A kind of User-centric fine-grained privacy protection management mechanism with location address transformation based on semantics

Nan Xiao

Xi’an Jiaotong University

xiaonanx@stu.xjtu.edu.cn

Keywords: Location; Privacy; Android; Geo-Indistinguishability; Laplace distribution

Abstract. This paper designs a fine-grained location privacy preserving system based on semantics. It is divided into two major modules which are privacy rule setting module and coordinate transformation module. The relationship between these two modules is that the users can use the rule language to express its privacy rules, and when the privacy rule is in the protection mode, the users use coordinate transformation method to protect the their true position information.

Introduction

As there are more and more mobile apps, in the course of using the mobile terminal, users inevitably leak part of privacy, such as home address, company address, etc. The existing mobile operating systems (for example, Android) are lack of the protection of privacy, and users can only protect their privacy by giving the applications permission. But for such applications with the user’s permission can easily have access to the user’s privacy information especially the location information such as home addresses, therefore, only have applications permission managed is not a effective way for protecting privacy.

In order to protect users ' privacy, this paper talks about a user-centric Android. location privacy management system, allowing users to customize privacy management strategies for their own needs. Users need to add a personalized program between the choice of "Yes" or "no" in order to get privacy protection at a certain time, a certain place and a certain scenario. Personalized program requires the user to enter the agreed format of the program rules and mark the area on the map for their protection need.

This paper realizes the user-centric privacy management mechanism on the Android platform. And the Software designs the ways of how users mark the areas in need of protection on the map, how rules are set, how to use the rules to mark out the area to protect privacy. Geo-Indistinguishability model belongs to the differential privacy. Its core idea is to join Laplace noise in the user's current position information. Since the actual application needs a limited position, it is necessary to carry out discretization of the continuous plane. Finally, it will remap the noise point to the user's nearest interest points.

System Design

User setup process. If users need to protect the privacy of the location, firstly, they needs to map out the the location needed to be protected on the map. Secondly, users need to manage the permission of all the third-party software which request the user’s location information, then select the specific software to provides location privacy protection, that is, the mobile phone will not provide the correct location information for such software. Finally, if users feel that the current protection of the privacy policy is not flexible enough, they can enter some rules for the protection privacy, and the software will give priority to these news rules to regulate the location information for the third-party software. if there are no new rules, it will use the default rules to manage the location information.

Workflow of privacy protection management mechanism. In the LocationManager class of the Android system framework, as long as there are third-party software requiring to access location information, this designed system will find all the rules entered by users. And according to the rules, it will determine whether to provide the software will the correct location information. If there are no proper rules, it will determine whether the current position is in the area needed to be protected. If not, then give the software the true location information, instead, it will determine whether the software needs the location privacy protection. If not, then give the software the real location information, otherwise, the software will get a certain transformation of the location information using a certain transformation rule.

Rules design. (1) Ipshield idea. Ipshield is a framework for protecting the privacy of mobile phones. Ipshield will protect all the sensors in the phone, and it will also allow users to control time, location, and when a third-party software access a sensor, the information that sensor provides to the third-party software.

According to the Ipshield idea, the software only need to control one sensor, and Ipshield control all the sensors. Therefore, it is a viable strategy.

(2) ACL idea. ACL(Access Control List) is a list of instructions for routers and switches, used to control the data packets across the port.

During the access control, there will be a problem of conflict, as there is a subnet mask. The longer subnet mask will have a short overlap with the shorter subnet mask, and if the two implementation strategy is different, it will produce a conflict.

The order of the data packets is determined by the order of ACL. It is very important to place a sequence of statements in ACL. When the router decides that a data packet will be forwarded or blocked, it will check the data following the order of the statements in the ACL and the judging conditions of each description statement. Once it find a matching condition, it will finish the comparison process, no longer check the other conditions.The most restrictive statement should be put in the first line of ACL statement. The most restrictive statement should be placed in the first-line of ACL statement or the statement near the front of the position, preventing letting off the data packet needed to be rejected or rejecting the data packet needed to pass. The new table entry can only be added in the end of the ACL.

Strategy design. The rule strategy is designed to determine the order of the rules. Rule is inserted by sequence when in the storage, and when in search, it will find the rule in sequence according to certain conditions, once found a matching condition will finish the comparison process.

This is the overall strategy, so when actual operate the strategy, it need a more detailed strategy. The rules are stored in accordance with the third-party software name. if there is a rule constraint a certain third-party software, then this rule will be stored into the corresponding data table of this software. When searching the rule, it will not search all the rules in data table in order, instead, it will search the corresponding data table of such software accessing current location. If a rule is searched, then end the search operation and use the such rule, otherwise, current location will not be restrained. At this point, it find whether current location is in the location privacy protection area, if not, then it can give the software the correct location information. Otherwise, it will give a changed location information.

Transformation Algorithm

Geo-Indistinguishability model. This paper requires that privacy level L is decided by radius R. In particular, it requires that the privacy level L and the radius R to be proportional. Therefore, if the user is protected at any position where range is R, he will have a protection levels of $l = \varepsilon r$.

Mechanism in two dimensional continuous plane. In a two dimensional continuous plane, when the user's true position is x_0, the probability of reporting a random point x by Laplasse distribution can be found below.

$$D_{\varepsilon}(x_0)(x) = \frac{\varepsilon^2}{2\pi} e^{-\varepsilon d(x_0 - x)} \tag{1}$$

And it can also be described by polar coordinates.

$$D_\varepsilon(r,\theta) = \frac{\varepsilon^2}{2\pi} r e^{-\varepsilon r} \tag{2}$$

Then go on to change this formula.

$$\begin{aligned} D_{\varepsilon,R}(r) &= \int_0^{2\pi} D_\varepsilon(r,\theta)d\theta = \varepsilon^2 r e^{-\varepsilon r} \\ D_{\varepsilon,\theta}(\theta) &= \int_0^{\infty} D_\varepsilon(r,\theta)dr = \frac{1}{2\pi} \end{aligned} \tag{3}$$

So the distribution function of r is to use integration of its probability.

$$C_\varepsilon(r) = \int_0^r D_{\varepsilon,R}(\rho)d\rho = 1-(1+\varepsilon r)e^{-\varepsilon r} \tag{4}$$

So it can regulate the rule of drawing a point (r,θ).

Design Optimization

Location Optimization. Using the privacy management mechanism to protect the privacy of the location of the time is the implementation of privacy management mechanism to protect the privacy of the steps of the time, then the direction of optimization should be the use of privacy management mechanism to protect the privacy of the steps to optimize.

In the process of protecting the privacy of the privacy management mechanism, it is judged whether each position is in accordance with the rules. If the rules are the same name, the location data is still in the name of the region. This increases the time cost.

In the experiment, the first position information in the implementation of the privacy management mechanism, the need to open the database, in the database, to the current third party software named after the data table to find, and then to a number of locations to find the data table, and finally close the database. Second location information also need to open the database, a number of tables in the database to find, and finally close the database. This also increases the time cost.

The database file is memory, not in memory, then the database read and write is time-consuming, because the memory much faster than memory speed. In the experiment, every privacy management mechanism to protect the privacy of the location of the test data, we need to make a complete database access, that is, the database file to read and write, then the time will become very large.

Optimization strategy. (1) A strategy for searching the place data table. When a position is beginning to find a rule that is in line with the time condition, it is required to record the location of the protected area within the value corresponding to the value of wheredidian in the current rule. Next, when the location of the search for the time condition of the rules, will first determine the current rules of the value of the wheredidian corresponding to the protection of the region has been recorded, if it has been recorded, it is directly used to record the value, do not need to find the location of the value of wheredidian in the current rules of the corresponding protection area.

This process can ensure a location, the most of the area to protect the user to protect the area traversal, but also can save the unnecessary cost.

(2) A strategy for the database to read and write repeatedly. View privacy management mechanism to protect the privacy of the steps, it can be found that the privacy management mechanism to protect the privacy of the steps of the database only read operations, and not to write operations. This can not have to consider a lot of problems, solutions can be relatively simple.

Since only need the database read operations, then can the database data all read into memory, let all the location information is no longer read memory database, but directly read in memory database information. Because all the read operation will not modify the database, all, directly read the memory database information, and will not affect the results.

Conclusion

The main work of this paper is the design of the user-centeric position privacy management mechanism, and the realization of the software with the location privacy management mechanism and the Geo-Indistinguishability model. The design rule strategy solves the conflict problem of the protection of the same area in different rules.

User-centric privacy management mechanism does not require the user to set the level of protection according to the sense of feeling, but need to clearly point out the location and time of the protection of privacy. The advantage of this design is that it will not affect the availability and user experience of the third-party software if there are some algorithm defects, and the disadvantage is high degree of dependence on the user, can not smartly protect the privacy of. Indeed, there is no algorithm or strategy can achieve not only to smartly protect the location privacy, but also to make all software available.

References

[1] Shankar P, Ganapathy V, Iftode L. Privately querying locationbased services with SybilQuery. Proceedings of the 11th international conference on Ubiquitous computing. ACM, 2009: 31-40.

[2] Pingley A, Zhang N, Fu X, et al. Protection of query privacy for continuous location based services. INFOCOM, 2011 Proceedings IEEE. IEEE, 2011: 1710-1718.

[3] Ghinita G, Kalnis P, Khoshgozaran A, et al. Private queries in location based services: anonymizers are not necessary. Proceedings of the 2008 ACM SIGMOD international conference on Manage ment of data. ACM, 2008: 121-132.

[4] Lee K C K, Lee W C, Leong H V, et al. Navigational path privacy protection: navigational path privacy protection. Proceedings of the 18th ACM conference on Information and knowledge management. ACM, 2009: 691-700.

[5] Sweeney L. kanonymity: A model for protecting privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2002, 10(05): 557-570.

[6] Andrés M E, Bordenabe N E, Palamidessi C. Geo-indistinguishability: differential privacy for location-based systems. Corr, 2013:901-914.

[7] Fawaz K, Shin K G. Location Privacy Protection for Smartphone Users. Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security. ACM, 2014:239-250.

[8] Gruteser M, Grunwald D. Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking. Proceedings of the 1st international conference on Mobile systems, applications and services. ACM, 2003:31--42.

[9] Duckham M, Kulik L. A Formal Model of Obfuscation and Negotiation for Location Privacy.Pervasive, 2005:152--170.

[10] Beresford A R, Stajano F. Location privacy in pervasive computing. IEEE Pervasive computing, 2003 (1): 46-55.