GFM: A Group Finance Manager Using Secret Sharing and Data Synchronization on Mobile devices

Department of Computer Science, Graduate School of Systems and Information Engineering
Meng Li 201620728
Supervisor Yasushi Shinjo
November 17th, 2016

1 Introduction

The conventional applications are based on a client-server mode, which requires a central server for storing shared data. That means the users of the web applications must fully trust central servers and their application providers. Once a server is attacked by hackers, user information may be revealed because data is stored on only one server. Finally, users may lose their data when service providers shut down their services.

Thus, we are going to implement a group finance manager application: GFM, which is not relied on central servers, that means user data cannot be cracked easily and service can be recovered after shutting down the old service. Most current popular way is using P2P(Peer to Peer) to transfer user data between devices. However, there is a obvious problem in such a no-server proposal, that synchronization can only be finished in condition that two devices is online in a same time. Another problem is that the number of P2P connections will be increased fast with the increasing of group scale because each device can create a connection to all of the other devices.

To address this problem without using P2P, we are considering to use multiple servers to build an untrusted server group. Data will divided to several parts and uploaded to different servers. Each server can only save a part of data temporarilywhich means this part will be deleted after a period time we specified. Those two rules explained above can ensure that user data cannot be cracked easily. In fact, we can regard the server group as a bridge between devices of group members. In addition, all devices of group members have saved a complete data set, data can be recovered theoretically even the server group breaks down. By this way, we think such a method can satisfy the requirements for an application which is not relied on central servers.

2 Related Work

DepSky[1] is a system that storage encrypted data on servers and run application logic on the client-side[8]. DepSky provide a storage service that improves the availability and confidentiality provided by commercial storage cloud services. In GFM, the system cannot run without either or both client and server. For the reason that server group in GFM undertake responsibility of temporarily data storage and message delivery with server-side computation, GFM also needs its own web service rather than using commercial storage cloud services.

Mylar[5] stores sensitive data encrypted on the server, and decrypts that data only in users browsers. Like GFM, application in Mylar can control how user data is shared[8]; unlike GFM, Mylar build its system on a B/S model while GFM using C/S model.

Of course, there are many applications with untrusted network and servers. Compared to them, GFM using secret sharing and temporary data storage to protect user data from malicious attacking.

3 Methodology

We consider that implementing such a distributed application GFM on mobile devices is easier than PC, due to the reason that mobile devices can access to Internet by cellular network everywhere and every time. For this reason, we implement our financial management application on iOS mobile devices.

3.1 Data Sync by Server Group

We implement GFM based on data synchronization by multiple servers rather than a central server. The main difference between our proposal and the traditional way using central server is that GFM is not relying on server. There are three core principles in our proposal.

The first is data transfer by server, not save on server. Most current popular B/S modal application storage user data on several central servers, user data will not be deleted unless user delete his account. GFM use server as a bridge for transferring data. Considering a group including three members Alice, Bob and Carol. Alice created a new record in her device, this record will be upload to our server group, and will not exist after Bob and Carol synchronized it from server
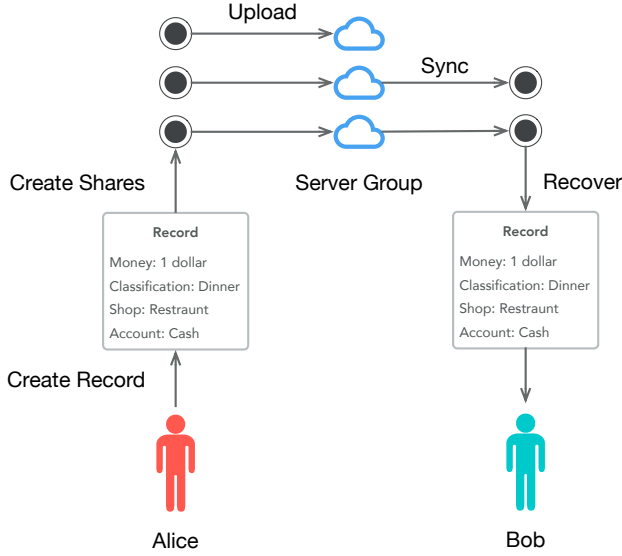
Figure 1: Flow of Synchronization

group. In other words, server group storage a record until it is be synchronized by all group members.

The second is server save data uploaded from mobile devices temporarily. We define a period of time in which data can be saved in a server. In this paper, we set this period to 1 hour for our example situations, that means the record Alice uploaded to server group can only be saved for 1 hour. After 1 hour since uploaded, this record will be deleted. Alice is required to resend this record to server group until all of other members has synchronized successfully. It is obvious that this period of time should be long while more members attend the group. However, the longer saving period means the higher risk of data reveal. The most suitable period is influenced by the number of group members, security requirement, user expectations and so on.

The third is server does not know the content of data. Only by saving data temporarily cannot ensure data security, because server knows the content of data in this temporary period. For this problem, developers usually encrypt data before uploading to servers, which need a private key to decrypt data in other devices. In order to distribute the private key, Sweets, which is a decentralized social networking service application using data synchronization on mobile devices, use QR Code to share the private key. Sweets can generate a QR Code contained the private key, while other devices can scan this QR Code to recover the private key. In this paper, we are going to use secret sharing, which can create a number of shares and distributing them to a set of participants[6], to share data between user devices and server group.

Based on the three core principles introduced above, our synchronization flow can be described as figure 1. In figure 1, Alice create a record including money, classification, shop and account. GFM create a number of

shares by secret sharing and upload those shares to server group. When Bob is online, GFM in Bob's device will pull data from server group and recover the new record uploaded by Alice. In our situation, GFM can recover the record after pulling more than two shares randomly. In this process, each server is separated from other server, and cannot access to other server, which means user data cannot be recovered unless you have permission access to more than two server in Alice and Bob's situation. In our proposal, only group members have permission to access server group. Access control will be introduced in the paragraph below.

3.2 Shamir's Secret Sharing

Secret sharing which can create a numbers of shares, plays an indispensable role in protecting user data from getting lost, destroyed, or into wrong hands. In this paper, we are going to use Shamir's secret sharing, a form of secret sharing proposed by Shamir and Blakley independently. In secret sharing scheme, a dealer securely shares a secret with a group of participants, by generating n shares using a cryptographic function[6]. At least k or more participants can reconstruct the secret, but $k-1$ or fewer participants can obtain nothing about the secret[4]. We describe this scheme as a function $f(k, n)$, n is the number of all shares, and k is the threshold to combine shares. A popular technique to implement threshold schemes uses polynomial interpolation ("Lagrange interpolation"). This method was invented by Shamir in 1979. Thus, we call it Shamir's Secret Sharing.

There are many implementations of Shamir's secret sharing by different programming languages. For developing an iOS app by Objective-C, we need an implementation without limitation of length, which supports Objective-C language and UTF-8 character set. C-SSS[3] is an implementation in original C code of Shamir's Secret Sharing by Fletcher T. Penney. C-SSS provide two main functions for generate shares and reconstructing them. Function *generate* can generate n shares by the string text with the threshold k. Function *extract* can recreate the text string after accessing to more than k shares.

3.3 Synchronize Completely

There is no doubt that GFM should provide a reliable synchronization service. For Alice, once she created a new record in her device, all of other members in a group should synchronize this record, even if this record may be deleted by server group after 1 hour. We call this problem *Synchronize Completely*.

Figure 2 described the situation data cannot be synchronized completely. Alice sent a new record to server group in 10:00 AM and Bob synchronized it successfully before 11:00 AM. Carol forgot to synchronize this
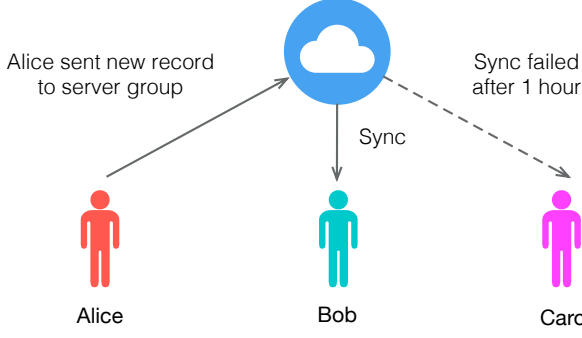
Figure 2: Synchronize Completely

Figure 3: Sync Table

record before 11:00 AM. She cannot synchronize after 11:00 AM because server group has delete the record in 11:00 AM. To solve this problem, Alice is required to resend her new record until Carol synchronize it successfully. However, by what way can Alice get the information that all of other members in her group have synchronized. In other words, it is an unreasonable requests for Alice to resend her record indefinitely. It is necessary for us to find a efficient way for notifying Alice that all members have synchronized successfully.

We are going to design a counter in clients as the solution to *Synchronize Completely* problem. Intuitively, this counter which can calculate the number of times other clients synchronized, works on the sender to decide resending record or not. The condition that sender device need not resend record can be described as equation 1.

$$\sum_{i=1}^{n} s_i = (m-1) \cdot k \qquad (1)$$

In equation 1, we suppose that there are n servers in our server group , they are server 1, server 2,..., server n. We can also look at n from a different perspective, it is first parameter, the number of all shares, in the scheme $f(k,n)$ of Shamir's secret sharing. Each server saves a share, so n servers saves n shares. Each server should know how many times the share has been downloaded by clients. Thus, we use s_i to represent the number of times in server i. The sum of s_i is calculated in sender client. In the left of equation 1, m represents the number of group members, and k is the threshold in scheme $f(k,n)$. Obviously, a client can reconstruct data after getting k shares. If $m-1$ clients have reconstructed data, $(m-1) \cdot k$ shares is downloaded. That means clients except sender has synchronized successfully, so sender need not resend.

To implement such a counter, two restrictions about security and connection should be considered. On the one hand, servers cannot communicate with each other. Actually, one server of the group does not know the others exist in the group. On the other hand, clients cannot communicate with each other. Clients do know IP address of other devices, so that they cannot send message with P2P connections. We decide
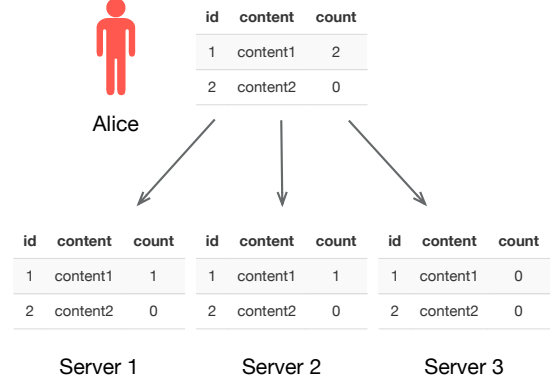
to design a sync table described as figure 3, which is running on both servers and client to statistic sync times. Sync tables dominates all the synchronization activity in GFM. Physical id, content, and count of sync times are main column in sync table. Physical id is an unique identifier created by SQLite database in client, it is same as those in servers. Content in client saves the JSON string converted by new record object. In server, it saves one of the shares generated by secret sharing scheme. Count, the core column in sync table, saves the content sync times in server. In client, count saves the sum of all counts from server. Of course, physical id of count in client and counts in servers are same. Suppose the secret sharing scheme is $f(2,3)$, and only Bob have synchronized Alice's new record within prescribed time-limit. In figure 3, the column corresponding to Alice's new record, whose id is 1, is equals to those columns whose ids are 1 in server 1 and server 2. This is to say Bob downloaded two shares from server 1 and server 2. Client of Alice will know that by HTTP response from server 1 and server 2 when it resend data. It will send for one more time, because it's count is 2, which is less than 4, the threshold Alice need not resend.

4 Architechture

GFM includes an iOS app for clients and a web service in server group.

4.1 Architechture in Client

In part 3 Methodology, we have introduced how to sharing a string with other deices via untrusted server group. In this paragraph, we concentrate persistent store and data synchronization before taking about architechture in client. With untrusted server group, GFM have to storage all data on mobile devices, which requires a object-oriented way. GFM use Core Data[2], a native iOS framework to manage the model layer objects, which provides generalized and automated solutions to common tasks associated with object life
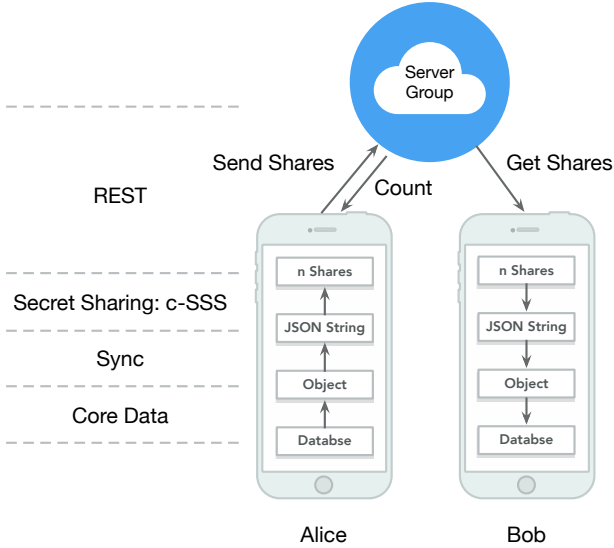
Figure 4: Architechture

cycle and object graph management, including persistence, for data storage and operation. Sync[7] by Elvis Nuez, a modern swift JSON synchronization to Core Data, can help we complete data synchronization easily in GFM by parsing a JSON response and getting it into Core Data. With Sync based on Core Date, we can concentrate on sharing with untrusted server group rather than data storage and synchronization.

c-SSS by Fletcher T. Penney can generate shares from a JSON string saved on sync table. GFM develop with Objective-C language, which is compatible with original C code. GFM will generate shares and send them by REST API to server group until this record need not to resend. Response of REST API contains the count of sync times in server group, for calculating the sum and deciding wether resend or not.

4.2 Necessity of Having a Web Service

Many applications with synchronization is based on cloud services like Amazon S3, Google Cloud or iCloud. Here, GFM does not synchronized via such cloud services for following reasons.

- It is better for web service of GMS to access data by object, while cloud service provide a way by document.
- It is more convenient to control when data should be delete by web service of ourself.
- Web service need collaborate with client to statistic sync times, which requires to add and modify a column as we want.

5 Future Work

An application running in client has been developed. In future, we are going to develop a web service run-

ning on untrusted server group to manage data temporary storage and provide REST API for mobile devices. This web service developed by Java EE should be install by group owner, in order that nobody except the group owner know all the IP address of servers. Thus, security of GFM can be guaranteed.

6 Conclusion

This paper introduced GFM, a group finance manager, using secret sharing and data synchronization on mobile devices. GFM is synchronized by untrusted server group, each server of which does not know existence of others. Each server only saves one of the shares generated by secret sharing temporarily, to ensure that user data cannot be cracked easily.

References

[1] Alysson Bessani et al. "DepSky: dependable and secure storage in a cloud-of-clouds". In: *ACM Transactions on Storage (TOS)* 9.4 (2013), p. 12.

[2] Apple Inc. *Core Data Programming Guide*. URL: `https://developer.apple.com/library/content/documentation/Cocoa/Conceptual/CoreData/index.html`.

[3] Fletcher T. Penney. *c-SSS, an implementation of Shamir's Secret Sharing*. Version 0.2.0. June 29, 2015. URL: `https://github.com/fletcher/c-sss`.

[4] Liao-Jun Pang and Yu-Min Wang. "A new (t, n) multi-secret sharing scheme based on Shamirfffdfffdfffds secret sharing". In: *Applied Mathematics and Computation* 167.2 (2005), pp. 840–848.

[5] Raluca Ada Popa et al. "Building web applications on top of encrypted data using Mylar". In: *11th USENIX Symposium on Networked Systems Design and Implementation (NSDI 14)*. 2014, pp. 157–172.

[6] Guillaume Smith, Roksana Boreli, and Mohamed Ali Kaafar. "A Layered Secret Sharing Scheme for Automated Profile Sharing in OSN Groups". In: *International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services*. Springer. 2013, pp. 487–499.

[7] Elvis Nuñez. *Sync*. Version 2.2.0. URL: `https://github.com/SyncDB/Sync`.

[8] Frank Wang et al. "Sieve: cryptographically enforced access control for user data in untrusted clouds". In: *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*. 2016, pp. 611–626.