Poster: P^{2}Hub
Private Personal Data Hub for Mobile Devices

Michael Haus[†], Vittorio Cozzolino[†], Aaron Yi Ding[†], Jörg Ott[†]

[†]Department of Computer Science, Technical University of Munich
{haus, cozzolin, ding, ott}@cs.tum.edu

ABSTRACT

Mobile and wearable devices like smartphones or tablets are data hubs of our digital life and contain a high amount of sensitive data, which makes them a potential target for attackers. The aim of our P^{2}Hub approach is to consider the privacy-by-architecture principle directly during the system design phase. We enhance the isolation of sensitive private information through a privacy-preserving module supported by novel, lightweight virtualization techniques. Thus, we inherently improve the system's security and privacy.

1. MOTIVATION

Extensive privacy leakage is possible due to mobile devices with insufficient data control and management [**?**]. Almost every application uses sensible data in mobile ambient environments, such as location, contacts, login data, pictures, video and audio. This private information can be exploited to reveal the identity or other private information about the user. Besides that, there is a shift of paradigm in computer science, moving away from stationary PCs to more dynamic devices, which are lightweight and powerful enough to be carried and used everyday [**?**]. As a result, wearable devices like smartphones and tablets are the central gadgets of our digital life containing a high amount of personal data. For this reason, mobile devices need a better data protection.

Multiple studies report that most users are concerned about the protection of personal data in mobile devices [**?**]. The clear majority worry about stealing personal information or identity (84 %) and the overall loss of privacy (83 %) [**?**]. Especially important is the transparency and control of the data, 49 % would be more comfortable if able to control with whom sharing information and to what extent. Therefore, mobile applications should only be allowed to access data in a privacy-preserving way depending on privacy levels for family, friends, colleagues and social networks. On the other hand, users might underestimate the threats of exposing their sensible information, leading to the opinion that privacy is an unnecessary abstraction. Hence, we need more effective tools able to explain the effects of possible data leakages in terms of privacy breaches.

Developing resource-efficient, fine-tuned applications is a primary concern in the mobile domain. Especially, mobile and wearable devices are resource constrained in terms of memory, computing power and available energy. Therefore, we propose a new, lightweight mechanism that combines the isolation granted by novel virtualization techniques with our privacy-preserving data layer to enhance the information encapsulation and protection. Moreover, the virtualization techniques allows to easily offload tasks among devices optimizing the energy cost.

2. SYSTEM DESIGN

We propose a built-in privacy framework to secure mobile and wearable devices, such as smartphones, tablets, smart watches or smart glass. It minimizes the collection of identifiable personal data, focusing on anonymization and client-side data storage and processing. Figure **??** illustrates the system design. We identify two main components: one device including the P^{2}Hub, which securely stores the information while the other device utilizes it. The first component stores all personal data in one place in combination with on board processing to provide a central data access management. Moreover, it communicates with the other device by short range wireless technologies as Bluetooth Low Energy (BLE) and Wi-Fi. In addition, without the P^{2}Hub device, no mobile application is able to access personal information (stateless privacy storage). However, we must ensure that a set of basic functionality is available to the end-user. The P^{2}Hub includes two main layers to access sensible information: Privacy Layer (PL) and Data Layer (DL).

The PL adjusts the sensitive, raw data into privacy preserving data, only exposing the minimum required data accuracy to maintain sufficient service quality of mobile applications. To achieve this goal, the PL includes three different software components. First, the *User preferences*, an interface that explains the implication of the corresponding privacy setting to the user. Second, the *Privacy policy*, a self-adjusting control about which application can access personal data to what extent. Finally, the *Privacy preserving-method*, which processes raw data and generates privacy preserving data for mobile applications. These three software components build the PL, which receives metadata by the *Context model* for automatic adjustment of the respective privacy component. Moreover, the privacy model includes the following attributes: *Transparency*, which informs users about source and destination of data transfer, *Visibility*, which informs users about which application requested what data and *Accountability*, if an application got once permission, it is necessary to further analyse its behaviour to detect later malicious patterns.

The DL is a software layer dedicated to extract raw information from a generic source and storing them inside a local database, where the information can be subsequently fetched by the PL. The DL is meant to retrieve the information in a secure, resource-efficient way by exploiting the advantages of unikernels, such as MirageOS. Unikernels are single-purpose

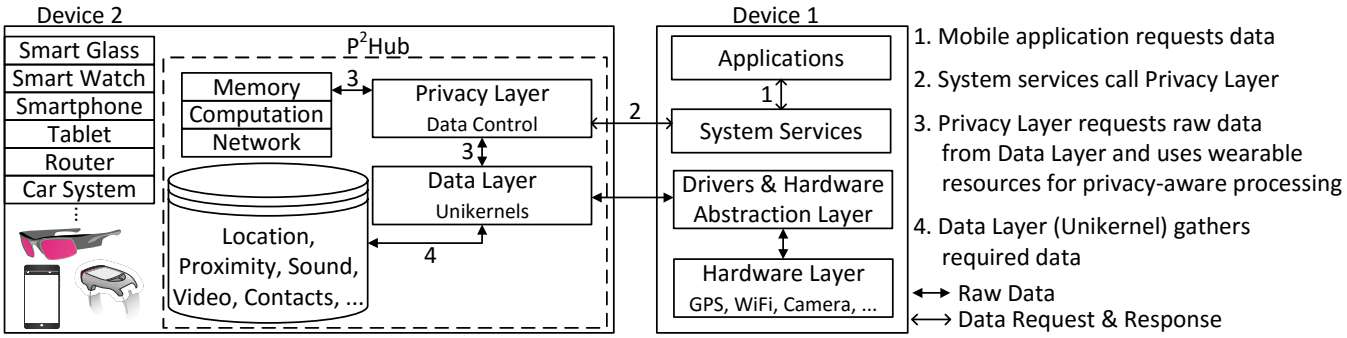

Figure 1: System design with processing flow. This approach is generic and can be implemented across various devices, such as smart glass, smart watches known as wearables or routers, tablets, smartphones, car systems, etc.

appliances that are compile time specialised into stand-alone kernels, and sealed against modifications when deployed [**?**]. Moreover, unikernels are lightweight and compact, making them perfectly transferable from one device to the other, allowing not only computational offloading but also functional ubiquity. Our approach integrates unikernels at the OS level and modifies standard system calls to encapsulate the sensible code inside a sand-box environment. Given the aforementioned characteristics, unikernels can be used to retrieve specific information from the mobile device guaranteeing a favourable trade-off between power usage and speed. The pre-compiled unikernels can be used on demand to fulfill a specific task independent of the executing hardware layer.

3. STATE-OF-THE-ART

Several approaches aim at building a personal data store (PDS) for better data control and security. The work of [**?**] proposed a framework known as openPDS to collect, store and manage third party access to personal metadata. However, it needs effort from the user to manage storage and data access to third parties and the design does not support user feedback. Haddadi et al. [**?**] shows a similar framework called Databox, which is a networked device that collects all personal data and provide data control and anonymization of sensitive information. These solutions do not address correlation-based attacks and unable to authenticate that the data request comes from an (un)compromised mobile application. In addition, Taintdroid detects inter-application privacy leaks and other recent tools, such as AppFence, Mockdroid and TISSA provide protection against any leakage of private data [**?**]. The Haystack system [**?**] aims at monitoring encrypted and non-encrypted network communication on mobile phones to provide insights, how the mobile applications handle private user information. But, so far the privacy-preserving mechanisms are not the focus of the work.

Most of these approaches have in common the usage of a network box at home, though it is impractical to carry it. Especially when we taking into account that the success story of mobile devices are mainly based on their high mobility to have it always at hand.

4. SUMMARY & OUTLOOK

While many other projects concentrate on monitoring data flows, P^{2}Hub is different: we propose a radical design to outsource the data management of private information on a software layer that can be implemented on different mobile devices. This approach enhances the data control, which is a major requirement in multiple studies regarding data privacy. Moreover, our approach decouples the data from the mobile device directly at the lowest level by virtualizing some operative system's functionalities. The user has even physical control. For instance, by disconnecting the P^{2}Hub, nobody is able to access the data.

Our system design aims to minimize energy consumption by exploiting unikernels and, at the same time, ensuring task isolation, which lowers the risk of potential attacks. Moreover, the benefits include small memory footprints and tiny computational requirements.

In the next steps, we will focus on the implementation of the Privacy Layer and the Data Layer, including the interface design and development of a mock-up Android application. Afterwards, evaluating the system from different points of view: computational overhead, energy consumption, memory occupation and privacy quantification in trade-off to service quality. Consequently, we will compare our metrics with other existing solutions.

5. REFERENCES

[1] Y.-A. de Montjoye et al. *openPDS: Protecting the Privacy of Metadata through SafeAnswers.* PLoS ONE, Vol. 9, No. 7, 2014.

[2] H. Haddadi et al. *Personal Data: Thinking Inside the Box.* arXiv, 2015.

[3] A. Razaghpanah et al. *Haystack: In Situ Mobile Traffic Analysis in User Space.* arXiv, 2015.

[4] M. Haris, H. Haddadi and P. Hui *Privacy Leakage in Mobile Computing: Tools, Methods, and Characteristics.* arXiv, 2014.

[5] Microsoft' *Location Based Services Usage and Perceptions Survey Presentation.* 2011.

[6] J. M. Urban, C. J. Hoofnagle and S. Li *Mobile Phones and Privacy.* UC Berkeley Public Law Research Paper, 2012.

[7] A. Madhavapeddy, et al. *Unikernels: Library Operating Systems for the Cloud.* ACM SIGPLAN Notices, Vol. 48, No. 4, 2013.

[8] V. Woods and R. van der Meulen *Worldwide Device Shipments to Grow 1.9 Percent in 2016.* http://www.gartner.com/newsroom/id/3187134, 2016 (visited on 26.04.2016).