On the Similarity of User-specific Features on Mobile Devices

Petr Matoušek, Radek Loub

matousp@fit.vutbr.cz, xloubr00@stud.fit.vutbr.cz

Faculty of Information Technology
Brno University of Technology
Czech Republic

Extended Abstract

Today, many people own one or more mobile devices, e.g., mobile phones, tablets, or smart watches, that they use for their work or personal usage. In case of cyber crime investigation, it is necessary to prove if the mobile devices that were involved in cyber crime activities belong to one person or not. This proof can be based on data obtained from these devices. In this paper, we propose a technique that is similar to mobile device fingerprinting. It is based on a selected set of features that represent personalized configuration and user-specific data stored on a mobile device. Unlike mobile device fingerprinting, we do not focus on hardware attributes but on user-specific data or personalized configuration that reflect behaviour and preferences of a user.

The idea of finding the similarity among mobile devices is based on assumption that each person uses a specific set of applications according to his or her personal priorities (e.g., web browsers, e-mail clients, VoIP or chat applications), prefers specific settings of a device, uses the same set of social networks with similar personal settings, and shares a contact list, phone directory, bookmarks, music/video player lists, etc. among all of his or her devices.

In this paper, we show what kind of user-specific data obtained from a mobile device is appropriate for similarity measurement. The data have to be hardware independent, transformed in a normalized form so that we can compare e.g., bookmarks from different web browsers, and have to hold high entropy so that they can be used to distinguish different users. Selection of our set of user-specific data is based on previous research [1-4], our own survey among mobile users who were asked what application, settings, or functions they usually use, and our own experiments. For each source of data a different comparison function that respects specific features of the data is employed. The result of comparison is expressed as a similarity function that sums partial results of comparison functions weighted by entropy of the data.

Our contribution includes specification of the set of user-specific data, a toolset for obtaining the data from mobile devices, computation of similarity, and evaluation of the approach. This presented approach is a part of the research of *Integrated platform for analysis of digital data from security incidents* supported by Czech ministry of Interior.

References

[1] Thomas Hupperich, Davide Maiorca, Marc Kührer, Thorsten Holz, and Giorgio Giacinto. 2015. On the Robustness of Mobile Device Fingerprinting: Can Mobile Users Escape Modern Web-Tracking Mechanisms?. In *Proceedings of the 31st Annual Computer Security Applications Conference* (ACSAC 2015). ACM, New York, NY, USA, 191-200. DOI: https://doi.org/10.1145/2818000.2818032.

[2] Kurtz, A., Gascon, H., Becker, T., et al. (2015). Fingerprinting Mobile Devices Using Personalized Configurations. *Proceedings on Privacy Enhancing Technologies*, 2016(1), pp. 4-19.

[3] Suranga Seneviratne, Aruna Seneviratne, Prasant Mohapatra, and Anirban Mahanti. 2014. Predicting user traits from a snapshot of apps installed on a smartphone. *SIGMOBILE Mob. Comput. Commun. Rev.* 18, 2 (June 2014), 1-8. DOI=http://dx.doi.org/10.1145/2636242.2636244

[4] Eubank, Christian, Marcela Melara, Diego Perez-Botero and Arvind Narayanan. “Shining the Floodlights on Mobile Web Tracking — A Privacy Survey.” (2013).