Inference of User Demographics and Habits from Seemingly Benign Smartphone Sensors

Manar Safi, Irwin Reyes, Serge Egelman
University of California – Berkeley & International Computer Science Institute

Introduction

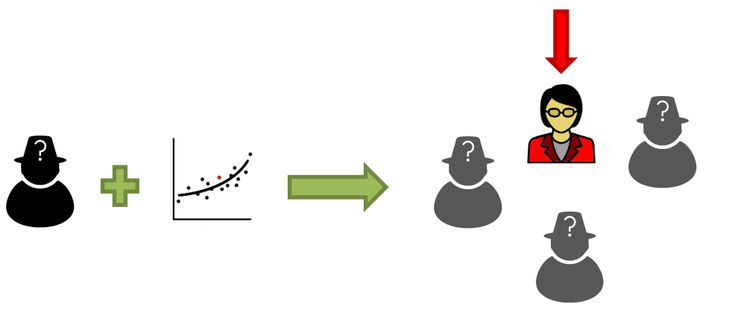
Smartphone permissions systems control access to private user data. [1,2]

Non-GPS sensors are not restricted.

The ad-tech industry builds individual profiles to better target ads [4] We explore methods to infer private data from these "benign" sensors.

Objectives

Research Question: Are sensor readings correlated with user traits? Is it sufficient for machine learning?



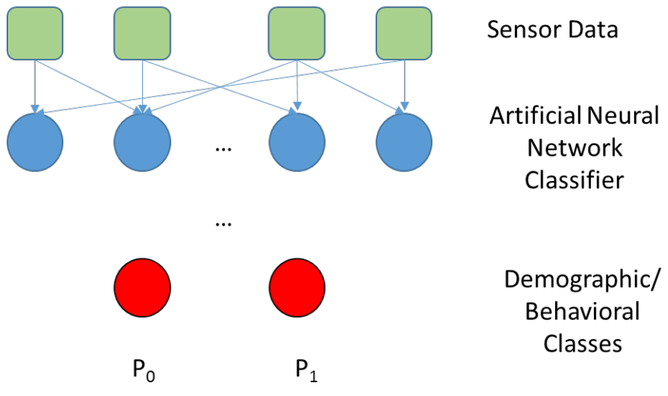
Gender	Weight	Fitness Habits
Income level	Height	Work Schedule

Methodology

Data Collection:

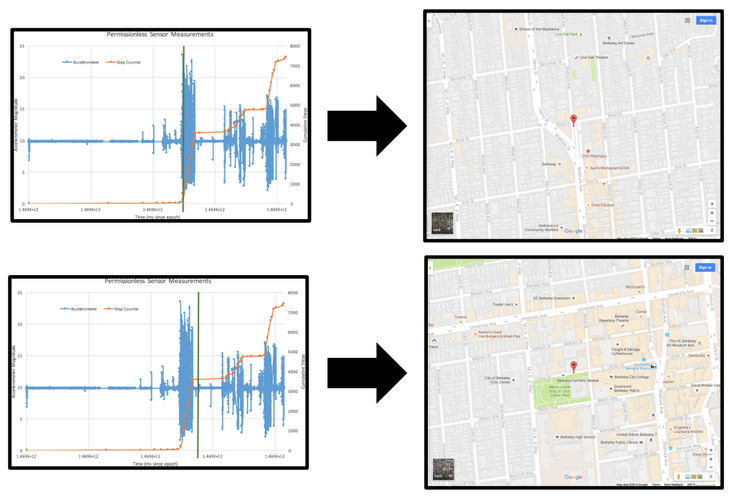
Population: MTURK, 1-week observation period, N = 100

Preprocess, segment, and classify



Preliminary Results

Intuition on Sensor Measurements



- Motion sensor activity correlates with one another (e.g., step counter activity reflected in accelerometer)
- Sensor activity + times could indicate habits, changes in location, etc.

Walking Motion and Gender

- Intuition: Men and women store their phones differently when walking (e.g., side pocket vs. back pocket vs. hand bag)
- Intuition: Different storage will have distinct motion characteristics

- Step counter data used to condition accelerometer readings to walking
- Limited analysis shows possible distinguishable features in time and frequency domains
- Needs broader investigation and sensitivity analysis

Conclusion

Key Takeaways

- Sensors can correlate with one another under various activities
- Conditioned on events, inferences about the user can be made
- Needs further investigation to determine generalizability

Questions For Further Analysis

- For gender inference using walking motion, how stable are the features within genders and across different walking sessions?
- What are the best ways to preprocess, segment, and formulate features from rich sensor data for demographic classification?
- Are there systemic differences in handset sensor hardware that can bias data and resulting inferences? Can those be leveraged?

References

1. Requesting Permission - Interaction - iOS Human Interface Guidelines. https://developer.apple.com/ios/human- interface guidelines/interaction/requesting-permission/
2. Working with System Permissions — Android Developers. https://developer.android.com/training/permissions/index.html
3. Sensor types — Android Open Source Project. https://source.android.com/devices/sensors/sensor-types.html.
4. How much is your personal data worth? - FT.com. http://www.ft.com/cms/s/2/927ca86e-d29b-11e288ed-00144feab7de.html#axzz4BDh1fipu
5. Y. Michalevsky, A. Schulman, G. A. Veerapandian, D. Boneh, and G. Nakibly. Powerspy: Location tracking using mobile device power analysis. In 24th USENIX Security Symposium (USENIX Security 15), pages 785–800, Washington, D.C., Aug. 2015. USENIX Association.
6. E. Owusu, J. Han, S. Das, A. Perrig, and J. Zhang. Accessory: Password inference using accelerometers on smartphones. In Proceedings of the Twelfth Workshop on Mobile Computing Systems & Applications, HotMobile '12, pages 9:1–9:6, New York, NY, USA, 2012. ACM.