MeWS-IT: A Mental Workload Based System for Interruption Timing

Shamsi T. Iqbal
Department of Computer Science
University of Illinois,
Urbana, IL 61801,USA
siqbal@uiuc.edu

ABSTRACT
Proactive systems executing in multitasking environments are increasingly interrupting user tasks. To mitigate the negative impact of ill-timed interruptions on task performance and users' affective state, researchers are exploring systems that reason about when to interrupt. In this paper I present my research on MeWS-IT - a system that leverages mental workload to identify opportune moments for interruptions. MeWS-IT offers the opportunity to evaluate how well the process of using mental workload can be automated and how this provides improved reasoning about when to interrupt, contributing to better interruption management in multitasking domains.

ACM Classification: H.5.2 [Information Interfaces and Presentation]: User Interfaces — evaluation/methodology, user-centered design, H.1.2 [Models and Principles]: User/Machine Systems – *Human Information Processing.*

General terms: Design, Experimentation, Human Factors, Measurement

Keywords: Attention, Interruption, Mental Workload, Pupil Size, Task Models, User Studies.

INTRODUCTION

One of the most challenging problems in multi-tasking environments is *interruption* [19]. Proactive systems executing in aviation cockpits, control rooms, in-vehicle displays and office environments are increasingly interrupting a user's primary tasks [2, 4, 5, 7, 17]. When tasks are interrupted at random moments, users take longer to complete tasks, commit more errors and experience increased annoyance, anxiety and frustration [2, 5, 14].

On the other hand, information conveyed by proactive systems are often beneficial for the user [6, 18]. To balance between maintaining information awareness and minimizing negative impacts of interruption, researchers have been investigating systems that reason about when to interrupt [7, 8, 10]. Existing systems use environmental cues to compute a cost of interruption (COI). However, to compute a more accurate COI, systems need to include a measure of a user's mental workload. Using mental workload provides a direct assessment of the user's internal state, allows the COI to be calculated on a continuum and firmly grounds the COI in psychological theories [20, 21].

In this paper I present my research on MeWS-IT - a Mental Workload Based System for Interruption Timing. I first discuss three preliminary studies investigating the feasibility of using mental workload as a predictor of opportune moments for interruption. I then present the system architecture and implementation details of MeWS-IT. Finally, I present studies evaluating how well the system automates identification of opportune moments.

RELATED WORK

In this section I discuss how effects of interruption can be mitigated, systems that reason about interruption and measures of mental workload.

Mitigating Effects of Interruption

Mitigating effects of interruption requires knowledge about a user's position in a task [20]. Resource theories [21, 22] state that if interruptions occur at lower workload moments, more resources would be available for the interrupting task, causing less disruption. Miyata and Norman [20] argue that lower workload moments occur at subtask boundaries. Studies show that scheduling interruption at boundaries improves task performance and reduces negative affect than when delivered at random times [2].

This work builds upon interruption theory to develop a system that defers interruptions until moments of low mental workload in a task sequence, automating a process shown to mitigate the effects of interruption.

Systems that Reason About When to Interrupt

In [10-12], researchers are constructing systems that reason about when to interrupt a user by weighing the value of information against the COI. The underlying models use external cues such as desktop activity, visual and acoustical analyses of the physical task environment, and scheduled activities of the user to compute the COI.

This work complements existing approaches by considering mental workload - a source of information about the user's internal state, increasing the accuracy of the COI.

UIST'05, October 23–27, 2005, Seattle, Washington, USA.
ACM 1-59593-023-X/05/0010...$5.00.

Measuring Mental Workload

Pupil size, eye movement, blink rate, and heart rate variance have been shown to be reliable sources of continuous mental workload data [16]. Pupil size additionally offers an *immediate* measure of workload, which simplifies data analysis. Under conditions of controlled illumination, research shows that pupil size is an effective and reliable measure of mental workload [3, 9], where increases in pupil size correlate with increases in mental workload. In my studies, I chose pupil size due to availability of and experience in using eye-tracking equipment.

MEWS-IT

MeWS-IT (pronounced as Muse-it) is a computational system which leverages mental workload induced during task execution to determine opportune timing for interruptions. In this section I describe the three phase-development process of MeWS-IT: preliminary investigation, design details of the system and plans for evaluation.

Preliminary Investigation

Based on [20], I conducted a series of three studies investigating how a user's mental workload changes during task execution and how workload can be leveraged to identify opportune moments for interruption.

In the first study [15] I explored the use of pupil dilation as a reliable measure of workload in *interactive* environments. Pupil size was measured as users performed easy and difficult tasks from each of four task categories. Results showed that for the one task category that required sustained mental effort, the more difficult tasks induced increased pupil dilation than the easier tasks. Other task categories did not demonstrate this effect. To investigate further, I decomposed the tasks and identified subtasks that had different cognitive demands between easy and difficult tasks. Including only these subtasks, a second analysis showed that the more difficult tasks caused increased pupil dilation. The results show that pupil dilation correlates with workload induced by interactive tasks and suggest that to understand how workload changes during execution of tasks with complex structures, pupillary response should be aligned to corresponding models of task execution.

Building on lessons from the first study, I conducted a second study to better understand how workload changes during task execution, focusing on subtask boundaries [13]. In this study, users performed interactive tasks while workload was monitored through the use of pupil dilation. I developed workload aligned task models by aligning workload (pupillary response) to corresponding GOMS models of the tasks. Analyzing the models, I empirically showed that low workload moments exist at subtask boundaries. The more remarkable finding was that workload varied across boundaries - boundaries *higher* in a task model had *lower* workload than boundaries lower in the model.

In the third study, I tested how opportune low workload moments were for interruption [14]. Results showed that interruptions at low workload boundaries had less negative impact than interruptions at other points – for example, users experienced 69% less resumption lag, 18% less annoyance and found interrupting applications to be 63% more respectful when interrupted at lower workload boundaries as compared to higher workload boundaries.

My findings validate the use of mental workload as an effective predictor of opportune moments for interruption and emphasize the need for building a system that leverages workload to reason about when to interrupt. If mental workload can be encoded into a continuous stream of COI, then systems can use the models to obtain the current COI and predict if a lower cost (more opportune) moment would occur within a short period, enabling more robust decisions about when to interrupt.

Design and Implementation of MeWS-IT

In this section I describe the design goals of MeWS-IT, the proposed system architecture and associated challenges.

Design Goals

The main design goal of MeWS-IT is to improve productivity by balancing information awareness with mitigation of disruption caused by interruptions by incoming information. Existing systems typically support one or the other but not both. For example, systems that use peripheral displays for incoming information use a negotiated strategy [19], placing the burden of monitoring information arrival on the user. This approach has the risk of reduced information awareness since the user may not access the information in a timely fashion. A possible solution is to indicate information arrival with an attentional cue. However, the attentional cue itself is an interruption and if ill-timed, will have similar negative impacts on the user.

To address this goal, my system will mirror the aforementioned strategy of making the information available on a peripheral display. However, instead of delivering the cue immediately, the system will mediate when to present the attentional cue. This will help the user maintain information awareness through a minimally disruptive cue.

Another design goal is to allow flexibility and extensibility in the system. Measuring workload requires specialized equipment not readily available and the system should be flexible to accommodate any physiological measure. Also, provisions for external cues should be made, further increasing decision accuracy about when to interrupt. We discuss how these goals are addressed in latter sections.

System architecture and Implementation

MeWS-IT has two major components – the Task Model Builder and the Interruption Manager (Figure 1). The Task Model Builder will build models of common user tasks offline and make them available for the Interruption Manager, which will operate in real time.

The Task Model Builder has two components – the Task Model Generator and the Cost Assigned Model (CAM) Generator. Users will perform common user tasks as their onscreen interaction is video recorded and coded as Task Models by the Task Model Generator. Mental workload

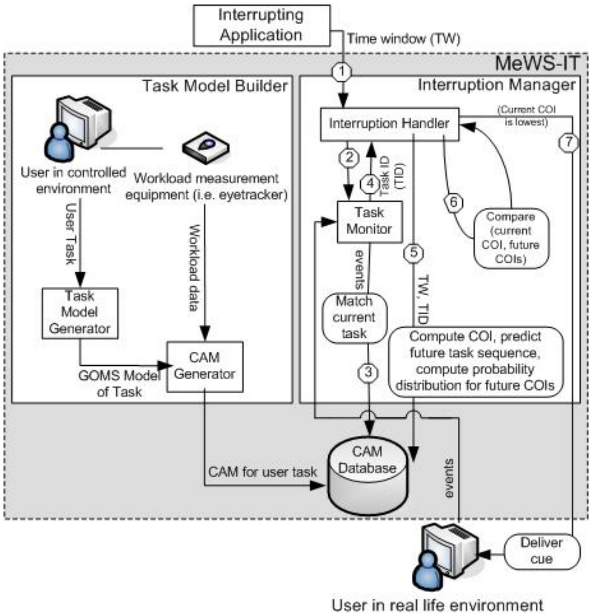

Figure 1: Schematic of the MeWS-IT architecture. Rectangles indicate components and rounded rectangles indicate operations. Numbers indicate sequence of operations that would be performed to identify opportune moments for interruption.

data will be collected through equipment measuring physiological data such as eye-trackers for measuring pupil dilation. Illumination, position of fixated object, fatigue etc will be controlled for as users perform tasks. This will ensure that changes in the physiological data can be attributed to the induced workload rather than environmental factors.

The task models and the workload data will be synchronized in the CAM generator, providing a cost of interruption (COI) for every point in the task model. The process of alignment is informed by studies in my preliminary investigation [13]. The generated CAM models will be stored in a persistent database for use by the Interruption Manager.

The Interruption Manager will interact with the interrupting application in real time and manipulate timing of attentional cues. It has two components – the Task Monitor and the Interruption Handler. Commonly used applications will be instrumented so that the Task Monitor is notified when events occur within those applications. The Task Monitor will attempt to recognize the current task by matching the events to events in the models stored in the CAM database.

The Interruption Handler (IH) will be invoked on arrival of an interruption so that it can reason about when to present the attentional cue. For now it is assumed that a separate mechanism will encode the relevancy and urgency of the interruption into a maximum allowable time window (TW). The sequence of operations that will execute on arrival of an interruption is shown in Figure 1. First, the TW associated with the interruption will be sent to the IH, which will query the Task Monitor which task is currently being executed. The Task Monitor will return a task ID to the IH. The IH will access the CAM database to obtain the CAM model for the current task. A decision making algorithm within the IH will then perform the following: 1) compute the current COI, based on the CAM model of the recognized task, 2) predict the future task sequence given the predictive model of task execution, 3) compute the probability distribution for the COI for the possible task sequences within the given time window, 4) compare the current COI against expected future COIs to decide whether to interrupt now or delay till a lower cost moment and 5) if the decision is to delay, then continue to step 1. The algorithm will terminate either when the lowest cost moment is found or when the time window expires. Termination of the algorithm will invoke immediate delivery of the cue.

Implementation of the Task Model Generator and the Task Monitor will be based on existing prototypes [1] and will be extended to support a wider variety of applications. This research will focus on developing the CAM generator and the Interruption Handler and integrating all components so that they execute together as an independent whole.

Addressing the design goals

To balance between information awareness and mitigation of disruption the IH will mediate delivery of the attentional cue based on the TW of the interruption, the current task and predicted task sequences obtained from the corresponding CAM model. The decision algorithm of the IH will decide when to interrupt by comparing the current COI (negative impact) to probable future COIs within the TW, ensuring that information is viewed in time but at the most opportune moment within that time window.

To achieve flexibility, the workload measurement equipment in the Task Model Builder can be replaced by any suitable measure, as long as a continuous stream of workload data (representing the COI) can be fed to the CAM generator. In fact, if workload cannot be measured, then a set of heuristics developed from theoretical understanding of workload changes can be used to compute an approximate COI. While these heuristics may not provide perfect results, they will still be better than what exists today.

To achieve extensibility, further sources of information (i.e. external cues) can be supplied to the Interruption Handler, so that its decision making algorithm can compute a more accurate COI. For example, if the system had access to COI based on external cues in addition to workload based COI, then the combined COI could be:

$$COI_{combined} = W_{wl} * COI_{wl} + W_{ec} * COI_{ec}$$

Where W = weight, wl= workload and ec= external cues. The weights can be manipulated based on the quality of the data source – providing a more robust COI estimation.

Challenges

Some of the challenges that have to be addressed in this research are 1) developing a set of meaningful heuristics for assigning COI in absence of workload measurement equipments, 2) allowing flexibility for variations in execution sequences and 3) choosing the best modality for presenting the cue. While the third challenge is context dependent, I am currently working on the first two. I am developing a cost function based on a set of heuristics obtained through theoretical understanding of workload changes and analyzing task models from prior work. I am also investigating machine learning techniques that would be appropriate for modeling task execution behavior.

Evaluation

The final step of this research is to evaluate how well the system automates the process of identifying opportune moments for interruption and how the identified moments balance awareness of information with mitigation of the negative impact. Based on McFarlene's suggested modes of interruption timing [19], I will compare four strategies for interruption timing –Immediate, Scheduled, Negotiated and Mediated.

For primary tasks, I will use tasks representative of common user tasks. Time windows will be hand-coded into interruptions. Information will be presented in a peripheral window while timing of the cues will be manipulated. For the Immediate strategy, the cue will be presented as soon as information arrival. For the Scheduled strategy, the cue will be presented at timed intervals. For the Negotiated strategy, no cue will be presented - the user will be responsible for monitoring the arrival of the information. Finally, for the Mediated strategy, the cues will be presented at moments picked out by the MeWS-IT system. Users will be instructed to attend to the peripheral display as soon as they are presented with the cue.

Measures will include awareness, resumption lag, errors and affective state. These measures have been successfully employed in the past in measuring disruptive impacts of interruption. The evaluation will be conducted as a controlled experiment as opposed to field deployment since the primary goal is to study the behavior of the system in detail and then refine it in preparation for field studies.

CONCLUSION

Consideration of task induced mental workload in reasoning about interruption timing offers a promising step towards balancing between information awareness and mitigation of disruption. My research will make a significant contribution in this direction through development and evaluation of a theoretically grounded system. This will not only benefit commercial products that disseminate information to the user preemptively but also the research community looking into maximizing productivity in workspaces.

REFERENCES

1. Bailey, B.P., P.D. Adamczyk, T.Y. Chang and N.A. Chilson. A Framework for Specifying and Monitoring User Tasks. *Journal of Computers in Human Behavior*, July/August, 2005.

2. Bailey, B.P. and J.A. Konstan. On the Need for Attention Aware Systems: Measuring Effects of Interruption on Task Performance, Error Rate, and Affective State. *Journal of Computers in Human Behavior*, July/August, 2005.

3. Beatty, J. Task-Evoked Pupillary Responses, Processing Load, and the Structure of Processing Resources. *Psychological Bulletin*, 91 (2), 276-292, 1982.

4. Cutrell, E., M. Czerwinski and E. Horvitz. Notification, Disruption and Memory: Effects of Messaging Interruptions on Memory and Performance. *INTERACT* 2001, 263-269.

5. Czerwinski, M., E. Cutrell and E. Horvitz. Instant Messaging and Interruption: Influence of Task Type on Performance. *OZCHI* 2000, 356-361.

6. Dey, A.K. and G.D. Abowd. Cybreminder: A Context-Aware System for Supporting Reminders. In *Proceedings of 2nd International Symposium on Handheld and Ubiquitous Computing*, 2000, 172-186.

7. Fogarty, J., S.E. Hudson and J. Lai. Examining the Robustness of Sensor-Based Statistical Models of Human Interruptibility. In *CHI* 2004, 207-214.

8. Fogarty, J., A.J. Ko, H.H. Aung, E. Golden, K.P. Tang and S.E. Hudson. Examining Task Engagement in Sensor-Based Statistical Models of Human Interruptibility, *CHI* 2005, 331-340.

9. Hess, E.H. and J.M. Polt. Pupil Size in Relation to Mental Activity During Simple Problem Solving. *Science*, 1190-1192, 1964.

10. Horvitz, E. and J. Apacible. Learning and Reasoning About Interruption. In *Proceedings of the Fifth ACM International Conference on Multimodal Interfaces*, 2003, 20-27.

11. Horvitz, E., A. Jacobs and D. Hovel. Attention-Sensitive Alerting. In *Conference Proceedings on Uncertainty in Artificial Intelligence*, 1999, 305-313.

12. Hudson, S.E., J. Fogarty, C.G. Atkeson, D. Avrahami, J. Forlizzi, S. Kiesler, J.C. Lee and J. Yang. Predicting Human Interruptibility with Sensors: A Wizard of Oz Feasibility Study. *CHI* 2003, 257-264.

13. Iqbal, S.T., P.D. Adamczyk, S. Zheng and B.P. Bailey. Towards an Index of Opportunity: Understanding Changes in Mental Workload During Task Execution.*CHI* 2005, 311-320.

14. Iqbal, S.T. and B.P. Bailey. Investigating the Effectiveness of Mental Workload as a Predictor of Opportune Moments for Interruption. *CHI* 2005, 1489-1492.

15. Iqbal, S.T., X.S. Zheng and B.P. Bailey. Task Evoked Pupillary Response to Mental Workload in Human-Computer Interaction. *CHI* 2004, 1477-1480.

16. Kramer, A.F. Physiological Metrics of Mental Workload: A Review of Recent Progress. In Damos, D.L. ed. *Multiple-Task Performance*, Taylor and Francis, London, 1991, 279 - 328.

17. Lee, J.D., J.D. Hoffman and E. Hayes. Collision Warning Design to Mitigate Driver Distraction. *CHI* 2004, 65-72.

18. Maglio, P. and C.S. Campbell. Attentive Agents. *Communications of ACM*, 46 (3), 47-51, 2003.

19. McFarlane, D.C. and K.A. Latorella. The Scope and Importance of Human Interruption in Hci Design. *Human-Computer Interaction*, 17 (1), 1-61, 2002.

20. Miyata, Y. and D.A. Norman. The Control of Multiple Activities. In Norman, D.A. and Draper, S.W. (eds.) *User Centered System Design: New Perspectives on Human-Computer Interaction*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1986.

21. Wickens, C.D. Multiple Resources and Performance Prediction. *Theoretical Issues in Ergonomic Science*, 3 (2), 159-177, 2002.

22. Wickens, C.D. Processing Resources and Attention. In Damos, D.L. ed. *Multiple-Task Performance*, Taylor & Francis, London, 1991, 3-34.