# The Costs of Web Advertisements while Mobile Browsing

Jeffrey van den Brande and Aiko Pras

University of Twente
The Netherlands
`j.vandenbrande@student.utwente.nl,a.pras@utwente.nl`

**Abstract.** Tablet PCs, iPads and mobile phones all include facilities to browse the mobile Internet. The costs of mobile Internet access may become extraordinary, however, when the data limit is exceeded or when the user is roaming abroad without a roaming data plan. Since users may see advertisements as unwanted traffic, they can be confronted with a bill of which a significant part is caused by data they did not ask for. The goal of this paper is to investigate the amount of data generated by advertisements when browsing, and the potential costs of such advertisements for average users. Our analysis shows that web advertisements take up 7-9% of the total data traffic, depending on Flash usage.

**Keywords:** mobile Internet, web advertisements, roaming, accounting

## 1 Introduction

Wireless Internet access is an important feature for users of mobile devices such as PDAs, mobile phones, tablet PCs, iPads etc. [1]. Like users of fixed PCs, mobile users receive advertisements on web pages, for which they may have to pay. Currently many providers offer relatively cheap flat-fee data plans for mobile Internet usage and some even provide some flat-fee plans for roaming. At this moment there are many discussions, however, since several providers want to reduce or even stop such offerings, as according to these providers such plans cost more money for investments in their infrastructure than the plans bring in [2]. Also, when mobile users access the Internet outside the provider's base country (roaming), the costs of accessing the Internet are usually not covered by the regular flat fee that was valid at home. It is only recently that the European Commission decided to force providers to limit their roaming costs inside the European Union [30]. Finally, even if the Internet is accessed in the provider's base country, users may have to pay additional money in case they exceed the limit of their data plan. Although plugins exists for browsers on fixed PCs, such as Adblock Plus and Opera's URL filtering, such plugins are generally not available for mobile devices. Therefore it is interesting to investigate whether the amount of data (and thus the costs) generated by web advertisements is an issue that is worth to consider, or can be ignored. Although many people may have an opinion on this, the scientific evidence to proof such opinion hardly exists.

The goal of this paper is therefore to perform such investigation, and give initial results. Since the number of pricing options is overwhelming, it is explicitly *not* the intention to present a complete analysis of all mobile Internet costs.

The research questions that will be investigated in this paper are as follows[1]:

1. Do web advertisements take up a significant amount of traffic?
2. Do these amounts vary per browser and over time?
3. Is there a significant difference when Flash is enabled?
4. Can we estimate the costs of web advertisements for roaming users?

The approach to answer these questions, is to set-up a measurement environment in which a configurable proxy server is used as advertisement filter. To mimic user behavior, an AutoBrowse program is used to retrieve hundred of the most popular URLs. With such set-up, the answer to the first question can be found by measuring the amount of web data with an active, and with an inactive proxy filter. To answer the second question, measurements will be repeated multiple times and with different browsers. By enabling and disabling Flash, is is possible to answer the third question. The approach to answer the fourth question is to search the web for pages showing the average (prepaid) costs of mobile Internet. Since the amount of web traffic generated by an average advertisement is known from the previous questions, we can calculate the costs of a single advertisement by multiplying the amount of web traffic with these average costs. Once we found the average number of pages downloaded per day, we can also calculate monthly costs. We focussed, as examples, on the Netherlands and the United States. Note that, due to the overwhelming number of pricing options, the often confusing information on roaming costs and the fact that mobile operators sometimes 'reshape' data (for example, to re-encode pictures), our answer to the fourth research question remains limited to an initial estimation; a complete answer to this question would be too intricated to fit within the scope of this workshop paper.

The structure of this paper is as follows. Section 2 introduces the general topic of Internet accounting and discusses the related work. The approach to tackle the research questions 1, 2 and 3 will be presented in Section 3; the answers to these questions is provided in Section 4. The last research question is analyzed in Section 5, whereas the conclusions can be found in Section 6.

## 2    Related Work

Internet accounting is a relative well-established research area, which received quite some attention with the invent of INTSERV [4] and DIFFSERV [5] technologies at the end of the nineties of the previous century [6]. At that time it

---

[1] It should be noted that an initial version of this paper has been presented at the fourteenth Twente Student Conference on Information Technology. However, that was an internal event of the University of Twente, of which the proceedings have not officially been published by any publisher [3].

was believed that the Internet would evolve into a network capable of providing different quality of service classes to its users; terms like 'gold', 'silver' and 'bronze' were used to denote these different classes. Depending on the quality required for a specific service, the user could select a specific quality and switch to another class, if needed [7]. To cope with the potential problem of service asymmetry, *i.e.* the client requests for data, whereas the server sends the data and will therefore be charged, well-known concepts like reverse charging were investigated for suitability within the Internet [8].

Due to the complexity of service differentiation within the Internet core, the idea of charging based on service usage and quality slowly diminished and nowadays many ISPs have adopted flat-fee schemes, also since users seem to prefer such schemes [9]. Content providers are still confronted with service asymmetry, however, but discovered advertisements as a viable approach to cover at least part of their costs.

Nowadays the usage of iPads, PDAs, and other mobile devices capable to access the Internet shows an impressive growth. The pricing schemes used by mobile operators are completely different from the flat-fee schemes used by ISPs for the fixed Internet, however, which means that also the research interest in (mobile) Internet accounting is again growing. Several studies already exist on generic accounting architectures, and how accounting data can be collected within networks [10]. There are also many papers that describe the design of accounting systems for the mobile Internet [11][12]. The behavior of mobile users browsing the Internet has been investigated [13], and there are many papers that discuss the effects and click-through rates of web advertisements. Such papers are often written from the perspective of the advertiser or website owner (to improve the design of web servers and effectiveness of the advertisements), however, and not from the perspective of the user. This paper takes the perspective of the mobile user, however, and builds upon an earlier paper that analyzed the hidden energy costs of web advertisements [14].

## 3   Approach to measure advertisement traffic

The best approach to answer the first three research questions, is by comparing the browsing behavior of mobile internet users without advertisements being blocked, and with such blocking. Since it is hard to obtain usage data from mobile operators, experiments with real users cannot easily be performed. An alternative approach is therefore to set up a test environment that automatically retrieves some of the most popular web pages. By retrieving these pages twice, the first time with and the second time without web advertisements, it is possible to calculate how much traffic is used for web advertisements. This chapter first discusses approaches to filter advertisements (section 3.1), followed by how users browse webpages (section 3.2), and ways how such browsing can be emulated (section 3.3). The chapter concludes with approaches to measure traffic (section 3.4) and a presentation of the specific measurement set-up used in the remainder of this paper (section 3.5).

### 3.1   Blocking advertisements

First of all a method to block web advertisements is needed. There are several methods available for this purpose. Most methods make use of at least URL filtering, which means advertisements are blocked based on their URL (prefix). The Opera web browser includes this feature by default, although it does not include default entries. Most advertisement blocking solutions rely on more sophisticated methods, such as CSS- and JavaScript-based filtering. The Adblock Plus plugin for Mozilla Firefox [15] is an example of such a solution. Since different types of web browsers should be investigated in this research, browser-specific advertisement blocking solutions should not be used.

Another approach is to use a proxy server. A proxy server is a device that is located between the web browser and the web server. The proxy receives the requests from the web browser, and forwards the requests to the web server on behalf of the web browser. Responses also go via the proxy. Since all data passes the proxy server, the proxy can simply measure the data that is being exchanged. The proxy can also modify this data, for example to filter advertisements.

Proximodo [18] is a kind of proxy server that can filter based on a URL blacklist. Unfortunately the program is no longer maintained and seems to be outdated. Privoxy [19] is a more advanced proxy server that can also filter based on the more sophisticated methods described above. Finally there is the possibility to convert an Apache web server to a proxy server, using the 'mod_proxy' module [20]. Such approach would take a lot of time and detailed knowledge on the operation of Apache web servers, however, and would for our purpose not bring any advantages compared to Privoxy. Therefore it was decided to use Privoxy as advertisement blocking method for the remainder of this research.

### 3.2   Browsing behavior

Since there are, to the best of our knowledge, no recent and openly available behavioral studies that specifically analyze how mobile users browse the Internet, we will assume within this paper that the browsing behavior of mobile users is comparable to that of 'normal' Internet users. In this paper we will therefore not consider that some websites also offer 'mobile' versions or special Apps, which are optimized for mobile devices with small screens. In this paper we also assume that roaming users are more reluctant to browse the mobile internet, and thus visit far less web pages than when they would not be roaming. The research of Infante and Vallejo [31] gives an indication of this, as they state that on average the international roaming market accounts for about 4% of the total revenues in the mobile services industry, while the international roaming rates are several factors higher than domestic data rates. Therefore, it can be concluded that roaming users produce only a small fraction, say 10%, of their normal traffic when they are not roaming.

To mimic the behavior of average Internet users, the URLs of the web pages that are visited should be known. For our purpose it will not be sufficient to

simply look at website rankings to find the most popular websites, as such rankings only include the home page of popular websites, which may contain many advertisements, while an average user may also open deep pages, which contain fewer advertisements.

According to Simons [14], a good way to obtain a list of frequently seen web pages is to analyze the most popular posts on Digg.com. This is a community website that allows users to post links to web pages with content that the users find interesting [16] and thus gives a good reflection of popular web pages loaded by average users. Digg provides an Application Program Interface (API), based on the PEAR package of PHP, for retrieving its content; optionally it is also possible to retrieve XML-files. For our research we have used this option to retrieve the 100 most popular stories. Extracting the URLs out of these stories results in a list of URLs average Internet users would browse to.

Next to the URLs, also the average time spent on these web pages needs to be discovered, as some "heavier" rich media advertisements might continuously reload. There has been some research on the average time a person stays on a web page. That research shows that a pageview lasts on average about 12 seconds [17], but the result includes pageviews of non-interesting web pages as well, to which a user has accidentally navigated to. These pages will be loaded, but viewed for only a short time. Since no numbers are available on the number of pages that were 'accidentally' loaded, our research assumes that all visited web pages are of interest to the user, and the user stays the same time on each page. Based on these considerations, the average time spent on each web page should be higher than 12 seconds.

### 3.3   Browsing emulation

For emulating a user's daily browsing behavior, the AutoBrowse program written by Simons [14] was used. This program is a JavaScript tool that runs on every modern web browser and that retrieves a given set of URLs. This set is the top 100 URL list mentioned before (see section 3.2); as interval to load new pages we chose 30 seconds.

### 3.4   Traffic measurements

Several tools exist to measure network traffic. Tcpdump can monitor network activity and offers some options for filtering the captured network traffic. Wireshark is based on Tcpdump, but offers a graphical interface and has better and more comprehensive filtering capabilities. Also it allows to easily generate statistics. To analyze wireless networks, Kismet is a good tool as well, and has the ability to export its traces to Wireshark compatible file formats.

For practical reasons, our research will not emulate actual wireless data transferral, but use wired ethernet connections instead. Therefore it was decided to use the packet sniffing program Wireshark on the proxy server to capture all traffic. For our analysis we filtered on the IP-address of the client PC on which the web browser and AutoBrowse programs were running.
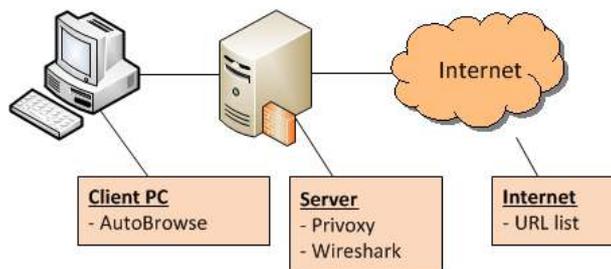
**Fig. 1.** Diagram of measurement set-up

### 3.5   Measurement set-up

The previous sections discussed all considerations to set-up our measurements environment. This environment consists of a client PC, a proxy server and a connection to the rest of the Internet (see Figure 1). The client PC has an Intel Core 2 Duo 2.40 GHz processor, 4.00 GB RAM and runs Windows 7 Professional. The proxy server is a dedicated Dell Latitude D830 laptop with an Intel Core 2 Duo 2.20 GHz processor, 2.00 GB RAM and also running Windows 7 Professional.

The client PC runs only a browser with the AutoBrowse program. To answer the question whether the results would be browsers dependent, two different browsers were used: Mozilla Firefox (version 3.6.12) and Opera (version 10.63). The proxy server runs the Privoxy program, as discussed in section 3.1; Wireshark is also run on the proxy. The cloud called "Internet" represents the connections to the web servers elsewhere, as identified by the URLs in our top 100 list.

## 4   Measurement results

Table 1 shows the measurements results. Each row in the table represents one measurement; in the first four measurements Flash was enabled, whereas in measurements 5 to 8 Flash was disabled. To see if results would change over time, the Firefox measurements were repeated multiple times. In each measurement 100 pages were loaded.

The table shows two numbers for the amount of network data captured. 'Normal' indicates the run in which the advertisements filter was disabled and 'Ads blocked' indicates the run in which the filter was enabled. The 'Difference per page' shows the differences between both amounts, using Equation 1. This difference is specified per page, and expressed in KB (instead of MB).

$$\text{Difference per page} = \frac{\text{Normal browsing} - \text{Ads blocked}}{\text{No. of pages loaded}} * 1024 \qquad (1)$$

The table shows an average difference of 109.41 KB per page when Flash is enabled; if Flash is disabled this difference is reduced to 71.47 KB per page. The

| No. | Browser | Flash used | No. of pages loaded | Total traffic (MB) | | Difference per page (KB) |
|---|---|---|---|---|---|---|
| | | | | Normal | Ads blocked | |
| 1 | Opera 10.63 | Yes | 100 | 133.38 | 120.63 | 130.51 |
| 2 | Firefox 3.6.12 | Yes | 100 | 117.70 | 108.10 | 98.99 |
| 3 | Firefox 3.6.12 | Yes | 100 | 144.61 | 132.55 | 123.52 |
| 4 | Firefox 3.6.12 | Yes | 100 | 113.18 | 104.92 | 84.62 |
| | | | **Average** | **127.22** | **116.54** | **109.41** |
| 5 | Opera 10.63 | No | 100 | 91.66 | 86.12 | 56.74 |
| 6 | Firefox 3.6.12 | No | 100 | 102.05 | 94.52 | 77.10 |
| 7 | Firefox 3.6.12 | No | 100 | 107.71 | 101.99 | 58.57 |
| 8 | Firefox 3.6.12 | No | 100 | 107.51 | 98.38 | 93.46 |
| | | | **Average** | **102.23** | **95.25** | **71.47** |

**Table 1.** Measurement results – Traffic with and without web advertisements

average normal traffic with Flash enabled (thus the normal case), is 127.22 MB; without advertisements the amount reduces to 116.54 MB. If Flash is disabled, these amounts are 102.23 MB and 95.25 MB respectively. From these numbers we may conclude that web advertisements increase data traffic with 7.3% (Flash disabled) to 9.2% (Flash enabled).

## 5 Costs

In the previous section we've discussed the percentage of network traffic generated by web advertisements. To estimate the costs of web advertisements for roaming users (our fourth research question), we will first present rates for mobile Internet usage (Section 5.1) and then use these rates to investigate whether the costs incurred by web advertisements is worth considering (Section 5.2).

### 5.1 Mobile Internet Rates

Before we can calculate the costs of web advertisements, we need to know the average rate per data unit. This rate is determined by analyzing the tariffs from popular mobile providers, from the US as well as the Netherlands (since the authors live in the Netherlands). Since we are interested in the additional costs caused by web advertisements, flat-fee plans for domestic as well as roaming use are not taken into account. Also, it is not the intention to present a complete analysis of all mobile Internet costs, as the number of pricing options is limitless, but to give a good indication of possible costs incurred.

For the Netherlands four major providers (KPN [21], Vodafone [22], Telfort [23] and T-Mobile [24]) have an average cost estimate of €0.22 per MB domestic and €3.88 per MB when roaming. The United States' major providers (Verizon Wireless [25], AT&T [26], Sprint [27] and T-Mobile [28]) average their costs per MB at $0.16 domestic and $17.48 when roaming. All numbers were collected in June 2012. Although these numbers are continuously changing and our tables

do not pretend to give a complete view, the numbers should still be usable to get an impression whether the additional costs of web advertisements can be significant, or not.

## 5.2   Web Advertisement Costs

To calculate the average costs of web advertisements for mobile Internet users, we still need to know the browsing behavior of average mobile users. Unfortunately little is publicly known regarding this behavior. However, Opera, which is used by roughly 56% of all mobile Internet users [32], publishes usage statistics for their mobile browser [29] on a monthly basis. In March 2011, for example, Opera measured per user an average number of daily pageviews of 13.9 in the United States, and 7.4 in the Netherlands.

These numbers can also be used to calculate the costs associated with mobile Internet browsing. Multiplying these average daily pageviews with the average data rates (Section 5.1) and the average amount of data generated per web advertisement (Section 4), gives us the average daily costs of web advertisements and thus the average monthly costs (see Table 2). In this paper it is assumed roaming users visit 90% less pages than domestic users, because they only browse to essential webpages, and thus this is taken into account in the calculations. As there is a noticeable difference between traffic with Flash enabled or disabled, both cases are shown separately.

|  | Monthly costs | |
|---|---|---|
|  | **Flash enabled** | **Flash disabled** |
| NL domestic | €5.32 | €3.48 |
| NL roaming | €9.27 | €6.06 |
| US domestic | $7.11 | $4.64 |
| US roaming | $78.94 | $51.56 |

**Table 2.** Average monthly costs of web advertisements

In our analysis it is quietly assumed that the websites downloaded on PCs, PDAs, tablet PCs and mobile phones are comparable. In reality there may be differences, but we do not have access to data to verify this. It is known, however, that several mobile operators modify content for mobile devices, for example to reduce the quality and resolution of JPEG pictures. It is unlikely, however, that such operators would distinguish between normal data and web advertisements. As such, the percentage of traffic generated because of advertisements may not be effected.

## 6    Conclusions

The goal of this paper was to investigate whether advertisements take up a significant part of the data traffic produced while browsing on mobile devices, and whether these advertisements lead to extra costs that can not be neglected.

Section 3 of this paper presented an approach to measure the average amount of network traffic generated by web advertisements. In this approach the data from hundred popular webpages was automatically retrieved via a proxy server; in the first run the proxy did not filter any advertisement, but in the second run the proxy did. By calculating the difference between both runs, the average amount of traffic generated by advertisements was found.

Our results, which were presented in Section 4, show that the amount of network traffic generated by web advertisements will be more than 100 KB per page if Flash is enabled, and around 70 KB without Flash. These numbers vary over time, based on the actual advertisement being offered by the website. Web advertisements therefore increase the total amount of network traffic with roughly 9% if Flash is enabled, and 7% without Flash.

In Section 5 these numbers were used to calculate the additional costs that mobile users would experience because of advertisements. In the US, average users without any data plan would be charged an additional \$4.5 to \$7 when browsing domestically, and between \$50 and \$80 when roaming. In the Netherlands these additional costs would be, without any data plan, between €3 and €5 when browsing domestically and about €6 to €9 when roaming. These costs need to be viewed as rough estimates, however, for mobile users not using any data plan.

Because the share of web advertisements in the total data traffic produced appears to be significant, it would be interesting to use this paper as motivation for further research in this area. Such research could focus, for example, on the distribution function of advertisement sizes, their variance, heavy-tailed characteristics, etc. Also further research is needed on the browsing behavior of mobile users.

## References

1. Eurostat.   Individuals  –  Mobile  Internet  access  (October  2010),  `http://epp.eurostat.ec.europa.eu/portal/page/portal/information_society/data/database`
2. Tweakers.net: providers draaien de klok terug (March 2012), `http://tweakers.net/reviews/1899/2/2010-providers-draaien-de-klok-terug-datalimieten.html`
3. van den Brande, J.: Mobile Internet costs of web advertisements. In: Proceedings of the fourteenth Twente Student Conference on Information Technology (2011)
4. Braden, R., Clark, D., Shenker, D.: Integrated Services in the Internet Architecture: an Overview. In RFC 1633 (June 1994)
5. Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., Weiss, W.: An Architecture for Differentiated Services. In RFC 2475 (Dec. 1998)

6. Pras, A., B.J. van Beijnum, van Beijnum, B. J., Sprenkels, R. A. M., Parhonyi, R.: Internet Accounting. In: IEEE Communications Magazine, Vol. 39, No. 5, 108 – 113 (May 2001)
7. Sprenkels, R. A. M., Pras, A., van Beijnum, B. J., de Goede, B. L.: A customer service management architecture for the Internet. In: Proceedings of the 11th IFIP/IEEE International Workshop on Distributed Systems: Operations and Management (DSOM), 71 – 82 (2000)
8. Sprenkels, R. A. M., Parhonyi, R., Pras, A., van Beijnum, B. J., de Goede, B. L.: An architecture for reverse charging in the Internet. In: Proceedings of the IEEE Workshop on IP-oriented Operations and Management (IPOM), 87 – 92 (2000)
9. Odlyzko, A.M.: Internet Pricing and the History of Communications. In: Computer Networks, Vol. 36, No. 5-6, 493 – 517 (2001)
10. Racz, P., Forster, V., Stiller, B.: Design and implementation of an integrated accounting architecture for distributed UMTS and WLAN networks. In: International Journal of Wireless and Mobile Computing, Vol. 2, No. 4, p.275 – 287 (Jan 2007)
11. Koutsopoulou, M., Kaloxylos, A., Alonistioti, A., Kawamura, K., Merakos, L.: Charging, Accounting and Billing Management in Mobile Telecommunication Networks and the Internet. In: IEEE Communications Surveys, Vol. 6, No.1 (2004)
12. Foell, U., Fan, C., Carle, G., Dressler, F., Roshandel, M.: Service-Oriented Accounting and Charging for 3G and B3G Mobile Environments. In: Proceedings of 9th IFIP/IEEE International Symposium on Integrated Network Management (2005)
13. Roto, V., Geisler, R., Kaikkonen, A., Popescu, A., Vartiainen, E.: Data Traffic Costs and Mobile Browsing User Experience. In: 4th MobEA Workshop on Empowering the Mobile Web, pp. 1–6. Department of Computer Science and Engineering Helsinki University of Technology, Edinburgh, Scotland (2006)
14. Simons, R.J.G., Pras, A.: The Hidden Energy Cost of Web Advertising. In: Proceedings of the twelfth Twente Student Conference on Information Technology (2010)
15. Palant, W.: Adblock Plus - for annoyance-free web surfing (March 2012), `http://adblockplus.org`
16. Digg. What is Digg? (March 2012), `http://about.digg.com/about`
17. Weinreich, H., Obendorf, H., Herder, E., Mayer, M.: Not quite the average: An empirical study of Web use. In: ACM Trans. Web 2(1), 1–31 (2008)
18. Boucher, A.: Proximodo (March 2012), `http://proximodo.sourceforge.net`
19. Privoxy. Privoxy (March 2012), `http://www.privoxy.org`
20. Apache Foundation. Apache Module mod_proxy (March 2012), `http://httpd.apache.org/docs/2.2/mod/mod_proxy.html`
21. KPN. Mobiel Internet abonnement voor je tablet of laptop (June 2012), `http://www.kpn.com/prive/mobiel/mobielinternet/abonnementen.htm`
22. Vodafone. Tarievenoverzicht Vodafone Prepaid (June 2012), `http://www.vodafone.nl/Vodafone/wg30/pdf/tarieven/tarievenoverzicht_prepaid_februari_2012.pdf`
23. Telfort. Productinformatie & Tarieven (June 2012), `https://www.telfort.nl/media/tarieven/2011/Tarieven_Telfort_Consument_mobiel_internet_1_Jul2011.pdf`
24. T-Mobile. Prijslijst T-Mobile Mobiel Internet (June 2012), `http://www.t-mobile.nl/persoonlijk/media/pdf/service-contact/nieuwe-prijslijsten/mobiel-internet.pdf`
25. Verizon. Data Connect to the things that matter. (June 2012), `http://www.verizonwireless.com/b2c/mobilebroadband/?page=plans`

26. AT&T. AT&T Data Plans (June 2012), `http://www.wireless.att.com/cell-phone-service/cell-phone-plans/data-connect-plans.jsp`
27. Sprint. Plans - Mobile Broadband (June 2012), `http://shop.sprint.com/mysprint/shop/plan/plan_wall.jsp?tabId=pt_data_plans_tab&flow=AAL&planFamilyType=null`
28. T-Mobile USA. Prepaid plans (June 2012), `http://www.t-mobile.com/shop/plans/prepaid-plans.aspx?WT.z_shop=Prepaid`
29. Opera Software ASA. State of the Mobile Web, March 2011 (June 2012), `http://www.opera.com/smw/2011/03/`
30. European Parliament. Industry committee sets out to lower mobile roaming rates (March 2012), `http://www.europarl.europa.eu/news/en/pressroom/content/20120227IPR39334/html/Industry-Committee-backs-cheaper-mobile-roaming-rates`
31. Infante, J., Vallejo, I.: Regulation of international roaming in the European Union - lessons learned. In: 39th Research Conference on Communication, Information and Internet Policy (2011)
32. Opera Software ASA. Opera use and user behavior white paper (2011), `http://media.opera.com/media/b2b/Opera_Mini_use_and_user_behavior_white_paper.pdf`