Discovery of Silhouette—a New Threat to Privacy—and Our Efforts to Counter It

Takuya Watanabe

Abstract

To prevent damage due to threats that are unknown to users and businesses, it is important to understand potential security issues in systems before the attackers do and to take preventive measures before an attack occurs. This article describes the mechanism of a new privacy threat called Silhouette, discovered in the course of this sort of empirical research on threats, together with a method for handling the threat, and initiatives to reinforce services and browser security functions around the world.

Keywords: web security, social web services, privacy threats

1. About Silhouette

Social networking services, video sharing, and other social web services (SWSs), which create content through communication between people, have continued to evolve since they were first introduced and have now become an essential part of our lives. A survey [1] of Internet users revealed that the average person maintains five or more different SWS accounts. On the SWS sites, users' profiles and postings can be seen based on the account name, so personal information such as the name, photographs, and the activities of the person are linked to each account.

The Silhouette privacy vulnerability discovered by NTT Secure Platform Laboratories (NTT SC Labs) makes it possible for a third party to identify the SWS accounts held by a user when the user accesses a website of the third party. For example, when a malicious website unrelated to any SWS is accessed by means such as through a search engine, advertising on a public site, or a link in an email message, the malicious website can communicate with an SWS that the user may have an account with and collect information to identify the account name. This can be done in the background without the user's knowledge.

For this to work, the user needs to have left the web browser of their personal computer or mobile terminal logged-in to an SWS that is vulnerable to the threat and must visit the malicious website. Generally with SWSs, users automatically remain logged in until they logout explicitly and the cookies* are deleted. For this reason, users that have used a vulnerable SWS even once in the past may be identifiable by the malicious party.

2. Mechanism(s) resulting in the threat

This threat is carried out by maliciously exploiting the user-blocking function widely available on SWSs (**Fig. 1**). The user-blocking function is intended to enable ordinary users to control whether other users, who may be undesirable, can view their page, thus protecting themselves from behaviors such as harassment or spam. NTT SC Labs has identified a latent security issue in the user-blocking feature, which can control whether a page can be viewed by both malicious and legitimate users alike.

The malicious third party must first create multiple accounts on the SWS. These are known as *signaling accounts*. Then they systematically block certain

* Cookie: A feature enabling web services to store information in the browser of a visiting user so that it can manage sessions, user settings, login state, and other information.

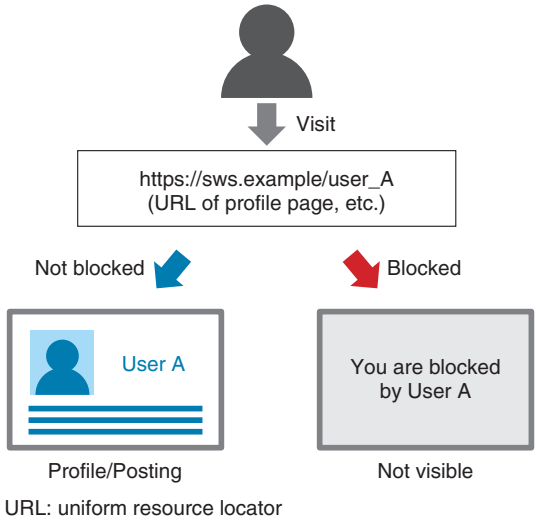

Fig. 1. User-blocking function.

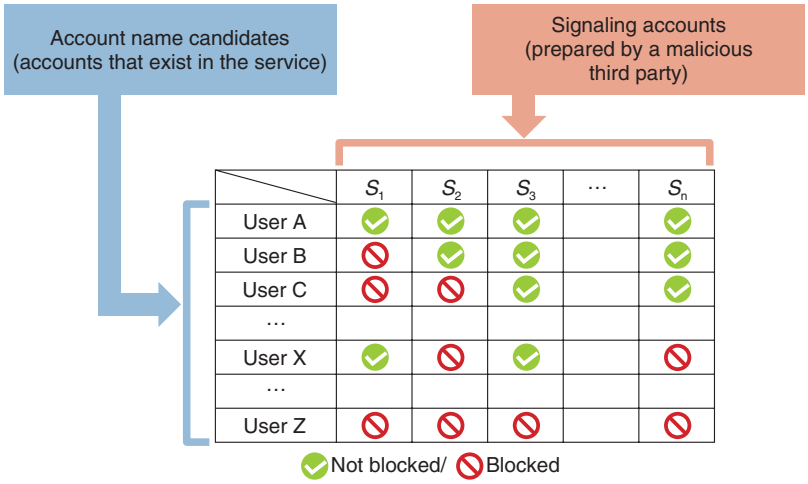

Fig. 2. Examples of patterns of blocked and non-blocked accounts.

users on the service to create combinations of blocked and non-blocked users. These patterns can then be used to uniquely identify user accounts (**Fig. 2**).

When a user visits the malicious website, the website contains a script for identifying the user account name, and the script makes requests for the pages of each of the signaling accounts. Browsers use the same-origin policy to protect sites, preventing data from leaking to other sites, so the third party is not able to obtain the content of the responses to this communication directly (**Fig. 3**). However, there is a statistical difference in response times for blocked and non-blocked requests. The malicious third party can use these differences to infer whether the visiting user is blocked from each of the signaling accounts. These results can then be used to identify the user's account name on the SWS by referring to the account blocking patterns prepared earlier.

Silhouette would be classified as a cross-site request forgery (CSRF) and side-channel attack. A CSRF is a type of web attack in which requests are sent to sites not intended by the user, to steal data or execute some malicious code. A side-channel attack is a generic name for attacks that use information

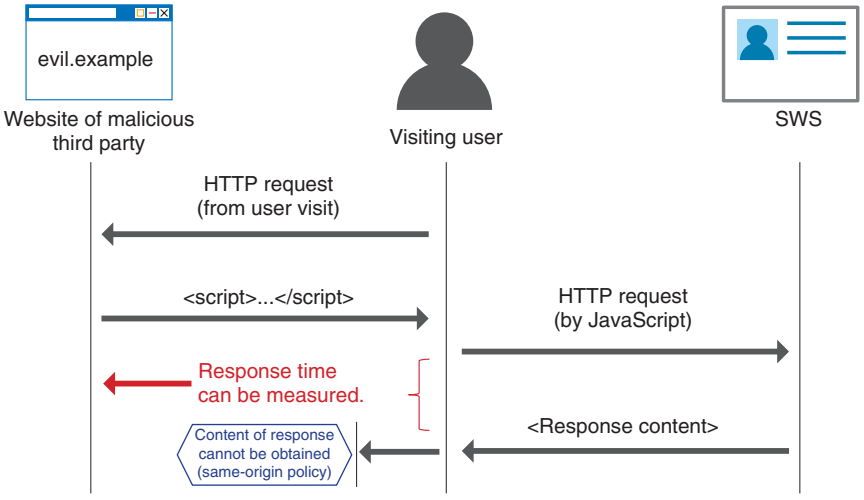

Fig. 3. Protecting response content using the same-origin policy.

from physical space, such as the response time or the power consumed, to infer sensitive information. This research has identified a security issue in how these services are designed, which enables threats to the privacy of legitimate users. This is done by making malicious use of the user-blocking function widely used by SWSs in an attack that combines a CSRF and a side-channel attack.

3. Countermeasures

In this section, countermeasures to this threat are described, which can be taken by both SWS operators and by users. Since the threat is a combination of CSRF and side-channel attacks, countermeasures can be implemented by preventing either of these attacks. Countermeasures for side-channel attacks require a specialized perspective on the characteristics of response timing, but there are well-known countermeasures for CSRF that only involve changes to the programming of the web service [2]. Below, countermeasures that focus on the CSRF component of the attack are introduced.

3.1 Assumptions

SWSs that are susceptible to this threat must have an account registration function and a user-blocking or similar function that enables a user to change whether another user is able to view the user's content pages (their profile etc.). Services without these functions are not vulnerable to this threat.

3.2 Countermeasures for the SWS

The first measure that an SWS can take is to use the cookie option called the SameSite attribute. A cookie with the SameSite attribute prevents requests from being sent to other sites by JavaScript or other means. As such, if this attribute is specified in the cookie used to manage the login state, a CSRF can be broadly prevented, including this threat. However, to use this feature, the user's browser must support SameSite, and the SWS must declare that it is using SameSite in the Hypertext Transfer Protocol (HTTP) header. As described below, the major browsers used around the world now support SameSite and can handle Silhouette thanks to the efforts of NTT SC Labs.

The second measure that can be taken is called request verification. In a CSRF, HTTP requests not intended by either the user or the service are generated. When this occurs, a well-known countermeasure [3] is to have the SWS or other service determine whether the request is legitimate by checking the referrer, which identifies the URL (uniform resource locator) of the website sending the request, or by checking a request parameter included as a CSRF countermeasure that contains a special code. Request verification is usually used for pages received using the POST method, such as posting to a web service, but it can also be used for pages received using the

GET method, such as a user profile. However, in this case, the verification fails when it is linked directly from a search engine or blog article, and such cases could be rejected as illegitimate requests.

To deal with this, the service can add a procedure in which it returns an intermediate page when the verification fails, and JavaScript on the intermediate page retrieves the content. This increases the number of requests needed to display the page, but it enables implementation of the countermeasure without obstructing access through direct links.

3.3 Countermeasures for users

One measure that can be taken by users is to use the private browsing mode in their browser. This mode has different names in different browsers, for example, Secret Mode, Private Window, or InPrivate, and when it is enabled, the browser does not use any prior cookie information and will delete any new cookie data stored when the session is ended. The threat of having the user's account name identified by visiting a third-party site can thus be prevented by enabling private browsing.

Another measure that users can take is to log out of the SWS. The threat can identify account names only if the user is logged in to the SWS. As such, the threat can be avoided by logging in to the service every time they use it and logging out as soon as they have finished using it.

4. Threat prevention efforts

NTT SC Labs has established a procedure to evaluate whether an SWS is vulnerable to Silhouette and has conducted a survey of all NTT Group SWSs and external SWSs that are well known around the world. As a result, we identified some well-known and influential international services that could result in account names being identified through this vulnerability. We have shared details of the vulnerability and countermeasures with these operators and collaborated in tests to verify the effectiveness of the countermeasures.

Through this effort, Twitter and other SWSs changed their specifications to improve security mechanisms and prevent the threat that makes it possible to identify account names. Major browsers including Microsoft Edge, Internet Explorer, and Mozilla Firefox also now support the SameSite cookie attribute using the method from this research or a similar method to prevent this vulnerability from being exploited.

Due to this contribution, the safety of most SWSs, used by more than 600 million users around the world, has greatly increased, and operators, including NTT, are able to use advanced functionality in designing secure web services. The results of this research have therefore led to a safer environment for using the Internet for users from short, medium, and long-term perspectives.

The paper [2] summarizing the discovery of this threat, its verification, and countermeasures, has also made an extremely great impact on improving web security. It was the first from Japan selected for the IEEE European Symposium on Security and Privacy, which is a prestigious academic conference, and was also selected for Black Hat Europe [4], a very influential international conference in the cybersecurity industry.

5. Future prospects

As part of research and development on cybersecurity at NTT SC Labs, we are developing methods for evaluating new threats, including the recently discovered Silhouette threat, and when an issue is discovered, we will work to implement countermeasures in collaboration with relevant organizations. By continuing to discover latent threats and develop countermeasures in the future, we will strive to be able to provide robust services, promote more secure web services and browsers, and facilitate safe and secure use of the Internet.

References

[1] Brandwatch, "121 Amazing Social Media Statistics and Facts." https://www.brandwatch.com/blog/amazing-social-media-statistics-and-facts/

[2] Information-technology Promotion Agency, Japan, "How to Secure Your Website," 5th Edition, 2011. https://www.ipa.go.jp/files/000017318.pdf

[3] T. Watanabe, E. Shioji, M. Akiyama, K. Sasaoka, T. Yagi, and T. Mori, "User Blocking Considered Harmful? An Attacker-controllable Side Channel to Identify Social Accounts," Proc. of 2018 IEEE European Symposium on Security and Privacy (EuroS&P), pp. 323–337, London, UK, Apr. 2018.

[4] T. Watanabe, "I Block You Because I Love You: Social Account Identification Attack Against a Website Visitor," Black Hat Europe 2018, London, UK, Dec. 2018. https://www.blackhat.com/eu-18/briefings/schedule/index.html#i-block-you-because-i-love-you-social-account-identification-attack-against-a-website-visitor-12912

Trademark notes

Takuya Watanabe
Researcher, NTT Secure Platform Laboratories.
He received an M.E. in computer science and engineering from Waseda University, Tokyo, in 2016. He joined NTT in 2016 and has been engaged in research and development of the cybersecurity project.