Estimation of Interruptibility during Office Work Based on PC Activity and Conversation

Satoshi Hashimoto, Takahiro Tanaka, Kazuaki Aoki, and Kinya Fujita*

Graduate School, Tokyo University of Agriculture and Technology
2-24-16 Nakacho, Koganei, Tokyo 184-8588, Japan
{takat,kazuak,kfujita}@cc.tuat.ac.jp

Abstract. The chances of being interrupted by online communication systems, such as email, instant messenger, and micro-blog, are rapidly increasing. For the adequate control of interruption timing, the real-time estimation of the interruptibility of the user is required. In this study, we propose an interruptibility estimation method using PC activity and conversational voice detection based on the wavelet transform. The offline estimation was applied to a dataset of 50 hours obtained from 10 users. The results indicated the feasibility of improving the interruptibility estimation accuracy by the automatic detection of the existence and end of conversations.

Keywords: interruptibility, availability, voice detection, office work, interruption.

1 Introduction

A wide variety of online communication systems, such as email, instant messenger, and micro-blog, have been growing in popularity with the growth of the Internet. Although such systems pose a risk of inconveniently interrupting the users, such interruptions are not controlled in present systems. Studies of human multitasking have suggested that the suspension and resumption of a working memory related to active tasks occurs when task switching is forced by an external interruption. This set of suspensions and resumptions causes a lag at task switching [1, 2]. Furthermore, it has been pointed out that a "resumption lag" can be potentially increased by interruption timing and the relationship between switching tasks [3, 4]. Therefore, frequent inter-ruptions regardless of the user's status risk the fragmentation of the user's working time and decrease intellectual productivity [5]. Therefore, automatic user status estimation and interruption timing control can be expected to reduce the chance of inappropriate interruptions and to enhance online communication quality.

One potential method to estimate a user's status is user activity monitoring by using sensors. For example, counting keystrokes and mouse clicks has been utilized during PC work [6], and various sensors have been installed in the work space or have

* Corresponding author.

S. Yamamoto (Ed.): HIMI/HCII 2013, Part III, LNCS 8018, pp. 297–306, 2013.

been attached to the users [7, 8]. These methods estimate interruptibility based on PC activities or recognized events, such as having a guest or a telephone call [9, 10]. These methods are expected to adequately estimate a user's status during tasks that have observable physical activity. However, the mental workload during a task or intellectual activity, such as deep thinking, has no measurable output. Moreover, it is practically difficult to install sensors on users or in the work space.

Another approach is to estimate the breakpoints of a task. Several studies have reported the relationship between the resumption lag and breakpoint. These studies have experimentally demonstrated that resumption lags at breakpoints are shorter, because the breakpoint is the end of a subtask, and therefore, suspension and resumption of a working memory is not required, even if the task requires significant intellectual activity [11, 12]. Iqbal and Bailey have also proposed a breakpoint detection method based on the structural analysis of the tasks [13].

However, this approach requires task-structure analysis to determine interruptibility levels at breakpoints. Therefore, a real-time breakpoint estimation method, which can be performed by an information system, is desired for automatic interruption timing control.

We consider that focused application switching (AS), which is the transition of the active application window, is a potential breakpoint in PC work. Our experimental results demonstrate that the interruptions at AS are significantly more acceptable than those during continuous work. Finally, we propose a user interruptibility estimation method at AS [14]. However, for office workers, there are social factors that affect their interruptibility, such as communication and collaboration during work. Although the interruptibility during communication would be low, the proposed method estimates high interruptibility in such a situation, because there are no, or very few, observable PC activities. Therefore, we need to consider social activities in addition to PC activities to improve the estimation accuracy.

In this study, first, we experimentally collected PC activity, environmental sound, and subjective interruptibility data of participants during work in a laboratory. Second, we analyzed the relationship between conversational status and interruptibility. Third, we proposed an interruptibility estimation method based on conversation status in addition to PC activity. Finally, we experimentally confirm that estimation accuracy is improved by considering conversational status.

2 Real-Time Estimation of User Interruptibility Based on PC Activity

We used focused AS, which is the transition of the active application window, as a breakpoint in PC work [14]. AS is considered as the user's intentional switching of the working space or working target. Therefore, the user's concentration at AS is expected to be instantaneously weakened compared with that during continuous work. Moreover, AS commonly and frequently occurs in PC work and is easily detected. Thus, AS is a potential source of information presentation timing with less risk of task disturbance.

To examine this assumption, we experimentally collected and analyzed PC operation records and subjective interruptibility scores. The experimental results demon-strated that interruptions at AS are significantly more acceptable for users than those during continuous work ($p < 0.01$, t-test). Moreover, the resumption lags caused by interruption at AS were significantly shorter than those during continuous work.

We analyzed the relationship between the interruptibility scores and features, which were calculated from the operation records and were expected to reflect the interruptibility at AS. Finally, we defined an estimation rule, which estimates user interruptibility at three levels on the basis of the occurrence probability of each index at a particular interruptibility level.

On the other hand, the frequency of AS during PC work depends on the user's task and situations. Therefore, the opportunity of information presentation may significantly decrease in some cases. Thus, we proposed an interruptibility estimation method during continuous work (Not Application Switching (NAS)) based on four PC-activity-related indices. The estimation method demonstrated a level of accuracy almost comparable with that of the estimation method at AS.

However, the accuracy was susceptible to change because of non-PC work and conversation during work. In both cases, the indices based on PC activity do not reflect the actual user activity; therefore, the proposed method estimated the numbers of low interruptibility status as high. In particular, 50% of these serious errors in the office environment were caused by conversation. We needed to consider the conversation state to improve the estimation accuracy.

3 Collection of Interruptibility Data during Deskwork

3.1 Experimental Setup and Conditions

To investigate the effect of conversation on the interruptibility of office workers while using PCs, we conducted experiments to interrupt the participants during work. The participants were requested to subjectively score their interruptibility at interruption. The configuration of the experimental system is shown in Fig.1. The experiment program remained in the "task-tray" and recorded the PC operation information and the sound at 500-ms intervals. The recorded data are keystroke, mouse click, mouse wheel, clipboard update, window number, and foreground application name.

The experiment program intermittently interrupted the participants when one of the preprogrammed interruption conditions was satisfied. The participants were requested to score their subjective interruptibility in five grades through a "dialog-box." The participants were instructed to assume that the interruption is the request of a five minute conversation with their colleagues. The interruption conditions are set as follows to obtain the interruptibility scores during conversation, after finishing the conversation, and in the quiet state:

1. Continuous sound detection (potential conversation): The rate of samples in which the sound pressure exceeded the threshold is between 25 and 90 percent in the last 20 seconds and more than 15 percent during the period between 10 and 20 seconds before the interruption.

2. End of continuous sound (potential conversation ending): Condition 1 had been satisfied at the previous sample and is not satisfied at the current sample.
3. Quiet (potential not conversation): Conditions other than 1 and 2.

Furthermore, the minimal interruption intervals, between 30 and 900 seconds, were imposed on each condition to avoid excessive interruption and bias in data. The participants were eight university students and two faculty members. All participated in the five-hour experiment at their own desks. No restrictions in activity were imposed. The observed PC activities were programming, document editing, data analysis, email, and net surfing; and the non-PC activities were reading from a printed paper, conversation, smartphone operation, and eating. The conversational contents were research-related matters, hobbies, and daily news.

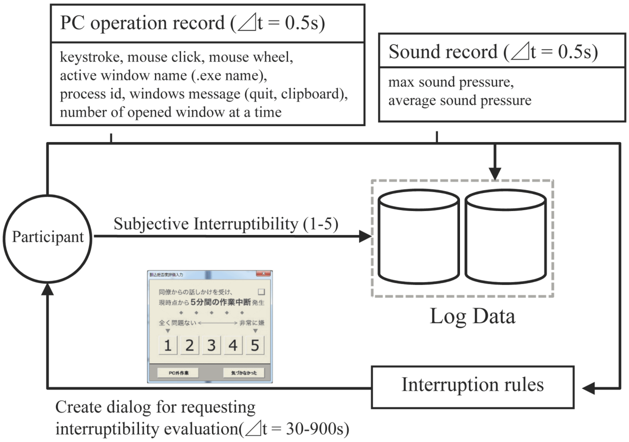

Fig. 1. Experimental system for collecting interruptibility data during deskwork

3.2 Detection of Conversational Statuses

We defined three types of status—"During conversation," "Conversation ending," and "Quiet"—to investigate the effect of conversation on interruptibility. The decision rules were as follows:

1. During conversation: human voice had been detected in the last 10 seconds.
2. Conversation ending: human voice had been detected in the last 20 seconds and had not been detected for 10 seconds including interruption timing. The cases in which a keystroke or mouse click is detected have been excluded, because PC operation implies the participants' resumption of the task.
3. Quiet: conditions other than 1 and 2.

The conversational voice was detected by the Wizard-Of-Oz (WOZ) method and an automatic method based on the wavelet transform [15]. In the WOZ method, the experimenter judged the conversational status from the recorded sound based on the described rules. Therefore, “During conversation” included both the conversations in which the participant was included and not included.

In the automatic method, the conversational voice was detected on the basis of the features in the frequency and time domains. In the frequency domain, voice has a pitch frequency and the first formant frequency. The first formant frequency is about two times the pitch frequency. In the time domain, the pitch frequency gradually changes during conversation. To detect these features, the following detecting rules were defined by using wavelet transform in reference to a method for speech detection [16].

- Frequency domain conditions:

  (a) The pitch frequency is between 80 and 360 Hz.
  (b) The first formant frequency is similar to two times the pitch frequency.
  (c) The power at the dip is sufficiently smaller than that at the pitch.

- Time domain condition:

  (d) The number of the divided-into-eight subsamples that have stable pitch frequency is greater than three. The stability index was the variation of the pitch frequency.

Samples that satisfied both conditions were regarded as conversational voice.

3.3 General Trends of Experimental Results

The general trends of the obtained data are shown in Table 1. Because the PC operation might influence the effect of conversation on interruptibility, the data was categorized by the existence of PC activity as well as the conversational status. PC activity was detected as a keystroke or mouse click within the last 30 seconds.

Table 1. Average subjective scores and frequencies at various interruption timings

	During conversation	Quiet	Conversation ending
PC activity	2.19 (159)	2.80 (150)	3.96 (25)
No PC activity	2.71 (156)	3.39 (76)	3.98 (51)

The result of multiple comparisons revealed that the interruptibility “During conversation” is significantly higher and that at “Conversation ending” is significantly lower than that during the quiet state, regardless of PC activity. Therefore, “During conversation” and “Conversation ending” are promising indices for interruptibility estimation. The interruptibility with PC activity is detected was significantly lower than that without PC activity.

The average interruptibility values “During conversation” and at “Conversation ending” were 0.61 lower and 1.16 higher, respectively, than that during the quiet state, when PC activity was detected. The differences changed to 0.68 and 0.59 in the cases without PC activity, respectively. These differences suggest that conversation ending has a larger effect on interruptibility.

4 Offline Estimation of Interruptibility

4.1 Interruptibility Estimation Equation

We introduced two conversational indices into the estimation equation during NAS in reference to a previous study [17]. The PC operation indices that were proposed in the previous study are shown in Table 2. Equation 1 represents the estimation equation. If the data meet the condition for each index, the index takes the value 1, otherwise 0. The value of function f ranges from 0 to 1. Interruptibility is classified into three levels using Equation 2.

Table 2. Indices based on PC operation

ID	Indices	Interruptibility
A	Keystroke in last 20 seconds.	Low
B	PC activity detection in more than 30% of the last 2 minutes.	Low
C	Use of both keyboard and mouse in the last 2 minutes.	Low
D	Transitioned from shell (desktop) within 5 min.	Low

$$f(x) = (2A + B + C + D)/5 \quad (1)$$

$$\text{Interruptibility} = \begin{cases} \text{Low} & 0.7 \le f(x) \le 1 \\ \text{Medium} & 0.2 \le f(x) < 0.7 \\ \text{High} & 0 \le f(x) < 0.2 \end{cases} \quad (2)$$

The added conversational indices in this study are shown in Table 3. Equations 3 and 4 show the estimation equations. In Section 3.3, it was suggested that the effects of "During conversation" and "Conversation ending" on interruptibility are different in the cases PC activity and No PC activity. Therefore, the coefficients in the two equations differ. The coefficients were experimentally decided with reference to the results of a multiple linear regression analysis. After calculating the function f, interruptibility was classified into three levels using equation 2.

Table 3. Added indices of conversational statuses

ID	Indices	Interruptibility
E	During conversation	Low
F	Conversation ending	High

$$\text{PC activity:} \quad f(x) = (2A + B + C + D + E + 2\bar{F})/8 \quad (3)$$

$$\text{No PC activity:} \quad f(x) = (2A + B + C + D + 2E + \bar{F})/8 \quad (4)$$

4.2 Interruptibility Estimation Results

Fig.2 (a) shows the estimation result using the already proposed PC operation indices. Figs.2 (b)–(d) show the offline estimation results with “During conversation,” “Conversation ending,” and both indices, respectively. The conversational voice was de-tected by the WOZ method. In Fig.2 (a), the precision for high interruptibility was 40%. The indices “During conversation” and “Conversation ending” increased the precision to 48% and 42%, respectively, as seen in Fig.2 (b, c). The use of both indices further improved the precision to 59%. Especially, the rate of high-risk error esti-mation, which means the unacceptable status was mistakenly estimated as interrupti-ble, decreased from 40% to 18% by applying the two indices. On the other hand, the improvement in precision for low interruptibility was from 66% to 69%. The im-provement of the total recall was from 36% to 40%. The conversational indices ap-pear to be more effective for the improvement of accuracy for high interruptibility.

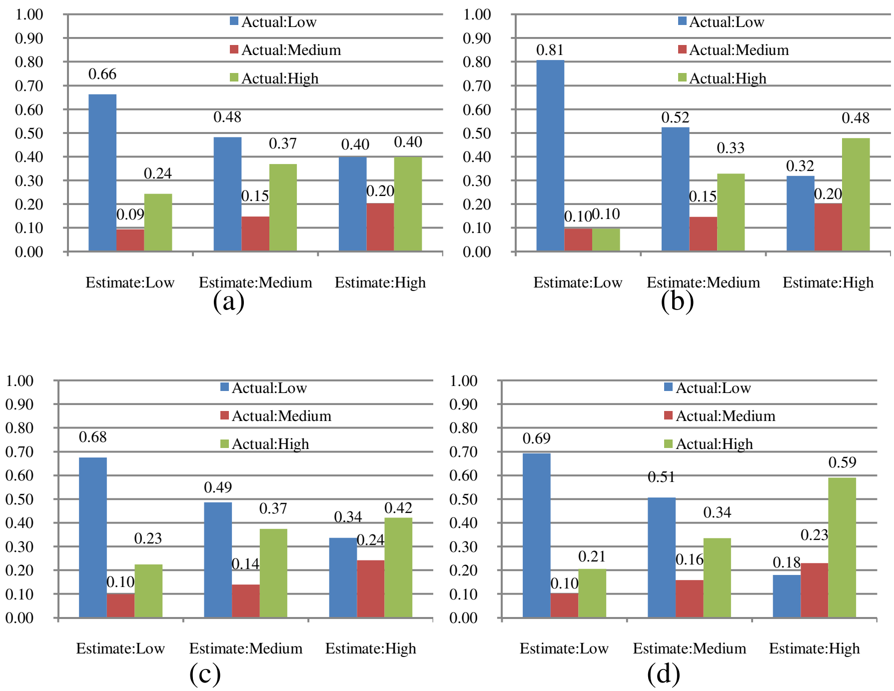

Fig. 2. Results of offline interruptibility estimation based on (a) PC activity, (b) PC activity and conversation, (c) PC activity and conversation ending, and (d) PC activity and both conversational indices

Fig.3 shows the estimation results based on automatic conversational voice detection. The precision for high interruptibility was slightly increased to 42% compared to the PC-activity-based method. The rate of high-risk error estimation was reduced to 34%. The precision of high interruptibility was also increased to 68%. Even though the current automatic voice detection algorithm still needs to be improved, the use of conversational information in combination with PC activity appears promising.

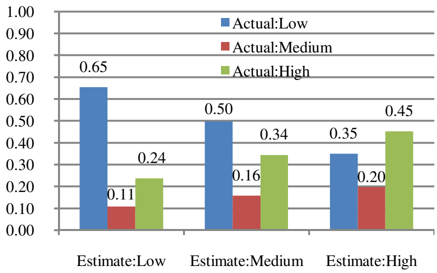

Fig. 3. Estimation results based on the automatic conversation detection

5 Discussions

As a result of the analysis of the correlation between interruptibility and conversation detection by the WOZ method, we confirmed that "During conversation" and "Conversation ending" affect interruptibility. Moreover, by adding these two indices to the previous estimation method based on PC operations, the precision of high interruptibility was improved, and serious errors were reduced. Further improvement is expected by using other indices that correlate with interruptibility. One promising index is "Conversation duration," which may correlate with the importance of conversation.

Table 4. Comparison of conversation status detection methods

		Automatic detection		
		During conversation	Quiet	Conversation ending
WOZ	During conversation	162	103	49
	Quiet	12	207	7
	Conversation ending	24	27	25

Among the "During conversation" samples that were judged by the WOZ method, 103 samples were categorized as "Quiet," and 49 samples were detected as "Conversation ending." It implies that the automatic detection method tends to fail for the detection of conversational voice. Therefore, an improvement of the detection rule is required. As for the computational time of the automatic method using the continuous wavelet transform, it took 0.86 s for 0.5 s sound data. For real-time estimation, it is necessary to reduce this computational time. The reduction of time and frequency resolutions and the use of the discrete wavelet transform may be potential solutions.

In this study, the introduction of two conversational indices improved the accuracy of the interruptibility estimation method based on PC operations. It suggests that the reflection of social factors in addition to task-related factors will enable a more accurate estimation. The task-related factors utilized in the present study are only PC activity features, suggesting that the target task of the present method is limited to office work with PC usage. However, other types of office activities exist, such as paperwork. The introduction of additional indices, reflecting the working density or user's working attitude during office works without PC, is expected to extend the applicable office activities. Head motion is one potential index, because it reflects the gaze that is tightly related with the worker's information acquisition behavior.

6 Conclusions

In this study, we analyzed the correlations between the conversational features and interruptibility during office work. Two conversational indices "During conversation" and "Conversation ending" improved the estimation accuracy of the interruptibility estimation method based on PC activity. The improvement of automatic conversation detection accuracy and the reduction of the computational time are required for adequate interruptibility control toward better online communication.

Acknowledgement. This work was partly supported by Grants-in-Aid for scientific research from the Japan Society for the Promotion of Science and the National Institute of Information Technology (NICT).

References

1. Altman, E.M., Trafton, J.G.: Memory for Goals: An Activation-based Model. Cog. Sci. 26, 39–83 (2002)
2. Salvucci, D.D., Taatgen, N.A.: Threaded Cognition: An Integrated Theory of Concurrent Multitasking. Psychol. Rev. 115(1), 101–130 (2008)
3. Cutrell, E.B., Czerwinski, M., Horvitz, E.: Effects of instant messaging interruptions on computing tasks. In: Extended Abstracts on Human Factors in Computing Systems, pp. 99–100 (2000)
4. Monk, C.A., Trafton, J.G., Boehm-Davis, D.A.: The Effect of Interruption Duration and Demand on Resuming Suspended Goals. J. Exp. Psychol. Appl. 14, 299–313 (2008)
5. Mark, G., Gonzalez, V.M., Harris, J.: No Task Left Behind? Examining the Nature of Fragmented Work. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 321–330. ACM, New York (2005)
6. Honda, S., Tomioka, H., Kimura, T., Oosawa, T., Okada, K., Matsushita, Y.: A Home Office Environment based on the Concentration Degrees of Workers: A Virtual Office System Valentine. Trans. Inform. Process. Soc. Jpn. 39(5), 1472–1483 (1998) (in Japanese)

7. Avrahami, D., Fogarty, J., Hudson, S.E.: Biases in Human Estimation of Interruptibility: Effects and Implications for Practice. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 50–60. ACM, New York (2007)
8. Fogarty, J., Hudson, S.E.: Toolkit Support for Developing and Deploying Sensor-based Statistical Models of Human Situations. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 135–144. ACM, New York (2007)
9. Danninger, M., Stiefelhagen, R.: A Context-Aware Virtual Secretary in a Smart Office Environment. In: Proceedings of the 16th ACM International Conference on Multimedia, pp. 529–538. ACM, New York (2008)
10. Lu, H., Pan, W., Lane, N.D., Choudhury, T., Campbell, A.T.: SoundSense: Scalable Sound Sensing for People-Centric Applications on Mobile Phones. In: Proceedings of the 7th International Conference on Mobile systems, Applications, and Services, pp. 165–178 (2009)
11. Borst, J.P., Taatgen, N.A., Van Rijn, H.: The problem state: A Cognitive Bottleneck in Multitasking. J. Exp. Psychol. Learn. Mem. Cognit. 36(2), 363–382 (2010)
12. Monk, C.A., Boehm-Davis, D.A., Trafton, J.G.: Recovering from Interruptions: Implications for Driver Distraction Research. Hum. Factors 46, 650–663 (2004)
13. Iqbal, S.T., Bailey, B.P.: Leveraging Characteristics of Task Structure to Predict Costs of Interruption. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 741–750. ACM, New York (2006)
14. Tanaka, T., Fujita, K.: Study of User Interruptibility Estimation Based on Focused Application Switching. In: Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work, pp. 721–724. ACM, New York (2011)
15. Juang, C.-F., Cheng, C.-N., Tu, C.-C.: Wavelet Energy-Based Support Vector Machine for Noisy Word Boundary Detection with Speech Recognition Application. Expert Syst. Appl. 36(1), 321–332 (2009)
16. Juang, C.-F., Cheng, C.-N., Che, T.-M.: Speech detection in Noisy Environments by Wavelet Energy-based Recurrent Neural Fuzzy Network. Expert Syst. Appl. 36(1) (2009)
17. Tanaka, T., Fukasawa, S., Takeuchi, K., Nonaka, M., Fujita, K.: Study of Uninterruptibility Estimation Method for Office Worker during PC Work. Inform. Process. Soc. Jpn. 53(1), 126–137 (2012) (in Japanese)