On the Use of TCP Interruptions to Assess User Experience on Web

Junaid Shaikh, Markus Fiedler, Patrik Arlos
School of Computing
Blekinge Institute of Technology
Karlskrona, Sweden
Email: {junaid, markus.fiedler, patrik.arlos}@bth.se

Denis Collange
Orange Labs Networks and Carriers
Sophia Antipolis, France
Email: denis.collange@orange-ftgroup.com

I. INTRODUCTION

The Quality of Experience (QoE) is becoming an increasingly popular research area among the ICT related industry and academia. While the major work was being done on the Quality of Service (QoS), the past decade has witnessed a vacuum in the area of QoE because QoE is rather a new notion. This is quite evident from the inconsistent definitions of QoE that we find in the literature. However, the area is now gaining more attention and maturity as there is increasing realization that 'the better the QoS' doesn't necessarily mean 'the more happy the users'. It is possible to have excellent QoS but poor user experience [1]. For service providers to compete in the market, they need to give more attention to how user perceives the service than just working on QoS. This further supports the argument that there is more than just technology-centric approach when it comes to satisfaction of users. We often observe that new applications and services that claim additional performance are rejected by the users. Consequently, there is a need of a user-centric approach for the design and monitoring of network applications and services.

It stimulates the need for an online passive mechanism to monitor the user activity in real-time. The mechanism may vary with the type of application. In this work, we keep our focus on Web browsing application as it is one of the most popular applications on the Internet. The aim is to let the users browse freely and monitor their activity instead of asking them subjectively about their usage experience. Monitoring of the TCP connection interruptions could be one of the ways to have indications about users’ feelings on the Web. This is based on the idea that, in case of bad performance, users break a Web browsing session by pressing a reset or stop button in the browser hence generating a TCP reset on TCP flow level. However, several other causes may explain the generation of TCP resets as described in [2]. In this work, we apply the TCP connection interruption criterion presented in [3] to filter the TCP resets generated by the users. We apply this criterion on the Web traffic of an operational mobile network.

In our previous work [4], we correlated the user session volumes with the network performance i.e. throughput and packet loss. We found that the volumes increase with the increasing throughput implying that the happy users surf more in the case of better performance. In this present work, we carry this analysis further to observe the user interruptions in relation to transfer sizes and durations. We want to see how the users react to bad performance and try to explain why we observe smaller mean transfer sizes. Do the users avoid launching large transfers in the case of bad performance or do they try to get the same files but stop the transfers that are becoming too long?

II. MEASUREMENT SETUP

Our study is based on a traffic trace collected on a mobile access networks, more precisely on a link between a GGSN (Gateway GPRS Support Node) and the router connecting it to the Internet as represented in Figure 1. Table I summarizes the main characteristics of the trace. We focus our work on HTTP connections, filtered using Wireshark [5] to detect HTTP headers. So all the connections we consider are HTTP connections, but we may not consider all of them.

We use the heuristic described in [3] to detect interrupted TCP connections. The actual output of this algorithm classifies the TCP connections in four groups, they may be: unfinished, packet loss. We found that the volumes increase with the increasing throughput implying that the happy users surf more in the case of better performance. In this present work, we carry this analysis further to observe the user interruptions in relation to transfer sizes and durations. We want to see how the users react to bad performance and try to explain why we observe smaller mean transfer sizes. Do the users avoid launching large transfers in the case of bad performance or do they try to get the same files but stop the transfers that are becoming too long?

II. MEASUREMENT SETUP

Our study is based on a traffic trace collected on a mobile access networks, more precisely on a link between a GGSN (Gateway GPRS Support Node) and the router connecting it to the Internet as represented in Figure 1. Table I summarizes the main characteristics of the trace. We focus our work on HTTP connections, filtered using Wireshark [5] to detect HTTP headers. So all the connections we consider are HTTP connections, but we may not consider all of them.

We use the heuristic described in [3] to detect interrupted TCP connections. The actual output of this algorithm classifies the TCP connections in four groups, they may be: unfinished,
normally finished, interrupted, or eligible (i.e. maybe interrupted). We consider in the next section the TCP connections classified as interrupted by this algorithm: their proportion and their traffic characteristics.

III. RESULTS AND ANALYSIS

The plots in figure 2 display the distribution of the interruption rates against the connection durations. These connections are differentiated according to their sizes. We observe different results for the different volumes of the connections. Interruption rates for the connections between 1 KB and 100 KB increase with the increasing durations. Surprisingly, we observe higher interruption rates for the biggest transfers at smaller download times.

Users are generally more sensitive to the performance of the short transfers e.g. a Web page containing some text and/or small image. They want these pages to be loaded within a certain time. So the probability of the interruptions of such transfers by the users grows with their increasing latencies. The observation of the higher interruption rate on biggest transfers could be explained by other reasons. Transfers larger than 100 KB could be different than simple text pages with small images. They may contain video for example and the user may lose his interest in this content after few seconds, or click on an other link.

We assume that users bothered by end-to-end performance are more likely to "interrupt" their transfers. However we must keep in mind that:
- The algorithm we use to detect interrupted TCP connections is only a heuristic, based on TCP flags exchanged in the last packets of the connection.
- Especially, the user may have different reactions in case of bad performance. For example, concerning the Web application, the user may stop the transfer, reload it, kill the browser... We have observed on active tests with various browsers that these different feedbacks appear differently depending on the implementation and they are not systematically considered as interruptions according to the heuristic. We have also observed during these same experiments that video transfers between IIS (the Microsoft Web server Internet Information Services) and Internet Explorer (the Microsoft Web client) appear systematically interrupted even without any action of the user.
- The end-to-end network performance is not the only possible cause that can lead a user to abort a transfer. There can be several contextual reasons behind it. For example, the user may stop a transfer if the content appearing with the first packets is less interesting than he expects. This may explain our observations about the large transfers.
- The user is not the only responsible for the interruption of a TCP connection. It may be caused by the application on the client or on the server side. It may be also be the usual type of end between some servers and some clients.
For all these reasons, we cannot expect more than statistical correlation between the probability of interruption and the bad performance observed by a given transfer.

IV. CONCLUSION

We have observed a clear positive correlation between the interruption rate and the duration for small transfers. But this correlation becomes negative for the biggest transfers. Thus, the use of the interruption rate to detect the user satisfaction need further work. For example, the type of information transferred should be taken into account. The interruption criterion should also be refined as it appears on the active tests. For all these reasons, we cannot expect more than statistical correlation between the probability of interruption and the bad performance observed by a given transfer. The objective of our analysis is then to find the end-to-end performance criteria showing the highest correlation factor with the interruption probability.

ACKNOWLEDGMENT

This work has been partly sponsored by the European 7th Framework Networks of Excellence EuroNF (216366), the Swedish Agency for Innovation Systems VINNOOVA (d-nr 2007/02505), and the COST project TMA (IC0703).

REFERENCES