A Mixed Content Awareness Tool for Safe Browsing

Yoshio Kakizaki*, Shiomi Nishibiro, Ryoichi Sasaki
Tokyo Denki University, Japan.

* Corresponding author. Tel.: +81-3-5284-5583; email: kakizaki@im.dendai.ac.jp
Manuscript submitted June 14, 2016; accepted January 7, 2017.
doi: 10.17706/jcp.13.1.90-99

Abstract: Web pages that contain both content that are served over HyperText Transfer Protocol (HTTP) and content that are served over HTTPS are called mixed content pages. Such pages are susceptible to eavesdropping and modification by third parties owing to the unsecured HTTP communication. Consequently, various countermeasures have been implemented in modern Web browsers with the aim of protecting users from such risks. However, the implemented measures have proven inadequate. In this study, we designed and implemented a tool that enables safe browsing by ensuring that users take action when mixed content pages are encountered. The tool provides five functions that can be applied along with a browser's native countermeasure to notify users of mixed content. The results of experiments conducted to evaluate the usability and effectiveness of our tool showed that it made most users aware when they encountered mixed display (passive) content. Thus, our tool effectively informs and protects users from mixed content risks.

Key words: Mixed content, safe browsing, awareness tool.

1. Introduction

HyperText Transfer Protocol (HTTP) is the foundation of data communication on the World Wide Web. However, because it does not provide encrypted communication, third parties can eavesdrop and modify the contents of communication over this protocol. Thus, when there is a need to communicate private information, HTTPS is usually employed. HTTPS facilitates communication over HTTP within a connection encrypted by Transport Layer Security (TLS) or Secure Sockets Layer (SSL). Third parties cannot eavesdrop and modify the contents of HTTPS communications because end-to-end encryption is employed.

If an HTTPS page contains HTTP content, which is unencrypted, then the communication with that page is only partially encrypted. We call a page with such content a mixed content page [1]. The unencrypted content can be overheard and also modified because the connection is not secure. A malicious attacker can modify the unencrypted content and may obtain the user's private information by fraudulent means.

In an effort to obviate the risks to users from mixed content, various countermeasures have been implemented in modern Web browsers such as Google Chrome and Mozilla Firefox [1]-[3]. However, according to a survey we conducted (see Section 4), most users do not understand the risks associated with mixed content. As a result, users tend to be lax in taking appropriate countermeasures against mixed content; for instance, changing browser setting to block mixed content, and not putting private information on mixed content pages. In other words, users are still susceptible to mixed content risks.

In this study, we designed and implemented a tool that detects mixed content and ensures that users take action in order to facilitate safe browsing. The tool informs and protects users from mixed contents.

Experiments conducted indicate that our tool is effective.

2. Related Work

The efficacy of SSL warnings has been researched from various viewpoints. Sunshine *et al*. surveyed over 400 Internet users to examine their reactions to and understanding of current SSL warnings [4]. Their warnings performed significantly better than existing warnings, but far too many participants exhibited dangerous behaviour in all warning conditions. Their results suggested that, while warnings can be improved, a better approach may be to minimise the use of SSL warnings altogether by blocking users from making unsafe connections and eliminating warnings in benign situations.

Huang *et al*. designed and implemented a tool that detects the occurrence of SSL man-in-the-middle attacks on Facebook [5]. Over three million real-world SSL connections to this website were analysed. The results obtained indicated that 0.2% of the SSL connections analysed had forged SSL certificates.

Akhawe and Felt empirically assessed whether browser security warnings are as ineffective as suggested by popular opinion and the literature [6]. They found that, in contrast to other warnings, users continued through 70.2% of Google Chrome's SSL warnings. This indicated that users' experiences with such warnings could have a significant impact on their behaviour.

Rachna *et al*. provided empirical evidence showing which malicious strategies are successful at deceiving general users [7]. They assessed hypotheses with a usability study in which 22 participants were shown 20 Web sites and asked to determine which ones were fraudulent. They reported that 23% of the participants did not look at browser-based cues such as the address bar, status bar, and security indicators—leading to incorrect choices 40% of the time. Pop-up warnings about fraudulent certificates were ineffective: 15 out of the 22 participants proceeded without hesitation when presented with warnings. These results illustrate that standard security indicators are not effective for a substantial fraction of users, and suggest that alternative approaches are needed.

Sascha *et al*. sought to better understand the potential security threats posed by benign Android apps that use the SSL/TLS protocols to protect data they transmit [8]. Consequently, they conducted an online survey to evaluate users' perceptions of certificate warnings and HTTPS visual security indicators in Android's browser. They reported that 419 of 754 participants had not seen a certificate warning before and typically rated the risk they were warned against as medium to low.

Chaitrali *et al.* pointed out that the drastic reduction in screen size and the accompanying reorganisation of screen real estate significantly change the use and consistency of the security indicators and certificate information that alert users of site identity and the presence of strong cryptographic algorithms in the mobile computing environment [9], [10]. They performed the first systematic and comprehensive comparison between mobile, tablet, and traditional browsers based on the best practices set forth in the W3C guidelines. They found that whereas desktop browsers largely conform to these guidelines, mobile and tablet browsers fail to do so in numerous instances.

Felt *et al.* designed a new SSL warning with the goal of improving comprehension and adherence [11], [12]. Browsers display SSL warnings when the encryption is too weak or the server cannot be authenticated. However, prior studies have shown that confusion about SSL warnings is widespread, and users overwhelmingly ignore some SSL warnings. They analysed large-scale field data from Mozilla Firefox which show that Firefox users adhere to nearly 70% of SSL warning impressions, which is a good adherence rate [6]. Conversely, they also found that users adhere to only 30% of Google Chrome SSL warnings [6]. They therefore designed a new SSL warning that informs (or, failing that, convinces) users, and released it as the new Google Chrome 37 SSL warning [11]. They also briefly ran a version of the Firefox SSL warning in Google Chrome, and found that Google Chrome SSL warning adherence rates increased to 44%.

3. Mixed Content

If an HTTPS page contains content that is retrieved through regular, cleartext HTTP, then the connection is only partially encrypted; the unencrypted content is accessible to sniffers and can be modified by man-in-the-middle attackers, so the connection is not secure. When a Web page exhibits this behaviour, it is called a mixed content page [1]. Using these resources, an attacker can often take complete control over the page, not just the compromised resource. This scenario is, unfortunately, quite common on the Web, which is why browsers cannot simply block all mixed requests without restricting the functionality of many sites [2].

There are two categories of mixed content: mixed passive/display content and mixed active content [1], [2].

3.1. Mixed Passive/Display Content

Mixed passive/display content is content served over HTTP that is included on an HTTPS webpage, but which cannot alter other portions of the webpage [1]. The HTML img, audio, video, and object elements are considered passive content. For example, an attacker can intercept HTTP requests for images on a site and swap or replace them; swap the save and delete button images, causing users to delete content without intending to; or even replace product pictures with ads for a different site or product [2].

3.2. Mixed Active Content

Mixed active content poses a greater threat than mixed passive content. An attacker can intercept and rewrite active content and thereby take full control of a page or even an entire website. This allows the attacker to change anything about the page, including displaying entirely different content, stealing user passwords or other login credentials, stealing user session cookies, or redirecting the user to a different site entirely [2]. HTML script, links, iframes, object elements, some Cascading Style Sheets (CSS), and XMLHttpRequest object are considered active content.

In the mixed active content case, a man-in-the-middle attacker can intercept requests for the HTTP content and rewrite the responses to include malicious JavaScript code. Malicious active content can steal a user's credentials, acquire sensitive data about the user, or attempt to install malware on the user's system [3].

3.3. Alert and Indication of Threats

Contents	IE	Edge	Chrome	Firefox
http				
mixed active			https:	
mixed passive			https:	
https			https:	

Fig. 1. Alert and indication of each browser by type of content.

To prevent threats from mixed content, each Web browser blocks such content and displays an alert. Fig. 1 shows alerts and indication of each browser by type of content. We investigated the following browsers on a Windows 10 Pro 10586 computer:

- Microsoft Internet Explorer 11.162.10586.0
- Microsoft Edge 25.10586.0.0
- Google Chrome 49.0.2623.87
- Mozilla Firefox 45.0.1

Microsoft Internet Explorer 11 and Edge 25 only give a brief indication that other content apart from

HTTPS content is present. Chrome and Firefox indicate according to contents in more detail. The indication of mixed passive content is different but their indications are almost the same. Chrome indicates the same indication as HTTP content although HTTPS is used. Edge, Chrome, and Firefox block mixed active content by default. Moreover, Chrome and Firefox alert users of mixed passive content by default. In this manner, modern browsers attempt to remove the threat to the user by blocking mixed active content.

3.4. Countermeasures

There are two kinds of countermeasures for mixed content: server side and client (browser) side. We enumerate those countermeasures below.

- Server side countermeasures:
  - Rewrite HTTP to HTTPS
  - Remove protocol scheme
  - Use content security policy (CSP)
- Client (browser) side countermeasures:
  - Alert warning
  - Blocking
  - Force HTTPS

Web developers can do correspondence that rewrite all HTTP to HTTPS, or remove the protocol scheme, such as http:// to //. Of course, the relative URL is a good measure. W3C recently defined the upgrade-insecure-requests CSP directive [13] and the block-all-mixed-content CSP directive [14]. Thus, all mixed content are improved by either delivering a Content Security Policy HTTP header or by embedding the policy in a meta-element. As explained in Section 3.3, Edge, Chrome, and Firefox block mixed active content by default. Further, Chrome and Firefox alert users of mixed passive content by default. In addition, Internet Explorer and Firefox can block mixed passive content via their advanced setting. Moreover, some add-ons or extensions can force the use of HTTPS.

4. Preliminary Survey

4.1. Method

We conducted a preliminary survey to determine how many users can identify mixed content. We created two Web pages—one with HTTPS only, the other containing HTTPS and HTTP contents (mixed passive contents). Participants accessed the two Web pages using the Web browser with which they were accustomed. Then, we determined whether the participants had noticed any difference between the two Web pages via a questionnaire that included the following questions.

- Q1-1. Did you notice any difference between the two Web pages?
- Q1-2. What exactly was different?
- Q1-3. Do you know mixed content?
- Q1-4. Please explain what you know about mixed content.

4.2. Result and Discussion

We obtained 113 valid responses to our preliminary survey. The results obtained are shown in Tables 1 and 2.

As can be seen in Table 1 (Q1-1), 92% of participants did not notice any difference between the two Web pages. Moreover, nine participants, who had answered 'Yes', stated the exact difference (Q1-2) as follows:

- Difference between HTTP and HTTPS
- Certificate is signed by an Authority or is self-signed
- Modern security or legacy security
- One may contain insecure content

Table 1. Q1-1. Did You Notice Any Difference between the Two Web Pages

	Yes	No
Answers (%)	9 (8%)	104 (92%)

Table 2. Q1-3. Do You Know Mixed Content

	Yes	Partially	No
Answers (%)	2 (1.8%)	38 (33.6%)	73 (64.6%)

As can be seen in Table 2 (Q1-3), only two participants stated that they know mixed content. We requested a brief description from the two participants. One participant stated that, "Some of the content was not encrypted". The other participant did not respond.

As regards the overall questionnaire, some participants answered 'No' to Q1-3 even though they had answered 'Yes' to Q1-1. In addition, a participant answered 'Yes' to Q1-3 even though he/she had answered 'No' to Q1-1.

Our results show that most participants do not understand mixed content and its risk. Generalising these results, it is safe to assume that many users do not understand the risks inherent in mixed content while browsing the Internet. As a result, users may not change the setting of their browser to block mixed content. Moreover, users do not know that eavesdropping is possible even though the page is HTTPS.

In short, the alert displayed by Web browsers warning about mixed content is insufficient. Therefore, it is necessary that an alert that warns the user about the risks posed by mixed content be displayed.

5. A Mixed Content Awareness Tool

Modern Web browsers display warnings to users when mixed content pages are encountered. However, as indicated by our preliminary survey, users may not notice these warnings. Thus, the challenges can be itemised as follows:

1. Web browser security warnings.
    A. Warning statements and display methods are different in each Web browser.
    B. Users fail to notice the warnings.
    C. Users do not know the location of the mixed content on the Web page.
2. Visibility of mixed content to users is low.

To ensure users are cognisant of the risks posed by mixed content, we developed a mixed content awareness tool that reveals the location of the mixed content on a Web page. To solve the above-mentioned problems, we defined our objective as follows:

1. The warning should clarify the whereabouts of the mixed content.
2. Warnings should increase user awareness.
3. Warnings should indicate the risk of mixed content being transmitted.

Problem 1-B corresponds to objective 2, problem 1-C corresponds to objective 1, and problem 2 corresponds to objective 3. Objective 1 includes not only visible contents such as images and movies, but also invisible contents such as CSS and JavaScript.

Achieving the three objectives above would make it easy for users to notice mixed content more than they do the Web browser's alert displays, and also for users to obviate the risk of mixed content. Our tool provides the following functions:

- Detection
- Warning
- Highlighting
- Blocking

- Enforced TLS

The detection function detects mixed content under TLS communication. The warning function displays the amount of mixed content detected, and explains its threat. Visible mixed content, such as images and movies, are plainly highlighted by the highlight function. The blocking function prevents the threat by filtering the communication of mixed content. The enforced TLS function rewrites HTTP, which is the protocol of detected mixed content, to HTTPS, to prevent modification by malicious attackers.

5.1. User Interface Consideration

As explained in Sections 3.3 and 4.2, the alert given by Web browsers about mixed content is insufficient. We determined that our tool should detect the existence of mixed content and effectively notify the user. To this end, we conducted a user experiment to ascertain what kind of warning display would be effective to users.

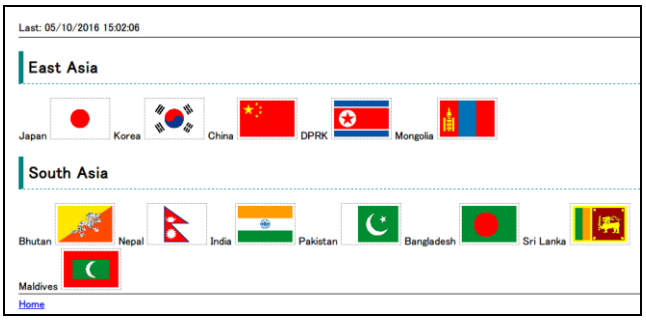

Fig. 2. Web page containing mixed display contents composed of three types of content.

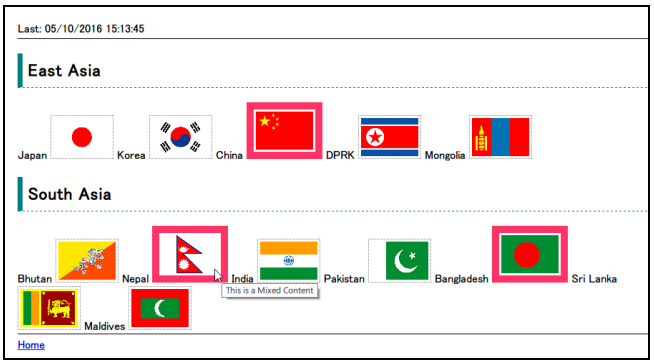

Fig. 3. Warning display Concept 1 (red frame).

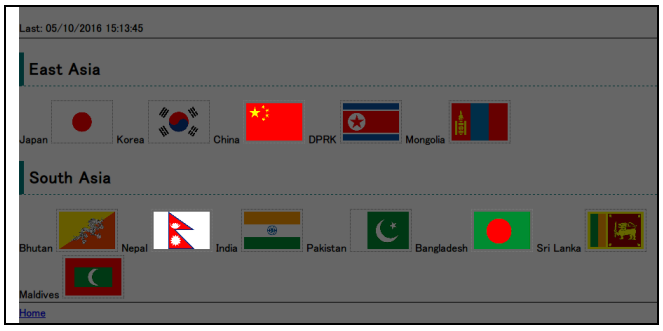

Fig. 4. Warning display Concept 2 (black out).

First, we explained the concept of mixed content to participants. Then, we asked the participants to access a mixed content Web page that contained three different types of content (shown in Fig. 2). Next, the participants tested two alarm display concepts, and were asked the question: "Which warning display is good for noticing mixed content?" We showed them warning display Concepts 1 and 2 in Figs. 3 and 4,

respectively. In Fig. 3, the mixed display content object is highlighted by surrounding it with a red frame. Moreover, a tooltip is displayed when the mouse cursor is over the mixed display content. In Fig. 4, the mixed display content is highlighted by darkening the entire screen but excluding the mixed display content.

We obtained 22 valid responses in this experiment. The results are shown in Table 3. Most of the participants stated that Concept 2 (Fig. 4) was good. Some participants gave reasons such as, "Concept 2 was an abnormal display, so I noticed it", and "Concept 1 was not noticed easily because it is a familiar design". Consequently, we decided to implement alarm display Concept 2 (Fig. 4) in our tool.

Table 3. Which Warning Display Is Good for Noticing Mixed Content?

	Concept 1 (Fig. 3)	Concept 2 (Fig. 4)
Answers (%)	2 (9%)	20 (91%)

5.2. Implementation

We implemented this mixed content awareness tool as a Google Chrome extension because, as at April 2016, Google Chrome had the most users globally [15]. The basic steps used in the algorithm employed in our tool were as follows:

1. Does the active page contain mixed content?
2. If this active page contains mixed content, extract all mixed contents.
3. If Enforced TLS option is enabled, rewrite all HTTP to HTTPS and exit.
4. Display the number of mixed active/passive content and their information in an alert pop-up.
5. Inform the user of the mixed passive content using the display method outlined in Section 5.1.

5.2.1. Alert pop-up

We implemented the alert pop-up indicating that a Web page contains mixed content as shown in Fig. 5. The pop-up is a modal window that prevents the Web page from being browsed if the 'OK' button is not pushed. Thus, the user is forced to acknowledge the alert pop-up. We also chose to indicate the mixed content found on the alert pop-up, because mixed active content is not visible content. When the mixed active content is included, the URL of the target object is shown in the alert pop-up, and the user notices it. We believe that objectives 1 and 2 can be satisfied by using this type of alert pop-up. On closing the alert pop-up, the warning display outlined in Section 5.2.2 is shown.

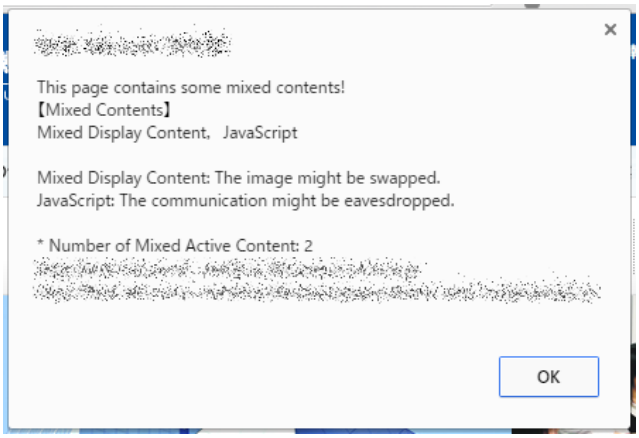

Fig. 5. The alert pop-up.

5.2.2. Warning display

We implemented the warning display in a Web page containing mixed content based on the results given in Section 5.1 (as shown in Fig. 4). After the warning display is closed, the user can click any arbitrary part of the screen to browse the Web page.

6. Evaluation and Discussion

6.1. Method

To determine whether our tool satisfied the objectives outlined, we conducted a user experiment in which we asked each participant the following questions corresponding to defined evaluation criteria:

- EC-1. User-friendly warnings (design): Did you understand the meaning of the warnings?
- EC-2. Clarity of meaning: Did you feel that the warning display was comprehensible?
- EC-3. Usability: Will you continue using this tool?

Evaluation criterion 1 corresponds to objectives 1 and 2, evaluation criterion 2 corresponds to objective 3, and evaluation criterion 3 corresponds to all objectives.

First, we explained the concept of mixed content to all participants and explained the operation and function of our developed tool. Then, we conducted two evaluation experiments with the aid of the participants using Google Chrome extended with our developed tool. A Web page containing four types of contents, two images and two JavaScripts, was used in the experiments. The participants were tasked with finding the mixed contents in the Web page. Evaluation experiment 1 was conducted using Google Chrome without our developed tool. Evaluation experiment 2 was conducted using Google Chrome with our developed tool as an extension.

After the two experiments, we ascertained how many mixed contents examinees had found using a questionnaire. The three evaluation criteria were assessed using a five-point Likert scale.

- Q2-1. How many mixed contents did you find using the Web browser only?
- Q2-2. How many mixed contents did you find using our tool?

6.2. Results

We obtained 52 valid responses in this evaluation experiment. The results obtained are shown in Tables 4 and 5. Table 4 shows that 43 participants (83%) did not find any mixed content using the Web browser only (Q2-1 and Q2-2). Only five participants stated that they had found some of the mixed content. Further, three participants stated that they had found more than four mixed content, even though the Web page contained only four mixed content. In contrast, 51 participants stated that they had found two mixed content using our tool, and the remaining participant stated that one mixed content had been found.

Table 4. How Many Mixed Content Did You Find

	0	1	2	3	4	>4
Web browser	43	3	2	1	0	3
Our tool	0	1	51	0	0	0

Table 5. Three Criteria Assessed on a Five-Point Likert Scale

	1	2	3	4	5
EC-1	0	2	15	18	17
EC-2	0	0	4	21	27
EC-3	0	4	24	21	3

Table 5 shows the mode values for EC-1, EC-2, and EC-3 as four, five, and three, respectively. Eighteen of the 52 participants gave EC-1 a value of four, and 67% of the participants gave an evaluation value of four or more. As for EC-2, the warning display of our tool was given a high evaluation, as expected from the evaluation results in Section 5.1. Almost 50% of the participants evaluated EC-3 as three.

6.3. Discussion

Table 4 (Q2-1) shows that very few of the participants found any mixed content using the Web browser only; in fact, some participants appeared to have been confused. In contrast, Table 4 (Q2-2) indicate that all

examinees were able to find mixed display content (images); however, none was able to find the two JavaScripts (mixed active content). We implemented our tool to present information on all mixed content using our alert pop-up outlined in Section 5.2. However, the participants did not receive the expected information.

On the other hand, the participants were cognisant of mixed display (passive) content with our tool. Fifty-one of the 52 participants were able to find all mixed display content, and the other participant was able to find only one. We believe that the participant found only the mixed content in the header part because that participant did not scroll down the experimental Web page.

From another aspect, the participants were unable to find four mixed content with Q2-2, even though EC-1 had been given a high evaluation. Most participants noticed only visible content, even though our tool showed that two mixed passive content and two mixed active content were included in the Web page; thus, it is clear that visualisation is important.

The evaluation given to EC-3 was low. Participants gave comments such as, "Pop-up is annoying" and "It is troublesome when displayed every time". This indicates that we need to refine our tool; for example, some functions could be turned off according to the user's desire and a simplified notification of the warning level could be displayed using the colour of the extension's icon. However, we surmise that users may not notice the colour of the extension's icon because many users did not notice the indication by the browser of insecure connection on TLS, such as in Fig. 4, as explained in Section 4.2. We may improve our tool by recording the Web pages the user browsed and present a warning to the user once a day, and also by using a whitelist.

Our tool has Enforced TLS function. When this function is enabled, all mixed content communicate with secure communication, even when the user has not noticed the mixed content. Web browsing will become more secure in modern browsers when server side developers correctly define the CSP directive [13], [14].

In summary, our tool can detect mixed display content and notify users but it is still inadequate for mixed active content. However, because most modern browsers block mixed active content by default, as explained in Section 3.3, we believe that this is not an issue. Our tool enables users to be more aware of mixed content than modern browsers. Finally, we can state that our tool achieves all the objectives enumerated in Section 5.

7. Conclusion

In this study, we designed and implemented a tool that promotes safe browsing by detecting, notifying, and ensuring that users take action on mixed content. Our tool was implemented as a Google Chrome extension, and in addition to the browser's native function, provided five functions to safeguard against mixed content. The results of evaluation experiments conducted to verify the usability and effectiveness of our tool indicate that it made most users aware of mixed display (passive) content. Further, when the Enforced TLS function of our tool is enabled, all mixed content communicate with secure communication, even without the user noticing the mixed content. Thus, our tool is in fact very useful. The results also showed that most users noticed only visible content, even though our tool showed mixed passive (display) content and mixed active content. This means that visualisation is important. We aim to implement this aspect in future work.

In conclusion, we can state that our tool informed and protected users from mixed content risks.

References

[1] Mixed Content (Security). Retrieved 2016, from https://developer.mozilla.org/en-US/docs/Security/Mixed_content

[2] *Google Developers, What is Mixed Content?* Retrieved 2016, from

https://developers.google.com/web/fundamentals/security/prevent-mixed-content/what-is-mixed-content.

[3] Mixed Content Blocking in Firefox. Retrieved 2016, from https://support.mozilla.org/en-US/kb/mixed-content-blocking-firefox.

[4] Joshua, S., Serge, E., Hazim, A., Neha, A., & Lorrie, F. C. (2009). Crying wolf: An empirical study of SSL warning effectiveness. *USENIX Security*, 399-416.

[5] Erling, E., Collin, J., *et al*. (2014). Analyzing forged SSL certificates in the wild. *IEEE Security and Privacy*, 83-97.

[6] Devdatta, A., & Adruenne, P. F. (2013). Alice in warningland: A large-scale field study of browser security warning effectiveness. *USENIX Security,* 257-272.

[7] Rachna, D., Tygar, J. D., & Marti, H. (2006). Why phishing works. *ACM SIGCHI 2006*, 581-590.

[8] Sascha, F., Marian, H., Thormas, M., Matthew, S., Lars, B., & Bernd, F. (2012). Why eve and mallory love android: An analysis of android SSL (In)security. *ACM CCS 2012*, 50-61.

[9] Chaitrali, A., Patrick, T., & Paul, C. van O. (2012). Measuring SSL indicators on mobile browsers: Extended life, or end of the road? *ISC'12, LNCS7483*, Springer, 86-103.

[10] Chaitrali, A., Patrick, T., & Paul, C. van O. (2015). An empirical evaluation of security indicators in mobile web browsers. *IEEE Transactions on Mobile Computing, 14(5),* 889-903.

[11] Adrienne, P. F., Hazim, A., & Sunny, C. (2014). Experimenting at scale with Google chrome's SSL warning. *ACM CHI 2014*, 2667-2670.

[12] Adrienne, P. F., Alex, A., Robert, W. R., Sunny, C., Somas, T., Alan, B., Helen, H., & Jeff, G. (2015). Improving SSL warnings: Comprehension and adherence. *ACM CHI 2015*, 2893-2902.

[13] Upgrade Insecure Requests. Retrieved 2016, from https://www.w3.org/TR/upgrade-insecure-requests/

[14] Mixed Content. Retrieved 2016, from https://www.w3.org/TR/mixed-content/

[15] Desktop Top Browser Share Trend–NetMarketShare. Retrieved 2016, from http://netmarketshare.com/

Yoshio Kakizaki received his B.S., M.S., and Ph.D. degree in engineering from Tokai University, Japan, in 2003, 2005, and 2008, respectively. Currently, he is an assistant professor in Tokyo Denki University, Japan. His research interests include information security and information system.

Ryoichi Sasaki received his B.S. Degree in health science and Ph. D. Degree in system engineering from Tokyo University in 1971 and 1981, respectively. Between April, 1971 and March, 2001, he was engaged in the research and research management on systems safety, network management and information security at Systems Development Laboratory of Hitachi Ltd. Now, he is a professor of Dept. of Information Systems and Multi Media at School of Science and Technology for Future Life, Tokyo Denki University, Japan. He is also an Advisor on Cyber Security for Cabinet Secretariat of Japan Government.