**Individual plan**

**PROJECT INFORMATION**

**Project title:** Hardened Registration on Mobile crowdsourcing

**Examiner:** Panos Papadimitratos E-mail: [papadim@kth.se](mailto:papadim@kth.se)

**Supervisor:** Cihan Eryonucu E-mail: [eryonucu@kth.se](mailto:eryonucu@kth.se)

**Date:**

**Keywords:** Mobile crowdsourcing, Android Root, Emulator detection, Sybil hunting

**BACKGROUND & OBJECTIVE**

It’s hard for 10-years-ago us to imagine how capable mobile devices are in 2021. Nowadays, mobile devices are totally like portable PC or smaller laptop with strong and powerful micro processors. These developments in mobile devices help bring Mobile crowdsourcing into life.

Mobile crowdsourcing takes the advantage of sensing ability of mobile devices, gathering surroundings noise, humidity, etc., and helps us know where we live better. The most well-known practice of Mobile crowdsourcing is a map app[1]. Based on users’ location info, a map app can help design routes which would cost less time or less money for users.

However, the openness of such system also brings in hazards. In the previous case, if some malicious users try to inject fake information, a map app may falsely believe there is a traffic jam happened while there is barely no car on the street[1]. There are some existing technologies help detect false data in the database but it’s inevitable that the database has been corrupted already.

This project aims at designing a hardened registration system which help eliminate those potential malicious devices or emulators at the very beginning and keep the data pool clean and accountable. The main enemy to our system would be those individual attackers with common devices including laptop and mobile devices. Those attackers colluding together is beyond consideration in this project.

**RESEARCH QUESTION & METHOD**

**Questions:**

1. In what cases, malicious users can upload fake, forged or erroneous data?
2. How can we detect potential malicious users and eliminate in the first place?

Breaking down the questions, we are building the hardened registration system in two aspects.

First, in order to send false data to the system, the hardware or software must be exploited. Otherwise, you could never see a well-performing mobile device sending false data. To prevent legitimate users sending erroneous data, the hardware(e.g. sensors) needs to be examined. To prevent malicious users from exploiting the software or evade their hardware erroneous, device root condition needs to be checked[2].

However, this is not easy since the race between root detection and evasion is asymmetric, favoring evaders[3].

For the reason of inevitable exploitation of software, another detection methods need to be introduced, corresponding to aforementioned question two.

In a recently published Sybil hunting method[4], Sybil attacks1 can be detected with the help of honest nodes taking part in the system. The idea of the paper is that Sybil nodes rely on real devices to broadcast WIFI signals. Though this Sybil hunting method seems to have solved the problem, it’s not strict in the assumption that one device can only broadcast on one node’s behalf in a round. Therefore I will introduce a new score system to complement the Sybil hunting method.

This self-designed score system is based on RSSI(Received Signal Strength Indicator), and can perfectly suit the hunting method. Therefore, it would introduce only little calculation burden on the server.

According to [4], the whole process of Sybil detection wouldn’t last more than a minute and require no historic information. Thus, it could be a user friendly as well as a secure and privacy-protected system.

This system so far also has its drawback and insufficiency.

1. If some nodes want to attend the participatory movement instantly, some other additional Sybil detection methods are needed.
2. The detection method requires enough quantity of users. With only few users in some suburb area, the accuracy of detection would be lowered greatly.

To prove the feasibility of such hardened registration system, a simulation model would be built. The model would include individual malicious attacker trying to join the system and the process of elimination.

**EVALUATION & NEW VALUES**

Percentage of detected Sybil nodes would be the major feature used to evaluate the capability of the system. Sybil nodes number should be less than the number of malicious access points in the system(i.e. one laptop and two mobiles are regarded as three access points)

In the project, the root detection method and the hunting Sybil method are both elaborated and more rigorous. Since Sybil attack is always a problem in distributed system, thus the project would be tiny improvement to the solution.

1. Faulty or hostile remote computing elements in a distributed system collude to impair the system[5].

**PRE-STUDY**

In the Participatory Sensing system model [1], users need to register themselves to the Group Manager in order to attend tasks. To filter out potential false or erroneous data source at the very beginning, a hardened registration system is needed.

In [2], a new model for this kind of process is raised. This paper emphasize the importance of root detection in mobile crowdsourcing and gave some basic solutions. However, the root detection process is too easy to be exploited. This part could be elaborated. Additionally, rather than stopping Sybil from entering the system, CAPTCHA can only increase the cost which is not even great.

In [6] and [3], multiple root evasion and 7 detection methods are mentioned. Since this is a competition that favors evaders, root detection would not be the main focus of the project.

**Sybil detection:**

In [7] and [8], distance based authentication methods are adopted. These kinds of methods predominantly come from the domain of VANET. One of the main drawback of such methods is that, millions of trusted units and antennas need to be pre-installed on the streets, making the practice in near term remains unclear.

Using centralized authentication co-operating with government may have solved the problem[9]. However, this may work in one country even if users’ privacy is not concerned, it’s never a sustainable solution world wide. Not to mention the trade-off between private information and users’ willingness to attend the PS tasks .

In [9] and [10], a nodes-relation based graph is used to categorize Sybil malicious nodes and eligible nodes. This is a well solution for distributed system especially social networks. Nonetheless, it’s hard to adopt such Sybil detection method while connections between nodes are not built[4]. To keep node’s own privacy, only Group Manager knows limited information. Therefore, it’s not easy identify the communities in a mobile crowdsourcing system. Nevertheless, the graph based idea is still novel.

Based on [4] and [8], a graph based, elaborated approach is raised. The detection method framework is the same as mentioned in [4], thus introducing no extra burden. In each detection round, a score system is introduced. After all detection rounds, the nodes with the highest score would be regarded as Sybil node. The key idea is: without Sybil nodes, each node would always get 0 points.

So far with simple simulation, the false negative(honest nodes regarded as Sybil nodes) and false positive(Sybil nodes regarded as honest nodes) are both under an acceptable range.

**CONDITIONS & SCHEDULE**

The project would be carried out in two aspects: device root and emulator detection; Sybil detection.

Mobile devices stand for Android devices in the project. Java and C++ are used to develop root detection software.

As for Sybil detection, Python simulation would be adopted. The simulation would present whole detection procedure and the reaction of every node taking part in. In the end, graph and other kinds of results will be used for evaluation.

I keep in touch with my supervisor Cihan and communicate with him about my progress and considerations. So far, this new idea about Sybil detection is under the help from Cihan.

**Milestone:**

**1.15 - 3.15:** literature study and drafting proposal

**3.15 - 4.08:** Android development, root detection and emulator development

**4.08 - 6.01:** Sybil detection model simulation and testing

**6.01 - 7.01:** Master thesis report

**REFERENCES**

1. Gisdakis, S., Giannetsos, T., & Papadimitratos, P. (2014, July). SPPEAR: security & privacy-preserving architecture for participatory-sensing applications. In *Proceedings of the 2014 ACM conference on Security and privacy in wireless & mobile networks* (pp. 39-50).
2. Borsub, J., & Papadimitratos, P. (2018, June). Hardened Registration Process for Participatory Sensing. In *Proceedings of the 11th ACM Conference on Security & Privacy in Wireless and Mobile Networks* (pp. 281-282).
3. Nguyen-Vu, L., Chau, N. T., Kang, S., & Jung, S. (2017). Android rooting: An arms race between evasion and detection. *Security and Communication Networks*, *2017*.
4. Verchok, N., & Orailoğlu, A. (2020, October). Hunting Sybils in Participatory Mobile Consensus-Based Networks. In *Proceedings of the 15th ACM Asia Conference on Computer and Communications Security* (pp. 732-743).
5. Douceur, J. R. (2002, March). The sybil attack. In *International workshop on peer-to-peer systems* (pp. 251-260). Springer, Berlin, Heidelberg.
6. Sun, S. T., Cuadros, A., & Beznosov, K. (2015, October). Android rooting: Methods, detection, and evasion. In *Proceedings of the 5th Annual ACM CCS Workshop on Security and Privacy in Smartphones and Mobile Devices* (pp. 3-14).
7. K. Mekliche and S. Moussaoui, L-P2DSA: Location-based privacy-preserving detection of Sybil attacks, 11th International Symposium on Programming and Systems (ISPS), Algiers, 2013.
8. M. Demirbas, Y. Song, An RSSI-based Scheme for Sybil Attack Detection in Wireless Sensor Networks, In Proceedings of International Symposium on a World of Wireless, Mobile and Multimedia Networks, 2006.
9. Yu, H., Gibbons, P. B., Kaminsky, M., & Xiao, F. (2008, May). Sybillimit: A near-optimal social network defense against sybil attacks. In *2008 IEEE Symposium on Security and Privacy (sp 2008)* (pp. 3-17). IEEE.
10. Yu, H., Kaminsky, M., Gibbons, P. B., & Flaxman, A. (2006, August). Sybilguard: defending against sybil attacks via social networks. In *Proceedings of the 2006 conference on Applications, technologies, architectures, and protocols for computer communications* (pp. 267-278).