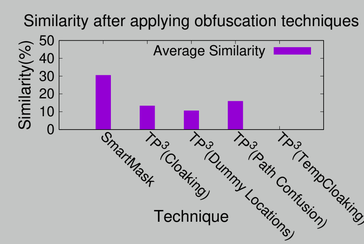
Figure: Example of social link exploitation attack

5. Privacy Techniques for urban data applications

- TP^3: Trajectory Privacy Protection in Practice
  - On-device model for user mobility
  - Captures social links that can be exploited
  - User tunable privacy preserving techniques for user trajectories
- On-device model
  - Aims at minimizing the amount of trajectory data that are stored in users phone
  - Performs computations on user's mobile phone and on the edge
  - Exploits locally stored social link data (e.g contacts on the phone, social graph of user etc.)
- Social link exploitation attack:
  - A third party data analyst can associate the user's patterns with groups of similar users while sharing trajectory data
- User tunable privacy preservation
  - Users can select across different levels of privacy preservation
  - 4 different privacy techniques tailored on their needs for accuracy in received results (e.g. recommendations)
- Evaluation
  - TP^3 outperforms competitors in terms of preserving trajectory privacy







Acknowledgments

- This research has been financed by the European Union through the FP7 ERC IDEAS 308019 NGHCS project and the H2020-ICT-688380 VaVeL project.

Contact Information

- Dimitrios Tomaras
  - Email: tomaras@aueb.gr
  - Web: http://www.aueb.gr/users/tomaras/
- Vana Kalogeraki
  - Email: vana@aueb.gr
  - Web: http://www2.cs.aueb.gr/~vana/
- Real-Time Distributed Systems Group
  - Web: http://rtds.aueb.gr