DECENTRALIZED SOCIAL NETWORK SERVICE USING THE WEB HOSTING SERVER FOR PRIVACY PRESERVATION

Yoonho Nam, Changhoon Lee, Youngman Jung, Woongryul Jeon and Dongho Won

Information Security Group, Sungkyunkwan University
300 Cheoncheon-dong. Jangan-gu. Suwon-si. Gyeonggi-do 440-746. South Korea

Email: yhnam@security.re.kr

Abstract

In recent years, the number of subscribers of the social network services such as Facebook and Twitter has increased rapidly. In accordance with the increasing popularity of social network services, concerns about user privacy are also growing. Existing social network services have a centralized structure that a service provider collects all the user's profile and logs until the end of the connection. The information collected typically useful for commercial purposes, but may lead to a serious user privacy violation. The user's profile can be compromised for malicious purposes and even may be a tool of surveillance extremely. In this paper, we remove a centralized structure to prevent the service provider from collecting all users' information indiscriminately and present a decentralized structure using the web hosting server. The service provider provides only the service applications to web hosting companies and the user should select a web hosting company that he trusts. Thus, the user's information is distributed and the user's privacy is guaranteed from the service provider.

Keywords: *decentralized, social network, user's information, privacy.*

Abstrak

Dalam beberapa tahun terakhir, jumlah pelanggan layanan jaringan sosial seperti Facebook dan Twitter telah meningkat pesat. Sesuai dengan meningkatnya popularitas layanan jaringan sosial, kekhawatiran tentang privasi pengguna juga berkembang. Layanan jaringan sosial yang ada memiliki struktur terpusat dimana penyedia layanan mengumpulkan semua profil pengguna dan log sampai akhir sambungan. Informasi yang dikumpulkan biasanya berguna untuk tujuan komersial, tetapi dapat menyebabkan pelanggaran privasi pengguna yang serius. Profil pengguna dapat dikompromikan untuk tujuan jahat, dan bahkan mungkin menjadi alat pengawasan yang sangat ketat. Dalam tulisan ini, kami menghilangkan struktur terpusat untuk mencegah penyedia layanan dari pengumpulan informasi semua pengguna, dan menyajikan struktur desentralisasi dengan menggunakan *web server hosting*. Penyedia layanan hanya menyediakan aplikasi layanan untuk *web hosting* perusahaan, dan pengguna harus memilih perusahaan *web hosting* yang ia percaya. Dengan demikian, informasi pengguna terdistribusi, dan privasi pengguna terjamin dari penyedia layanan .

Kata kunci: *desentralisasi, jejaring sosial, informasi pengguna, privasi.*

1. Introduction

Nowadays, since SNS (Social Network Service) such as Facebook and Twitter gained great popularity, the various SNS began to emerge. According to the companies, a billion people, 500 million people joined the Facebook, Twitter, respectively [1][2]. SNS mainly is used for formation of personal relationship and information exchange. Recently, various services are provided in conjunction with SNS. The user may compete in games with SNS friends and post favorite clothes from online shopping sites on his wall. In other words, it means that the user unwittingly leaks his information in various paths. Many people even have more than one SNS account.

As the popularity of SNS grows, on the other hand, concerns about user privacy get attention for the issue. The existing SNS has a centralized structure generally that all users' profile information is stored in database of service providers. The information collected by service provider is mainly used in the investigation or for commercial purposes. However, if service provider

had malicious purposes, the information might have been stolen. In addition, SNS that contains information of millions of people has become a good target for hackers. For example, passwords of 6.5 million users of LinkedIn were hacked in 2012 [3]. It may affect all friends of users hacked, so a centralized structure is accompanied by a greater risk.

Thus, in this paper, we remove a centralized structure in order to prevent the service provider from collecting all users' information indiscriminately and present a decentralized structure using web hosting server. The service provider provides only the service applications to web hosting companies and the user should select a web hosting company that he trusts. Then, each web hosting server manages its user's homepage, so the service provider cannot collect any information. Thus, the user's information is distributed and the user's privacy is guaranteed from the service provider.

In section 2, we examine exiting decentralized SNS and DHT (Distributed Hash Table). In section 3, Section 4, the structure and operation process of the proposed system, respectively. In section 5, we analyze security requirement. Finally, we analyze the proposed system, conclude the paper and present future work in section 5.

2. Related Work

2.1 Decentralized Social Network Service

Recently, some researches about decentralized SNS are underway. Previously proposed models are divided broadly into two categories, web-based model and P2P (Peer-to-Peer)-based model in the following table I [4].

TABLE I
CLASSIFICATION OF EXISTING MODELS

Model name	Categories
Diaspora	Web-based
FoaF	Web-based
Likir	P2P-based
PeerSoN	P2P-based
Safebook	P2P-based

Web-based model typically uses Oauth, cloud host, web-hard and private homepage. Typical web-based models are Diaspora and FOAF (Friend-Of-A-Friend) [5][6]. In the case of Diaspora, the users may make their own homepage using the open source provided by service provider and use aspect that new concept used to control the access in the system to protect their privacy. In FOAF, the users exchange their FOAF files that contain their information in a consistent format. Web-based model has some advantages such as high availability and ease of management.

P2P-based model usually build a private server on own PC or mobile. Typical models are PeerSoN and Safebook [7][8]. In PeerSoN, the users build a personal server on their PC or mobile and use DHT to find location of available friend's device. In Safebook, the users make replica including their information and keep them on their friends' nodes dispersedly. However, personal PC or mobile cannot be turned on at all times, so sometimes uses a mix with web-hard or cloud service for improving the availability.

2.2 Distributed Hash Table

DHT is a technique usually used in P2P network for lookup service [9]. Regardless of the entry point, the location of the key is found easily using the routing algorithms such as Chord and Pastry on overlay network. DHT has some advantages of quick search speed, suppression of network load and search range of billions nodes. Also each node includes only the key, but do not contain any information.

In the proposed system, keywords such as e-mail address and user's name specified by the users are used for lookup and search result is mapped to the web hosting server that contains the user. DHT is used only the first time when the users search their friends.

3. The Proposed System

The proposed system is web-based model and has a decentralized structure as shown in the Figure 1.

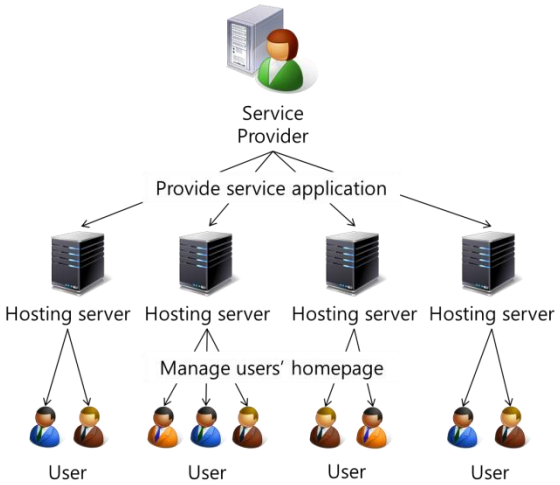

Fig 1. Structure of the system

Common web hosting server lends web space to individual who do not have a server computer. Thus, we assume two things. The first, web hosting server in proposed system would be for not only

this system and the second, there would be revenue model between the service provider and web hosting companies. That is, web hosting server may have other clients who do not use this SNS and the service provider's income is generated through such as advertising API.

The service provider provides service applications to web hosting servers and updates them consistently. The user selects trusted servers for the homepage to establish. The user's information is stored on the web hosting server, so the service provider cannot access it.

The inter-user communication is achieved through communication between web hosting servers directly. In the general P2P model, the users should manage their private server by themselves. However, most people do not know how to manage server, so the users' homepages managed by the web hosting companies, which is better. The proposed system also provide automatic interface to establish homepage.

4. Operation Process

There are four steps for the proposed system. Homepage establishment, register a node on DHT, add friends and message exchange. The following table shows the notation.

TABLE II
NOTATION

Symbol	Meaning
$USER.S_{ID}$	user's hosting server ID
$USER.U_{ID}$	user ID
$USER.S_{IP}$	IP address of user's hosting server
$USER.U_{IP}$	IP address of user's homepage
$USER.K$	search keyword that user sets up

4.1 Homepage Establishment

First of all, the user should choose trusted hosting server. The service provider posts the list of affiliated web hosting service on service homepage. After the user choose it, create $USER.U_{ID}$ through e-mail authentication. Each web hosting server is assigned $USER.S_{ID}$ from the service provider. Since $USER.S_{ID}$ is unique in the system because it is hashed $USER.S_{IP}$ so it is used for the authentication among the hosting servers.

After then, the user establishes a private homepage. The service provider provides service applications that consist of functions such as wall, photo and friend list. The user is able to choose services what he wants to receive and to set them up. However, web hosting server has to apply some advertising APIs on its homepage. That's why the service provider provides all the service APIs for free and doesn't require the users' information from each web hosting server.

4.2 Register a Node on Distributed Hash Table

After homepage establishment, the user chooses $USER.K$ to register his node on the DHT. The keyword would be general or detailed. It is used for someone else to find the user, so the user had better set it more specific like e-mail within the range of acceptable public.

DHT in the proposed system uses Chord algorithm. The user sends hosting server the keyword and then hosting server register a node on the DHT using Chord. Each node has finger table that includes successors which means the next node would be routed on the identifier circle. Lookup procedure is performed in a binary tree scheme, but we do not describe how to process specifically in this paper.

4.3 Add Friend

For adding friend, the user should search with keyword. Generally SNS is an extension of the relationship in real world, so the user is able to friend's keyword offline directly. If user A finds user B through the DHT, DHT allows mapping to the hosting server that contain user B and then the hosting server allows mapping to user B's home page along the red line as shown in the Figure 2.

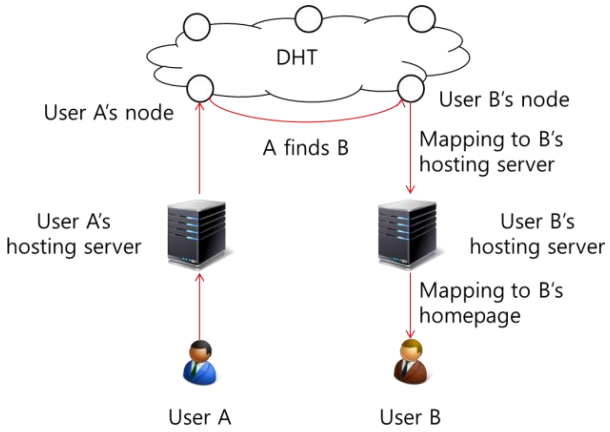

Fig 2. The path that user A finds user B

After user A finds user B, user A sends request message. If user B accept the request, user B sends acceptance message, but if not, refusal message. This process follows the following message exchange procedure.

4.4 Message Exchange

After the first time user A found user B using DHT, communicate through a direct connection between user A's hosting server and B's hosting server without going through the DHT anymore along the blue line as shown in the Figure 3.

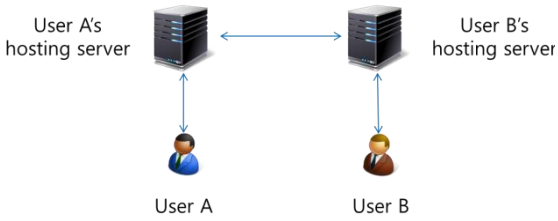

Fig 3. The path that user A sends a message to user B

All messages are encrypted with SSL, access to the homepage is based on both user IP and web hosting server IP.

5. Security Analysis

5.1 Confidentiality

All communication among the users and web hosting server should be protected totally. Storing all data encrypted may have too much load, so access control policy prevents third party from accessing the data. Over message exchange, all messages should be encrypted with SSL, so confidentiality is assured.

5.2 Privacy

All data should be first set to private. Any personal information should not be disclosed to third parties who do not have access. In existing SNS, the service provider has enrolled permission to access to any data. However, the service provider in the proposed system cannot even know whether the user subscribes. Other web hosting servers also do not know where the user's date is stored. Thus, the proposed ensures user privacy.

5.3 Authentication

Existing SNS provides e-mail or mobile authentication, but it is not perfect. If someone has more than one e-mail address, he can create multiple accounts, even using alias. The proposed system also does not provide user authentication, but data origin authentication is done using web hosting server IP. Before the service provider provides service application, all web hosting server are assigned unique ID that is based IP.

5.4 Access Control

All data should be specified by the user, depending on the level of disclosure (e.g. public, private and partially public to certain groups). In the proposed system, only keyword matched data can be known the presence and no one can even know the existence of the data.

5.5 Availability

Web hosting server should be connected anytime and anywhere. In general, web hosting server runs 24 hours except for inspection time. Flexibility against with DoS (Denial of Service) attack should be prepared and ongoing management and control should be required. The proposed system ensures high availability because of web-based model.

5.6 Data Integrity

Over any message exchange, the original message authentication and tampering detection should be available. In the proposed system, SSL protocol is used over all message exchange, so integrity is guaranteed with origin authentication and message encryption.

6. Conclusion

Many people have an interest in protecting their privacy. Since existing SNS still have some security vulnerabilities, a variety of counter measures have been researched. We point out the problem is the centralized structure of existing SNS. The service provider has an absolute permission to access to any data and collects all users' data. It is a clear violation of user privacy.

Thus, we attempt to decentralize SNS for preserving user privacy in this paper. To remove the centralized structure, we present the system using the web hosting server to restrict the service provider's access. Each web hosting server manages only their clients and never provides any user's information to third parties who do not access. In addition, the system satisfies availability because of web-based model and is easier to manage the user than P2P-based model. It also assures authentication, confidentiality and integrity because of using SSL. However, the proposed system also provide only e-mail authentication, so we will present stronger authentication method in the future work.

As people gradually increase the interest for the user privacy, the distributed SNS model has attracted a great attention. Future studies should be continued to protect user privacy while maintain the advantages of the existing SNS.

Acknowledgment

This research was funded by the MSIP (Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2013.

References

[1] Facebook, "http://www.facebook.com"
[2] Twitter, "http://www.twitter.com"
[3] C. Zhang, J. Sun, X. Zhu, Y. Fang, "Privacy and security for online socialnetworks:

challenges and opportunities", IEEE Netw. 24 (4) (2010).

[4] Paul, T.; Buchegger, S.; and Strufe, T, "Decentralized social networking services", InTrustworthy Internet. Milan: Springer.

[5] Diaspora, "diasporaproject.org"

[6] Yeung, C. A.; Liccardi, I.; Lu, K.; Seneviratne, O.; and Berners-Lee, T,"Decentralization:The Future of Online Social Networking", W3C Workshop on the Future of Social NetworkingPosition Papers. Vol. 2. 2009.

[7] Buchegger, S., Schi¨oberg, D., Vu, L.H., Datta, A,"PeerSoN: P2P social networking – earlyexperiences and insights", In: Workshop on Social Network Systems. 2009

[8] Cutillo, L.A., Molva, R., Strufe, T,"Safebook: a Privacy Preserving Online Social NetworkLeveraging on Real-Life Trust", IEEE Communications Magazine 47(12), 94 – 101 2009

[9] B.O. Kim, I.W. Lee, H.J. Park, "Trend of Distributed Hash Tables-Based P2P", Electronics and Telecommunications Trends vol.21(6), 2006