Privacy-aware: data dissemination in mobile networks

Romain SOMMERARD

February 3, 2016

Abstract

Sometimes, scientists need a large amount of data to make experimentations. Because data can validate or not a research, they are fondamental for scientists to make the best work. It exists many ways to collect data like going in the street and ask questions to people to take their point of view. This example works if you have time to spend for it and if the data size is not huge. When you need a lot of data, this approach is not possible.

For data coming from the mobile world, it exists APISENSE. This is a platform which provides mobile device data from users and send it to researchers. It is very useful because scientists can configure the kind of data they want. To reach their goal, APISENSE provides forms, collects devices data like GPS locations, sensors data, etc ... Users subscribe to campaign and send their data. Of course, they agree with campaign rules and accept terms and conditions.

Sharing data for scientific studies is a good act but, as a human, users want to be sure of data anonymity. Because privacy is a big issue today, they want warranty that data don't reveal informations about them. It is relevant to all data which can be collected (e.g., health data, location data, etc ...). Generally, users send their data directly to the end server. These data can contain personnal informations and the system needs to anonymize it after receiving it. The problem is that users have no warranty about anonymization. They don't know if server makes it for real. Moreover, we can easily imagine that system can be hacked, thus leaks users data.

A solution to this problem is to use a collaborative approach where each user work with other to anonymize data before sending it to the server. Because data anonymization is made before, the user privacy is kept and we can make sure that no data can be leaked or misused.

We produce a mobile software for the Android platform. It is used to exchange data between devices by a collaborative approach. The first prototype bases its interest on exchanging data between devices through Android peer-to-peer protocol to hide and keep secret the data producer. In that way, the server receives data but can't know where does the data come from. It doesn't know which device has producted the data. This tool could be used to test and compare different approaches and anonymization methods. It could improve this research area and help companies to build software that include these principles in their applications. Companies could apply it to take care of users privacy and keep anonymity and don't expose informations in case of hack.

The solution could be improve with different anonymization methods (e.g., using the TOR protocol combined with k-anonymity algorithm). This work

is the first step of mobile data anonymization which make anonymization a collaborative work. With it, privacy could be improved for the benefit of people and make the technologic world a better place. Sharing data can be safe and easy, and the benefits are for science.