REGULATION OF ACCESS TO WEB-RESOURCE BASED ON POST-ANALYSIS OF HTTP-REQUESTS

K.I. Budnikov, A.V. Kurochkin, A.A. Lubkov, A.V. Yakovlev

Institute of Automation and Electrometry of Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

Abstract. A method for regulation of access to web-resources by filtering HTTP requests at the packet level, using post-analysis of the HTTP requests that have passed through the filtering device is proposed. In this method, the request is analyzed not before it is sent to the Internet, but after, at the time when it comes by the Internet communication lines to the web-server on which the requested resource is placed, web-server forms the response, and this response comes back to the filter. In the results of checking, the response, obtained from the web-server, is passed by filter device to the user or locked. Such approach can reduce the time of waiting for the request to the resource in comparison to methods using preliminary analysis of the request before sending it to web-server.

Keywords: filtering of HTTP-traffic, network packets analyzing, regulation of access to web-resource

Citation: Budnikov KI, Kurochkin AV, Lubkov AA, Yakovlev AV. Regulation of access to Web-resource based on the HTTP-requests post-analysis. CEUR Workshop Proceedings, 2016; 1638: 284-289. DOI: 10.18287/1613-0073-2016-1638-284-289

Introduction

In recent years, development of the Internet is accompanied by appearance of a large number of information resources, access to which requires a limitation by different criteria: age, moral and ethics, compliance with requirements of security, copyright, labor regime, etc. This task can be solved using different methods such as restricting access by IP-address, URL address, by changing requests to DNS-servers, using a proxy server or packet filtration. These approaches have their advantages and disadvantages [1]. The most balanced ratio of advantages and disadvantages can be found in a method of request to the resource filtering by its URL address. This approach allows filtration of a specified resource. The method consists of intercepting user's request by filter, extraction of requested resource address, searching it in lists of banned addresses, and formation of actions corresponding to obtained result. If the address is not prohibited, the request is passed to the Internet, comes to the server with the necessary resource that returns the response with the requested information. If access to the resource of interest is prohibited, the request is blocked by filter.

Filtering the URL can be performed for a single Internet access device (PC, smart phone, tablet), or a group of devices. In the first case, the filtering process is usually performed by a specially installed program. and in the second case by a filtering device having access to the Internet from one side and computers of users connected to it from another side (Fig. 1). The first approach is proposed, for example, in patents [2, 3]. Patents [4, 5] are examples of the second approach.

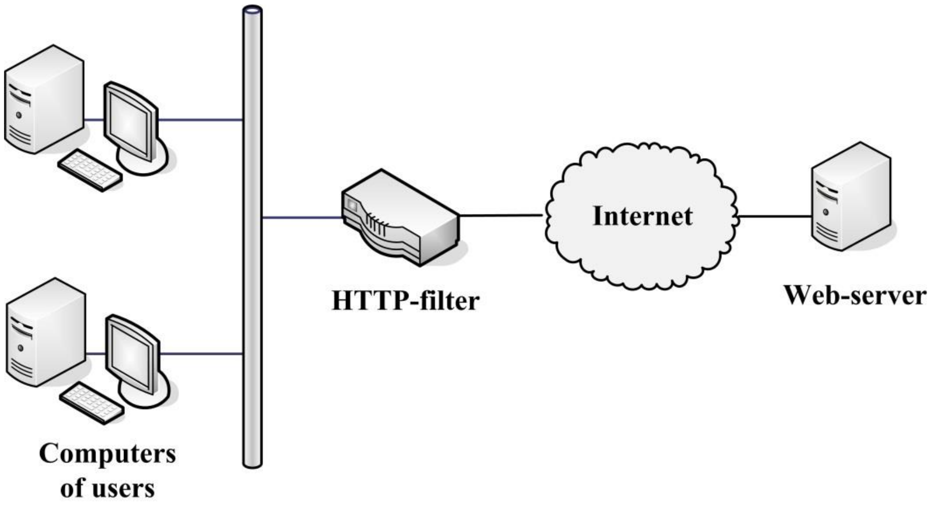

Fig.1. Diagram of the filtering device connection

With the relative ease of implementation of the first approach (it can be implemented by software means), it has such a significant disadvantage, as the potential capability for the user to disable the filtering software at his own discretion and thus to circumvent the filtering process. Working capabilities of devices that implement approach of the second type cannot be regulated by users connected to them. Information necessary for the decision making concerning addresses for filtering such devices receive from external servers called filtering or reputational servers via network connection. For users filters are transparent and represent only a delay line on the way of user's request to a resource of interest. The faster request will pass through the filter device, the less noticeable its presence to user and more requests are able to pass through the filter.

The filtering algorithm in such devices (see e.g. [4,5]) involves preliminary checking of request in the input of the device, and only according to the checking result the request will pass the device or will be locked. The inspection holds the time tied with the request interception, URL extraction from request and searching it in lists of banned addresses. For this time the request is delayed by the filter. On the duration of the verification procedure the duration of the request to the filtering server also significantly influences. The need for request to the filtering server arises in the case of lack of information on the URL in the local lists of banned addresses. Delay of request by filter device increases the response time. Reduction of a delay for user request passing is possible due to correction of processing algorithm for packets, the use of the method of post-analysis of requests to web-resource instead of preliminary analysis. This approach enables to provide acceptable response time for a larger number of Internet users whose requests pass through the filter.

Method of post-analysis of requests to web-resource

Proposed method of filtering consists of the following [6]. A filter that uses post-analysis of HTTP-requests always moves all packets arriving to the input of the device, including those ones containing user's HTTP-request, to the output of the device without delay and change. For the analysis of the request the device creates a copy of passed HTTP-packets. HTTP-request is verified not before it is sent to the Internet, as it occurs in the filtering devices using HTTP-request pre-analysis, but after it is send to the Internet, at the time while the request by communication lines comes to web-server where requested resource is located, the server forms the response, which comes back to the filter, i.e. in the mode of post-analysis of the request. According to verification results the response received from the web-server is either passed to the user or locked.

Device model and algorithm of its work

Introduced filtering method can be illustrated by the example of working of simplified model of packet filter shown in Fig. 2. Filtering device is installed in the gap between LAN with computers of users and the Internet with web-servers that provide resources via HTTP protocol connection, as shown in Fig. 1.

Filter model consists of a user network interface (UNI), Internet network interface (INI), two selectors (S1 and S2), the analyzer-corrector (AC), the storage of the current state of controlled TCP-sessions (SCS) and storage of banned resource identifiers (BRI).

Selectors extract packets of HTTP protocol from common traffic. S1 transmits all traffic coming from the UNI to the INI immediately and without any changes, and copies of HTTP-packets it sends to AC. S2 transmits all incoming from INI traffic to UNI, except of HTTP-packets that are sent to AC for verification. AC, using information from BRI, checks request for access right to requested resource and, if necessary, locks the response from the web-server to user with forbidden content.

AC contains a unit of formation and analysis of sessions (SFA). SFA forms TCP-sessions from network packets of HTTP protocol. During these sessions users send requests to certain web-resources. SFA stores information about these sessions in SCS and on demand provides status of web-request, whether it is banned or not.

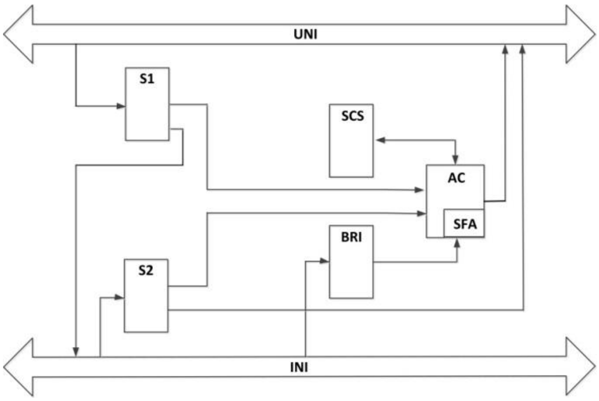

Fig.2. Filtering device model implementing method of post-analysis of requests to web-resource

The model works as follows. Packet stream which contains a user request on the way to the web-server reaching filtering device enters the user network interface UNI, and then to the first selector. S1 sends packets that contain request to the Internet network interface INI (i.e., to the web-server), and copies of packets - to the analyzer-corrector, which forms TCP-session, extracts URL from the request and verifies access rights to this web-resource. Thus, checking of a user request copy takes place at the same time with the transmission of the original user request from the filtration device to the web-server and a response from the web-server for user to the filtration device.

The response from web-server, reaching filtering device enters user network interface and onward the second selector. S2 separates packets of TCP-sessions of HTTP protocol and sends them to analyzer-corrector for subsequent processing. Unit of formation and analysis of TCP-sessions determines for incoming packet corresponding TCP-session in the list of monitored TCP-sessions that is found in the storage of the current state of controlled TCP-sessions and checks whether the current request is allowed for this TCP-session or not. If the request is granted, the packet is sent to the user network interface without modification. Otherwise actions associated with a particular algorithm of response locking (package deletion, modification of data, posting warnings about blocking, etc.) are performed.

The gain in transit time of the user request passing through a filtration device using post-analysis in comparison with a preliminary analysis consists of the time taken to determine TCP-session for each packet, the formation of the user request from packets, extracting the URL of the requested resource, and verification of the request for access right to the requested resource in internal lists of banned URLs.

Computer modeling of filtering device

For experimental evaluation of time gain for user query passing through filtration device using the method of post-analysis, compared with the preliminary analysis, computer simulations of filtration devices were conducted. In order to eliminate influence of the network infrastructure on the operation of the model, user and Internet network interfaces were simulated by software, and all network data streams flowed inside memory of modeling computer.

The time gain in the simulation depends on several factors, including the power of computer used for this purpose, software implementation of a filter model, the degree of the model's workload, the composition and intensity of traffic passing through the filtering device model, etc. During modeling a process of continuous sending of requests for web-resource through the filter by group of users and answers by web-server passing through the device was simulated.

Computer simulation has shown a decrease to 14% of the average time of passing for user request to the web-resource through the emulated filtering device that works in the post-analysis of requests mode in comparison with a device that works in preliminary analysis of requests mode.

Conclusion

To filter requests to a web-resource, a method based on the use of post-analysis of requests to web-resource was proposed. Unlike traditional methods, the request analysis in filtering device occurs not before the request is sent to the Internet, but after that, at a time when request by communication lines comes to web-server on which the requested resource is hosted, web-server forms the answer and the response comes back to the filter. According to the check results the response, received from the web-server, is either passed to the user or locked. Such approach can reduce the response time for a request to a resource in comparison with the approach using a preliminary analysis of the request before it is sent to the web-server.

Applied method considered more broadly as a method of monitored data post-analysis in real time may be used in other areas and systems [7-10].

References

1. Apetyan S, Kovalev A, Veybach A. Filtering content in Internet. Analysis of international practice. Foundation of Civil Society Development, 22 May, 2013. URL: http://civilfund.ru/Filtraciya_Kontenta_V_Internete_ Analiz_Mirovoy_Praktiki.pdf (reference date: 02/06/2016) [in Russian].
2. Osipov GS, Tihomirov IA, Sochenkov IV. Method and system for web content filtration. Patent RU 2446460 C1. IPC G06F 21/20 (2006.01), published 27.03.2012. Bull. number 9 [in Russian].
3. Bellinson C, Evans Ch, Fravert H, Taylor W. Content filtering for web browsing. Patent US20040006621 A1, IPC G06F17/30, G06F13/00, G06F15/00, G05B1/00, G06F17/00, published 08.01.2004.
4. Bloch E, Mohan Sh, Pagaku RR, et al. Apparatus for monitoring network traffic. Patent US 7849502 B1, Int Cl G06F 15/16 (2006.01), G06F 11/00 (2006.01), Pub. Date: Dec. 7, 2010.
5. Balasubrahmaniyan J, Daftary K, Venkateswara Rao Yarlagadda, Kumar K. System and method for URL filtering in a firewall. Patent US 20060064469A1, Int. Cl.G06F 15/16 (2006.01), Pub. Date: Mar. 23, 2006.
6. Budnikov KI, Kuruchkin AV. Method of HTTP-packet flow filtration based on post-analysis of requests to Internet resource and filtering device for its realization. Patent Ap-

plication RU 2015114186/08 (022237) IPC G06F15/00, G06F15/16, G06F 21/00, G06F 21/50. Priority: date of filing 16/04/2015.

7. Kazanskiy NL, Popov SB. The distributed vision system of the registration of the railway train. Computer Optics, 2012; 36(3): 419-428.
8. Yakimov PYu. Preprocessing of digital images in systems of location and recognition of road signs. Computer Optics, 2013; 37(3): 401-405.
9. Epifantsev BN, Pyatkov AA, Kopeykin SA. Multi-sensor systems for monitoring access to restricted areas: capabilities of the intrusion detection video analytical channel. Computer Optics, 2016; 40(1): 121-129. DOI: 10.18287/2412-6179-2016-40-1-121-129.
10. Fedoseev VA. A unified model for information hiding systems. Computer Optics, 2016; 40(1): 87-98. DOI: 10.18287/2412-6179-2016-40-1-87-98.