User Interruptibility Estimation based on Focused Application Switching

Takahiro Tanaka and Kinya Fujita
Tokyo University of Agriculture and Technology
{takat, kfujita}@cc.tuat.ac.jp

Abstract

In this study, we propose a method to estimate the interruptibility of a user on the basis of focused application switching (AS) during working on a personal computer. We experimentally demonstrated that the interruptions of AS are more acceptable than those during continuous work. Therefore, we constructed an algorithm that estimates the interruptibility of AS at three levels based on 19 selected indexes. The feasibility of the interruptibility estimation of AS was demonstrated by an estimation experiment using another dataset of 13 office workers.

1. Introduction

In recent years, the possibility that users might be interrupted by information systems has been increasing along with the popularization of the Internet and the ubiquity of the computing environment. However, the timing of interruptions has not been controlled in most systems. A preceding study [1] has shown that frequent interruptions that do not reflect the user status can fragment the user's working time and decrease intellectual productivity.

In this study, we concentrated on focused application switching (AS), which is a change in the use of applications. AS is considered to be the breakpoint not only of work on a personal computer (PC) but also of intellectual activity. We present the experimental data and propose an interruptibility estimation rule on three levels for AS.

2. Related Work

2.1 Estimating user status and busyness

There are several studies that aim in estimating user status using various techniques such as counting keystrokes or mouse clicks, and using various sensors that are set in the working space or attached to the users [2-4]. These methods estimate interruptibility from amount of pc activities or recognizing events of room like a having a guest and telephone call, and they are expected to adequately estimate the user status during tasks that have observable physical activity indexes. However, intellectual activity such as deep thinking has no observable output. Moreover, an environment with various sensors would be expensive to prepare.

On the other hand, some studies reported a relationship between interruptibility and the breakpoint of the work [5-7]. At the breakpoint, the user's interruptibility is expected to increase instantaneously, even if the task requires intellectual activity. Furthermore, it was suggested that the breakpoint level affects interruptibility [5]. A task structure that categorizes operation analysis is therefore required to determine this level. The multi-tasking aspect of PC usage also makes it potentially difficult to apply the task-structure-based method. Therefore, distinguishing the interruptibility level of a breakpoint remains challenging [7].

2.2 Our approach

We concentrated on AS, which is the transition of the active application window, as a breakpoint in PC work because AS is considered to be the user's intentional switching of his or her 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 during PC work and is easily detected, which implies that AS is a potential source of information presentation timing with a low risk of task disturbance. In this study, we tried to estimate the interruptibility of AS as a breakpoint using indexes from PC operation records.

3. Experiment I: Collecting Records

To examine the assumption, we experimentally collected PC operation records and subjective interruption evaluation logs and analyzed them.

CollabTech 2012 , August 27-29, 2012, Hokkaido, Japan.

3.1 Method

Figure 1 shows the logging tool we developed that records users' PC operations every 500 ms. The tool collects a user's keystrokes, click counts, mouse wheel usage, the active window name (.exe name), process id, window message (quit and clipboard), and the number of windows open at one time.

Moreover, the tool interrupts subjects at automatically selected ASs, or every 5 min during continuous work, i.e., not-focused application switching (NAS),, and requires the subjects to subjectively evaluate if and how they want to be interrupted. The scores were scaled as a five-grade evaluation: "1" indicates the user was absolutely uninterruptible, "2" means the user was uninterruptible, "3" means neither uninterruptible or interruptible, "4" means interruptible, and "5" means absolutely interruptible. We instructed the subjects to evaluate the interruption timing without any reference to the frequency of interruptions.

The evaluation logs, which consist of eight sets of daily PC activity over 1 h, were collected from eight university students, who did their own daily works such as data arranging, programming, and writing reports.

3.2 Summary of results

Table 1 shows the interruptibility comparison between ASs and NASs, the frequencies with which they occurred, and their averages. The experimental results demonstrate that interruptions at an AS are significantly more acceptable for users than interruptions during continuous work ($p < 0.01$, t-test). We analyzed the relationship between the interruptibility during NASs and several PC operation activity indexes, as was conducted by previous works [2]. A weak correlation between interruptibility and the activity data accumulated during 5 min, which is the summation of keystrokes, mouse clicks, and mouse-wheel use, was also observed within the NAS dataset. Therefore, we especially focus on AS for more acceptable interruption.

Even the general tendency suggests that interruptions at ASs are more acceptable; less interruptible ASs exist, as seen in Table 1. As the previous study pointed out, the interruptibility of a breakpoint varies with the level of the breakpoint in the task structure [8]; the interruptibility level of an AS appears to have a similar task-related factor. That means the information systems need to distinguish the ASs that are more interruptible on the basis of the acquired information.

Therefore, we analyzed the relationship between the interruptibility scores and the indexes that were calculated from the operation records with no particular sensors. The relationship is expected to reflect the interruptibility at an AS.

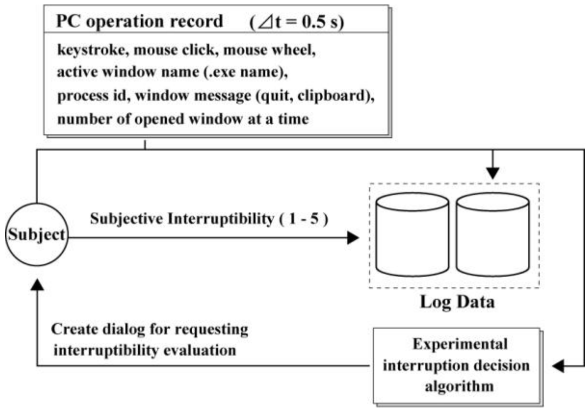

Figure 1.Experimental logging system.

Table 1. Subjective interruptibility scores at AS vs. NAS.

	Subjective Interruptibility					Freq.	Ave.
	Low		Mid	High			
	1	2	3	4	5		
AS	36	35	62	65	77	275	3.4
NAS	25	19	20	27	14	105	2.9

4. Selection of Indexes and Estimation Rule Generation

In general, the opening of an application window is thought to be a beginning of a new task, and the closing of it will be a sign of the end of a task. Furthermore, the tight coupling of the tasks that are done using two applications might make an AS uninterruptible. The indexes that might reflect the application coupling, such as "copy and paste" and "parent-child relationship," were also examined. Moreover, physical activities such as keystroke and mouse operations are expected to reflect the interruptibility, as reported in previous studies [2]. These types of indexes are expected to reflect the work discontinuity.

In this study, we analyzed a log dataset from university students, as explained in section 3, for selecting indexes that reflect interruptibility at AS.

4.1 Indexes that affect Interruptibility at an AS

We analyzed the 12-h operation records that were collected from eight university students, according to the following three viewpoints.

1. Work discontinuity.
2. Application coupling.
3. Physical activity.

We examined the indexes categorized in the above three groups and selected the indexes that showed significant differences. Each index has a binary value and expresses the existence or absence of the future. The threshold value for each index was experimentally decided.

4.1.1 Work discontinuity

In general, users perform their work by running multiple applications at the same time. Therefore, we presumed that increases or decreases in the number of open windows might reflect the beginning or end of tasks.

We analyzed the effect of the difference in the number of open windows after an AS by one-way analysis of variances that made the number a factor. The result revealed that the number significantly reflects interruptibility at an AS ($F = 3.08$, $p < 0.01$). In particular, the situation when the number of open windows increased had significantly lower interruptibility than the situation when the number of open windows did not change ($p < 0.01$). However, when the user repeated the opening and closing of a window, i.e. searching a specific data file, the difference in the number of open windows did not indicate the start or finish of a task. Therefore, we examined the increase or decrease in the number of open windows compared with the average numbers during the previous 2 min. The comparison showed a strong relationship with interruptibility by one-way analysis of variances ($F = 6.02$, $p < 0.01$). From the results, we confirmed, by multiple comparison, that the situation when the number of open windows increased had significantly lower interruptibility than the situation when the number of open windows decreased ($p < 0.05$), and the situation when the number of open windows decreased had higher interruptibility than the situation when the number did not change. Furthermore, we confirmed by Welch's t-test that the situation when the number decreases had a high interruptibility when a quit message was observed at that time ($p < 0.01$).

Table 2 shows the indexes for work discontinuity during three situations: when the number of open windows increased (Increased Situation), when the number of open windows decreased (Decreased Situation), and when the number did not change (Not Changed). In Table 2, "Low" means that the AS with the index tends to be low interruptibility. "High" means that the AS with the index tends to be high interruptibility. The mark "-" means that the index had no effect on the interruptibility.

Table 2. Indexes that affect interruptibility about work discontinuity.

id	Indexes	Window Number		
		Inc.	Not	Dec.
A	Increase in open windows.	Low	-	-
B	Decrease in open windows.	-	-	High
C	Increase in open windows compared to average of last 2 min.	Low	-	-
D	Decrease in open windows compared to average of last 2 min.	-	-	High
E	Window message (quit).	-	-	High

4.1.2 Application coupling

The relationship between how two applications are used before and after an AS might reflect the shift of the current task and changes to the interruptibility. In previous work, the AS with a clip-board message had lower interruptibility than without it during Not-Changed situation [8]. In this study, we analyzed the transitions between parent and child processes, the reuse of the same application, and the shell (Explorer in Windows OS) transitions.

We examined the transition between a parent process and a child process and found the correlation between it and interruptibility. We confirmed by Welch's t-test that the parent-child transition during Increased Situation had significantly lower interruptibility than other situations ($p < 0.01$). It appears that the transition to a child-window means the start of a subtask such as program debugging. On the other hand, the child-parent transition during Decreased Situation had higher interruptibility.

Similarly, we presumed that frequently using the same application might not be a breakpoint. Therefore, we examined the reuse of the same application during 2 min. It showed that the reuse had significantly lower interruptibility than using a new application (Welch's t-test, $p < 0.01$) during Increased Situation but not Not-Changed situation.

Furthermore, we examined the relationship between the transition to and from the shell (Explorer) and the

Table 3. Indexes that affect interruptibility about application coupling.

id	Indexes	Window Number		
		Inc.	Not	Dec.
F	Window message (clipboard).	-	Low	-
G	Parent-window to child-window transition.	Low	-	-
H	Child-window to parent-window transition.	-	-	High
I	Reuse of the same application within 2 min.	Low	Low	-
J	Transition to the shell (Explorer).	-	-	High
K	Transition from the shell (Explorer).	High	-	-
L	Re-transition to the shell within 2 min.	-	Low	-
M	Re-transition from the shell within 2 min.	-	Low	-

interruptibility of each of the three situations using a one-way analysis of variances in which the Explorer transition was a factor. The result revealed that the transition also affected the interruptibility. The transition to the shell had a significantly high interruptibility during Decreased Situation, and the transition from the shell had a significantly low interruptibility during Increased Situation ($p < 0.05$). In both cases, transitioning to and from the shell was done via the desktop. On the other hand, the re-transition to and from the shell within 2 min had low interruptibility during Not-Changed situation ($p < 0.05$). These results were observed when the user tried to find the target file using the shell. Table 3 shows the indexes for work discontinuity during three situations.

4.1.3 Physical activity

In related works, the PC activity, keystrokes, and mouse operations were related to the interruptibility during tasks. Therefore, we examined the effect of activities to the interruptibility of an AS.

At first, we used Welch's t-test to examine the relationship between the frequency of ASs and interruptibility. The results revealed that an AS that had a duration time over 2 min was significantly interruptible than one with a shorter duration, no matter in which of the three situations it occurred ($p > 0.05$). The continuous use of a single application for more than 2 min tends to show higher interruptibility. On the other hand, an AS that had a duration time less than 15 s was significantly interruptible during the Increased Situation ($p < 0.01$). In those cases, the user concentration might be relatively low.

Second, we examined the correlation between the interruptibility and PC activities just before an AS. If the user was typing within 20 s before an AS occurred, the interruptibility became significantly low during Increased Situation or Not-Changed situation (Welch's t-test, $p < 0.01$). Moreover, if there was mouse activity within 20 s before an AS operation, the results showed the same effect to the interruptibility during Not-Changed situation ($p < 0.05$). On the other hand, during Decreased Situation, the index had no effect to the interruptibility. We presume that the activity observed before an AS during Decreased Situation was done to close a current task, save data, etc.

In related works, the researchers were concerned about activities during the minutes to estimate interruptibility. Therefore, we examined the relationship between the interruptibility and the operation time during the last 2 min before an AS. The results revealed that an AS that had more than 10% operating time during the last 2 min was significantly uninterruptible ($p < 0.01$), except during Decreased Situation. Moreover, we presumed that the tasks that involved complex operations tended to be uninterruptible. Therefore, we examined the relationship between the interruptibility and an AS when the user was and was not using both the keyboard and the mouse during the last 2 min. During Increased Situation and Not-Changed situation, the AS had significantly lower interruptibility than using the keyboard or mouse ($p < 0.05$). On the other hand, the interruptibility tended to be high during Decreased Situation ($p < 0.1$). We presumed that the complex activity observed before an AS during Decreased Situation was done to close a current task. Table 4 shows the indexes for physical activity during the three situations.

Table 4. Indexes that affect interruptibility about physical activity.

id	Indexes	Window Number		
		Inc.	Not	Dec.
N	Continuous use of one application over 2 min.	High	High	High
O	Continuous use of one application within 15 s.	Low	-	-
P	Typing activity within 20 s before AS.	Low	Low	-
Q	Mouse activity within 20 s before AS.	-	Low	-
R	More than 10% operating time in last 2 min.	Low	Low	-
S	Use of both keyboard and mouse in last 2 min.	Low	Low	High

4.2 Interruptibility estimation rule

Next, we investigated a co-occurrence relationship between 19 indexes to define the estimation rule. We anticipated that the effect of the indexes varied depending on the status of the task: beginning of task, end of task, or continuity of task. Therefore, we divided ASs according to the three situations based on the difference of the number of windows opens after an AS. According to the results of our analysis, we defined three estimation equations for "increased," "decreased and "not-changed" situations, as follows. Each equation consists of indexes that significantly affected the interruptibility. The coefficients of a, b, and c in equations (2) and (3) were set to 2. The coefficients and threshold values were experimentally decided.

During Increased Situation, the parent-child transitions and reuse within 2 min decreased the interruptibility. Indexes from PC activities also had a lowering effect on interruptibility. During Decreased Situation, a quit message, a transition to a shell, and a child-parent transition indicated that the user had finished a task, and these increased the interruptibility at an AS. Moreover, in this situation, PC activity had almost no effect on the interruptibility. In contrast, the indexes from PC activity affected the interruptibility during Not-Changed situation. Similar to previous work [2], it appears the consideration of the activity is effective for estimating interruptibility during continuous work.

Finally, we compressed the interruptibility into three levels, from a practical viewpoint, using equation (4); interruptible, neither, and uninterruptible. The subjective scores “5” and “4” were converted to “High,” the score “3” was converted to “Medium,” and the scores “2” and “1” were converted to “Low.”

$$F(x) = \begin{cases} f_{inc}(x) & \textbf{increased situation} \\ f_{dec}(x) & \textbf{decreased situation} \\ f_{neu}(x) & \textbf{no change situation} \end{cases}$$

$$f_{inc}(x) = \{ Ax + Gx - Nx + Ix(1 - Kx) + (Cx + Ox)Kx \} + \{ Rx + (Px + Sx)(1 - Kx) \} \quad (1)$$

$$f_{dec}(x) = - \{ Bx + Ex \cdot Jx + Dx + Hx(1 - Kx) + Nx \} + \{ 2Rx - 2Sx \} \quad (2)$$

$$f_{neu}(x) = \{ Fx + Ix - Nx + Kx \cdot Mx + Jx \cdot Lx \} + \{ 2max(Px, Qx) + 2max(Rx, Sx) \} \quad (3)$$

$$\textbf{Interruptibility} = \begin{cases} High & F(x) < 0.5 \\ Med & 0.5 \leqq F(x) < 0.8 \\ Low & F(x) \geqq 0.8 \end{cases} \quad (4)$$

Figure 2 shows the results of the estimation using the trial experiment dataset that was corrected to select the indexes. Although the average of the precisions was less than 50%, the precision of high interruptibility was more than 60%. In addition, the probability of the estimations in which the error is less or than or equal to 1 was 94%. It suggests that the proposed method greatly reduces the risk of interruption during unacceptable moments.

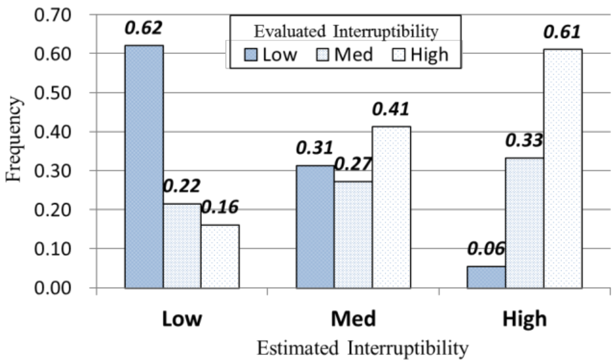

Figure 2. Results of the analysis of interruptibility estimation.

5. Experiment II: Evaluation

We evaluated the interruptibility estimation rule at ASs using two newly corrected datasets. One was a 50-h log dataset of the work activities of 11 university students and another was a 930-h log dataset of the work activities of 13 office workers.

5.1 Estimation using the dataset from university students

We collected 50 h of data of the work activities of 11 university students using the experimental logging tool (Figure 1). The tool interrupts subjects at automatically selected ASs and NASs, and requires the subjects to subjectively evaluate interruptibility.

Figure 3 shows the results of the estimation using the proposed method with the dataset from the university students. The precision of the high interruptibility was 53% and the probability that the estimated error is less than or equal to 1 was 87%. Similarly, the precision of the low interruptibility was 54% and the probability that the estimated error is less than or equal to 1 was 79%. These results were almost the same as the results in Figure 2.

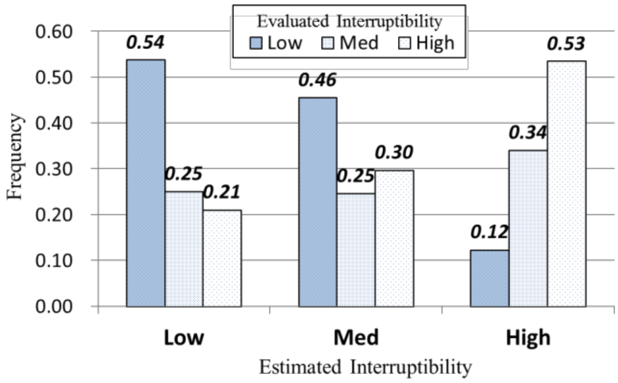

Figure 3. Results of interruptibility estimation using the dataset from university students.

5.2 Estimation using the dataset from office workers

To evaluate the usability of the method in a real environment, we collected 930 h of data from 13 office workers during their daily business operations for 2 weeks. The workers included nine R&D workers and four clerical workers. We collected the log data using the logging tool in the same way as we did for the university students. In addition, we required the subjects to report a detail of their daily work in order to select PC work logs from these data.

Figure 4 shows the results of the estimation using the method with PC work data from office workers. The precision of the high interruptibility was 55% and the probability that the estimated error is less than or equal to 1 was 73%. On the other hand, the precision of the low interruptibility was 45% and the probability that the estimated error is less than or equal to 1 was 65%. The high interruptibility precision was almost the same as the results in 5.1. However, the low interruptibility precision was low compared with the other two sets.

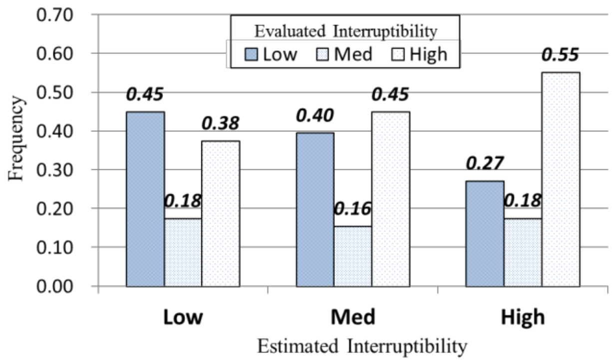

Figure 4. Results of interruptibility estimation using the dataset from office workers.

6. Discussion

According to the results of the evaluation, the average precisions of high interruptibility for the student and office worker datasets were both approximately 54%. Although the proposed method did not provide high accuracy, it demonstrated the feasibility for estimating tasks that were highly interruptible with more than 50% precision. In addition, the high-risk estimation rate, which is the average misestimating of low interruptibility ASs to high ones, was approximately 20%. Therefore, the precision improves almost to 80% if the user allows interruptions at a medium interruptibility level. Furthermore, these results were observed for not only the university-students dataset but also on the office workers dataset under real office conditions.

Meanwhile, the accuracy of low interruptibility for the office workers dataset was decreased by approximately 10%, compared with the university-students dataset. Currently, our proposed method reflects a performance of current task only. However, for office workers, there are some biases that affect their interruptibility, such as a delivery deadline of a current task, the importance of the task, and the number of tasks they have been assigned to complete. We need to add to our method indexes that reflect these biases in order to improve the low interruptibility accuracy.

The proposed method appears to allow information systems to reduce the risk of serious work disturbance by interruption. The other features of this method, which could serve as practical advantages, are: (1) it works without any sensors, and (2) the estimation rule requires no restriction in PC usage. We need to improve the rule to estimate the highly interruptible ASs with greater accuracy. Furthermore, there are various tasks in office work that should be considered aside from PC work, like desk work and mixed work. We need to increase the kind of work applicable to estimation by our method.

7. Conclusion

We experimentally confirmed that application switching is a relatively good opportunity to interrupt PC work. Furthermore, we proposed an interruptibility estimation rule and demonstrated the feasibility of the estimation.

References

[1] G. Mark, V. M. Gonzalez, and J. Harris, No task left behind? Examining the nature of fragmented work: Proc. CHI'05, pp. 321-330, 2005.

[2] S. Honda et al. A home office environment based on the concentration degrees of workers: A virtual office system valentine. Trans. Info. Processing Society of Japan, 39(5), pp. 1472-1483, 1998 (in Japanese).

[3] J. Fogarty and S. E. Hudson. Toolkit support for developing and deploying sensor-based statistical models of human situations. Proc. of SIGCHI conference on Human factors in computing systems, pp. 135-144, 2007.

[4] D. Avrahami, J. Fogarty, and S. E. Hudson. Biases in human estimation of interruptibility: Effects and implications for practice. Proc. of SIGCHI conference on Human factors in computing systems, pp. 50-60, 2007.

[5] S. T. Iqbal and B. P. Bailey. Leveraging characteristics of task structure to predict the cost of interruption. Proc. of SIGCHI conference on Human Factors in Computing Systems, pp. 741-750, 2006.

[6] M. Czerwinski, E. Cutrell, and E. Hirvutz. Instant messaging: Effects of relevance and timing; people and computers XIV. Proc. of HCI2000, Vol. 2, pp. 71-76, 2000.

[7] S. T. Iqbal and B. P. Bailey. Effects of intelligent notification management on users and their tasks. Proc. of SIGCHI conference on Human factors in computing systems, pp. 93-102, 2008.

[8] T. Tanaka, K. Matsumura, and K. Fujita. Operation Record Analysis for Information Display Timing Control based on Application-Switching. Trans. Info. Processing Society of Japan, 50(1), pp. 314-322, 2009 (in Japanese).