A PRIVACY-PRESERVING ENCRYPTION SCHEME FOR AN INTERNET REALNAME REGISTRATION SYSTEM

Fei Xu, Ken Yau, Ping Zhang, Kam-Pui Chow

Chapter 7

A PRIVACY-PRESERVING ENCRYPTION SCHEME FOR AN INTERNET REAL-NAME REGISTRATION SYSTEM

Fei Xu, Ken Yau, Ping Zhang and Kam-Pui Chow

Abstract Internet real-name registration requires a user to provide personal identification credentials including his/her real name to an online service provider when registering for an account. In China, real-name registration has been implemented since 2010 for purchasing train tickets and mobile phone SIM cards. In 2013, the Chinese government announced that real-name registration would be implemented for Internet users to protect against cyber crimes, including cyber bullying, rumor spreading and identity theft. When real-name registration is in place, law enforcement agencies can obtain the real identity of a user who attempts to leverage the anonymity provided by the Internet to conduct criminal activities. However, real-name registration also potentially infringes on online privacy.

This chapter presents a privacy-preserving Internet real-name registration approach based on Shamir's secret sharing scheme. The approach helps protect the identities of online users while enabling law enforcement to obtain the identities of users involved in criminal activities on the Internet.

Keywords: Internet crimes, real-name registration, privacy preservation

1. Introduction

Online services offer users with unparalleled selections of applications, products and opportunities. However, they also provide numerous opportunities for malicious users to commit cyber crimes. When a cyber crime is committed, it is difficult to trace the real criminal because the association between the online user identity and the real identity is hidden or is otherwise unavailable. Due to the increase in Internet crime, real-name registration systems have been proposed to trace criminals

more effectively. Such a system requires a user to provide identification credentials and his/her real name to an online service provider when registering for an account (e.g., for a blog, website or online discussion forum). During a criminal investigation, the real identity of a user could be provided to a law enforcement agency upon provision of the appropriate legal documents. However, unless the system is designed very carefully, real-name registration could pose serious threats to user privacy.

This chapter presents a privacy-preserving real-name registration approach based on Shamir's secret sharing scheme. The approach provides privacy protection to users while enabling law enforcement to obtain the identities of users involved in criminal activities on the Internet.

2. Background

This section discusses the notions of privacy and its implications with regard to name-registration systems in South Korea and China.

2.1 Privacy

Privacy has long been understood as "the right to be let alone" [11]. Any person should be able to keep information about himself/herself out of the public domain. An individual who is unable to control the dissemination of his/her private information has lost the right of privacy.

The right of privacy has become the center of attention due to the ubiquity of the Internet. Some users have leveraged the anonymity of the Internet to spread gossip or reveal sensitive or private information about individuals. The affected individuals often complain that anonymous posts are defamatory or infringe on their privacy. On the other hand, many Internet users argue that they are entitled to enjoy the freedom of expression and communication, as well as their right to privacy. The solution is to design and implement a real-name system that offers privacy protection to online users while enabling law enforcement to obtain the real identities of users involved in criminal activities upon provision of the appropriate legal documents.

2.2 Real-Name Registration in South Korea

South Korea was the first country to implement a real-name registration system for Internet users. The system, which was implemented in 2007, required users to supply their real names and resident registration numbers for web portals accessed by more than 300,000 visitors per day, the goal being to reduce and help prosecute online criminal activities. Since June 28, 2009, 35 South Korean websites have implemented real-

name registration systems according to the newly amended Information and Communications Network Act (Choi Jin-Sil Law) [3].

Real-name registration in South Korea was driven by the suicide of Ms. Choi Jin-Sil after malicious comments about her were posted on Internet bulletin boards. Ms. Choi, who was one of the most famous actresses in South Korea, committed suicide on October 2, 2008 because she was rumored to have been involved in the suicide of another actor, Mr. Ahn Jae-Hwan [2]. In 2009, the Government of South Korea amended its Information and Communication Network Act to require real-name registration for all websites with more than 100,000 visitors per day. Moreover, the law required online service providers to disclose personal information of alleged offenders in cases involving libel or infringement of privacy.

The South Korean Resident Registration Number (RRN) is a thirteen-digit number that uniquely identifies each resident. The RRN is divided into several segments, each representing a piece of personal information. For example, the first digit segment represents the date of birth. In addition to the date of birth, gender and birthplace can be discerned from the RRN [12].

South Korean law required users to provide their real names and RRNs when creating online accounts. The collection of this personal information about millions of South Korean residents was a grave concern, especially because many websites did not have adequate mechanisms for protecting the sensitive information they collected. In March 2010, the South Korean police arrested individuals who were trafficking personal information belonging to about 20 million South Korean Internet users, which was leaked or stolen from 25 websites [7].

As a result of several incidents involving personal data leakage, the South Korean government changed its requirement, requesting web service providers to adopt i-Pin IDs to identify users instead of RRNs. In June 2008, five credit information providers were certified to issue i-Pin IDs and provide web services with interfaces that validate the IDs [7]. Thus, users were able to create web accounts using i-Pin IDs instead of RRNs. However, according to one study [7], using a pseudonym such as an i-Pin does not protect individuals from phishing attacks. In addition, previous research indicates that the availability and confidentiality of personal information cannot be maintained by the name-registration system as it is currently implemented.

On August 23, 2012, the Constitutional Court of South Korea [4] ruled unanimously that the real-name requirements were unconstitutional. In fact, the court maintained that the provision violated freedom of speech in cyberspace:

> The system does not appear to be beneficial to the public. Despite the enforcement of the system, the number of illegal and malicious postings online has not decreased. Instead, users have moved to foreign websites and the system has become discriminatory against domestic operators. It also prevents foreigners who do not have resident registration numbers from expressing their opinions online.

After the 2012 Constitutional Court decision, real-name registration was terminated in South Korea. At this time, users are not required to supply personal identification credentials when registering for accounts with South Korean online service providers.

The lesson from the South Korean experience is that, while real-name registration may enable a government to regulate and investigate anonymous activities in cyberspace, it endangers personal privacy and infringes on freedom of speech. Other laws related to data privacy and surveillance may also potentially be violated due to the massive amounts of personal information necessarily collected, stored and used by a real-name registration system.

2.3 Real-Name Registration in China

On April 29, 2002, Professor Xiguang Li of Tsinghua University stated on Guangzhou TV that, "The People's Congress should pass legislation that prohibits anyone from anonymously publishing things on the Internet." Since then, government agencies in China have taken legal and technical steps to prevent online anonymity. In 2003, government authorities required people to register with their ID cards before surfing the web at Internet cafes. In 2004, the Internet Society of China published a draft standard for web-based public email services that included a real-name registration requirement. In 2005, real-name registrations for website administrators, QQ group creators and administrators were implemented. By 2012, many online service providers, including `sina.com`, `sohu.com`, `163.com` and `blog.qq.com`, implemented real-name registration. In 2013, the Chinese government announced that real-name registration would be implemented by June 2014 [6].

On January 21, 2013, the Standardization Administration of the Ministry of Industry and Information Technology released a document entitled "Information Security Technology – Guidelines for Personal Information Protection within Public and Commercial Information Systems" [10], which discussed various principles related to personal information protection. It did not clearly identify the relationship between the implementation of the real-name system and the personal information protection guidelines, such as how inconsistencies should be resolved. In 2014, the Chinese State Administration of Press, Publica-

tion, Radio, Film and Television required users to use their real names when uploading videos to websites [8].

At the time of writing this chapter, the main concern of Internet users in China is the security of their personal data collected by the real-name registration system. This chapter discusses techniques for protecting personal data identification information when implementing a real-name registration system.

3. Real-Name Registration Requirements

The basic principle of real-name registration is "real-name at the back, but pseudonym online." Every online user should be able to use an online identity, or pseudonym, but the user's real identity should not be revealed. The real-name of a user may be released only when the user requests its release or if a crime has been committed and the user's real identity is required during the investigation.

An online user may always disclose his/her identity whenever he/she desires to do so. Moreover, an online user should be able to prove that he/she is the owner of an online identity or pseudonym. In a criminal investigation, a law enforcement agency may obtain the real identity of an individual of interest by following the appropriate procedures (e.g., obtaining a court order). Since an individual may be unwilling to disclose his/her real identity in such a situation, a real-name registration system should be able to release a user's identity without the approval or support of the user.

The requirements of a real-name registration system are:

- A user (pseudonym owner) is identified by an online pseudonym; the real identity of the user is hidden.
- A pseudonym owner can prove he/she owns the pseudonym.
- A law enforcement agency can obtain the real identity of a pseudonym owner without the approval of the pseudonym owner upon using appropriate legal procedures.

4. Real-Name Registration Overview

In order to satisfy the real-name registration requirements, the first issue is to select a registration authority (i.e., the entity responsible for conducting the registration process). Clearly, it would be inappropriate for a law enforcement agency to serve as the registration authority. It would also be inappropriate for an online service provider to serve as the registration authority because of the massive amount of personal data

involved and because the online service provider may not have adequate resources to protect the data. Other issues that must be addressed are identifying the entity responsible for collecting personal data and where and how the data would be stored.

The registration scheme proposed in this chapter is designed to avoid potential misuse of personal data and to handle potentially massive volumes of personal data. Also, multiple entities are involved to prevent any one entity from having unrestricted access to personal data. Thus, in addition to a user, the entities involved in online pseudonym or webname registration are a registration center, independent authority and personal data storage center.

When a user registers for an online pseudonym, he/she is required to provide his/her real-name and identity credentials to the registration center. Upon receiving the request from the user, the registration center asks the user to submit personal authentication information, which is a collection of attribute-value pairs. The user's real-name and identity credentials and the binding between the pseudonym and the real-name should be strongly protected.

The registration center encrypts the personal authentication information received from the user using public key encryption and sends the encrypted information to the personal data storage center. Each user is assigned a public/private key pair. The personal data storage center has two major components, a data storage server and a key server. The data storage server is responsible for storing the encrypted personal authentication information supplied by users. The key server has three functions: (i) generating key pairs for users; (ii) partitioning user private keys into four parts; and (iii) storing one part of the private key. Each of the four parts of a user private key is distributed to a different entity: (i) user; (ii) registration center; (iii) private data storage center; and (iv) independent authority.

The proposed encryption scheme uses public key cryptography. The public key is used to encrypt the user's personal authentication information and the private key is used to decrypt the information. The public key is also employed by the user during the authentication step of the online pseudonym registration process. Note that the public key is used differently from how it is used in a public key infrastructure in that it is not made available to the public through a directory service.

Shamir's secret sharing scheme [9] is used to partition the private key into four parts such that combining any three parts yields the private key. Thus, without the part from the user or the part from the independent authority, the private key cannot be reconstructed to decrypt the

personal authentication information, even if the registration center and private data storage center provide their portions of the key.

The independent authority, which holds one part of the private key, can assist in private key reconstruction when a user is unwilling to provide his/her part of the private key. At the end of the registration process, the registration center must ensure that the plaintext form of the personal authentication information is destroyed. Thereafter, the personal authentication information only exists in encrypted form in the private data storage center.

In order to prove pseudonym ownership, the user submits a request to the registration center along with his/her part of the private key. Upon proper authentication, the registration center submits its part of the private key and the user-submitted part to the private data storage center. The private data storage center then combines its part of the private key with the user and registration center parts to reconstruct the private key, and retrieve and decrypt the encrypted version of the personal authentication information stored at the data storage server.

After an online pseudonym is identified during a criminal investigation, a law enforcement agency can follow the prescribed process to retrieve the associated personal authentication information. This typically involves obtaining a court order and requesting the independent authority and registration center to release their parts of the private key associated with the online pseudonym. With proper authorization, the law enforcement agency can then request the private data storage center to retrieve the real-name corresponding to the pseudonym by submitting the private key parts received from the independent authority and registration center. The private data storage center can then combine its part of the private key together with the independent authority and registration key parts to retrieve and decrypt the personal authentication information stored at the data storage server.

User real-names and online pseudonyms have different roles in the real-name registration system. The next two sections present two registration processes in detail: (i) user real-name registration; and (ii) user web-name (online pseudonym) registration. In user real-name registration, the user submits his/her personal authentication information to the registration center, which authenticates the user and creates the binding between the user's real-name and web-name. In user web-name registration, the user sends an online service request along with the web-name to the web service provider and private data storage center. The private data storage center authenticates the user to confirm that the user is, in fact, the owner of the web-name.

5. Privacy-Preserving Real-Name Registration

The proposed privacy-preserving real-name registration scheme employs public key cryptography to secure personal authentication information. For each user, the key server of the private data storage center generates a public-private key pair. The public key is used to encrypt the personal authentication information and the private key is used to decrypt the encrypted information. The public key is also used to encrypt the user's personal authentication information to authenticate the user during the web-name registration process. When a user needs a web-name for an online service, he/she uses the public key to encrypt his/her personal authentication information, which is used to prove that he/she is the real owner of the web-name at the private data storage center. Instead of the user retaining the private key, the key is partitioned into four parts using Shamir's secret sharing scheme and each part is distributed to one of four entities: user, independent authority, registration center and private data storage center. As mentioned above, combining any three parts yields the entire private key, which can then be used to decrypt the encrypted personal authentication information. A user can always request the registration center to retrieve his/her real-name and release it to other entities. However, during a criminal investigation, a law enforcement agency can request the court to order the independent authority and registration center to retrieve and provide the real-name of a person of interest without his/her approval.

Since the registration center, independent authority and private data storage center are crucial components of the name-registration infrastructure and are, as such, high-value targets, they must be secured to the maximum extent. The components should conform to the ISO 27000 family of standards. In particular, they should adhere to ISO/IEC 27001 [5], the best-known standard in the family, which provides requirements for information security management systems.

The next two sections discuss the user real-name registration and user web-name registration processes in detail.

5.1 User Real-Name Registration Process

In user real-name registration, a user makes an in-person appearance at a registration center with documents that prove his/her identity. Upon successful authentication of the user, the registration center requests the user to submit predefined personal authentication information that will be used to authenticate the user in the future. The predefined personal authentication information is a collection of attribute-value pairs (e.g., user ID number, mobile phone number, etc.). The

registration center should publish the required personal authentication information in advance so that the user can collect the necessary information before the time of registration. The registration center encrypts the personal authentication information during the registration process and the plaintext version of this information is destroyed at the end of the registration process.

When the registration center receives the personal authentication information from the user, the registration center requests the key server at the private data storage center to generate a public-private key pair for the user. The key server sends the public key to the registration center, which uses the key to encrypt the personal authentication information submitted by the user. The encrypted personal authentication information is then sent to the data storage server for storage.

The key server then partitions the private key into four parts using Shamir's secret sharing scheme. In particular, the (4, 3) threshold secret sharing scheme is employed, in which the private key is shared with four entities such that any three parts can be combined to yield the private key and thus decrypt the encrypted personal authentication information. One part of the private key is kept at the key server while the other three parts are distributed to the user, registration center and independent authority. The registration center also generates a web-name for the user, which can be based on the user's preference. Upon successful registration, the web-name is provided to the user.

The real-name registration process involves the following steps (see Figure 1):

1. The user requests real-name registration in person at the registration center. After authenticating the user, the registration center assigns a web-name to the user and requests the key server to generate a public-private key pair for the user.

2. The key server creates the public-private key pair (pk_{ID}, sk_{ID}) and a two-degree polynomial interpolation $f_{ID}(x) = a_2x^2 + a_1x + a_0 \; mod(p)$ for the user such that $f_{ID}(0) = sk_{ID}$ where p is a prime number such that $p > a_0$. The key server then generates four pairs (x_{ik}, y_{ik}) where $k = 0 \ldots 3$ and $f_{ID}(x_{ik}) = y_{ik}$ [9]. The pair (x_{i0}, y_{i0}) is kept by the key server; the other three pairs are sent to the user, registration center and independent authority. The key server then destroys the private key sk_{ID}.

3. The user submits his/her personal authentication information to the registration center.

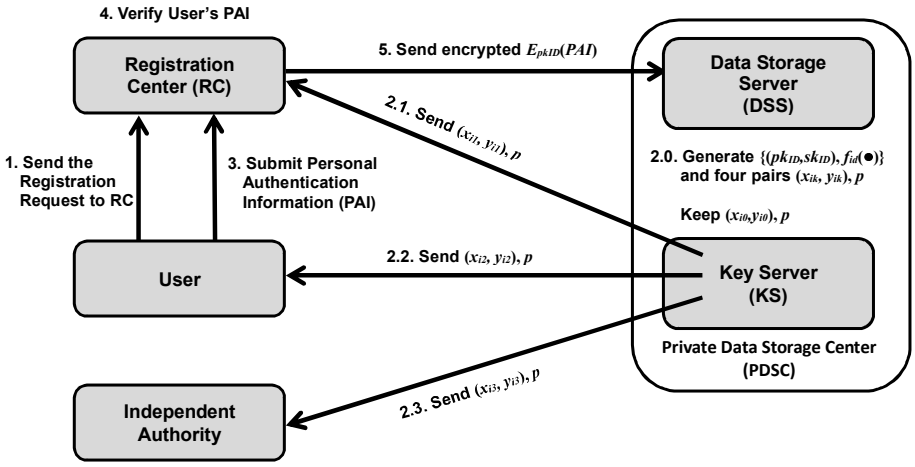

Figure 1. User real-name registration.

4. The registration center verifies the user's personal authentication information.

5. If the user's personal authentication information is valid, the registration center encrypts the user's personal authentication information with pk_{ID} and sends the web-name, pk_{ID} and encrypted $E_{pkID}(PAI)$ to the data storage server.

6. The registration center returns the web-name and public key pk_{ID} to the user and then destroys the plaintext version of the personal authentication information.

7. The registration center maintains the web-name and the f_{ID} along with (x_{i1}, y_{i1}).

5.2 User Web-Name Registration Process

The user web-name registration process is executed when a user wishes to use an online service and has to register his/her web-name with the online service provider. The process requires the online service provider to authenticate the user's web-name using the private data storage center. It is assumed that the user already has received a web-name from the registration center upon completing the real-name registration process.

In web-name registration, the user sends his/her web-name and encrypted personal authentication information to an online service provider. Meanwhile, the user also sends an authentication request with his/her web-name to the private data storage center. After receiving the user

request, the online service provider forwards the user's web-name and encrypted personal authentication information to the private data storage center for authentication.

Having received the user's web-name and encrypted personal authentication information from the online service provider and the authentication request from the user, the private data storage center retrieves the user's encrypted personal authentication information based on the supplied web-name from the data storage server. This enables the private data storage center to verify the user's encrypted personal authentication information sent by the online service provider. If the user's encrypted personal authentication information is valid, then the private data storage center sends a challenge message (random number R) to the user. The user computes and sends the response $E_{pkID}(R)$ back to the private data storage center for authentication. This challenge-response exchange is required to prevent a malicious entity from impersonating the user during the web-name registration process.

If the challenge-response exchange is successful, the private data storage center informs the online service provider that the user's web-name authentication was successful; otherwise, the online service provider is informed that the authentication failed. The online service provider accepts the user web-name registration request if the authentication was successful and the user can use the web-name to access the online service; otherwise, the web-name registration is rejected.

The web-name (WN) registration process involves the following steps (see Figure 2):

1. The user sends a web-name registration request to the online service provider with his/her web-name and $E_{pkID}(PAI)$. Meanwhile, the user sends an authentication request to the private data storage center.

2. The online service provider forwards the user's web-name and $E_{pkID}(PAI)$ to the private data storage center.

3. After the private data storage center receives the authentication request from the user and the web-name and $E_{pkID}(PAI)$ from the online service provider, the private data storage center retrieves the encrypted personal authentication information of the user from the data storage server based on the web-name and compares it with $E_{pkID}(PAI)$ received from the online service provider. If the two encrypted messages match, the private data storage center sends a random number R as a challenge to the user; otherwise, the authentication fails.

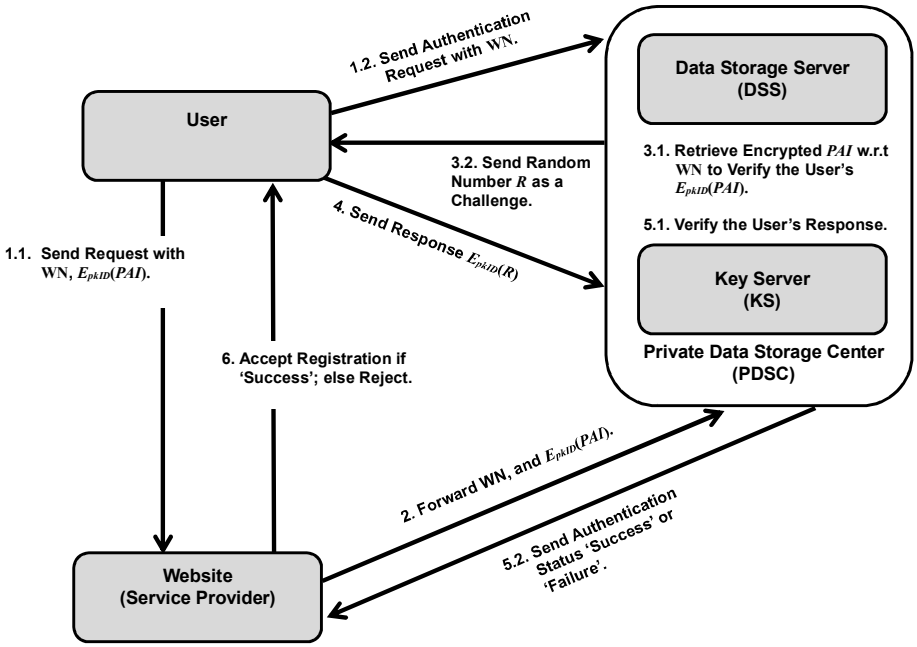

Figure 2. User web-name registration.

4. The user sends the challenge-response $E_{pkID}(R)$ to the private data storage center for authentication after receiving the challenge message R.

5. The private data storage center validates the user's response. If $E_{pkID}(R)$ is valid, then the private data storage center confirms to the online service provider that the authentication was successful; otherwise, the authentication fails.

6. The online service provider accepts the user's web-name registration request if the authentication was successful; otherwise, it rejects the registration request.

5.3 Privacy-Preserving Properties

In the real-name registration process, no personal authentication information is stored as plaintext at the registration center or private data storage center. Only the encrypted version of the personal authentication information is stored at the data storage server of the private data storage center and it can only be decrypted by the user's private key, which is partitioned into four parts and distributed to four entities.

In the web-name registration process, the online service provider does not access the user's personal authentication information as plaintext and is not involved in decrypting the personal authentication information for verification. A challenge-response sequence is used between the private data storage center and the user to ensure that impersonation does not occur during the web-name registration process using the user's web-name and encrypted personal authentication information.

In order to decrypt the encrypted personal authentication information maintained at the data storage server, two entities from among the user, registration center and independent authority are required to supply their portions of the user's private key. Without the participation of the independent authority, only the user can decrypt his/her own personal authentication information upon receiving a request from the registration center.

When a web-name is identified as being of interest in a criminal investigation, the law enforcement agency has to go through a well-defined process that involves obtaining a court order and requesting the independent authority to supply its part of the private key. When the independent authority and registration center supply their parts of the private key corresponding to the web-name, the encrypted personal authentication information associated with the web-name can be decrypted and the real-name corresponding to the web-name can be obtained. The law enforcement agency can then continue its investigation with the real-name in hand.

6. Conclusions

The privacy-preserving real-name registration approach presented in this chapter helps protect the identity of online users while enabling law enforcement to obtain the identities of users involved in criminal activities. Shamir's secret sharing scheme is used to partition a user's private key in four parts, with a different part sent to each of four entities: user, registration center, private data storage center and independent authority. The user's real identity cannot be revealed unless at least three of the four parts of the private key are combined; therefore, except for the user, no other entity can obtain the user's real identity independently. During an investigation, when a person of interest is unwilling to provide his/her real identity, a law enforcement agency can obtain a court order and then request the independent authority to provide its part of the secret key to reveal the person's real identity.

Future research will attempt to implement the real-name registration approach. Also, efforts will be made to evaluate the performance, and

to identify and address the challenges that manifest themselves in a real-world environment.

References

[1] L. Adleman, P. Rothemund, S. Roweis and E Winfree, On applying molecular computation to the Data Encryption Standard, *Journal of Computational Biology*, vol. 6(1), pp. 53–63, 1999.

[2] S. Choe, Web rumors tied to Korean actress's suicide, *New York Times*, October 2, 2008.

[3] S. Choe, South Korea links web slander to celebrity suicides, *New York Times*, October 12, 2008.

[4] Constitutional Court of South Korea, Constitutional Court Decision 2010Hun-Ma47, Seoul, South Korea (`koreanlii.or.kr/w/index.php/2010Hun-Ma47`), August 23, 2012.

[5] International Organization for Standardization, ISO/IEC 27001 – Information Security Management, Geneva, Switzerland, 2013.

[6] O. Lam, The business behind China's Internet real name registration system, *Global Voices Advocacy*, June 10, 2013.

[7] Y. Oh, T. Obi, J. Lee, H. Suzuki and N. Ohyama, Empirical analysis of Internet identity misuse: Case study of the South Korean real name system, *Proceedings of the Sixth ACM Workshop on Digital Identity Management*, pp. 27–34, 2010.

[8] Reuters, China orders real name register for online video uploads, January 21, 2014.

[9] A. Shamir, How to share a secret, *Communications of the ACM*, vol. 22(11), pp. 612–613, 1979.

[10] Standardization Administration of China, Information Security Technology – Guidelines for Personal Information Protection within Information System for Public and Commercial Systems, GB/Z 28828-2012, Beijing, China, 2012.

[11] United Nations General Assembly, Promotion and Protection of Human Rights: Human Rights Questions, Including Alternative Approaches for Improving the Effective Enjoyment of Human Rights and Fundamental Freedoms, Report of the Third Committee, A/67/457/Add.2, General Assembly, Sixty-Seventh Session, New York, December 8, 2012.

[12] Wikipedia, Resident Registration Number (`en.wikipedia.org/wiki/Resident_registration_number`), 2015.