Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 129 (2018) 331–340

Procedia Computer Science

www.elsevier.com/locate/procedia

2017 International Conference on Identification, Information and Knowledge in the Internet of Things

A Privacy Preserving Data Aggregation Scheme to Investigate Apps Installment in Massive Mobile Devices

Tianhao Mao[a], Chang Cao[b], Xiangru Peng[c*] ,Weili Han[b*]

[a]*School of Computer Science, Fudan University, 825 Zhangheng Road, Shanghai, 201203, China*
[b]*Software School, Fudan University, 825 Zhangheng Road, Shanghai, 201203, China*
[c] *The Third Institute of Public Security, 339 Bisheng Road,Shanghai, China*

Abstract

Currently, mobile devices often try obtaining the list of installed applications to survey the install ratio for advertisers and app owners to evaluate the market share of their applications. However, adversaries could yet infer or even directly learn the user's interests and personal information from his/her mobile application installations. Furthermore, this investigation, which could lead to user privacy leakage, now calls for more attention than ever before. Motivated to resolve the above issue, this paper presents a scheme to investigate the install ratio of mobile applications for app owners and advertisers while at the same time preserving the users' privacy. The scheme is executed with one round communication between a server and mobile devices. According to experiment results, the proposed scheme can balance privacy protection and publishing the application install ratio accurately. Furthermore, the proposed scheme also satisfies differential privacy. This paper has also given the curve fitting of the amount of mobile devices and a privacy budget. It may instruct users how to choose a privacy budget for a given amount of devices to investigate.

Selection and peer-review under responsibility of the scientific committee of the 2017 International Conference on Identification, Information and Knowledge in the Internet of Things (IIKI2017).

Keywords: Privacy Preserving Data Aggregation; Differential Privacy; Mobile Applications

* Corresponding author. Tel.: +86-021-5135-5388
E-mail address: wlhan@fudan.edn.cn, pengruxiang@stars.org.cn

Selection and peer-review under responsibility of the scientific committee of the 2017 International Conference on Identification, Information and Knowledge in the Internet of Things (IIKI2017).
10.1016/j.procs.2018.03.084

1. Introduction

With the popularity of mobile phones, most of the users are using mobile applications and become dependent on them [1][2][26]. The investigation of the installation ratio for some applications has been an important procedure in the design and marketing of applications [3].

App owners usually want to investigate installation ratios of different types of applications, so that they evaluate which type of applications should be developed. Advertisers also want to check the effects of the advertisements in applications [4][5]. Then they can push personalized advertisements to users to promote products. For example, for users who installed a lot of games, advertisers could push game ads to them, while for users who installed applications about stock; advertisers could push ads about stock and investment to them.

Getting an application installation list is one of the common Android applications' actions [6]. Using the package management in development, developers can easily design codes to get all the information of installed application on a mobile device [7].

Many users, however, concern their privacy where adversaries or attackers may infer or even directly learn from the information of app installation. A certain amount of personal information could be inferred from the application installation list of someone's mobile devices. For example, for users installed an application which is used to provide sex trade (Good2go[1]), it could be speculated that he is the consumer of sex service. For users installed application about religion (Halal Guide[2]) or homosexual social (Blued[3] /spicy love[4]), they may not want others to know their belief or sexual orientation. Applications about specific theme or club also expose its users' hobbies and interests (Boqii Pets[5] /Gapday[6]). Attackers can use these disclosed privacies for social engineering and prepare for further attack. Advertisers can also push advertisements according to user privacy from installation list of applications. Therefore, some users are reluctant to accept the requirement of acquiring application list on the mobile device because of the risk of privacy divulgation [8]. Hence there is the requirement of a mechanism to balance application installation ratio measurement and privacy protection.

This paper presents a scheme of privacy preserving data investigation of mobile application installations. The mechanism to implement the proposed scheme uses a server to execute the investigation of all the mobile devices' application lists. Advertisers and app owners have access to the statistical results produced by the server. The data procession algorithm executes on mobile devices. So, it is hard for the server and eavesdroppers to get the raw data to speculate which application was installed for any user even when they have data which clients have posted to the server. The mechanism can protect user privacy, while at same time it provides the accuracy of statistics.

2. Related Works

2.1. Privacy Preserving Data Aggregation

Privacy preserving data aggregation requires that the aggregator should get the sum of the data but should not get the data of each user. It is because that users' personal data could be used to infer user privacy.

Usual schemata to achieve this target are based on cryptography, such as encrypting these data and then transmit them. The literature [9] gave the cryptography definition of privacy preserving aggregation of meter data. It put forward a scheme using trusted third party to add random noise to achieve the privacy protection on cryptography. The literature [10] used a crypto system to achieve private protection by using symmetric cipher. It also achieved data integrity by asymmetric cipher to realize digital signature. The literature [11] combined the advantages of two schemes before, by adding noise with Gamma distribution to satisfy differential privacy, and then using

[1] https://play.google.com/store/apps/details?id=com.good2go
[2] https://play.google.com/store/apps/details?id=com.unated
[3] https://play.google.com/store/apps/details?id=com.blued.international
[4] https://play.google.com/store/apps/details?id=com.jaumo.lesbian
[5] https://play.google.com/store/apps/details?id=com.massvig.ecommerce.boqi
[6] https://play.google.com/store/apps/details?id=com.gapday.gapday