

Collecting Subjective Ratings in Enterprise Environments

Kathrin Borchert*, Matthias Hirth*, Thomas Zinner*, Anja S. Göritz†

*Institute of Computer Science, University of Würzburg,

Am Hubland, 97074 Würzburg, Germany

Email: [kathrin.borchert|matthias.hirth|zinner]@informatik.uni-wuerzburg.de

†Occupational and Consumer Psychology, Freiburg University, Freiburg, Germany

Email: goeritz@psychologie.uni-freiburg.de

Abstract—Similar to many modern applications, enterprise applications like SAP are often implemented in a distributed fashion and consequently suffer from network degradations resulting in impairments like increased loading delays. While the influence of these impairments on the perceived quality of users is well researched for consumer applications and network services, their impact in a business environment is still unclear. To address this gap we develop a non-intrusive software tool for continuously collecting subjective ratings on the performance of an enterprise application from a large number of employees. Based on the feedback from two field studies in a company we briefly discuss challenges of QoE monitoring in the context of enterprises. As a first step towards building QoE models, we combine the subjective ratings with technical monitoring data and observe a negative correlation between the user satisfaction and the overall load of the server infrastructure.

I. INTRODUCTION

Employees in several business areas have to work with network based applications and services to fulfil their day-to-day work. These applications run on top of distributed systems, like thin client architectures. Thus, the application performance may suffer from congestion and performance issues in the network as well as in the data center. This leads to noticeable delays and other impairments at the user side. The decreased application performance may influence the perceived application quality of the employees.

To build an objective model based on the technical parameters of an application, the impact of these parameters on the quality of experience (QoE) of the users has to be investigated. However, several challenges arise when performing QoE measurements in enterprise environments. This paper tackles these challenges by designing and implementing a non-intrusive survey tool for enterprise environments and demonstrating its feasibility. The tool for assessing the perceived quality of enterprise applications considers key requirements deduced in cooperation with a large company. In two large pilot studies the applicability of the approach is demonstrated.

The remainder of this work is structured as followed. In Section II we briefly discuss related work. Section III gives a short overview over the requirements of QoE monitoring in enterprise environments and their realization in the developed

tool. Section IV presents results of the user studies, and Section V concludes with an outlook.

II. RELATED WORK

Numerous studies focus on the identification and quantization of factors influencing a customer's or end-user's satisfaction with the performance of applications and services. One of the main influence factors in network based applications and services is the transmission delay and the resulting application behaviour. Several studies investigate, e.g., the influence of loading delays [1] on the QoE of web site users. A first indication that the QoE is also affected by a delayed application performance in the context of business software is given by Bonhag et al. [2].

One approach to assess the perceived system quality in an enterprise is using the feedback and information given by employees via existing communication channels, e.g., messages of a ticketing system [3]. Approaches that result in more fine grained data usually involve active user feedback during or immediately after using the service or applications, e.g., Smith et al. [4]. While this method is feasible for a software that is only used from time to time, it cannot be applied for applications that are used throughout the whole workday, like e.g. SAP systems.

III. QOE MONITORING REQUIREMENTS AND REALIZATION

Several new challenges arise while conducting QoE studies in an enterprise environment. Those include for example the minimization of costs and the seamless integration into the day-to-day business. Moreover, the tool for collecting the subjective feedback needs to consider security aspects and technology constraints in the production system of the company. Despite these requirements and design limitations, the assessment methodology needs to be scientifically valid and stick as close as possible to established standards for subjective evaluations. Additionally, the tool design should follow common best practices, e.g., no implicit or explicit incentives for the employee to give a certain rating.

Based on these requirements, we designed and implemented a simple tool for monitoring the QoE of employees using active feedback. The monitoring is realized as a short survey

which is shown via a pop-up to the participating employees. First, the user assesses the current system performance by clicking on a green or a red smiley. This opens automatically the second step. Here, the user may explain his rating by selecting one out of several predefined reasons. Each predefined set of reasons is optimized for the requirements of the participants. The survey is automatically completed when a reason is selected.

To prevent the interruption of critical working processes or conversations with customers the pop-up closes automatically after a few seconds if the user does not react. These ratings are marked as missing. If the user does not select a reason after rating the performance, the pop-up is also closed after a specific amount of time. In this case only the reason is marked as missing.

The pop-up opens automatically once an hour, while the interval between two pop-ups varies between 15 min and 119 min. This prevents the effect that the participants 'expect' the pop-up and prepare themselves to rate the performance. For the survey conduction we applied a communication concept which includes the invitation of the participants, the provision of instructions and background information as well as the collection of feedback.

IV. USER STUDIES

We evaluate the applicability of our tool by two user studies in cooperation with a company with more than 15000 employees. The feedback has been collected during two working weeks in study A from 618 participants in December 2015 and in study B from 723 employees in January 2016. All of the participants work exclusively or most of their time with an SAP system. In total the survey has been shown in study A 33225 times with 16339 ratings marked as missing and in study B 47113 answers have been collected with 23525 marked as missing. This indicates that it was not always possible for the employees to answer the survey at any time during the daily working process. Despite missing answers, about 97% of the participants submitted at least one survey during the total survey period.

In order to evaluate the additional effort imposed to employees we investigate the overall response duration for study A. Missing ratings and ratings without reasons are omitted. For missing ratings no response times can be measured, as the user did not interact with the pop-up at all. Ratings without reasons can only occur due to software issues or if a user quits during the answering process. Thus, in this case the observed response times are not meaningful. The median response time for the remaining pop-ups is 6 seconds. Due to outliers with a response time of 132 seconds, the mean response time is significantly higher with 9.6 seconds. Further, there are no significant differences between rating times for positive ratings and the rating times for negative ratings. The importance of a time efficient assessment process can easily be demonstrated by considering the total time t_t spent on the submissions. Summing up of all response times results in $t_t \approx 38$ hours.

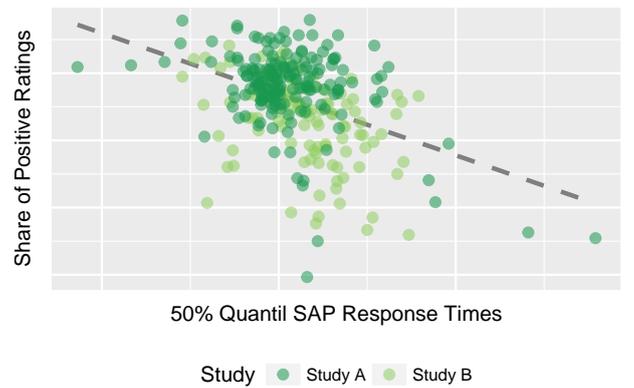


Fig. 1. Correlation of ratings and hourly aggregated response times.

As a first analysis toward building an objective model, we correlate the ratings with technical data from an internal monitoring systems of the enterprise. To correlate the data on an hourly base, the response times of the SAP transactions are aggregated per hour as a simple indicator for the overall system load. Figure 1 shows the share of positive ratings against the median of the system response time per hour. The scales on the x and y axis are intentionally blanked, as the concrete values are restricted to company internal use only. Despite the varying time interval between two ratings and the aggregation of the technical data, the results show a negative correlation of -0.42 . Similar results can be observed for the 25% and the 75% quantile of the system response time.

V. CONCLUSION

In this work we briefly discussed the specific requirements for QoE monitoring in an enterprise environment. We introduced a survey tool for rating the performance of enterprise applications which considers these requirements. The practicality of the tool and its seamlessly integration into the day-to-day work is evaluated by two large user studies with hundreds of participants. The results show that technical monitoring parameters and subjective employee ratings are connected. However, more and more fine grained data is required to construct QoE models for specific enterprise applications. This can be achieved with a long time deployment of the designed measurement tool.

Nevertheless, some constraints of the survey tool have to be taken into account, e.g., the varying time interval between two ratings may result in a varying amount of interactions with the applications and thus with the technical system. However, some of these effects can be mitigated, e.g., by weighting response times of the transactions depending on the distance to the user ratings [5] or by enabling push-based user ratings.

ACKNOWLEDGEMENT

The authors express their gratitude to André Rentsch, Peter Karg and Harald Wehnes for the support during the course of this work. This work is supported by Deutsche Forschungsgemeinschaft (DFG) under Grants HO TR 257/41-1 and kubus IT. The authors alone are responsible for the content.

REFERENCES

- [1] S. Egger, T. Hoßfeld, R. Schatz, and M. Fiedler, "Tutorial: Waiting Times in Quality of Experience for Web based Services," 2012.
- [2] W. Bonhag, D. Feindt, S. Olschner, and U. Schubert, "Wie schnell ist schnell bei Business-Software? Analyse zur Performance bei der Nutzung von Business-Software," *Mensch und Computer 2015–Usability Professionals*, 2015.
- [3] T. Zinner, F. Lemmerich, S. Schwarzmann, M. Hirth, P. Karg, and A. Hotho, "Text categorization for deriving the application quality in enterprises using ticketing systems," in *International Conference on Big Data Analytics and Knowledge Discovery*. Springer, 2015, pp. 325–336.
- [4] R. Smith and L. A. Kilty, "Crowdsourcing and gamification of enterprise meeting software quality," in *Utility and Cloud Computing (UCC), 2014 IEEE/ACM 7th International Conference on*. IEEE, 2014, pp. 611–613.
- [5] K. Borchert, M. Hirth, T. Zinner, and D. C. Mocanu, "Correlating qoe and technical parameters of an sap system in an enterprise environment," in *Teletraffic Congress (ITC 28), 2016 28th International*, vol. 3. IEEE, 2016, pp. 34–36.