Estimating the Driver's Workload

Using Smartphone Data to Adapt In-Vehicle Information Systems

Christina Ohm and Bernd Ludwig

Chair of Information Science, Universität Regensburg, 93040 Regensburg, Germany
{christina.ohm,bernd.ludwig}@ur.de

Abstract. The use of in-vehicle information systems has increased in the past years. These systems assist the user but can as well cause additional cognitive load. The study presented in this paper was carried out to enable workload estimation in order to adapt information and entertainment systems so that an optimal driver performance and user experience is ensured. For this purpose smartphone sensor data, situational factors and basic user characteristics are taken into account. The study revealed that the driving situation, the gender of the user and the frequency of driving significantly influence the user's workload. Using only this information and smartphone sensor data the current workload of the driver can be estimated with 86% accuracy.

Keywords: Driver's Workload, Workload Estimation, In-Vehicle Information Systems.

1 Motivation

Currently, many in-vehicle information systems provide assistance and entertainment to the driver but still can be a source of distraction and cognitive load since they require an input action or at least the attention of the user. For this reason it is necessary to determine the driver's mental state in order to prevent dangerous situations and optimize her driving performance. The primary driving task itself is cognitive demanding since the motorist has to develop different levels of skills [20]:

– Control level skills: handling of the car
– Maneuvering level skills: reaction behavior to the traffic situation and other road users
– Strategic level skills: planning the trip and defining goals

In addition to these basic activities, the operator has to fulfill secondary tasks which are not directly related to the actual transportation goal but are required to ensure safety. This could be e.g. turning on the windscreen wiper or the upper beam head-lights. Using in-vehicle information systems belongs to the tertiary driving task as they are mainly used for information, communication

I.J. Timm and M. Thimm (Eds.): KI 2013, LNAI 8077, pp. 130–139, 2013.

or entertainment purposes [3]. Thus, driving is a particularly difficult situation for human computer interaction because the operator is supposed to master primary and secondary tasks before she can actually pay attention to any kind of in-vehicle information system. If the user's current workload is not optimal, she should not be bothered with even more output of an assistance or entertainment system.

In this paper a study is presented that examines which factors significantly influence the driver's workload. Furthermore, the aim was to estimate the workload using sensor data collected by a smartphone. In addition, basic characteristics of the driver are taken into account. Being able to forecast the mental load of the user in a specific situation can be used to adapt a system's output to facilitate an optimal driving performance and therefore enhance the driver's safety. In contrast to existing approaches [8, 19, 32] the concept of using just a smartphone does not depend on sensor data of the car. Consequently, this technique provides a low-cost and simple possibility to integrate the prediction into smartphone services which could be for instance phone calls that can be blocked for a short period of time if the workload is high at the moment. This is also the intent of other systems which try to predict workload [8, 19]. The following section describes the concept of workload generally as well as in the context of driving. Afterwards measurement and estimation methods are depicted. Subsequently, we present our workload self-assessment tool, the study design and our results. Finally, we draw conclusions considering the implications of our findings.

2 Workload

Basically workload can be defined as a "demand placed upon humans" [29]. However, this task-oriented definition does not mind all aspects of the concept. Especially user-centered features such as her cognitive abilities or knowledge base are not minded. This is taken into account by different definitions [4, 7, 12, 15, 16, 29], and can be summed up as follows: An operator has to fulfill a task under given situational circumstances which demands information-processing capabilities. Therefore, the task causes cognitive load. Workload is the required capability to accomplish the task satisfactorily and achieve a certain level of performance. Meanwhile, the operator has a subjective experience which results from situational factors, the requirements of the task and the user's characteristics [24]. The latter includes criteria like her cognitive, sensory or motor skills, knowledge base, behavior, personality, age, sex etc. [14, 29]. Considering the context of driving this may also include driving experience [10]. [28] identifies several important factors that influence the driver's workload which are the age of the driver, the driving context (e.g. driving on a rural road) or the traveling daytime.

In the context of driving [21] distinguish three states which derive from the relation of the situational demands and the driver's condition. If the requirements of the task exceed the skills of the user, she has to endure an Overload experience which is a suboptimal cognitive state and can lead to loss of control or at least the feeling to do so. Contrary to this is the Underload state, which can be understood as a feeling of boredom or distraction. In this case the situational demands do not challenge the driver because she has sufficient skills. Thus, if the factors are in balance, the driver achieves optimal performance. This approach to classify driver experience is quite similar to the flow theory discussed in [6]. [9] develops a Task-Capability Interface Model which also examines the connection of the task demands and the capability of the driver. The task demands include the traffic situation, the behavior of other road users or the actual traveling speed. If the skills of the driver exceed the demands, the user is in a state of control. If this is not the case, most of all situations result in a lucky escape because other road users react to an error of the operator. Otherwise, a collision impends which indicates the importance of knowing the current driver state.

A more detailed categorization can be found in [29]. Once again, in this model workload is the consequence of the relation of the driver's performance and the task demands. The author distinguishes six different states which include one state where the workload is high and the performance is low because the driving demands are not challenging. Moreover, three states can be grouped as situations where the driver is in control but either is at the edge to Under- or Overload. In the remaining two states the driver has an experience of high workload or even loss of control as the task demands are received as too high. Both Under- and Overload can lead to cognitive load which affects the health of the user and his ability to drive safely and therefore has to be avoided.

3 Workload Measurement and Estimation

At first cognitive load has to be measured in order to enable an estimation and consequently avoid a suboptimal workload level. This can be achieved using a diverse range of available tools. It is possible to measure the operator's physiological state using for instance an electroencephalogram to determine cerebral activity or an electrocardiogram to measure myocardial contraction. An overview of physiological measurements is given in [29]. However, data collected with biosensors is considered to be arguable, since it is sometimes ambiguous and hard to analyze for someone who is not an expert in medicine. Moreover, external factors, which do not indicate workload, can influence the measurement [5].

Another possibility is to use tools that relay on user self-reports which can be one- or multidimensional. An example for an one-dimensional measurement is the Rating Scale Mental Effort (RSME) [33]. The test person is asked to indicate on a 0 to 150 cm line how much effort it took to fulfill the task. Several statements like "rather much effort" are located along the scale, which are supposed to clarify the level of current effort.

One of the most established multidimensional method in this context is the NASA-Task Load Index (TLX) [4] which takes different sources of workload into account. These are in particular mental, physical and temporal demand as well as the subjectively experienced performance, effort and frustration level of the user. The test person has to rate these dimensions pairwise referring to their impact on the task's workload which leads to overall 15 comparisons. Subsequently, a weight is calculated for every dimension. Afterwards, every dimension is additionally rated on a 0 to 20 scale. The Overall Workload Index (OWI) is calculated as follows with w_i as the weight and x_i as the rate multiplied by 5 [14]:

$$\mathrm{OWI} = \frac{1}{15}\sum_{i=1}^{6} w_i x_i$$

An overview of other multidimensional scales is for example given in [4]. In order to adapt in-vehicle information systems to the current workload of the driver, her state has not only to be measured but also estimated. Different approaches use car sensor data like the steering wheel angle or acceleration gauging to forecast the operator's load [8, 31, 32].[32] additionally use eye tracking techniques and reach up to a 65% correct driver-independent workload estimation with test data collected in a simulator. [8] take biosensors like the heart rate and environmental factors like traffic density into account.

4 Study

The overall goal of the study is to adapt in-vehicle information systems to the current workload of the driver. At first, several hypotheses were tested for validity in order to identify situational factors and user characteristics which significantly influence the driver's workload. [11, 13] show that every driving situation acquires diverse driving skills, so that it can be assumed that the level of workload differs according to this factor. Furthermore, [25] claim that women experience higher workload levels than men while driving. Accordingly, the following hypotheses were proposed:

– H_1: The workload during the driving situations "freeway", "rural road" and "city" differs.
– H_2: Women and men experience different workload while driving.

Since the test persons do not differ significantly in driving style (see Section 4.1), which is an important influencing factor on workload according to [30], other user characteristics were taken into account. [1, 30] distinguish young drivers (<25 years) from older ones and identify that these persons are likely to experience higher workload levels. This leads to the third hypothesis:

– H_3: The workload level differs considering the user's age.

[21] assume that persons who drive regularly usually experience lower workload levels so that the following hypothesis is as well tested for validity:

– H_4: The workload level differs considering the user's driving frequency.

Additionally, smartphone sensor data was collected to estimate the driver's workload.

4.1 Participants and Test Route

Eight female and twelve male students participated in the study. A detailed overview of the user characteristics is showed in Table 1.

Table 1. Overview of user characteristics separated by gender and driving frequency (DF)

	Age			Driving Experience (in years)		
	Mean	Standard deviation	Range	Mean	Standard deviation	Range
Total	24.9	2.1	21-28	6.9	2.4	1-10
Females	24.3	2.3	21-28	5.9	3.0	1-10
Males	25.3	1.9	22-28	7.5	1.8	5-10
DF "often"	25.7	2.3	21-28	7.8	2.2	5-10
DF "rarely"	24.0	1.4	22-26	5.3	2.5	1-8

All in all, the entire test group consisted of young and rather inexperienced but no complete novice drivers. Their driving style was additionally assessed by means of the Multidimensional Driving Style Inventory [26]. Most of the participants named to have a patient or careful driving style so that the test group can be considered as homogeneous concerning this factor.

None of the participants drove with the test car before. The test route consisted of three sections. After a familiarization with the test automobile (6.3 km) users drove on a rural road for 6.3 km. Afterwards they drove on a freeway section for 12 km and finally in the city of Regensburg (4.9 km). It took about 15 minutes to complete one section of the test route. None in-vehicle information systems were used as the main aim was to detect cognitive load of the primary and secondary driving task to adapt the output of assistance and entertainment systems. Moreover, there are several studies which prove that in-vehicle information systems and especially the usage of cell phones cause cognitive load [16–18, 22, 23, 27].

4.2 Measurements

The subjective workload level after accomplishing every route section was measured using a smartphone-based representation of the NASA-TLX (Fig.1). The form was filled in while parking. In addition, a self-assessment tool was implemented which enables users to rate their current workload during the driving process (Fig.1). The tool is based on the RSME. The Underload state adapted from [29] was added and the scale was simplified due to space limitations of the smartphone screen. Several potential designs could be used to visualize the workload levels. Four prototypes with different layouts were evaluated in a usability test with 10 participants. They were instructed that the application was intended to be used while driving. Qualitative as well as quantitative data was collected to determine the best design. Participants were observed by the test supervisor and asked to "think aloud". In addition they filled in the System Usability Scale (SUS) [2]. Most of the test persons agreed in one best design. This was also shown using a single factor variance analysis with repeated measurement adjusted according to Bonferroni for the SUS ($p < 0.05$). This resulted in the design showed in Fig.1. Colors and a scale are used to visualize the workload level and the whole screen can be clicked. During the actual test, participants were asked to indicate their current workload level every time it had changed. The smartphone was adjusted to the front screen so that only one short look and click was enough to indicate the workload level.

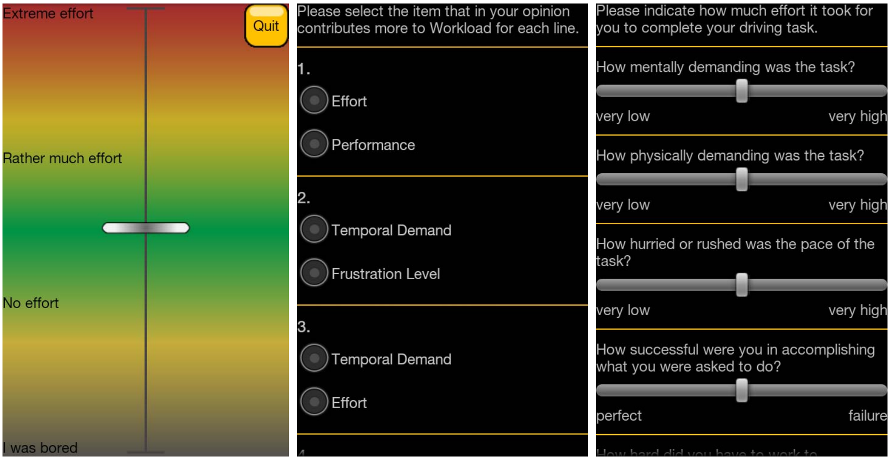

Fig. 1. Screenshots of the Android application which was used to rate the participants' workload while driving (left) and representation of the NASA-TLX (middle and right).

In addition to this, smartphone sensor data was collected to estimate the user's workload with data mining approaches. The lateral and longitudinal acceleration as well as the current speed is detected ten times per second, whereas the current workload level is assigned to this data. In a pre-test highly significant correlations with the equivalent car sensor data were measured ($r > 0.85$; $p = 0.000$). In addition, the lateral acceleration correlates highly significantly with the angle of lock ($r = 0.95$; $p = 0.000$). The sensor data described above is considered to have a significant influence on driver workload [8, 32]. Moreover, user data like gender and frequency of driving was assessed.

4.3 Results

H_1 could be confirmed considering the OWI for the different driving situations "rural road", "city" and "freeway" using a single factor variance analysis with repeated measurement. Since no sphericity could be assumed ($p < 0.1$), Greenhouse-Geisser results were consulted ($F = 6.178$; $p = 0.009$). Conducting a Bonferroni post-hoc test, more detailed findings could be gained: The OWI is significantly higher for "freeway" ($p = 0.039$) and "city" ($p = 0.013$) compared to the "rural road" situation.

After confirming normality using a Kolmogorov-Smirnov-Test ($p < 0.05$) and analyzing the histograms H_2-H_4 were tested.

H_2 could be confirmed using a T-test ($T = 2.314$; $p = 0.024$). Generally women experience slightly higher workload levels than men so that gender can be considered as an influencing factor on workload.

Age significantly correlates with driving experience in years ($r = 0.913$; $p = 0.000$), so that older test persons had higher experience levels. However, no significant difference could be detected for this factor.

No significant correlation between "gender" and "driving frequency" could be confirmed using Fisher's exact test.

The OWI of participants who named to drive rarely is significantly higher than of those who at least drive once a week ($T = 2.173$; $p = 0.037$) so that H_4 could be confirmed.

According to these findings, the driving situation, gender and driving frequency were taken into account for the estimation of workload in addition to the smartphone sensor data. On the one hand workload was categorized in the three states "Underload", "Optimal" and "Overload". On the other hand the workload levels were as well classified similar to [29] except the level of extreme Overload which results in five different states.

Table 2. Estimation accuracy in % for different classifications of workload

	Decision Tree (C4.5)	Sequential Minimal Optimization (SMO)	AdaBoost	Naive Bayes	Neural Network
Three states	85.70	63.3	62.30	63.92	70.51
Five states	72.92	46.27	46.13	46.13	54.81

Different classification algorithms were taken into account using an 80/20 split of the sample data (Table 2). The results show that the decision tree performs significantly better than the other techniques ($p < 0.05$).

The study revealed that the current cognitive load can be estimated with an accuracy about 86% using a decision tree. Operator-specific forecasts reach up to 96% correct predictions. If only sensor data is used, the estimation shows up to 76% accuracy. Moreover, the ROC-values of the decision tree method exceed 0.9 so that a good diagnostic accurateness can be assumed.

For a more detailed classification of the current workload level the estimation accuracy decreases to 73% with a ROC-value of 0.749.

5 Conclusion

All in all, there are many factors which influence the driver's workload since the driving situation and the characteristics of the user are very multifaceted.

However, the study presented in this paper shows that it is possible to estimate the driver's workload with very simple methods. Even if only smartphone-based data is used, cognitive load can be calculated with 76% accuracy. Moreover, if user data is taken into account, the prediction accuracy increases to 86%. Consequently, it would be very advantageous to shortly collect user characteristics, i.e. gender, driving frequency and age. In this study the latter could not be examined due to the rather homogeneous age of the participants. This should be a topic of future research.

Another improvement could be achieved through detecting the current driving situation like weather conditions or road type using e.g. data of the navigation system or the car sensors. Furthermore, taking into account whether in-vehicle information systems are used in the specific driving situation could improve the detection of the current workload level since several studies show that using this systems increases the cognitive load of the user (see above). User-dependent predictions can reach up to 96% accuracy so that a system which uses driver feedback to improve the estimation could minimize estimation errors.

The study also showed that it is important to determine the driver's workload several times per second since even if the OWI of some participants was rather low nearly everyone experienced very high or low workload for a short period of time while actually driving. Yet, it is just these moments which require workload detection to avoid dangerous situations.

Knowing the driver state can improve the user experience and safety if in-vehicle information systems are involved. As mentioned at the beginning phone calls or an output of a navigation system can be blocked for instance. If the workload level is too low music could be recommended to the user. Other application areas will surely follow.

References

1. Biermann, A., Eick, E.M., Brünken, R., Debus, G., Leutner, D.: Development and first evaluation of a prediction model for risk of offences and accident involvement among young drivers. Driver Behaviour and Training 2, 169–178 (2005)
2. Brooke, J.: SUS - A quick and dirty usability scale. Redhatch Consulting, United Kingdom (2011)
3. Bubb, H.: Fahrerassistenz primär ein beitrag zum komfort oder für die sicherheit? VDI-Berichte, pp. 25–44 (2003)
4. Cain, B.: A review of the mental workload literature. Tech. rep., DTIC Document (2007)
5. Cherri, C., Nodari, E., Toffetti, A.: Review of existing tools and methods. Tech. rep., AIDE Deliverable D2.1.1 (2004)
6. Csikszentmihalyi, M.: FLOW. Das Geheimnis des Glücks. Klett-Cotta, Stuttgart (2008)
7. Eggemeier, F., Wilson, G., Kramer, A., Damos, D.: Workload assessment in multi-task environments. In: Damos, D. (ed.) Multiple Task Performance, pp. 207–216. Taylor & Francis, London (1991)
8. Ford, https://media.ford.com/content/fordmedia/fna/us/en/news/2012/06/27/ford-research-developing-intelligent-system-to-help-drivers-mana.html

9. Fuller, R.: The task-capability interface model of the driving process. Recherche-Transports-Sécurité 66, 47–57 (2000)
10. Fuller, R.: Towards a general theory of driver behaviour. Accident Analysis & Prevention 37(3), 461–472 (2005)
11. Gaczek, D.: Entwurf und Regelung eines Verbrauchsassistenten. GRIN Verlag (2009)
12. Gopher, D., Donchin, E.: Workload - an examination of the concept. In: Boff, K., Kaufman, L., Thomas, J. (eds.) Handbook of Perception and Human Performance. Cognitive Processes and Performance, vol. 2, pp. . 41:1–41:49. Wiley, New York (1986)
13. Hale, A., Stoop, J., Hommels, J.: Human error models as predictors of accident scenarios for designers in road transport systems. Ergonomics 33(10-11), 1377–1387 (1990)
14. Hart, S.G., Staveland, L.E.: Development of nasa-tlx (task load index): Results of empirical and theoretical research. Human Mental Workload 1(3), 139–183 (1988)
15. Jex, H.R.: Measuring mental workload: Problems, progress, and promises. Advances in Psychology 52, 5–39 (1988)
16. Lysaght, R.J., Hill, S.G., Dick, A., Plamondon, B.D., Linton, P.M.: Operator workload: Comprehensive review and evaluation of operator workload methodologies. Tech. rep., DTIC Document (1989)
17. Ma, R., Kaber, D.B.: Situation awareness and workload in driving while using adaptive cruise control and a cell phone. International Journal of Industrial Ergonomics 35(10), 939–953 (2005)
18. Matthews, R., Legg, S., Charlton, S.: The effect of cell phone type on drivers subjective workload during concurrent driving and conversing. Accident Analysis & Prevention 35(4), 451–457 (2003)
19. Mayser, C., Ebersbach, D., Dietze, M., Lippold, C.: Fahrerassistenzsysteme zur unterstützung der längsregelung im ungebundenen verkehr. In: Conference Aktive Sicherheit durch Fahrerassistenz (2004)
20. Michon, J.A.: A critical view of driver behavior models: what do we know, what should we do? Springer (1986)
21. Oron-Gilad, T., Ronen, A., Shinar, D.: Alertness maintaining tasks (amts) while driving. Accident Analysis & Prevention 40(3), 851–860 (2008)
22. Pauzié, A.: Evaluating driver mental workload using the driving activity load index (dali). In: Proc. of European Conference on Human Interface Design for Intelligent Transport Systems, pp. 67–77 (2008)
23. Pauzié, A., Manzano, J.: Evaluation of driver mental workload facing new in-vehicle information and communication technology. In: Proceedings of the 20th Enhanced Safety of Vehicles Conference (ESV20), Lyon, France, vol. 10 (2007)
24. Recarte, M.A., Nunes, L.M.: Mental workload while driving: Effects on visual search, discrimination, and decision making. Journal of Experimental Psychology Applied 9(2), 119–133 (2003)
25. Schweitzer, J., Green, P.: Task acceptability and workload of driving city streets, rural roads, and expressways: Ratings from video clips (2007)
26. Taubman-Ben-Ari, O., Mikulincer, M., Gillath, O.: The multidimensional driving style inventory scale construct and validation. Accident Analysis & Prevention 36(3), 323–332 (2004)
27. Tsimhoni, O., Green, P.: Visual demand of driving and the execution of display-intensive in-vehicle tasks. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 45, pp. 1586–1590. SAGE Publications (2001)

28. Verwey, W.B.: On-line driver workload estimation. effects of road situation and age on secondary task measures. Ergonomics 43(2), 187–209 (2000)
29. de Waard, D.: The measurement of drivers' mental workload. Groningen University, Traffic Research Center (1996)
30. Wundersitz, L., Burns, N.: Identifying young driver subtypes: relationship to risky driving and crash involvement. Driver Behaviour And Training 2, 155 (2005)
31. Zeitlin, L.R.: Micromodel for objective estimation of driver mental workload from task data. Transportation Research Record: Journal of the Transportation Research Board 1631(1), 28–34 (1998)
32. Zhang, Y., Owechko, Y., Zhang, J.: Learning-based driver workload estimation. In: Prokhorov, D. (ed.) Computational Intelligence in Automotive Applications. SCI, vol. 132, pp. 1–17. Springer, Heidelberg (2008)
33. Zijstra, C., Doorn, R.V.: The construction of a scale to measure perceived effort. Tech. rep., Department of Philosophy and Social Sciences, Delft University of Technology (1985)