Mental Workload Management as a Tool in e-Learning Scenarios

André Pimenta¹, Sergio Gonçalves², Davide Carneiro¹, Florentino Fde-Riverola², José Neves¹ and Paulo Novais¹

¹Departament of Informatics, University of Minho, Braga, Portugal
²Informatics Department, University of Vigo, Ourense, Spain

Keywords: Mental Workload, Mental Fatigue, Machine Learning, e-Learning, Fatigue Management, Human Performance.

Abstract: In our daily life, we often have a sense of being exhausted due to mental or physical work, together with a feeling of performance degradation in the accomplishment of simple tasks. This is in part due to the fact that the working capacity and the performance of an individual, either physical or mental, generally decrease as the day progresses, although factors like motivation also play a significant role. These negative effects are especially significant when carrying out long or demanding tasks, as often happens in an educational context. In order to avoid these effects, initiatives to promote a good management of the time and effort invested in each task are mandatory. Such initiatives, when effective, can have a wide range of positive effects, including on the performance, productivity, attention and even mental health. Seeking to find a viable and realistic approach to address this problem, this paper presents a non-invasive and non-intrusive way to measure mental workload, one of the aspects that affects mental fatigue the most. Specifically, we target scenarios of e-learning, in which the professor may not be present to assess the student’s state. The aim is to create a tool that enables an actual management of fatigue in such environments and thus allows for the implementation of more efficient learning processes, adapted to the abilities and state of each student.

1 INTRODUCTION

In our day-to-day we often feel a sense of tiredness during mental or physical work, generally known as fatigue. The term fatigue is used to describe a series of manifestations that range from drowsiness or loss of concentration to lack of physical strength or agility (van der Linden et al., 2003). Thus, it is a very broad and subjective term that may include symptoms such as loss of performance (loss of attention, slowed reaction and response times, impaired decision-making, and poor performance on tasks that generally reflect the good performance) as well as more subjective ones such as sleepiness and tiredness (Williamson et al., 2005; Perelli, 1980).

In seeking to formalize it, fatigue may be defined as a degree of failure of physical or mental factors associated with loss of physical or mental performance, hindering the natural or spontaneous accomplishment of a usual activity. Bartlett provides one of the clearest definitions of fatigue with respect to day-to-day tasks (Bartlett, n.d.) :

Fatigue is a term used to cover all those determinable changes in the expression of an activity which can be traced to the continuing exercise of that activity under its normal operational conditions, and which can be shown to lead, either immediately or after delay, to deterioration in the expression of that activity, or, more simply, to results within the activity that are not wanted.

I.D. Brown, on the other hand, conceptualized mental fatigue as: (...) the subjective experience of individuals who are obliged (...) to continue working beyond the point at which they are confident of performing their task efficiently (...) [Fatigue is] the subjectively experienced disinclination to continue performing the task at hand. [The] main effect of fatigue [is] a progressive withdrawal of attention from the task at hand. [This withdrawal] may be sufficiently insidious that [operators] are unaware of their impaired state and hence in no position to remedy it (Brown, 1994).

Beyond the well-known effects on mood or energy, fatigue is also the cause or partial cause of errors and accidents. Often, this happens because fatigued individuals are unaware of the degree of their
impaired mental state (Miller, 2013). These errors and accidents assume particular importance when we consider the domain of high-risk jobs that involve operating vehicles as well as the military, firefighters or medical personal, just to name a few. Beyond these immediate problems, fatigue can also lead to health problems in the long term, such as chronic fatigue syndrome or depression.

The negative effects of fatigue are thus clear. Moreover, they are also broad in the sense that affect many of our cognitive abilities. Learning is one of the functions that becomes impaired when under fatigue. Hence the importance of addressing this issue, especially in a time in which the teacher and student are growing apart due to the increasing use of electronic tools for learning. Indeed, due to the separation imposed by technology, it results more and more difficult for teachers to be sensible to the state of their students, impairing their ability to adapt both the contents and the teaching strategy accordingly.

This work details a tool for fatigue management in e-learning scenarios, with or without the teacher’s presence, through the assessment of mental workload quantified in terms of the interaction patterns of the user with the computer. Through the use of behavioural biometrics, specifically Keystroke Dynamics and Mouse Dynamics, we analyse the type of task performed by each user, the time spent performing it, as well as the mental workload of the task. With this information we train classifiers that are able to distinguish situations in which users show signs of fatigue or high mental workload.

This approach can be considered both non-invasive and non-intrusive, since it is based solely on the observation of the use of the mouse and keyboard, which allow for an assessment of the user’s performance. This approach opens the door to the development of fatigue management initiatives in the context of e-learning, allowing teachers to not only have a better notion about their student’s state but also to more efficiently adapt and above all personalize teaching strategies.

1.1 The Need for Monitoring in the Context of e-Learning

Electronic instruction, more commonly known as e-learning, is increasingly used as a method of teaching. E-learning differs from classroom-based training in several ways. Thus, the transition from a traditional course to a course supported by e-learning can be complex and difficult. There is the need for a good course planning and an increased effort in monitoring and controlling all participants in all the different moments of the course while at the same time focusing on getting feedback that may allow to better steer the course (Hamburg et al., 2008).

Without the obligatory physical presence of a teacher, the process of e-learning is exposed to some deficiencies that may result in poor student learning. Specifically, the teacher is not able to observe the students in search of signs evidencing problems such as doubts, frustration, stress or fatigue, preventing teachers from taking action in such scenarios. The setting up of appropriate monitoring mechanisms in the context of e-learning is therefore very important in achieving an efficient learning process.

As shown in the literature, for an efficient monitoring of the student to take place, it is crucial that the e-learning system allows for a personalized study strategy and is able to show the needs and strengths of each student (Cantoni et al., 2004). Thus it becomes possible to track the progress of students, as well as improving their learning, by providing better personalized learning methods. The identification of learning problems and the cause of those problems is another advantage that can be achieved through the e-learning context, and via monitoring systems, such as the tool proposed in this paper.

1.2 Including Subjective Measures of Workload

In this paper we look at the monitoring and managing of mental workload as a way to improve the quality of information of the e-learning environment, especially to improve the teacher’s decision making abilities. One of the important parts of this work is a previously developed approach, deemed non-invasive and non-intrusive, for the analysis of the students’ interaction patterns.

Indeed, it was established in preliminary work that one’s patterns of interaction with the computer, measured in terms of the use of the keyboard and mouse, change when under fatigue as well as in periods of increased mental workload or even stress. Moreover, it were also found behavioral differences in performing different kinds of tasks, allowing to analyze patterns of attention in the students who participated in the experimental studies (Pimenta et al., 2015; Pimenta et al., 2014).

However, the work developed so far has the shortcoming of not considering mental workload, which is an important aspect when it comes to determining the actual level of fatigue. It is also important, for example, to distinguish between scenarios of boredom or excess of work (which, in a first instance, are both characterized by slowed performance).
This aspect is now included in this paper, thus representing a step forward in the development of more accurate fatigue assessment approaches (Figure 1) that encompass the type of task, the time on task, and the mental workload of the task. Indeed, workload levels can help isolate the causes affecting performance at a given time, improving fatigue management initiatives. To this end, and besides the metrics derived from the use of the mouse and keyboard, subjective measures of mental workload are also used.

Obtaining mental workload levels during task performance may be a challenging procedure. Moreover, the workload level experienced by an individual can affect task performance twofold: either through excessive or reduced mental workload. To this end, subjective measures are often used, some of them detailed in (Reid et al., 1982).

The two instruments most often used in research were developed in parallel in the 1980s, one at the NASA-Ames Research Center in California and the other within the U.S. Air Force human factors research group at Wright-Patterson AFB, Ohio.

The NASA Task Load Index (NASA-TLX) is a multidimensional assessment tool (Hart and Staveland, 1988). The main seven-point scale is: Overall Performance: How successful were you in performing the task? How satisfied were you with your performance? The TLX has five seven-point subscales that help identify difficult task characteristics. The subscales are:

- Mental Demand: How much mental and perceptual activity was required? Was the task easy or demanding, simple or complex?
- Physical Demand: How much physical activity was required? Was the task easy or demanding, slack or strenuous?
- Temporal Demand: How much time pressure did you feel due to the pace at which the tasks or task elements occurred? Was the pace slow or rapid?
- Frustration Level: How irritated, stressed, and annoyed versus content, relaxed, and complacent did you feel during the task?
- Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance?

Although these measures require a manual data entry, they are extremely useful to validate the interaction-based performance. It must thus be made clear that we do not intend for a final version of a fatigue management tool to include such indicators. Nonetheless, at the moment we look at such indicators as a viable way to assess the validity of the developed approach, when both are used in parallel, as described further ahead in this paper.

2 FATIGUE MANAGEMENT AS A TOOL FOR IMPROVING LEARNING PERFORMANCE

As stated before, this paper presents a tool for monitoring fatigue in e-learning scenarios using non-intrusive and non-invasive techniques. To this end it proposes an analysis and classification of the interaction patterns of users with the computer while using the mouse and keyboard. It is thus based on the students’ behavioural changes regarding the computer.

Similar approaches have been used in the past to estimate performance, albeit with more limited or different final aim. Cart et al. (Card et al., 1980) presented, in 1980, one of the earliest works on the topic,
aimed at the development of better interaction mechanisms with computers.

The proposed approach distinguishes from existing work twofold: (1) the application area and (2) the features considered, as detailed further ahead in this section. The tool collects data about the user’s interaction with the computer and stores it in the form of a log. This log contains each particular interaction event, their timestamp and other important information such as coordinates or key code, when applicable.

The following events are considered:

- **MOV**, timestamp, posX, posY - an event describing the movement of the mouse, in a given time, to coordinates (posX, posY) in the screen;
- **MOUSE_DOWN**, timestamp, [Left—Right], posX, posY - this event describes the first half of a click (when the mouse button is pressed down), in a given time. It also describes which of the buttons was pressed (left or right) and the position of the mouse in that instant;
- **MOUSE_UP**, timestamp, [Left—Right], posX, posY - an event similar to the previous one but describing the second part of the click, when the mouse button is released;
- **MOUSE_WHEEL**, timestamp, dif - this event describes a mouse wheel scroll of amount dif, in a given time;
- **KEY_DOWN**, timestamp, key - identifies a given key from the keyboard being pressed down, at a given time;
- **KEY_UP**, timestamp, key - describes the release of a given key from the keyboard, in a given time;

From these events, that fully describe the interaction of the user with the mouse and keyboard, we extract a set of features, based on notions of behavioural biometrics:

**KEY_DOWN TIME** - the timespan between two consecutive **KEY_DOWN** and **KEY_UP** events, i.e., how long was a given key pressed.

**TIME BETWEEN KEYS** - the timespan between two consecutive **KEY_UP** and **KEY_DOWN** events, i.e., how long did the individual took to press another key.

**VELOCITY** - The distance travelled by the mouse (in pixels) over the time (in milliseconds). The velocity is computed for each interval defined by two consecutive **MOUSE_UP** and **MOUSE_DOWN** events. Let us assume two consecutive **MOUSE_UP** and **MOUSE_DOWN** events, \( \text{map} \) and \( \text{mdo} \), respectively in the coordinates \((x_1, y_1)\) and \((x_2, y_2)\), that took place respectively in the instants \( \text{time}_1 \) and \( \text{time}_2 \). Let us also assume two vectors \( \text{posx} \) and \( \text{posy} \), of size \( n \), holding the coordinates of the consecutive **MOUSE_MOV** events between \( \text{map} \) and \( \text{mdo} \). The velocity between the two clicks is given by \( \frac{r\text{_dist}}{(\text{time}_2 - \text{time}_1)} \), in which \( r\text{_dist} \) represents the distance travelled by the mouse and is given by equation 1.

\[
\begin{align*}
\text{r\text{_dist}} = & \sum_{i=0}^{n-1} \sqrt{(\text{posx}_{i+1} - \text{posx}_i)^2 + (\text{posy}_{i+1} - \text{posy}_i)^2} \\
\end{align*}
\]

**ACCELERATION** - The velocity of the mouse (in pixels/milliseconds) over the time (in milliseconds). A value of acceleration is computed for each interval defined by two consecutive **MOUSE_UP** and **MOUSE_DOWN** events, using the intervals and data computed for the Velocity.

**TIME BETWEEN CLICKS** - the timespan between two consecutive **MOUSE_UP** and **MOUSE_DOWN** events, i.e., how long did it took the individual to perform another click.

**DOUBLE CLICK DURATION** - the timespan between two consecutive **MOUSE_UP** events, whenever this timespan is inferior to 200 milliseconds. Wider timespans are not considered double clicks.

**AVERAGE EXCESS OF DISTANCE** - this feature measures the average excess of distance that the mouse travelled between each two consecutive **MOUSE_UP** and **MOUSE_DOWN** events. Let us assume two consecutive **MOUSE_UP** and **MOUSE_DOWN** events, \( \text{map} \) and \( \text{mdo} \), respectively in the coordinates \((x_1, y_1)\) and \((x_2, y_2)\). To compute this feature, first it is measured the distance in straight line between the coordinates of \( \text{map} \) and \( \text{mdo} \) as \( \text{s\_dist} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \). Then, it is measured the distance actually travelled by the mouse by summing the distance between each two consecutive **MOUSE_MOV** events. Let us assume two vectors \( \text{posx} \) and \( \text{posy} \), of size \( n \), holding the coordinates of the consecutive **MOUSE_MOV** events between \( \text{map} \) and \( \text{mdo} \). The distance actually travelled by the mouse, \( \text{real\_dist} \) is given by equation 1. The average excess of distance between the two consecutive clicks is thus given by \( \frac{\text{r\_dist}}{\text{s\_dist}} \).

**AVERAGE DISTANCE OF THE MOUSE TO THE STRAIGHT LINE** - in a few words, this feature measures the average distance of the mouse to the straight line defined between two consecutive clicks. Let us assume two consecutive **MOUSE_UP** and
MOUSE_DOWN events, mup and mdo, respectively in the coordinates \((x_1, y_1)\) and \((x_2, y_2)\). Let us also assume two vectors \(\text{posx}\) and \(\text{posy}\), of size \(n\), holding the coordinates of the consecutive MOUSE_MOVE events between \(\text{mup}\) and \(\text{mdo}\). The sum of the distances between each position and the straight line defined by the points \((x_1, y_1)\) and \((x_2, y_2)\) is given by 2, in which \(\text{ptLineDist}\) returns the distance between the specified point and the closest point on the infinitely-extended line defined by \((x_1, y_1)\) and \((x_2, y_2)\). The average distance of the mouse to the straight line defined by two consecutive clicks is given by \(s_{\text{dists}}/n\).

\[
s_{\text{dists}} = \frac{1}{n-1} \sum_{i=0}^{n-1} \text{ptLineDist}(\text{posx}_i, \text{posy}_i)
\]  

**Distance of the Mouse to the Straight Line** - this feature is similar to the previous one in the sense that it will compute the \(s_{\text{dists}}\) between two consecutive MOUSE_UP and MOUSE_DOWN events, \(\text{mup}\) and \(\text{mdo}\), according to equation 2. However, it returns this sum rather than the average value during the path.

**Signed Sum of Angles** - with this feature the aim is to determine if the movement of the mouse tends to "turn" more to the right or to the left. Let us assume three consecutive MOUSE_MOVE events, \(\text{mov1}, \text{mov2}\) and \(\text{mov3}\), respectively in the coordinates \((x_1, y_1)\), \((x_2, y_2)\) and \((x_3, y_3)\). The angle \(\alpha\) between the first line (defined by \((x_1, y_1)\) and \((x_2, y_2)\)) and the second line (defined by \((x_2, y_2)\) and \((x_3, y_3)\)) is given by \(\text{degree}(x_1, y_1, x_2, y_2, x_3, y_3) = \tan(y_3 - y_2, x_3 - x_2) - \tan(y_2 - y_1, x_2 - x_1)\). Let us now assume two consecutive MOUSE_UP and MOUSE_DOWN events, \(\text{mup}\) and \(\text{mdo}\). Let us also assume two vectors \(\text{posx}\) and \(\text{posy}\), of size \(n\), holding the coordinates of the consecutive MOUSE_MOVE events between \(\text{mup}\) and \(\text{mdo}\). The signed sum of angles between these two clicks is given by equation 3.

\[
s_{\text{angle}} = \sum_{i=0}^{n-2} \text{degree}(\text{posx}_i, \text{posy}_i, \text{posx}_{i+1}, \text{posy}_{i+1})
\]  

**Absolute Sum of Angles** - this feature is very similar to the previous one. However, it seeks to find only how much the mouse "turned", independently of the direction to which it turned. In that sense, the only difference is the use of the absolute of the value returned by function \(\text{degree}(x_1, y_1, x_2, y_2, x_3, y_3)\), as depicted in equation 4.

\[
s_{\text{angle}} = \sum_{i=0}^{n-2} |\text{degree}(\text{posx}_i, \text{posy}_i, \text{posx}_{i+1}, \text{posy}_{i+1})|
\]  

DISTANCE BETWEEN CLICKS - represents the total distance travelled by the mouse between two consecutive clicks, i.e., between each two consecutive MOUSE_UP and MOUSE_DOWN events. Let us assume two consecutive MOUSE_UP and MOUSE_DOWN events, \(\text{mup}\) and \(\text{mdo}\), respectively in the coordinates \((x_1, y_1)\) and \((x_2, y_2)\). Let us also assume two vectors \(\text{posx}\) and \(\text{posy}\), of size \(n\), holding the coordinates of the consecutive MOUSE_MOVE events between \(\text{mup}\) and \(\text{mdo}\). The total distance travelled by the mouse is given by equation 1.

After the collection of the data, it is processed and converted into a set of behavioural biometric features, able to classify the behaviours of the user in terms of fatigue and level of attention to the task, described in detail in (Pimenta et al., 2015). This approach to assess performance has been developed, validated and used previously. However, after its development, we concluded that its most significant drawback is that it looks at performance of interaction alone. Fatigue is a complex phenomenon and performance measures by themselves may not be enough for an accurate measurement. In particular, and as stated before, mental workload is a very important aspect when it comes to characterizing fatigue.

We are thus now extending this approach with the acquisition of contextual information about the user, including the type of task being performed, the time spent on each different task, as well as the mental workload felt while performing the task (Figure 1). This means that this tool will now also be able to analyse the level of attention of a user to each task (e.g. distinguish between time spent on tasks related to e-learning activities against time spent on other tasks). In the overall, the new tool results in a more complete approach by including these important contextual factors.

Figure 2 depicts the process through which the system operates, where it is possible to observe the different classifications of information in order to allow, in the end, the management of fatigue. Initially the system captures the mouse and keyboard inputs. These data are further processed, stored and then used to calculate the values of the behavioural biometrics. In the learning phase the system shows a questionnaire in order to evaluate the subjective feeling of fatigue of the user, as well as the mental workload. When the system has a large enough dataset that allows to make classifications with precision, it will classify the inputs received into different mental fatigue and mental workload levels in real-time. At this
point, the system can start to be used by the people involved, especially the teacher who can better adapt and personalize his teaching strategies.

3 CASE STUDY

In order to assess the validity of the approach described in the previous section a case study was implemented with the aim of collecting data over a period of time that encompassed different sessions of e-learning, and thus test if it is possible to monitor fatigue through the use of behavioural biometrics and mental workload.

For this purpose, twenty four students volunteered (19 men, 5 women), all students of the course of Physical Sciences at the University of Minho. Their age ranged between 18 and 30. Participants were provided with the application for logging the events of the mouse and keyboard during the duration of the class. This application started automatically in the background, when the Operating System started, requiring no specific action from the part of the students. The previously mentioned list of features was extracted from the use of the mouse and keyboard for the whole period.

3.1 Methodology

The methodology followed to implement the study was devised to be as minimally intrusive as the approach it aims to support. Participants were provided with an application for logging the previously mentioned events of the mouse and keyboard. This application, which maintained the confidentiality of the keys used, needed only to be installed in the participant’s computer and would run in the background, starting automatically with the Operating System. The only explicit interaction needed from the part of the user was the input of very basic information on the first run, including the identification and age.

The course takes place physically in a classroom and comprises a teacher who is responsible for teaching a programming language (in this case MatLab) to a class of students. Each class has a duration of three hours, which always follows the same “protocol”: some theoretical concepts are introduced at the beginning of the class and the rest of the session is spent practising and solving exercises using the computer and a specific IDE. During each session the system, while in a learning phase, presents the user with a questionnaire (Figure 3) based on the NASA TLX for measuring mental workload.

Thus, in each session all inputs resulting from the interaction of the user with the computer using the mouse and keyboard are collected, together with the subjective values of cognitive load, acquired from the NASA TLX.

3.2 Results

Using the data collected in the classroom over the two weeks, a classification model was trained based on the K-Nearest Neighbour (KNN) algorithm. It is a method of classification based on closest training samples in the feature space.

A model was built based on a dataset with 74 instances, each instance being constituted by the average values of all features during periods of one our. Each instance is also assigned a label, which represents the response of students to the questionnaire for measuring workload, provided while using the com-
Figure 3: NASA TLX questionnaire used to collect information of mental workload.

Figure 4: Results of different models trained with different kernels and number of neighbours (K).

Several tests with different numbers of neighbours (K) and with different heuristics to the distance between neighbouring (rectangular, triangular, epanachicov, gaussian, rank, optimal) were performed. With a maximum of 50 neighbours, the solution having a lower mean squared error (MSE), was found with K = 30 and using the rectangular kernel, as shown in Figure 4.

The trained model was then used to predict the mental workload in data collected in the second period, in a total of 78 instances resulting from the interaction of users with the computer. According to the classification carried out, 64 out of 78 (83%) of the instances were in accordance with the subjective opinion of the user about the mental workload of the task that was to perform, i.e., were correctly classified. It is also important to note that the remaining 14 instances (17%) were classified as adjacent values.

Table 1: Results of the validation of the classification model (KNN). 83% of the instances were correctly classified (green cells). The 17% misclassified instances where nonetheless classified as neighboring values (red cells).

4 CONCLUSIONS

This paper describes a prototype of a tool for managing fatigue. Its main innovative aspect is that, for the first time, it considers the mental workload of a user while performing a task as an important component of fatigue assessment. The main objective is to detect patterns of behavior at different levels of mental workload. Measurement of levels of mental workload are obtained through the NASA TLX instrument, which is based on a subjective self-evaluation. These subjective measures, paired with measures of performance and context of the task being performed by a user, allow to train a classifier as the one depicted which achieved fairly good results. In the described case study, the tool was used in several classes during the period of two weeks, which allowed not only to test it in a real scenario.

The results achieved from the implementation of the case study show that it is indeed possible to analyse and quantify mental workload through the use of the mouse and keyboard, and this allows not only to measure cognitive load but also to improve the process of monitoring mental fatigue.

Although at the moment we aim to support the teacher’s decision making process, the long-term goal of this work is to develop environments that are...
autonomous and take actions concerning their self-management. These actions will be guided by several objectives, one of them being to manage cognitive load, minimize fatigue and increase performance and well-being of an individual or group of individuals through an appropriate selection of tasks and task durations.

ACKNOWLEDGEMENTS

This work is part-funded by ERDF - European Regional Development Fund through the COMPETE Programme (operational programme for competitiveness) and by National Funds through the FCT (Portuguese Foundation for Science and Technology) within project FCOMP-01-0124-FEDER-028980 (PTDC/EII-SII/1386/2012) and project PEst-OE/EEI/UI0752/2014.

REFERENCES


