The evaluation of AdBlock technique implementation for enterprise network environment

Mohd Iskandar Bin Samsuddin[1], Mohamad Yusof Darus[2], Shamsul J Elias[3],
Abidah Hj Mat Taib[4], Norkhushaini Awang[5], Roshidi Din[6]
[1,2,3,4,5]Faculty of Computer and Mathematical Sciences,University Technology of MARA, Malaysia
[6]School of Computing, College of Arts and Sciences, UUM, Malaysia

Article Info

Article history:

Received Oct 8, 2018
Revised Dec 6, 2018
Accepted Dec 18, 2018

Keywords:

DNS AdBlock
Enterprise network
Web browsing traffic

ABSTRACT

This paper presents the evaluation of AdBlock technique implementation for enterprise network environment. This study has presented the impact of web browsing activities where it is the most active traffic where is consumed the highest inbound bandwidth usage in enterprise network environment. We can conclude that DNS AdBlock is the best solution for enterprise network environment in term of blocking advertisement compare to extension adblock. Adblock technique also reduce network data request by comparing front-end solution (browser extension AdBlock) at client web browser and networks level adblock. The parameters such as HTTP request, TCP connection and network bandwidth are being examined to measure the effectiveness of blocking online advertisement. Both techniques perform the reduction of traffics and bandwidth utilization. The result shows that DNS AdBlock is the most effective in blocking online advertisement using the examined parameters. DNS AdBlock can sustain the usage of web browsing activity for enterprise network and also generate substantial saving across several fonts. This study has identified current web browsing trends traffic in enterprise network where it consumed 50 percent in average. This number increased when industries are moving to cloud web-based consumption. However, industries such as educational sector, web browsing traffic is one of connectivity that enterprises network should be investing to support openness and heavy traffic from educational users.

Corresponding Author:

Mohamad Yusof Darus,
Faculty of Computer and Mathematical Sciences,
University Technology of MARA,
Shah Alam, Malaysia.
Email: yusof@tmsk.uitm.edu.my

1. INTRODUCTION

Internet browsing is becoming essential part of everyday life often as it often used to gather information. At somehow, web contents delivered to the end user browser with online web advertisements (ads). Daily online activity of thousands of users in one network environment give an impact of online advertisement which can increase the traffic as well as increase bandwidth consumption. As World Wide Web (WWW) makes it more intelligent, the implantation of online web ads in website is one of the marketing strategies [1]. Web technologies are currently moving from Web 2.0 to Web 3.0, online ads contents are becoming one of the elements and it play a big role in web-eco-system [2]. A new semantic format of embedding online ads which would encourage large publishers to add them to their web sites [3]. Besides, online web ads became the basic revenue source for webpage publishers, where they just simply copy the code provided by the ad network provider then paste it into their HTML file. Moreover, with Web 4.0 and Web 5.0; online ads and e-commerce innovations are already under way [4].

Online ads will keep on growing as stated in [5] report shows that the 53 % revenue is being made from display embedded format. Display embedded formats for online ads which includes banner, audio, and video that displayed at webpages. A study had been done where they evaluate 500 URLs of popular websites and the results showed that 63 % of them displayed some form of ads [6]. From user perspective, they have bad browsing experience because of the extra code that being downloaded and displaying ads automatically [7]. Besides, it does not provide any meaningful control for user [8]. In addition, it can nonetheless be intrusive and represent a significant drain on network resources while web browsing [9]. When a webpage is loaded on web browser, it will download all the information of the website and also online ads. Furthermore, the process of fetching, rendering and displaying the web contents require a lot of traffic being generated which included extra code for ads [2]. When number of broadcasting increased, network congestion will exist due to excessive amount of broadcast messages occupied the bandwidth of communication channel thus the message could not be delivered properly due to packet loss during transmission [10]. Traffic congestion creates more negative impacts effecting our daily activities, safety and quality of life, thus attract many different scientific interests [11]. Assume that the daily online activity of thousands of users in one network environment; mean that the impact of online ads can be significantly increase the traffic as well as increase bandwidth consumption.

According to [12], the increased adoption of cloud services has greatly driven up the amount of Internet traffic. A study had been done at Simon Frasers University, as public sector such as universities are increasingly moving to online learning platform; they concern on the numbers of daily online users might impact draining on network resources where the present of web information is included with online ads contents [9]. In [13] show the statistic usage of daily bandwidth for educational sector is among the highest. With the growing numbers of online ads blooming for multiple purposes, AdBlock technology is being introduced with come in variety of install options [14]. There are several AdBlock technologies that have been developed which are AdBlock application for smartphone, browser extension AdBlock, proxies ad-blocker and DNS AdBlocker.

This study concerns on web browsing activities in enterprise network, it will be better to look over the current web browsing traffic consumption for enterprise network environment [7]. According to [14], they had comparing the webpage load time with and without AdBlock implementation; and they able to detect a few hundred milliseconds of delay in webpage load. The relationship between number of requests made and TCP connection may lead to more bandwidth consumption [2]. The increase in traffic and data exchange can create the complication in congestion control once the transmissions occurred between one user and multiple users [11]. Thus, as consequences it impacts to the increase of network traffic in web browsing if online ads traffic is not being controlled. This study also be the improvement of previous research conducted by [9] where they have evaluated browser extension Ad-block (AdBlock Plus) and asking to compare the result across another ad-block technique.

Previous AdBlock technique aims to block ads across the client devices where it applied as plug-in extensions in web browser such as Mozilla Firefox, Google Chrome etc. However, having such plug-ins working in the background requires extra activities and computational power [2]. According to [14], the default setting of AdBlock Plus; it has the future of allowing certain “acceptable ads” to be display on the web. Besides, having such as browser extension AdBlock might bring fewer difficulties form the perspectives of network administrator since the growing number of client own devices such as laptop, tablet, and smartphone are connected wirelessly access to the network [9]. A study conducted by [15] has evaluated multiple AdBlock tools effectiveness in blocking online ads during downloading popular websites however, their study scoped down to browser extension technique and it did not differentiate the traffic as will be conducted in this study. Besides, some research has done a study on the effectiveness of available Ad-Block; for example, they have conducted a test on most popular browser extension “AdBlock Plus” and well-known hardware solution “AdTrap” [14]. However, the effectiveness result does not distinguish between different ad-block techniques in term of type of traffic reduction. Although many researches may have an opinion on this, the scientific evidence to proof such opinion hardly exists.

This paper presents the Evaluation Of AdBlock Technique Implementation For Enterprise Network Environment. This study selected two types of adblock technique to be evaluate which are Browser Extension Ad-Block and DNS Ad-Block which is same method being done by [14] to compare and evaluate the effectiveness of adblock. An open-source application which are AdBlock Plus for Browser Extension and Pi-Hole for DNS are tested and analysed. In order to conduct comparison of Ad-block effectiveness, this study created three network scenarios of experiment where Scenario 1: is network without Ad-Block, Scenario 2 network with Browser Extension Ad-Block and Scenario 3: network DNS Ad-Block. According to [15] and [16] the best way to test the effectiveness of adblock is to test on most popular websites. In each test scenario; there are three areas of investigation that this study focusing which are “HTTP Request == GET”, “TCP Connection” and “Bandwidth Consumption”. Each area of investigation is analysed and result is compared with other scenario. In [2] has created a sequence of downloading webpage where it shows that "HTTP Request == GET" is the traffic being used to fetch online ads. In [2] also mentioned that TCP connections are needed for online advertising purpose. TCP is a connection-oriented transport layer protocol that being used to carry ongoing data exchange. As stated by [2]; the more TCP connection to established, the more bandwidth consumed. Therefore, by comparing the result, it will possible to calculate the reduction of network traffic and bandwidth consumption. As a result, this study able to propose a suitable AdBlock technique to be implement in enterprise network. Thus, this study attempts to investigate, analyse and evaluate the AdBlock techniques for enterprise network environment and documented the findings in reduction of network traffic and bandwidth consumptions with respect to web browsing activities.

2. RESEARCH METHODOLOGY

In this research, the two major types of variables are independent and dependent variables been identified. An independent variable is a variable that affect the dependent variable while dependent variable is the variable a researcher is interested in. The components of variables are shown in Figure 1.

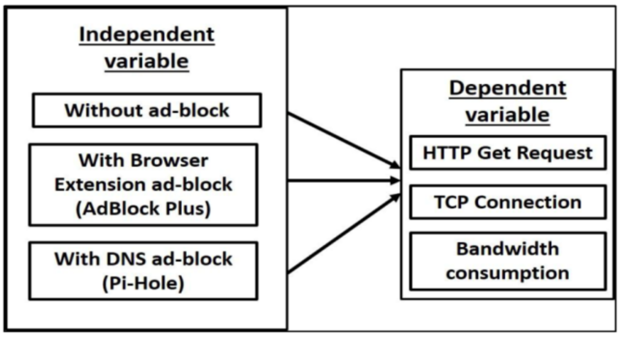

Figure 1. Experimental variables

In this research, there are three independent variables that been manipulated which are network without AdBlock (Scenario 1), network with browser extension AdBlock (Scenario 2), and network with DNS AdBlock (Scenario 3). From all independent component, there are three dependant areas that been investigate which are HTTP "get" request, TCP connection and bandwidth consumption.

2.1. Measurement Framework

The framework that been used for this study is similar with previous research conducted by [15] where they used it to investigate the amount of data generated by advertisements when browsing as shown in Figure 2.

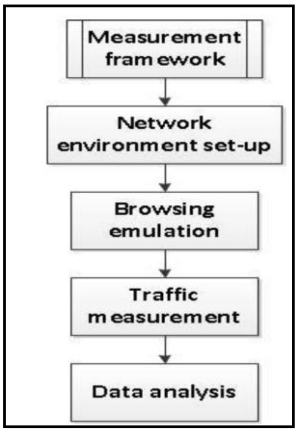

Figure 2. Experimental design approach

Activities conducted in this experiment when we used network packet data to investigate the amount of data generated by advertisements when browsing. This experiment needs to set-up a network environment where it will be tested with three scenarios as mention before. Web browsing emulation is conducted in each client PC for each scenario. Data collection of the web browsing traffic is measured using packet sniffing program “Wireshark”. Further discussion for each steps are explained in next sub-topics.

2.2. Network Environment Set-Up

Figure 3 shows the network environment set-up for this study. There are three types of scenarios where the network configured without AdBlock for Scenario 1, Scenario 2 with browser extension and with DNS AdBlock for Scenario 3.

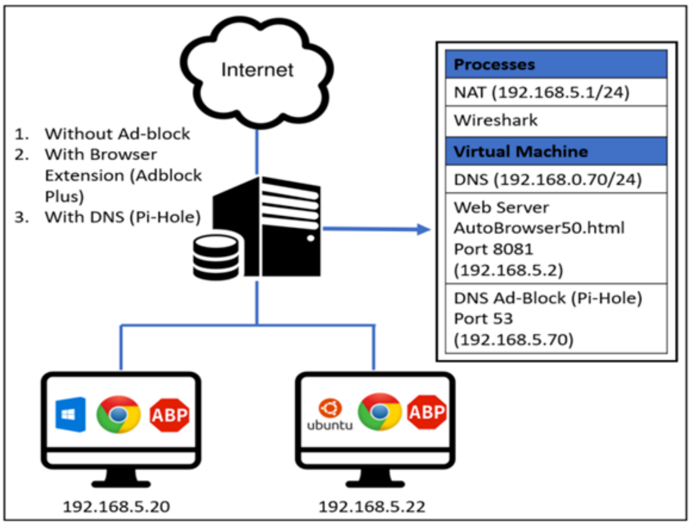

Figure 3. Testing architecture

To simulate the network environment, an isolate network it setup using network address translation (NAT) and this study set-up a network which client PCs connected to the switch and the all PCs will be configured as static IP. Virtual machine installs with Debian operating system to run AutoBrowse at Apache server port 8081. Network traffic is captured using Wireshark software.

In Scenario 1, normal DNS server is implemented as DNS forwarder to same upstream DNS (GoogleDNS) to give a fair and consistent measurement of test traffic in network without AdBlock. In Scenario 2, AdBlock Plus is implemented in chrome web browser and it point to the same DNS as in Scenario 1. While in Scenario 3, client DNS is pointed to Pi-hole as it future have the ability to perform DNS forwarder together with Ads Blocking.

2.3. Web Browsing Emulation

In this reserach, we adapted the mimic surfing using AutoBrowse program written by [17]. This program is in JavaScript and runs on the PCs web browser that retrieves a given set of URLs. There are fifty (50) lists of URLs that has been identified that needed to be loaded from Malaysia and other country website. In order to set up a consistent concurrent request from client web browser, this reserach is modified by adding a new function “gettime()” in AutoBrowse and run it in Apache server. An experimental test time is set at AutoBrowse source code by modifying the parameters shown in Table 1. Therefore, all PCs are requested for AutoBrowse and retrieve it with the current time in Apache server. All the URLs are loaded for each of the scenario.

Table 1. Gettime () Description

Parameters	Descriptions
Value1	Hours start
Value2	Minutes start
Value3	Seconds start

2.4. Traffic Measurement

In this reserach, we decided to use the packet sniffing program Wireshark to capture all traffic from each scenario. Since all data passes through NAT, Wireshark can simply measure data that is being exchanged only for the network. In scenario 1 and 2, all client PC is pointed to local DNS 192.168.0.70. Thus, those two scenarios have the same network flow and forward to same upstream DNS server. While in scenario 3, all client PC is pointed to 192.168.5.70 local DNS configured with Pi-Hole and the same upstream DNS as in scenario 1 and 2. All traffic from each IP address is investigated in Wireshark in to filtering the traffic. Wireshark is filtered the traffic according to the rules.

According to [9], to allow them to have uniform base of comparison across all the controls and all the subsequence they conducted the test one a single day. Therefore, this research is conducted tests for all scenarios within twenty-four (24) hours. All scenarios have to follow procedure below [17] in order to get the reliable result:

a) DNS queries point to same upstream DNS (GoogleDNS).
b) All web browser request to AutoBrowse and start and end concurrently.
c) Clear web browser cache before every test start.
d) Each scenario must be conducted a 3 times test.

2.5. Data Analysis

After capturing traffic from all scenarios, we conducted an offline investigation on the traffic generates from web browsing emulation. For data analysis, this experiment filtered on the IP address of the client PC and Autobrowse programs. This research measured each scenario at least three (3) times to get average result. Compare the results of network traffic from all scenarios and get the number of HTTP get request, TCP connection and bandwidth consumption while implemented the AdBlock.

3. RESULTS

After conducted the experiment, we have the output during data gathering and testing phase. The result was tabulated in the form of tables and graph and is discussed later in this topic. Figure 4 shows graph that demonstrate network without AdBlock after 25 minutes of web browsing activities is captured which including web content and online advertisement. A zigzagged lines and remained constants line show after spike occurs.

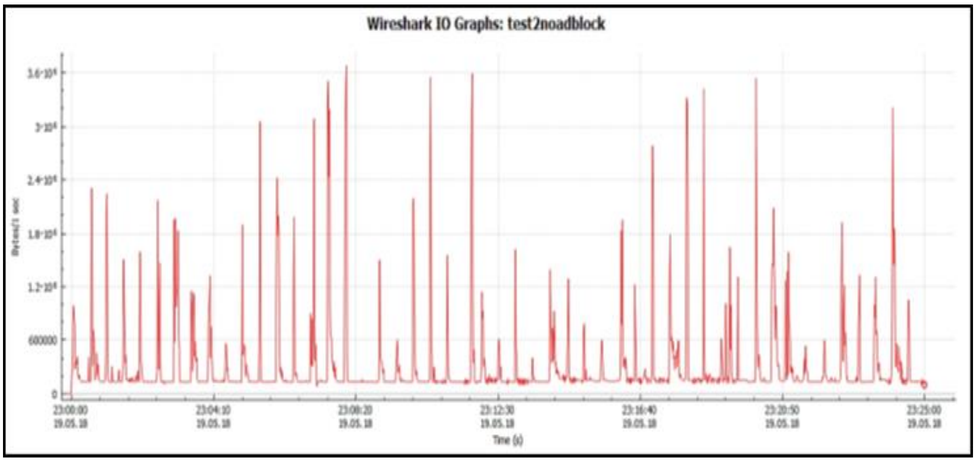

Figure 4. Graph network without adblock: scenario 1

Figure 5 shows a graph that demonstrate network with browser extension AdBlock. This scenario is demonstate where at end device is deployed using browser extension AdBlock. The graph shows a change of behavior where there are less zigzagged lines.

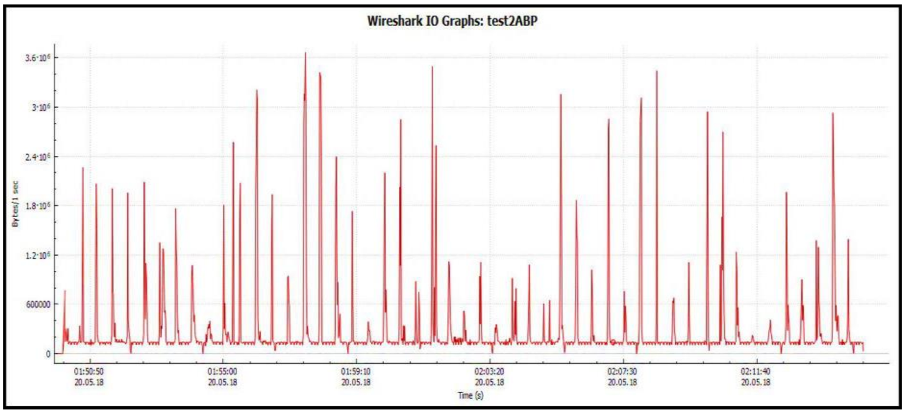

Figure 5. Graph network with browser extension adblock: scenario 2

Figure 6 shows a graph that represent network with DNS AdBlock where this scenario AdBlock was deployed at DNS level (Pi-Hole). The graph shows the similarities of spike occured as in scenario 1 and 2. In contrast, this graph demonstrated a change in behavior where almost none of zigzagged lines showing after a spike occured. Moreover, compare to graph Figure 4 and 5; this graph demonstrates after a spike occurs and it reached a low almost zero (0) byte per second.

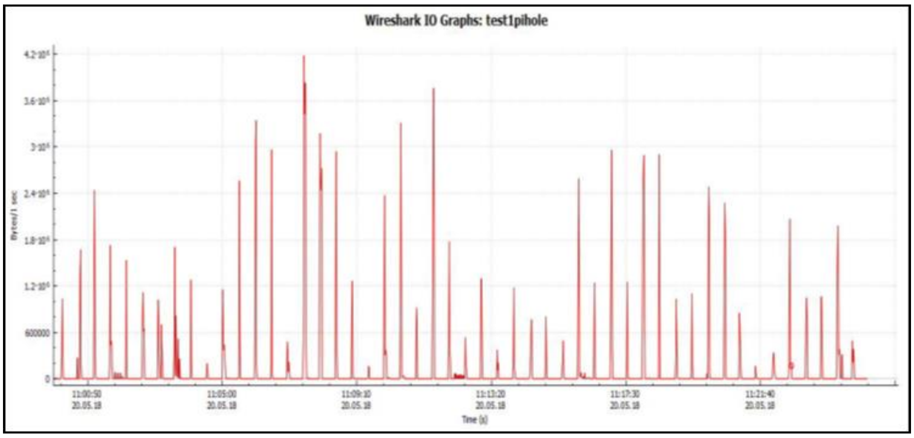

Figure 6. Graph network with DNS adblock: scenario 3

3.1. Wireshark Packet Analysis

In this analysis, we simulate two end devices where each of it running Ubuntu Linux and Windows operating system (OS) both concurrently requesting for the same URL using Google Chrome web browser. Table 2 shows test case result for 60 second. This the experiment, we found that different OS have different number of HTTP Request == "GET" and TCP Traffic. However, in DNS AdBlock, HTTP Request == "GET" is the same in both OS. It is important to mention here that web browser, AdBlock and OS are being setup as by default where there is no modification at end devices and the method of data collection are being captured at network level.

Table 2. Test Cases and Time Frame

Test Cases	Time Frame
Test 1 (Network without AdBlock)	(frame.time >= "May 19, 2018 23:00:00" && frame.time <= "May 19, 2018 23:01:00")
Test 2 (Network with Browser Extension AdBlock)	(frame.time >= "May 20, 2018 12:45:00" && frame.time <= "May 20, 2018 12:46:00")
Test 2 (Network with DNS AdBlock)	(frame.time >= "May 20, 2018 12:45:00" && frame.time <= "May 20, 2018 12:46:00")

As shown in Table 3, the implementation of AdBlock consume almost half of traffic in network without AdBlock. However, the deployment of different AdBlock technique generates a difference result (HTTP Request == "GET" and TCP Traffic Inbound). HTTP Request == "GET" for browser extension AdBlock generates less request compare to DNS AdBlock while TCP inbound connection for browser extension AdBlock generates more connection compare to DNS AdBlock.

Table 3. Packet Analysis

Area of Investigation	Without AdBlock	Browser Extension AdBlock	DNS AdBlock
HTTP Request == "GET"	319	102	152
TCP Traffic Inbound	4,575	2,687	1,918
Bandwidth (MB)	8	4	4

This happen because of TCP connection used the selective repeat protocol (SRP) with positive acknowledgments (ACKs) and time-out. Each byte sent is numbered and been acknowledged. A number of bytes and ACKs can be sent in the same packet where indicates the sequence number of the next byte expected by the receiver. ACKs carrying sequence number that carried acknowledgment. If there is packet lost, receiver send a duplicate ACKs for a subsequent received packet.

4. CONCLUSION

This research outcome shows the impact of web browsing activities. The active traffic consumed the highest inbound bandwidth usage in enterprise network environment. With the number of daily online activity of user in enterprise network, online advertising contents might impact high data demand from web browsing activity. As prediction by 2018, 80% of web contents are delivered in media platform [5]. This study has identified current web browsing trends traffic in enterprise network where it consumed around 50 percent in average. This statistic is claimed from a study conducted by Malaysian Communications and Multimedia Commission (MCMC) stated that the trends for years 2017, 87% of their respondents used web browser to retrieve information and 67 percent used for formal and informal for study.

This study agrees that both AdBlock techniques perform a reduction of traffic and bandwidth usage. However, the best solution in enterprise network for AdBlocking technique is DNS AdBlock. By implementing DNS AdBlock in enterprise network environment can sustain the usage of web browsing activity for enterprise network and also it has the potential to generate substantial saving across several fonts. Besides, the implementation of browser extension AdBlock proved that the process of displaying online advertisement drains a significant amount of energy usage. Therefore, when scaled to large network, DNS AdBlock is an effective solution to control of end devices for process of blocking online advertisement.

However, browser extension AdBlock has the limitation for enterprise network. As number of device connected is increased, each device need to installed AdBlock. A discussion had been made by Simon Fraser University where by deploying browser extension AdBlock presents some difficulties from the perspective of a network administrator. A study that has been conducted by [18], shown that the limitation in data collection where researcers did not know the exact number of users with browser extension enabled. In this researcher study proved that there is a gap in TCP connection where massive web activity is generated since it used WebRequest API. On a large network, massive amount of bandwidth is consumed in network with browser extension network compare to DNS AdBlock. Using a DNS AdBlock acts as DNS resolver when devices connected to the network environment. For online advertisement, DNS AdBlock controlled at network level. Thus, this solved the difficulties of a network administrator in managing the user network activities.

ACKNOWLEDGEMENTS

The authors would like to thank the Faculty of Computer and Mathematical Sciences (FSKM) Universiti Teknologi Mara (UiTM) for sponsoring this research.

REFERENCES

[1] I. Pawełoszek, “Web 3.0 Applications in Enterprise Strategy,” Publisher of the University of Economics in Katowice, vol. 234, pp. 129–139, 2015.

[2] A. Albasir, *et al.*, “Experimental Study of Energy and Bandwidth Costs of Web Advertisements on Smartphones,” *Proceedings of the 6th International Conference on Mobile Computing, Applications and Services*, pp. 90–97, 2014.

[3] E. Thomas, *et al.*, “Semantic Advertising for Web 3.0,” Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 6152 LNCS, pp. 96–105, 2010.

[4] M. Tekdal, *et al.*, “Developments Of Web Technologies And Their Reflections To Education: A Comparative Study,” *Journal Of Educational And Instructional Studies In The World*, vol/issue: 8(1), 2018.

[5] D. Silverman, “IAB Internet advertising revenue report,” 2017 first six months results. Pricewaterhouse Coopers LLP, 2017.

[6] H. Jelodar, *et al.*, “Provide a solution to isolate and identify web of advertising Persian: Web annoying,” *Procedia Computer Science*, vol. 57, pp. 411-417, 2015.

[7] I. Jalb and A. Olteanu, “Customized Ad Blocking,” *RoEduNet Conference: Networking in Education and Research*.

[8] M. H. Mughees and Z. S. Z. Qian, “Detecting Anti Ad-blockers in the Wild,” *Proceedings on Privacy Enhancing Technologies*, vol. 3, pp. 127–142, 2017.

[9] A. Parmar, *et al.*, “AdBlock Plus Efficacy Study. Technical Report,” Simon Fraser University, 2015.

[10] Darus M. Y., *et al.*, “Optimizing Congestion Control for Non Safety Messages in VANETs Using Taguchi Method,” in Alfred R., *et al.*, “Computational Science and Technology,” *ICCST 2017. Lecture Notes in Electrical Engineering*, vol. 488, 2018.

[11] S. M. Hatim, *et al.*, “VANETs and Internet of Things (IoT): A Discussion,” *Indones. J. Electr. Eng. Comput. Sci.*, vol/issue: 12(1), pp. 218–224, 2018.

[12] Aryaka Networks Incorporated, “The State of the WAN 2017,” Arkaya, San Mateo, CA, USA, 2017.

[13] Fortinet Incorporated, “Fortinet Threat Report Q2 2017,” Fortinet, Kifer Road Sunnyvale, CA, USA, 2017.

[14] E. L. Post and C. N. Sekharan, “Comparative Study and Evaluation of Online Ad-blockers,” *Proceeding 2nd International Conference Information Science and Security (ICISS)*, 2015.

[15] C. E. Wills and D. C. Uzunoglu, “What ad blockers are (and are not) doing,” *Proceedings - 4th IEEE Workshop on Hot Topics in Web Systems and Technologies, HotWeb 2016*, pp. 72–77, 2016.

[16] G. Merzdovnik, *et al.*, “Block Me If You Can: A LargeScale Study of Tracker-Blocking Tools,” *Proceedings of the 2nd IEEE European Symposium on Security and Privacy (IEEE EuroS&P)*, 2017.

[17] D. B. Van and A. Pras, “The costs of web advertisements while mobile browsing,” Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 7479 LNCS, pp. 412–422, 2012.

[18] R. J. Simons and A. Pras, “The Hidden Energy Cost of Web Advertising,” *Proceedings of the Twelfth Twente Student Conference on Information Technology*, 2010.