# Filtering Avoidance Using Web Translation Service and its Countermeasures

Ryota Suzuki, Atsuo Inomata, and Ryoichi Sasaki
Tokyo Denki University
5 Senjuasahicho, Adachi-ku, Tokyo-to 120-8551, JAPAN
suzuki@isl.im.dendai.ac.jp

## Abstract

Recently, damage by targeted attacks has been increasing and has also become diversified. A targeted attack is any malicious attack targeted toward a specific individual or organization. It has the characteristic that damage is likely to expand because it is hardly noticeable. Therefore, the assumed countermeasures, such early detection and damage reduction, are important factors for the prevention of targeted attacks. In this paper, we propose attack methods that are able to avoid filtering by using web translation services, and then we propose countermeasure methods. Also, to evaluate our proposals, we surveyed other attack methods, including those used in combination, such as shortened URL services and web archive services.

## Keywords

Targeted attack, Malware, C&C communications, Web translation service, URL Shortener

## 1 Introduction

Recently, damage by targeted attacks has been increasing and has also become diversified. A targeted attack is any malicious attack that is targeted toward a specific individual or organization. Therefore, the assumed countermeasures, such as early detection and damage reduction, are important factors for the prevention of targeted attacks. The larger the scale of an organization, the more difficult is the prevention of infection. That is, it becomes more difficult for the countermeasure in a large organization to prevent the targeted attack [1].

Figure 1 shows targeted attacks. An attacker sends malware to target (1). The malware infects the victim's computer, then the malware in the victim PC communicates with the command and control C&C server (2). And downloaded software to expand the function of the malware (3). The expanded malware spreads to other PCs or servers (4). In this case, the malware usually sends internal information to C&C server using HTTP protocol. (5).

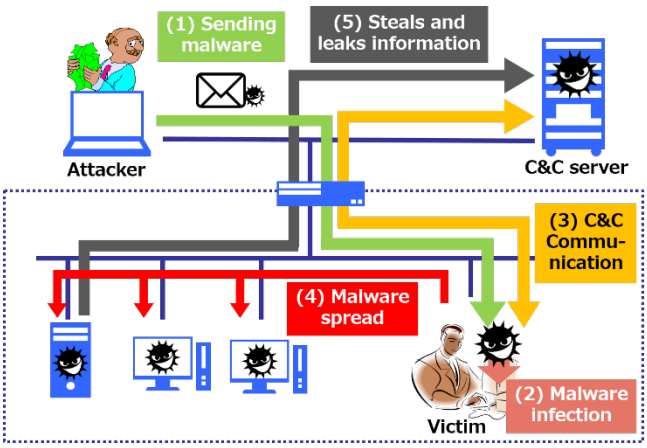


**Figure 1. Targeted attack**

Therefore, the assumed countermeasures, such as early detection and damage reduction, are important factors for the prevention of targeted attacks. [2]. In the present study, we propose attack methods that are successful because they avoid filtering by using web translation services. We assume an infection in order to consider the countermeasures.

In this paper, we first present attack methods that are able to avoid filtering by utilizing web translation services. Next, we show related work and our investigation of translation services.

ISBN: 978-1-941968-37-6 ©2016 SDIWC

Then, by conducting experiments, we show that it is possible to attack various translation services. We also evaluate the methods of attack. Finally, we conclude our paper.

In some web translation services, such as Google Translate and Excite Translator, anybody can easily translate the contents of the input web page into another language by entering the URL of the web page. Many web translation services provide translation functions of the target web page. By using a translation function, it is also possible for anybody to retrieve the contents of a target web page through a web translation service without the client establishing direct communication to the target web page. Figure 2 shows the filtering avoidance method that allows an attacker to impersonate a destination of communication from malware to the C&C server via the web translation service. Thus, filtering by such proxy servers in the organization allows communication with C&C servers that are prohibited and thus complicates the discovery of suspicious communications.

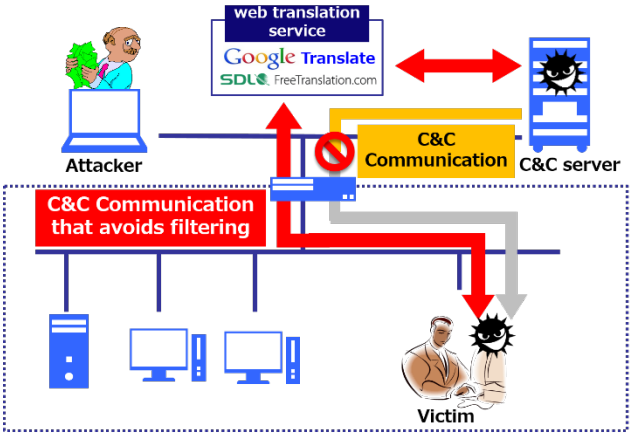


**Figure 2. Filtering avoidance method**

Some researchers have known about this feature of filtering avoidance, but they have not been able to verify whether it is actually possible. In order to evaluate the possibility, we designed an experimental scenario that shows filtering avoidance is indeed possible. In addition, we conducted an experiment with shortened URL services and other services such as the web archive service attack method, and again verified the possibility of a successful attack. Consequently, we discuss the techniques related to the attack methods. By understanding the techniques behind the attack methods, we can consider countermeasures preventing the attacks.

## 2 Related Work

One method of using web translation services for an attack is Jikto [3]. Jikto uses Google Translate for port scan in JavaScript. BarracudaLab [4] proposed a method in which the sender of spam e-mail uses a counter for spam filtering that combines Google translation and the URL shortening service. In this approach, URL filtering in an e-mail utilizes the fact that the URL filtering is verified by the evaluation of linked domains. First, the URL is shortened by a URL shortener to redirect a page to spam. The shortened URL is displayed in the web page translation function for Google Translate and then the URL of the generated page is sent by e-mail. Thus, the recipient is able to confirm the link of the Google domain and then filtering becomes difficult. All of these also relate to attacks using Google Translate, but they are not countermeasures to avoid the filtering of C&C communications.

## 3 Aims of this paper

We investigate the translation services that execute the translation for a web page and display external web pages that would be available for a filtering avoidance attack. Table 1 shows viewable external web page services.

**Table 1. Viewable web translation services**

| Service |
|---|
| Google Translate [5] |
| Bing Translator [6] |
| Excite Translator [7] |
| Yahoo! translation [8] |
| Infoseek translation [9] |
| So-net translation [10] |
| WorldLingo [11] |
| SDL (freetranslation.com) [12] |

ISBN: 978-1-941968-37-6 ©2016 SDIWC

From the results of the investigation, we found that most web services can translate and display an external web page with an original translation function. Furthermore, in addition to the web translation services in Table 1, we found that some web services can save and display a web page from an arbitrary time such as Internet Archive Wayback Machine [13] and Web Fish Print [14], so we think that these services should also be considered available to an attacker.

## 4 Experiments

In this paper, we set functions for determining whether the proposed method of filtering avoidance is applicable for targeted attacks. Specifically, in a state where communication to the target server by filtering is prohibited, we confirmed that the attack method is feasible by the following functions, which we will examine in Sections 4.1, 4.2, and 4.3:

(1) Command reception method
(From C&C server to malware)
(2) Files reception method
(From malware to C&C server)
(3) Response transmission method
(From malware to C&C server)

Our experimental environment is as follows:

- OS: Windows 10 Education
- Proxy server: Squid 3.5
- Web browser: Mozilla Fire Fox
- Experimental program: Java

To prevent the display of a page by the cache, we conducted an experiment with no-cache on a proxy server and web browser. Table 2 shows the filtering types in this experimentation.

**Table 2. Filtering types**

| | |
|---|---|
| IP address filtering | Block a communication to a host of a specific IP address |
| Domain filtering | Block a communication to a host for a particular domain name |
| URL filtering | Block a communication when a detected target string is in a URL. (String of the detected target specifies the IP address and the domain name of the target host.) |

### 4.1 Command reception method

We investigated whether the transmission of commands from the C&C server to malware is possible in the filtering environment. For the transmission method of commands, the attacker writes to a page on the web server and then the malware receives it by connecting via the translation service. In this experiment, we established a normal connection via the translation service and compared the results. We confirmed that an attacker can retrieve information from a server by avoiding the filtering via the translation service.

**Table 3. Filtering avoidance results by service for the command reception**

| Communication method | IP address | Domain | URL |
|---|---|---|---|
| Normal | × | × | × |
| Google Translate | ○ | ○ | ○ |
| Excite Translator | ○ | ○ | × |
| Yahoo! translation | ○ | ○ | × |
| Infoseek translation | ○ | ○ | × |
| So-net translation | ○ | ○ | × |
| WorldLingo | ○ | ○ | × |
| SDL | ○ | ○ | ○ |
| Internet Archive | ○ | ○ | ○ |
| Web Fish Print | ○ | ○ | × |

○: Acquisition successful ×: Blocked

The results of the experiment Table 3 showed that the communication with all filtering formats for normal connection detection are blocked. Thus, this filtering is considered to function properly. In contrast, in any web translation service, successful transmission of information from the C&C server was possible by filtering avoidance by the IP address and the domain. However, on many web translation services, when URL filtering was performed, the filtering functions by the proxy server prevented the transmission of information from the C&C servers. In the transmissions of storing the URL of the web page to be translated in the GET

parameter, the URL on the web translation service contained the C&C server domain name. The C&C domain name is underlined in Table 4. For other translation services, the URLs of web pages to be translated are sent by the GET method, so we think that URL filtering by using GET is considered to be an effective filtering avoidance method using translation services.

**Table 4. URLs when translated**

| Service | URL |
|---|---|
| Excite Translator | http://www.excite-webtl.jp/world/english/web/?wb_url=http%3A%2F%2Fweb.dendai.ac.jp%2F&wb_lp=JAEN |
| Google Translate | https://translate.google.co.jp/translate?hl=ja&sl=auto&tl=en&u=http%3A%2F%2Fweb.dendai.ac.jp%2F |

However, in the case of communication through Google Translate, SDL, and the Internet Archive, even if URL filtering is performed, we found the successful transmission of information from the C&C server by filtering avoidance. As shown in Table 4, in the case of Google Translate and Excite Translator, the translated target URL was sent by the GET method, and so the C&C server's URL in the translated URL is included. However, when the Google Translate and SDL communicate with the client, the GET parameters are encrypted because of the cryptographic secure communication by https [15]. Therefore, we think that the filtering failed because it was not able to confirm the content of the GET parameter from the proxy server.

### 4.2 Files reception method

Then we investigated whether the transmission of files from the C&C server is possible in our environment. We consider two file retrieval methods via the web translation services by malware:

A) In the web translation services, specify the URL of the file directly for downloading a file.
B) Get a page that contains a link to a file and download the file from the link on the page.

For approaches (A) and (B), we confirmed that it is possible to download a file from the server on a prohibited connection by using the translation service. As a result, saving the file failed for all translation services for all types of filtering. The reasons for approaches (A) and (B) failing are as follows. For (A), the reason why the approach failed is the return error for specifying the URL of the file type, e.g., not html or php; the file type is impossible to translate, and so the file cannot be downloaded directly. For (B), the approach failed because, for the web translation service, a targeted text page of html or php is the translation transmitted, but a file such as jpg or exe is not translated directly to the file existing on the translated original server. From these results for (A) and (B), we found that the transmittable file via a web translation service was only the targeted translation text data, as shown in Figure 3.

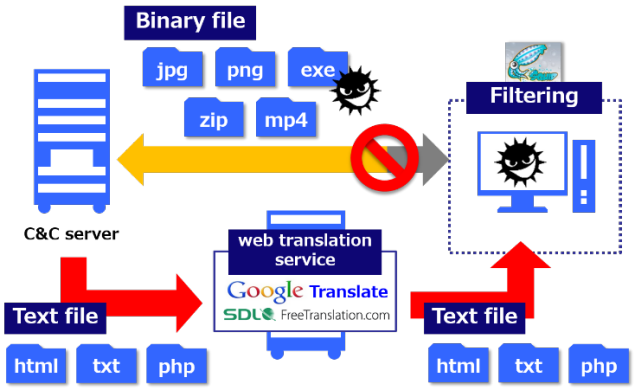


**Figure 3. Transmitted file formats**

Thus, we consider that sending a file via a web translation service as text data embedded into a file such as an html file is possible for filtering avoidance. As the method of embedding other files into html, a technique exists to encode the image by Base64 and embed the file as a tag into an html file. We confirmed that the Base64 encoded image file can be downloaded by filtering avoidance via the web translation service. In this experiment, while the filtering is performed, we executed the download via the web translation service on the browser and in the program for saving. The results are shown in Table 5 for both the display of the image file on

the browser and the download by the program.

**Table 5. Filtering avoidance results by service for the file reception**

| Service | IP address | Domain | URL |
|---|---|---|---|
| Google | ○ | ○ | ○ |
| SDL | ○ | ○ | ○ |
| Other service | ○ | ○ | × |
| Internet Archive | ○ | ○ | ○ |
| Web Fish Print | ○ | ○ | × |

○: Acquisition successful ×: Blocked

From Section 4.1, a service not using encrypted communication is clearly blocked by URL filtering, but we found that the download avoided the filtering via the web translation service to embed the file into html. By using the same technique to embed malicious software into an html file and store it by malware, we think that it is capable of transmitting an attack file from the C&C server. In addition, Internet Archive carries out the download of files, even if the file is archived; that is, Internet Archive succeeded in successful acquisition in either a direct specification of the URL or a link. However, the file failed to save as in case of a web translation service, because the Web Fish Print service does not archive a binary file. A file format type such as jpg or exe can be saved to the archive on the server (without a file type such as html or php) in Internet Archive. Therefore, when malicious software for attacking is used, it is also possible to realize file transmission in an environment in which avoidance filtering is performed when web archive services are used.

### 4.3 Response transmission method

Finally, we investigated whether the response method from malware is possible in our environment. In targeted attacks, malware must not only receive commands from the C&C server, but also response to the C&C server. We confirmed that it is possible to reply via the web translation services. On the other hand, we determined that the web archive services are excluded, because archive service does not access to original pages each time. In this experiment, we set up a page that can use the GET and POST methods on the web server, and then we focused on the filtering domain of the web server. With respect to the subject of the page, we sent a write request via the web translation services from our implemented program and investigated whether data were written. From Section 4.1, the filtering format was domain URL filtering in a translation service with or without using encrypted communication. In the results of the experiment, for all of the web translation services, the writing using the GET method was successful in avoidance filtering. Otherwise, the writing using the POST method failed.

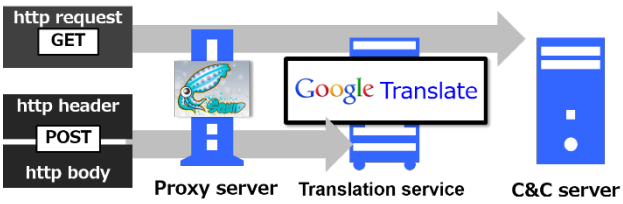


**Figure 4. Methods available for Google Translate**

As shown in Figure 4, in the GET method, the data are included in the http request part of the URL. This part reaches the translated original server. On the other hand, in the POST method, the data are included in the http header and body part and do not exist in the URL. We think that the information is lost at the time it is sent to the translation service. From this result, for the response from the malware to the C&C server, we needed to send the data by using the GET method. However, the GET method has disadvantages: the amount of data that can be transmitted is limited, and the content of the response is recorded on the proxy server as a log file of the URL.

### 4.4 Result Summary

From the results of the experiment in Sections 4.1-4.3, we found that receiving a command, a file, and a response in the C&C server were possible via the web translation service. Therefore, we think that the filtering avoidance method using web translation services can be applied for the communication of a targeted

attack. Furthermore, for all of the web translation services, it was possible to detect and block the communication of a targeted attack by sending the URL of the targeted web page with the GET parameter, because URL filtering cannot be avoided. However, web services on encrypted communication such as Google Translate, SDL, and Internet Archive encrypt the content to send, so a countermeasure is needed for decryption of the encrypted data before filtering is required, in addition to the URL filtering.

## 5 Experiments using a URL shortener service

### 5.1 Purpose

As described in the relevant research [4], it is effective to use a URL shortener for the avoidance filtering with web translation services. The URL shortener generates a short URL specifying a targeted URL. If access to a shortened URL is given, it is moved to the web page of the unshortened full URL by the redirect. By using the URL shortener, the URL of the C&C server is shortened, and access to the shortened URL from the web translation service is done with filtering avoidance. We conducted an experiment to investigate whether avoidance filtering together with a URL shortener can be applied for a targeted attack. In this experiment, we used the bit.ly [16] as the URL shortener.

### 5.2 Command transmitting method

We confirmed that the transmission of a command is possible in an environment of filtering combined with a URL shortener. In this experiment, we created a shortened URL for the filtering target and then accessed the shortened URL via a web translation service from a web browser. We found that filtering avoidance was possible.

**Table 6. Filtering avoidance results by translation service with URL shorter**

| service | IP address | Domain | URL |
|---|---|---|---|
| Normal | × | × | × |
| Google Translate | ○ | ○ | ○ |
| Excite Translator | ○ | ○ | ○ |
| Yahoo! translation | ○ | ○ | ○ |
| Infoseek translation | ○ | ○ | ○ |
| So-net translation | ○ | ○ | ○ |
| WorldLingo | ○ | ○ | × |
| SDL | ○ | ○ | ○ |

○: Acquisition successful ×: Blocked

We found that using a URL shortener with the avoidance of URL filtering without an encryption service is possible as compared with the result without the URL shortener. We think that the URL shortener hides the targeted domain and avoids detection when sending the URL to a web translation service in the form of a shortened URL. However, in the case of using WorldLingo, the filter avoidance failed. From these results, we think that even when using a web translation service, it is possible to avoid URL filtering with the combination of a URL shortener and a web translation service.

### 5.3 File transmission method

For the combinations of a URL shortener and a web translation service, we confirmed that it is possible to send a file. As in the Section 5.2 experiment, we executed the following three methods: specifying the URL of a file directly, downloading a file from a link, embedding a file into html, and sending. As in the case of the experiment in the web translation service only, the transmission of the file was successful only for the method of embedding the file. Furthermore, as in Section 5.2, avoidance filtering using the web translation service without encryption was successful.

### 5.4 Response method

With the combinations of the URL shortener and a web translation service, we confirmed that it was possible to respond to C&C servers in an environment of filtering. From the results in Section 5.3, we think that the URL with the GET parameter via the web translation service is needed for the response to the C&C server. Therefore, for transmitting a response by using a URL shortener, a URL is sent to the URL shortener for every response and then to the web translation service after executing the shortening.

From the experiment, we found that this approach was successful for responding to C&C servers for all translation services except WorldLingo. We think that a reason for the failure in WorldLingo is it is possible to filter with a URL Shortener, as shown by the result of Section 5.2.

### 5.5 Result Summary

From the results of the experiments in Sections 5.2-5.4, we found that the techniques combining a URL shortener and a web translation service were possible for transmission of a command, transmission of a file, and replying to the C&C server. Therefore, we think that the method with a URL shortener can be applied to the communication of a targeted attack. When using the filtering avoidance method in combination with a URL shortener, it was possible to avoid the URL filtering that was an effective countermeasure when using only a web translation service, and then to communicate about the targeted attack. However, in the combination with the URL shortener method, we found that it has a load disadvantage from the URL shortening for every response to the C&C server.

## 6. Countermeasures

From the Section 4 results, we found that URL filtering was effective for the avoidance method using a web translation service without encrypted communication. Some kinds of translation services with encrypted communication, such as Google Translate and SDL, need URL filtering with an encryption countermeasure for decoding on a proxy server. As a countermeasure for encrypted communication, we need filtering software [17] [18] calling the function of https decode [19] that stores a proxy server certificate at a terminal and then decodes the encrypted communication by using this certificate. For web archive services, it is possible to combine the countermeasure for encrypted communication and URL filtering. For the filtering avoidance method with a URL shortener, we found that it cannot use URL filtering, as shown by the result in Section 5. However, a countermeasure with content filtering is possible against the use of a URL shortener. This is because malware cannot generate a shortened URL unless it sends the original URL to the URL shortener. In the case of executing the content filtering, the countermeasure of encrypted communication is needed because there are many services for encrypted communication using a URL shortener. Furthermore, we think that a countermeasure could prohibit communication of the translation service or URL shortener from the perspective of operation and management. These countermeasures have the advantage that they can be realized only by filtering settings but the disadvantage of inconvenience for a user. These measures are shown in Table 7.

**Table 7. Countermeasure ranges and approaches**

| Countermeasure | Translate | https | URL Shortener |
|---|---|---|---|
| 1. URL filtering | ○ | × | × |
| 2. https decode | × | ○ | × |
| 3. Content filtering | × | × | ○ |
| 4. Ban URL shorteners | × | × | ○ |
| 5. Ban https use in translation services | × | ○ | × |
| 6. Ban translation services | ○ | ○ | ○ |

Countermeasures 1, 2, and 3 in the table are technical countermeasures that can solve the problems. However, they need more sophisticated and high-function filtering software, so we have to consider the increased costs and the burden of processing. Countermeasures 4, 5, and 6 in the table are countermeasures from the perspective of operation and management, and they increase the burden of inconvenience. To combine these countermeasures together with an environment of organization, it is possible to have countermeasures against the attack method of avoidance filtering with a web translation service.

## 7. Conclusion

In this paper, we confirmed that it is possible to perform an attack and avoid filtering with web translation services when an attacker communicates with the C&C server for the targeted attack. In addition, we showed that using a method in combination with a URL shortener is another approach for a targeted attack. As a countermeasure, we found that it is effective for a web translation service without encrypted communication to apply URL filtering. In addition, a countermeasure of filtering with encrypted communication of a web translation service is effective. For a countermeasure technique for a combination of a URL shortener and a web translation service, we also confirmed that content filtering is effective. Furthermore, we showed a countermeasure to prohibit communication to an available service for filtering avoidance as the countermeasure of operation and management. In future work, we will address the cost of the countermeasures and their performance and user convenience, and we will discuss the best solution for every organization.

## Reference


[1] Ryoichi Sasaki, Tetsutaro Uehara, & Takashi Matsumoto. (2013). Present Status and Future Direction Network Forensics against Targeted Attack. Computer Security Symposium, 2013(4), 155-162.

[2] Masahiro Yamada, Masanobu Morinaga, Yuki Unno, Satoru Toru, & Masahiko Takenaka. (2013). A Detection Method against Activities of Targeted Attack on The Internal Network. IPSJ SIG Technical Report, 1-6.

[3] Hoffman, B. (2008). JavaScript Malware for a Gray Goo Tomorrow.

[4] “Spammers disguise links using Google translate “, https://barracudalabs.com/2013/03/spammers-disguise-links-using-google-translate/, (access 2015-11).

[5] “Google Translate “, https://translate.google.co.jp/, (access 2016-04).

[6] “Bing Translator“, https://www.bing.com/translator(access 2016-04).

[7] “Excite Translator “, http://www.excite.co.jp/world/, (access 2016-04).

[8] “Yahoo! Translate “, http://honyaku.yahoo.co.jp/, (access 2015-11).

[9] “Infoseek multi Translate“, http://translation.infoseek.ne.jp/web.html, (access 2015-11).

[10] “Translate | So-net“, http://www.so-net.ne.jp/translation/, (access 2015-11).

[11] “WorldLingo“, http://www.worldlingo.com/ja/products_services/worldlingo_translator.html, (access 2015-11).

[12] “SDL FreeTranslation“, https://www.freetranslation.com/ja/, (access 2015-11).

[13] “Internet Archive: Digital Library of Free Books, Movies, Music & Wayback Machine“, https://archive.org/index.php, (access 2015-11).

[14] “Web Fish Print “, http://megalodon.jp/, (access 2015-11).

[15] D. G. J. D. Shorter, “Effectiveness of Internet Content Filtering.,” Journal of Information Technology Impact.

[16] "Bitly | URL Shortener and Link Management Platform," [Online]. Available: https://bitly.com/. [Accessed 4 2016].

[17] “i-FILTER SSL Adapter」｜i-FILTER“, http://www.daj.jp/bs/i-filter/old/option_relation_ssl_adapter, (access 2015-11).

[18] “Counter SSL Proxy “, http://www.swatbrains.co.jp/csp.html, (access 2015-11).

[19] E. Akbaş, “Next generation filtering: Offline filtering enhanced proxy architecture for web content filtering.,” ICIS 2016, 2008.