IN HTTP(S) POTENTIAL TRAFFIC OVERHEAD FOR MOBILE DEVICES

Volodymyr Kazymyr, Andrii Mokrohuz

Abstract: *A mobile device became a powerful computing device, which can provide the same functionality as a personal computer. Increasing amount of different mobile devices leads to increasing of data transferred by application installed on mobile devices. Nowadays, everybody uses a mobile device, due to the fact, that mobile devices are convenient and have big variety of applications available to install. Applications can send data by using the Internet access, which in most cases is provided by Wi-Fi, 3G or 4G networks. This paper considers the most popular application level protocol HTTP (HTTPS) in the Internet. The experiment was carried out to collect HTTP traffic and investigate amount of technical and potentially overhead data that increases overall traffic. The HTTP traffic was grouped by content type. Assumption that headers of HTTP packets contain overhead data was made. Data types with the greatest header to body ratio were identified. High header to body ratio can be non-acceptable for devices with some limitations such as mobile devices. Amount of traffic transferred between a mobile device client and a server plays a significant role in mobile device performance. Therefore, some limitations of mobile devices, in terms of data transferring, were described in the paper as well.*

Keywords: *HTTP, mobile device, packet, header, application level.*

ACM Classification Keywords: *C.2.2. Network Protocols*

Introduction

Mobile devices connected to the Internet are part of our life. A mobile device is a small computing device which has an operation system, runs various applications (software), offers reach functionality, supports interactivity with users, works on battery and has access to the Internet or to any local network. 3G and 4G networks provide anytime access to the Internet. A good example of mobile device is a smartphone or a tablet pc. Any mobile device is a set of hardware and software, which works together.

Mobile device is a small computing device with relatively powerful computing abilities, which makes possible to use it everywhere. The comScore report shows that number of mobile devices' users achieved the number of desktop computers' users in 2014 as shown on Figure 1 [Lella, 2014].

The trend forecasts increasing of mobile devices' user as it is shown on the Figure 1. Mobile devices with the Internet access or network access are considered in the paper. Statistics shows information on

the global mobile traffic per month for mobile devices, which can be seen on the Figure 2 [Statista, 2015].

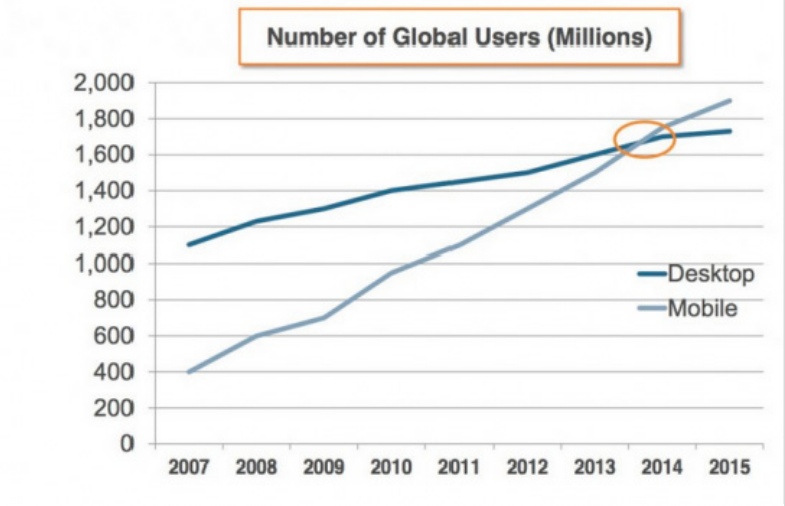

Figure 1 – Number of desktop and mobile devices' users.

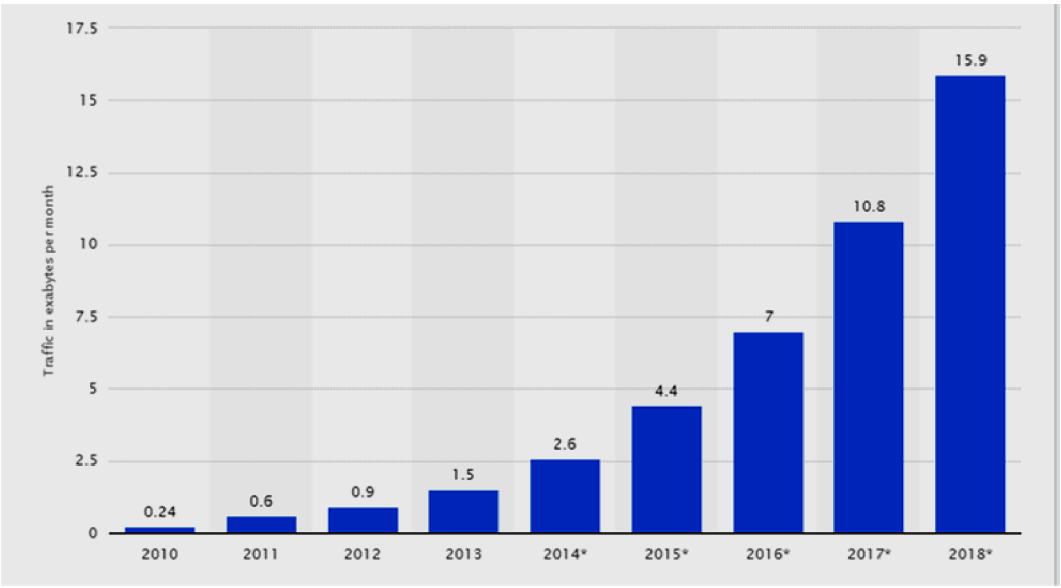

Figure 2 – Global mobile traffic forecast.

In 2015, mobile devices sent and received 4.4 exabytes ($4.4 * 10^{18}$ byte) per month. In 2016, devices will send and receive 7 exabytes ($7 * 10^{18}$ byte) of data per month and this value will grow in the future. Growing amount of mobile devices' users and amount of traffic transferred requires consideration of protocols and mobile devices' limitations (hardware and software) in order to handle increasing amount of data. The paper concentrates on software of mobile devices; especially, it concerns protocols of communication between client and server and touches hardware only in relation with software. We consider these types of data (information) which can be sent and received by mobile devices and be useful to end user:

1) Graphical is the most ancient type of information, which has been implemented for a way to store information about the world in the form of rock paintings, and later in the form of paintings, photographs, diagrams, drawings on paper, canvas, marble and other. The material depicting picture of the real world.
2) Sound (acoustic) type of the information is a kind of musical information. For this type, special method has been devised, which makes storing sound information similarly to graphical information.
3) Text is a method of encoding human speech by using special characters such as letters. Different people have different languages and use different sets of characters to display the text.
4) Numerical type is a quantitative measure of objects and their properties in the surrounding world. It is similar to the text information and special characters (digits) are used to display numbers using the method of encoding. The coding system can be different as well.
5) Video data is a way to save the "live" pictures of the world; it appeared with the invention of cinema.

HTTP for mobile devices

Users of mobile devices and the Internet or network send these types of information from a client to a server and back. Obviously, they send information between pieces of software or applications. Thus, special rules are required to send data from one application to another application. Good example of such rules is a protocol. There are different levels protocols in terms of network communication. These levels are presented by OSI model as shown on Figure 3 [Schwab, 2015].

Data is wrapped into headers and trailers as it seen on Figure 3. This can significantly increase size of a packet transferred via a network. Size of useful data can be smaller than size of technical data on each layer which is required to deliver the packet. The article considers only Application layer protocols. No doubt, one of the most popular protocols is HTTP (HTTPS) in the Internet. Every smartphone or tablet pc has a browser application which uses HTTP to obtain information from a server and send requests on a server.

HTTP protocol allows transferring of these five types of information, which are presented above. The structure of the HTTP protocol is relatively simple, as can be seen on Figure 4. HTTP works in client-server architecture environment: client sends HTTP request and server sends HTTP response back [Fielding, 1999].

Http header defines information, which is required for a client (server) to identify type of a data in a body. Header contains MIME header, which identifies a type of information in HTTP request/response. HTTP protocol can work with all types of information: images, video, text, numbers, sound. Different MIME types exist and they point what information is in a body of HTTP request/response [Fielding, 2014].

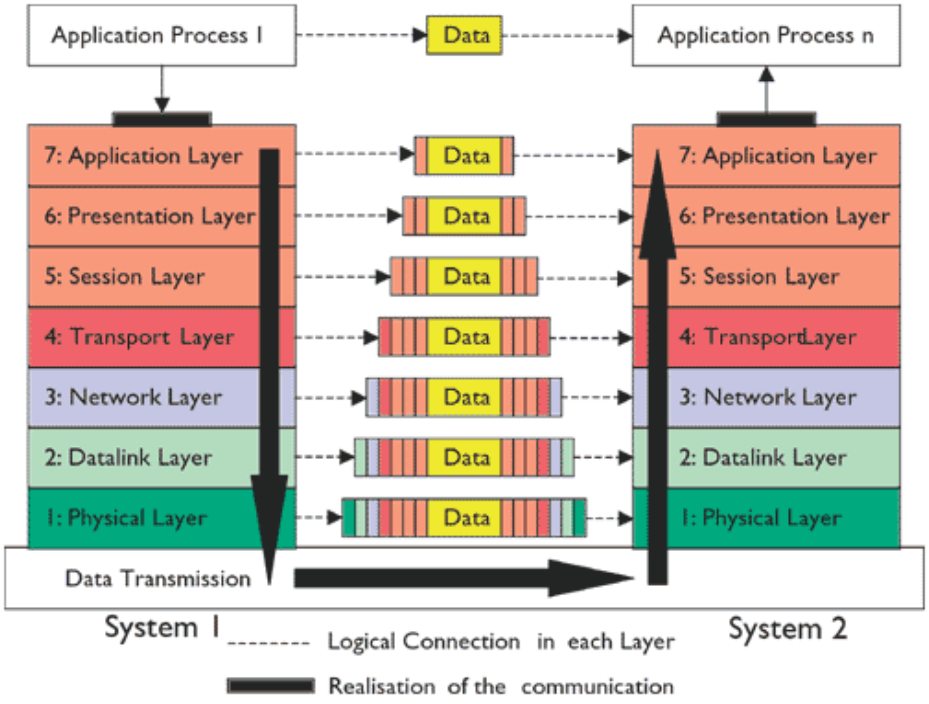

Figure 3 – OSI model and network process communication.

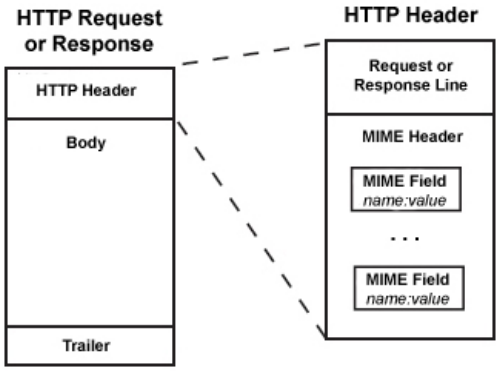

Figure 4 – Structure of HTTP protocol.

Protocol HTTP acts the same way on mobile devices as on normal PCs. However, a mobile device is generally a computing device with some limitations. These limitations make mobile device "special" in terms of transferring data via a network. We consider these limitations for mobile devices.

1. Range and Bandwidth of the Internet connection. 3G or 4G is slower than Wi-Fi or direct cable connection. Many mobile devices cannot have direct cable connection because of its construction. Wi-Fi can be used, but it has limited range. Slower and limited internet connections dictate obvious rule about information transferred from or to mobile device. The

amount of data transferred should be smaller, but this does not mean that loss of data is allowed. Costs of data transferred via 3G or 4G networks are relatively high as well.

2. A mobile device works on a battery power most of the time. 3G and 4G consume battery with bigger intensity. So working time of a mobile device depends on battery life.
3. Usually screens of mobile devices are smaller than devices of normal PC, so developers need to think what information to send to mobile devices' clients more carefully.
4. Security is one of the important issues for communication especially when one of the parties is a mobile device. Security assumes additional overhead of the traffic, so this is another limitation for mobile devices. Because more traffic means less battery life, so functionality of mobile devices is decreased as well.

Client applications, which work on mobile devices, should be implemented in a way, which keeps in mind the mobile devices' limitations. Mobile devices send a lot of traffic as it can be seen on Figure 2. However, this number shows overall traffic including headers and technical requests. In this paper, we have identified amount of technical data in HTTP traffic by conducting the experiment.

Description of the experiment

HTTP packets contains useful for user data and some technical data to make packet deliverable to its destination. Experiment was carried out to identify amount of technical data in different types of data and different content types. Fiddler web debugging tool was used to collect the http traffic on user's computers. Users' PCs were running Windows 7, 8.1 or 10. There were no special rules for experiment. Users lunched the Fiddler application on their computers and were acting as usual during working day (7 hours). Traffic has been provided by users from Ukraine and the United Kingdom. More than 30 000 HTTP packets were captured during the experiment. This experiment was not targeting to capture any specific type of data.

Results of the experiment

All captured data types (image, sound, text and video) are presented in the Table 1.

Table 1. General results of the experiment

Content type (data type) of packets	Header size bytes sent	Body size bytes sent	Header size bytes received	Body size bytes received
Without content type technical Data	1763533	751	2094350	331315

application/javascript	434312	166	179746	7795996
application/json	444664	54535	175779	1030410
application /ocsp-response	34774	4822	25791	115011
application/octet-stream	175870	0	101811	10712353
application/pkix-clr	25398	0	21834	1245605
application/x-javascript	548334	0	307252	12850516
Application/zip	406	0	549	0
audio/mpeg	49939	0	31540	661966617
font/woff	35900	0	25138	1978328
image/gif	1868903	55500	926093	41100607
image/jpeg	1753451	0	1283130	67497881
image/png	1101317	0	550675	23231135
Image/svg+xml	30685	0	15960	247044
image/x-icon	54926	0	24232	223253
text/css	347328	0	152731	6843145
text/html	2613740	1006415	1536377	14644210
text/javascript	5492106	1252222	1628782	7476888
text/plain	1657117	190501	696317	59600930
video/mp4	3688	0	2739	23156083
Total	**18436391**	**2564912**	**9780826**	**342047327**

The pie chart with the distribution by amount of the captured packets is presented on Figure 5.

All the data was group according to these data types: text, sound, video and image. This allowed decreasing amount of HTTP data types and presenting the data in a way that is more convenient. HTTP data types were grouped as they are presented in Table 2. Tomasz Bujlow et.al proposed similar traffic classification, however authors concentrated on traffic classes and we split the traffic according to the data types which recognizable by a human. [Bujlow, 2012]

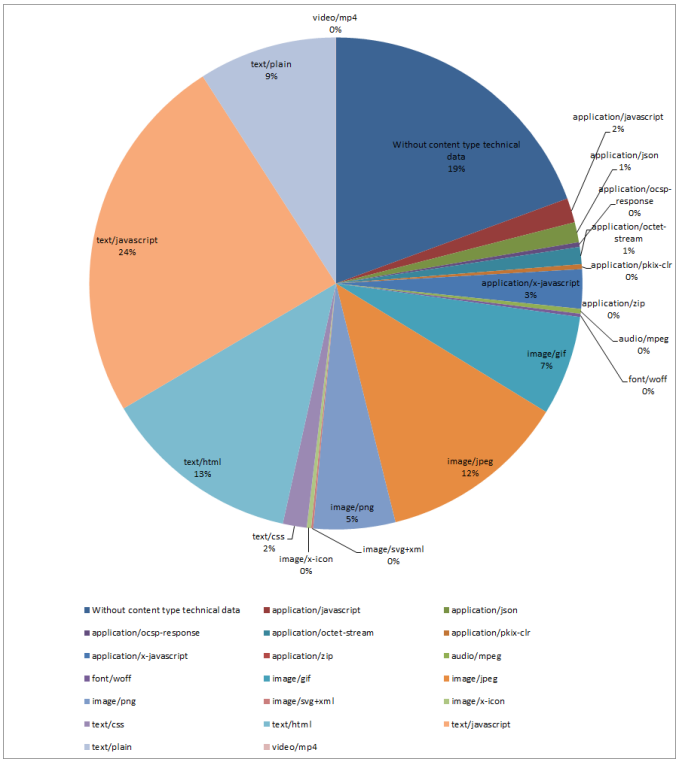

Figure 5 –Distribution of http data by amount of packets.

Table 2. Grouped HTTP data types

General type	**HTTP type**
Text	Without content type technical data, application/javascript, application/json, application/ocsp-response, application/octet-stream, application/pkix-clr, application/x-javascript, text/css, text/html, text/javascript, text/plain, font/woff
Image	image/gif, image/jpeg, image/png, image/svg+xml, image/x-icon
Sound	audio/mpeg
Video	video/mp4

Distribution by amount of packets, header size bytes sent, body size bytes sent, header size bytes received and body size bytes sent is shown on Figure 6. Text and image headers bytes sent and received fill in most of the overall headers traffic. Amount of packets with image and text data takes 99% of overall amount. Despite the sound packets' amount is small, body size received by the client takes 70 % of the overall body size traffic.

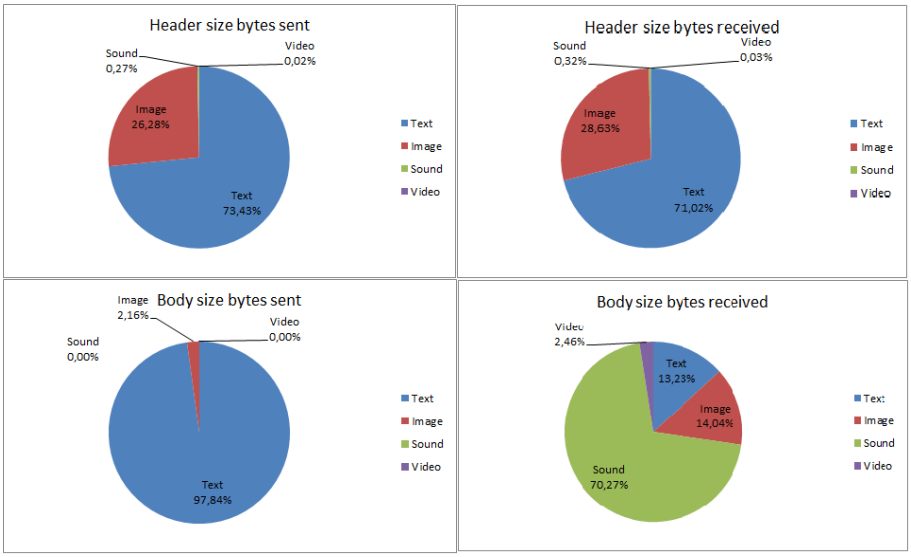

Figure 6 – Graphical results of data distribution by united type of information in bytes.

Distribution of the data types by amount of packets is presented on Figure 7.

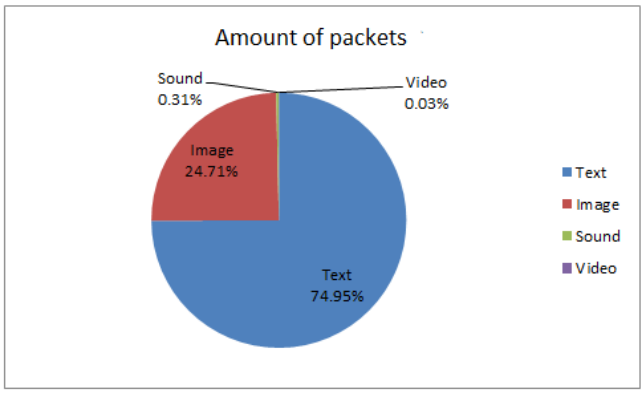

Figure 7 – Data types' distribution by amount of packets.

Another research, which has been carried out by Pengcheng Jiang et.al, has shown similar distribution by amount of packets. Authors analysed much more network traffic and received results presented on Figure 8. The data was collected from provincial network of well-known ISP (internet service provider) in China [Pengcheng Jiang, 2014].

Type	*Times of occurrence*	*Proportion*
Text	92702334	48.45%
Application	52749385	27.57%
Image	43817095	22.90%
Video	1377169	0.72%
Audio	244219	0.13%
Other	452809	0.24%

Figure 8 – Proportion of different data types.

We have united Text and Application data types into one Text data type and received 74.95% proportion. Pengcheng Jiang et.al received similar results if Text and Application are united 76.02 %. Proportion of image data type is similar as well: 24.71% and 22.90% accordingly [Pengcheng Jiang, 2014].

HTTP headers potentially contain information, which is not necessary to be sent to clients' applications. End users do not need the information, which is inside headers. However, the «useful» information cannot be delivered without headers. HTTP(S) is an application layer protocol, thus software developers usually define what information which client software sends. Sometimes developers use frameworks or libraries provided by third party developers, which create the HTTP requests and responses including headers. Headers for text and image data type take significant part of overall bytes sent and received in comparison to video and sound. The percentage of headers in packets is presented on Figure 9.

Figure 9 – Ratio of packets' body size to header size by type of information.

The percentage of headers in overall traffic is not as high as in text information as this can be seen on Figure 10.

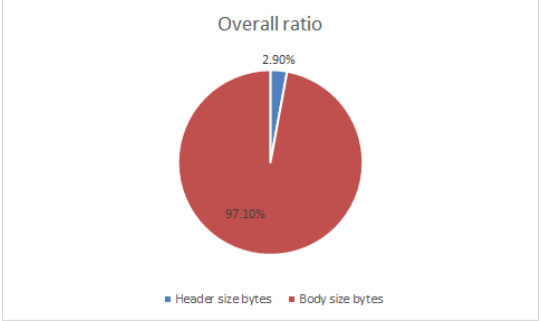

Figure 10 – Body size to header size ration for the data types.

Sound and video packets contains very small amount of headers bytes in comparison to body size. However, text and image headers size in bytes is noticeable. Unfortunately, headers cannot be fully eliminated, but some optimization can be carried out.

Conclusion

Experiment was conducted to identify data types which contain the biggest headers percentage in bytes in HTTP(S) packets. According to the experimental data, Text and Image information contains more technical information than other data types. Mobile devices' limitations were also considered. Mobile devices' operational time depends on battery life, mobile networks consumes battery life and decrease time which device can work without charging. The more data client needs to transfer the less time a mobile device can work without charging. Text and image information can decrease the battery life because of potential headers' overhead.

Results of the experiment allow future development and research in different directions. One of these directions can be consideration of alternative application level protocols. Another one can be improvement technologies and technics directed to decreasing amount of headers' data. Application level protocols depend on developers, in most of cases a developer decides which data to send. Therefore, some recommendation for developers can be implemented to improve amount of data transferred via HTTP as the most popular application level protocol. These recommendations can be also applied to develop software which does not use HTTP.

Bibliography

[Lella, 2014] Statistics on mobile usage and adoption to inform your mobile marketing strategy. comScore Whitepaper. August 21, 2014.

[Statista, 2015] Global mobile data traffic from 2015 to 2020 (in exabytes per month). Inernet: http://www.statista.com/statistics/271405/global-mobile-data-traffic-forecast.

[Schwab, 2015] Looking behind the automation protocols. Internet: http://www.iebmedia.com/index.php?id=4582& parentid=63&themeid= 255&showdetail=true.

[Fielding, 1999] Hypertext Transfer Protocol -- HTTP/1.1. Internet: http://tools.ietf.org/html/rfc2616.

[Fielding, 2014] Hypertext Transfer Protocol (HTTP/1.1): Semantics and Content. Internet: http://tools.ietf.org/html/rfc7231#section-3.1.1.5.

[Bujlow, 2012] T. Bujlow, T. Riaz and J. M. Pedersen, "Classification of HTTP traffic based on C5.0 Machine Learning Algorithm," Computers and Communications (ISCC), 2012 IEEE Symposium on, Cappadocia, 2012, pp. 000882-000887.

[Pengcheng Jiang, 2014] Pengcheng Jiang; Fang Liu; Huan Wang; Chenyu Li, "Characterizing HTTP Traffic of Mobile Internet Services in Provincial Network," in Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2014 Sixth International Conference on , vol.1, no., pp.78-81, 26-27 Aug. 2014.

Authors' Information

Volodymyr Kazymyr *– Dr. Sc., Prof. Chernihiv, National University of Technology, 95, Shevchenko street, Chernihiv-27, Ukraine, 14027; vvkaymyr@gmail.com*

Major Fields of Scientific Research: computer science, information technologies, complicated computer systems.

Andrii Mokrohuz *– Ph.D. Student, Chernihiv National University of Technology, 95, Shevchenko street, Chernihiv-27, Ukraine, 14027; myworkingadr@gmail.com*

Major Fields of Scientific Research: client-server architecture, client-server applications communication, application level protocols.