Towards a Conceptual Model for Provoking Privacy Speculation

Norbert Nthala, Emilee Rader
Michigan State University – Behavior, Information, Technology Lab

Background

- Ubiquitous computing systems *combine and link data* from different sources to produce new *inferences* in service of the platform creators' goals.
- An **Inference** is data derived through a structured process of reasoning that draws on existing information recorded about someone or something and has a high probabilistic chance of being true.
- Most system users consent to sharing data with limited or no understanding of the range of data that is collected or can be inferred, and potential privacy implications thereof [1].
- Provoking speculation about data that is collected, and possible inferences can improve the knowledge , mental theories, privacy decisions, and behaviors of users [2,3].
- Speculation involves curiosity which prompts users to explicitly question system behavior and search for information to guide their understanding and decisions.
- Privacy speculation, therefore, creates an opportunity for users to understand possible uses of their data and to negotiate allowable use of it.

How Can Researchers Create Inferences that Provoke Privacy Speculation among System Users?

- We argue that privacy speculation can be provoked by creating inferences showing surprising and unexpected system behavior.
- For instance, data from *a vehicle diagnostics and location tracking device can be processed using different approaches and triggers to create surprising inferences*.

Approaches for creating inferences

- Combine various data types collected from a user. (*e.g. data collected by a vehicle diagnostics device)*
- Compare data from various sensors and users.
- Integrate data from disparate sources.

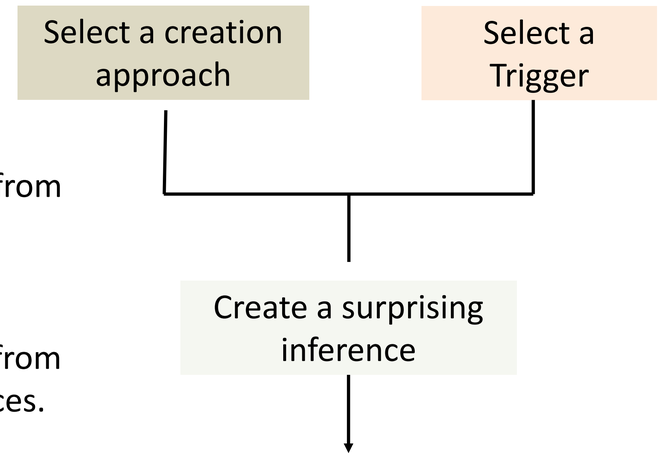

Fig. 1: Model for provoking privacy speculation

Triggers for Privacy Speculation

Frequency of recording data

We can infer how many, when, and to where a user made overnight trips away from home using location and timestamp data from the vehicle diagnostics device.

Retention period of recorded data

From longitudinal tracking of location data, we can infer if a user has or stays with a kid if s/he often visits K-12 schools or day-cares.

Kinds and amount of data recorded

Integrating location data with data from place APIs, we can infer the names and addresses of places where a user's relatives or friends live based on the user's weekend and holiday trips

Summary

- The model can be used to trigger privacy speculation among system users and study their expectations and reactions to various possible uses of their data.

References

1. Predrag Klasnja, Sunny Consolvo, Tanzeem Choudhury, Richard Beckwith, and Jeffrey Hightower. 2009. Exploring Privacy Concerns about Personal Sensing. In *Pervasive Computing*,
2. Emilee Rader and Janine Slaker. 2017. The importance of visibility for folk theories of sensor data. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*.
3. Ben Weinshel, Miranda Wei, Mainack Mondal, Euirim Choi, Shawn Shan, Claire Dolin, Michelle L. Mazurek, and Blase Ur. 2019. Oh, the Places You've Been! User Reactions to Longitudinal Transparency About Third-Party Web Tracking and Inferencing. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security (CCS '19)*.

Acknowledgements

This material is based upon work supported by the National Science Foundation Grant No. CNS-1524296.