Recovering Access to Account Using Location Data

Shuji Yamaguchi and Hidehito Gomi
Yahoo Japan Corporation

INTRODUCTION

Account Recovery

- Backup mechanism to reclaim a user's lost account that the user has been able to access (e.g., Password and Biometrics)
- Practical method for account recovery is required in terms of security and usability.

- Typical Case: Knowledge-based recovery methods
  Service provider managing user accounts encourage users to provide their personal knowledge questions which are known as “secrets or challenges”.

Problems of Account Recovery

- Knowledge-based recovery methods are known as vulnerable to attacks that compromise user accounts.

- Backup credential needs to be associated with the user account in preparation for an incident in which he/she cannot recall his/her password.

Challenge

- We propose an account recovery system using two sets of **location data** that have been collected independently via different devices of users **without any explicit registration process**.

SYSTEM OVERVIEW

Architecture

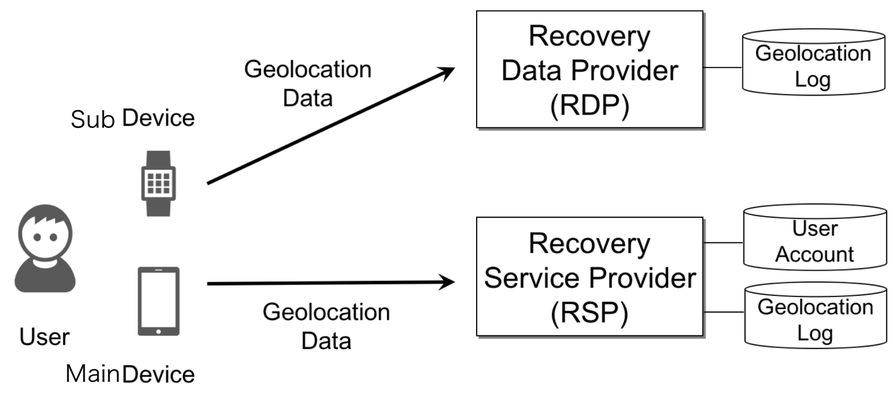

Figure 2: Overview of the proposed system.

- **Recovery Service Provider (RSP)** provides an account-recovery service for users holding accounts in account-loss incidents. Location data collected by the GPS sensor of the main device (“RSP Device”) are stored at the RSP.

- **Recovery Data Provider (RDP)** provides data for account-recovery. Sub device’s (“RDP Device”) location data are also sent to be stored at RDP.

- User’s location data sensed by the RSP and RDP devices are individually stored at the RSP and RDP, respectively without his explicit actions in preparation for an incident when he/she cannot access his account at the RSP.

Recovery Procedure

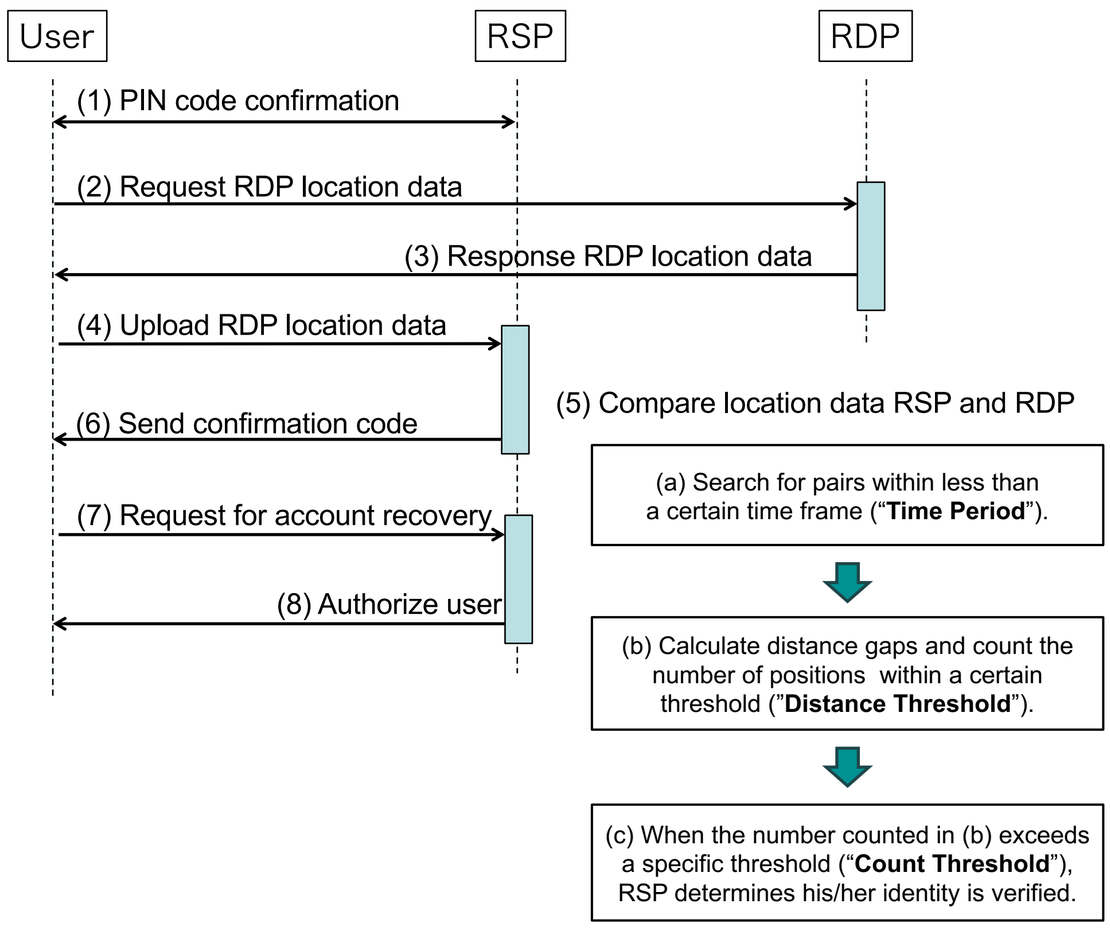

Figure 3: Recovery Procedure.

EXPERIMENTAL STUDY

Dataset

- Target data:
  Collect location data of Yahoo! JAPAN smartphone application according to our privacy policy and user agreement.

- Target users:
  285 users who accessed via two different devices which correspond to both the RDP and RSP.

- Each set of location data including [timestamp, latitude, longitude]

- The number of sensed positions for RSP and RDP Devices average 2106 and 466.

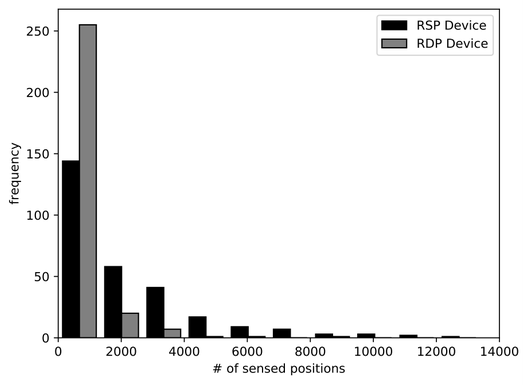

Figure 4: Histogram of location dataset

FRR, FAR and EER Calculation

- Calculated the performance measures, false rejection rate (FRR) , false acceptance rate (FAR) and equal error rate (EER) to evaluate the proposed method.
- Settings: Time Period = 10 (min) and Count Threshold = 3 (counts)

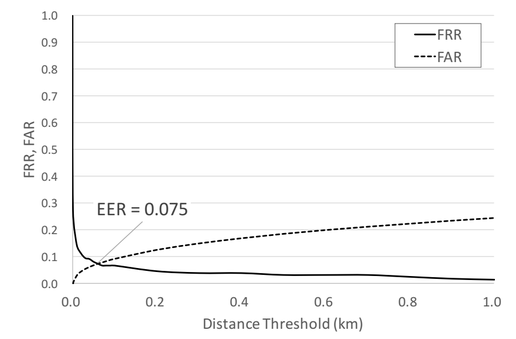

Figure 5: FRR and FAR results vs. Distance Threshold (from 0 to 1 km).

Results

- EER is equal to **0.075** for a **0.05** (km) Distance Threshold, where we obtained the best performance for the proposed system.

Discussion and Future Work

First experiment to use location data for account recovery.

- EER result is noteworthy to reveal the potential of location data.
- Usability of zero-interaction for proving identity using location in the proposed approach is a great advantage over traditional knowledge-based ones.
- Expect that this approach is applied to implicit or continuous authentication by further improving its accuracy as well as to identity proofing.

Threats.

- Possible threats: Guessing attacks and Stalking.
  - Attacker may move around by anticipating the whereabouts of a user or physically track.
  - User’s families, roommates, neighbors, and workmates are possible attackers if they are frequently close to the user.
- Although it is difficult for an attacker in either case to impersonate a user by completely copying the user’s behavioral log for a long period, such attacks may be successful if we use limited number of positions to determine whether two location datasets of a user match.
- We will further consider using different sets of positions within multiple contexts of the user to reduce the above security risk.

Limitations.

- This study assumes that users almost always have two GPS-enabled devices with them for a large portion of their base.
- Moreover, users may meet situations in which GPS signal is unavailable or unstable when they go into a building indoor or cities with skyscrapers.
- These situations may significantly increase the error rate of the proposed system because of the lack of location data for proving a user’s identity.

Privacy.

- Location data need to be carefully maintained because they relate to privacy.
- Although we obtain users’ consents to the use of such data in accordance with the privacy policy for our services, collecting location data excessively and persistently will enable the service provider to analyze the dataset and anticipate a user’s physical location.
- Thus, we alternatively consider a method for keeping raw location data inside the RDP Device locally without uploading to a server, which will be our future work.

CONCLUTION

- We proposed account recovery system verifying the similarity of two types of location datasets that have been collected via different devices of users without any explicit registration process.
- With the proposed method, we conducted an experimental study to evaluate performance measures such as false acceptance and false rejection using the location data of 285 users and obtained a 0.075 EER.

ACSAC’19, December 2019, San Juan, Puerto Rico