A Content-blocking Framework for Harmful Media Regulation

*Sanghoon Lee, *Namkyung Lee
*Electronics and Telecommunications Research Institute
*sanghoon@etri.re.kr, *nklee@etri.re.kr

ABSTRACT

We explored the blocking framework for regulating harmful content flooding the Internet. The procedure for obtaining content using the Internet was analyzed, and the technology for blocking content that can be applied at each stage of the acquisition process was investigated. Also, the characteristics and limitations of each blocking method were analyzed.

1. Introduction

As high-speed Internet and mobile communication become popular, the propagation speed of content is increasing rapidly. As a result, media can be easily consumed regardless of place and time, but harmful information is distributed indiscriminately. In conclusion, it is necessary to regulate the injurious media to create overall social soundness. In this paper, we explored the content blocking framework for regulating harmful content. Various content-blocking technologies that can be applied to the entire cycle of Internet users seeking and obtaining information were analyzed.

2. Background

Harmful media circulating through the Internet is rapidly spreading in a short period, resulting in serious social problems. For example, in the case of illegal filming, one of the harmful media, it seriously violates the human rights of the victim, who is subject to the shooting, while at the same time seriously undermining national sentiment. The number of digital sex crimes related to this is increasing steeply every year, and social costs to deal with are also increasing gradually.

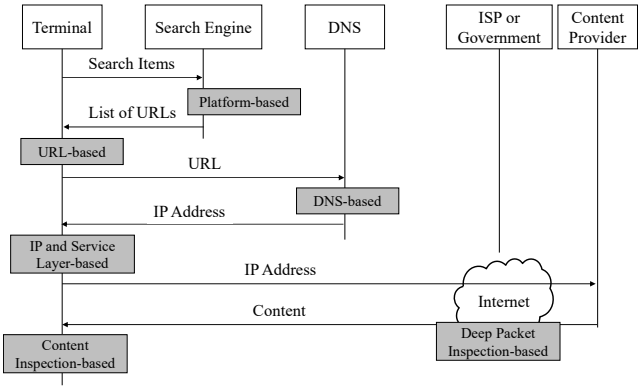

Figure 1 Typical content acquisition procedure of Internet users and content blocking methods.

One of the difficulties of blocking harmful media through the Internet is that content providers can use other Internet service providers (ISPs) or be in other countries. Since the definition of harmful content may vary from country to country or operators, there is a problem that content providers cannot be unconditionally illegal. In addition, due to the nature of the Internet, content delivery involves multiple entities. For example, there are entities such as search engines used by users, Domain Name Server (DNS) servers that convert URLs to IP address, and content hosting servers, each of which may exist in different countries. Each identity is used step by step, and as a result, it can be seen that blocking techniques are required that can be applied at each step.

3. Harmful Media Blocking Framework

The typical content acquisition procedure of Internet users is summarized in Figure 1. The order is as follows: A user 1) requests search items to a search engine using a web browser on a terminal, 2) receives a list of URLs from the search engine, 3) selects one URL and requests IP address to DNS, 4) obtains an IP address from DNS, 5) requests content to a content provider using IP Address, and 6) receives text, image, and audio content from the content provider. The blocking techniques that can be applied between each stage of the content acquisition procedure are as follows.

A. Platform-based Blocking

Platform-based blocking is a method that blocks content mostly used by service providers. Examples include content filtering through major search engines or social media platforms. The primary provider or country may manage a separate list of what to block and what not to block. Platform-based blocking methods are mainly operated based on blacklist or whitelist. The blacklist contains harmful sites that are blocked, and operators may not allow users to access a particular website if the address is on the list. In contrast, the whitelist contains secure sites, and users can visit only those included in the list. There can be various ways to build black or whitelist. [1] used the connection between websites, and [2] used the hazard prediction algorithm based on machine learning. On the other hand, platform-based blocking does not directly delete content, but there is a disadvantage in blocking the link to the content.

B. URL-based Blocking

URL-based blocking is one of the most popular methods, a blocking method for URL lists received from search engines. This approach is applicable to both individual terminal and network equipment. The requested target URL is compared to the blocking list URL contained in the local or remote database. This method is applicable even if the IP of the content provider changes, but if the URL changes, blocking may fail. URL-based blocking can also be implemented in several ways: [3] established a blocking list within the local proxy server inside the Android terminal to check URLs. [4] blocked URLs similar to URLs in the blocking list through URL pattern analysis.

C. DNS-based Blocking

The primary role of DNS is to receive URLs from users and provide IP addresses. DNS-based blocking is also operated by managing the blocking list. If the received URL exists in the block list within DNS, the user is notified that the content is blocked or the server does not exist. The disadvantage of this method is that, as with other network-based blocking, users can avoid blocking if they use another DNS server. To solve this problem, we need help from firewalls or other devices that can intercept all DNS queries from users. Another disadvantage of this approach is that it cannot respond adequately to changing the content provider's domain name. In addition, if harmful content exists only on a few servers within the same domain, there is a downside to all controlled servers.

D. IP and Service Layer-based Blocking

In IP and service layer-based blocking, traffic to listed IP addresses and port numbers in firewalls installed on terminals or Internet gateways is blocked. It also uses a method in which users do not indirectly use the content by limiting the speed of traffic instead of completely blocking it. However, it is challenging to maintain blocking performance if content providers frequently change IPs. It is also difficult to apply effectively if the content provider uses Content Delivery Networks (CDNs). Because CDN uses the same IP for different users and content, blocking can cause unintended service interruptions. It is also not useful when content is distributed across multiple datacenters. [5] sent IP spoofing packets and induced users to give up their own connections. [6] uses the access control list and sinkhole routing to block unspecific harmful traffic.

E. Deep Packet Inspection(DPI)-based Blocking

DPI is a string matching engine that can detect a particular string in a data stream at high speed, and the target includes not only the header of the packet but also the payload. Security purposes, bandwidth management, content regulation, etc. may be implemented under the supervision of the Internet service provider or the government. The German federal government used the DPI as a network blockade against child pornography [7], and several countries were also used to monitor all information that might pose a threat to the government [8], [9].

F. Content Inspection-based Blocking

In content inspection-based blocking, it analyzes content

directly or analyzes keywords, file names, usage patterns, or app types. When applied, it is common to install on a user's terminal because of its high computational complexity compared to other test methods. In [10], color distribution in video content was analyzed with SVM and used for detecting harmful images. In [11], [12], the intent monitoring and log system functions of Android OS were used to block harmful applications, respectively. However, the content-based approach has the disadvantage of being difficult to operate when traffic is encrypted. And because all traffic to the enduser is inspected, it is likely to violate the user's privacy.

4. Conclusion

Along with the rapid spread of multimedia content, harmful content is also flooding. In this paper, we explored the characteristics of content blocking methods that can be applied to Internet users' entire cycle of content acquisition. Each step-by-step blocking approach involves both advantages and disadvantages and will be able to achieve more effective media regulations than when one or more technologies are in harmony with each other.

Acknowledgment

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT). (No. 2019-0-00287, Development of the System to Analyze and Detect the Obscenity in Media Contents Utilizing Artificial Intelligence Technologies.)

References

[1] Junghoon Shin, "Methods for discriminating harmful web sites using link relations between web sites," Ph.D. dissertation, Soongsil Univ., Seoul, Aug. 2014.

[2] S. Boungjin Kim, "Improvement of methods for discriminating harmful web sites by using link relations between web sites and constructing white-list," pp. 506-510, Oct. 2019.

[3] M. Yang, "A study for blocking harmful contents via a local proxy on android," Ph.D. dissertation, Dongguk Univ., Seoul, Dec. 2013.

[4] Ji Sun Park, "Preventive measures against harmful sites based on url pattern analysis," Ph.D. dissertation, Konkuk Univ., Seoul, Nov. 2019.

[5] S. Paek, "A study on internet traffic control: Blocking of harmful information based on ip spoofing," Journal of the Korea Academia-Industrial cooperation Society, no. 5, pp. 447-453, 2004.

[6] Moon-Soo Chang, Jeong-Il Lee, and Chang-Suk Oh, "Harmful traffic control using sink hole routing," The Korean Society Of Computer And Information, no. 14, pp. 69-76, Apr. 2009.

[7] T. Kleinz, "Bundesregierung treibt Netzblockaden gegen Kinderpornografie voran," heise online, Jun. 2020.

[8] Christopher Rhoads and Loretta Chao, "Iran's Web Spying Aided By Western Technology," Wall Street Journal, Jun. 2009.

[9] B. Wagner, "Modifying the Data Stream: Deep Packet Inspection and Internet Censorship," Social Science Research Network, Rochester, NY, SSRN Scholarly Paper, Oct. 2008.

[10] Chang-Seok Lee, "Adult image detection based on the skin region distribution using svm," Ph.D. dissertation, Hanbat National Univ., Daejeon, Aug. 2011.

[11] SeHwan Yeo, "A method of blocking harmful application by monitoring intent in the android platform," Ph.D. dissertation, Hanyang Univ., Seoul, Feb. 2013.

[12] Jaeyeon Lee, "A study on harmful application blocking method on the android platform," Ph.D. dissertation, Hanyang Univ., Seoul, Aug. 2012.