Predicting Uninterruptible Durations of Office Workers by Using Probabilistic Work Continuance Model

Shota SHIRATORI†, Yuichiro FUJIMOTO††, Nonmembers, and Kinya FUJITA††, Member

SUMMARY In order not to disrupt a team member concentrating on his/her own task, the interrupter needs to wait for a proper time. In this research, we examined the feasibility of predicting prospective interruptible times of office workers who use PCs. An analysis of actual working data collected from 13 participants revealed the relationship between uninterruptible durations and four features, i.e., type of application software, rate of PC operation activity, activity ratio between keystrokes and mouse clicks, and switching frequency of application software. On the basis of these results, we developed a probabilistic work continuance model whose probability changes according to the four features. The leave-one-out cross-validation indicated positive correlations between the actual and the predicted durations. The medians of the actual and the predicted durations were 539 s and 519 s. The main contribution of this study is the demonstration of the feasibility to predict uninterruptible durations in an actual working scenario.

key words: interruptibility, probabilistic prediction, office worker, work continuance model

1. Introduction

Various sources of interruption such as telephone calls and visits by co-workers disrupt office workers performing their tasks. A survey on interruptions in the workplace revealed that senior developers experienced on average more than 120 interruptions during eight working hours [1]. Although interruptions are inevitable especially when performing collaborative tasks, frequent and uncontrolled interruptions potentially distract workers and fragment their working time [2], [3]. Previous studies have suggested that switching to an interrupting task forces a worker to temporarily store and retrieve information on the original task from his/her working memory, thus increasing the cognitive workload [4], [5]. Consequently, interruptions that occur without reflection upon the state of workers are likely to decrease their intellectual productivity [6].

Several approaches have been proposed to deal with this issue. One is to automatically estimate the interruptibility of a worker and prevent interruptions until a more appropriate time [7], [8]. This sort of approach can be used for managing the notification timings for asynchronous communications such as e-mail or SMS (short message service). Here, interrupters using asynchronous communication tools do not expect an immediate response from the interruptee. This allows the systems to delay notifications to some extent. On the other hand, requests for synchronous (real time) communication from co-workers, either in person or telephone calls, should not be delayed except when the interruptee is busy.

As a way of facilitating non-disruptive communications, previous studies have proposed to share their interruptibility visually by using a color-coded contact list in the messenger application [9], [10] or color-coded physical lamp [11]. These systems allow the interrupter to judge whether his/her colleague is interruptible or not and avoid undesirable interruptions. However, even if the interrupter notices that his/her colleague is not interruptible, he or she needs to keep on checking until the colleague becomes interruptible. Having to actively check and wait for a convenient time for an interruption could reduce the work efficiency of the interrupter.

On the other hand, if the interrupter knew the expected time when the interruptee will become interruptible in some way, he or she would be able to devote the waiting time to work. Consequently, such an arrangement would improve the work efficiency of both the interrupter and interruptee.

Although there have been no studies on predicting prospective interruptible times of office workers, various studies have attempted to predict a person’s future state, e.g., destinations of pedestrians [12], next visited place and dwell time [13], [14], and next activity [15]. Several studies proposed probabilistic human models and used them for prediction. Begole et al. proposed a probabilistic model of work rhythm [16]. Miyagi et al. modeled a human in a calculation task by using probabilistic transitions between working and non-working states [17]. Tanaka et al. predicted the presence of users using recorded probabilities [18]. Yamagoe et al. also used past statistics for predicting the time to return from lunch [19]. Along the line of these studies, we have chosen to develop a probabilistic model for predicting uninterruptible durations of office workers.

One useful piece of information with which we can predict interruptible times is the task breakpoint. Regardless of the content of a task, its breakpoints are likely to be interruptible [20]. Therefore, a prediction of when an ongoing task is completed or suspended might be a good substitute for a prediction of an interruptible time.

The task completion time will naturally depend on the size of the task being performed. The time until the worker...
takes a break may also be related to the degree of his or her engagement in the task. In addition to those speculations, our preliminary study implied that the uninterruptible duration has some relation with PC operation activity [21]. Thus, this study further examines the relation between the uninterruptible duration in an actual working scenario and four PC-related features, i.e., type of application software, rate of PC operation activity, activity ratio between keystrokes and mouse clicks, and switching frequency of application software. Section 2 summarizes the potential factors affecting uninterruptible duration, while Sect. 3 discusses the validity of the proposed features.

Then, this paper proposes a probabilistic work continuance model whose probability changes according to the four features shown in Fig. 1. It furthermore proposes a method to predict uninterruptible durations based on the time-dependent decrease in cumulative continuance probability. The details of the model and the prediction results are shown in Sects. 4 and 5.

The main contributions of this study are as follows:

1. It reveals the statistical relation between the four features of PC operation and uninterruptible durations of office workers through a long-term user study in a real office environment in a university.
2. It describes a model of work continuance probability that depends on the four PC-related features.
3. It demonstrates the feasibility of using the proposed model to predict uninterruptible durations in an actual working scenario.

2. Uninterruptible Duration and Affecting Factors

2.1 Methods for Estimating Interruptibility

We need to know the actual moment when a worker becomes interruptible in order to discuss the feasibility of predicting the end of his/her uninterruptible duration. However, asking workers to report their interruptible moments increases their cognitive workload. Instead, we chose to use a real-time interruptibility estimation system to define the uninterruptible duration.

Various interruptibility estimation methods have been proposed. Hudson et al. focused on the situation of the worker and proposed a method using the seating position, voice, and PC operation [22]. Tanaka et al. used head motion as an indication reflecting the worker’s degree of engagement in a task [23]. Tani and Yamada demonstrated the effectiveness of using tabletop pressure [24]. On the basis of the knowledge that cognitive workload decreases at task breakpoints [7], [20], several methods have been proposed, such as detection of transitions between application software in PC work [25] or self-initiated intermissions [26]. Another line of studies is based on psycho-physiological relationships. Züger et al. used electroencephalograph (EEG) data and electrodermal activity (EDA) [27]. Katidioti et al. focused on pupil dilation as a measure of cognitive workload [28].

The interruptibility estimation methods based on physiological signals are principally better at reflecting the user’s mental state. However, most of these methods require users to wear a sensor and conduct a specific calibration periodically. Thus, they would be difficult to use in the long-term office scenario that we target. Instead, we chose an interruptibility estimation method based on PC operation information [25] that does not require the user to wear sensors.

2.2 Identification of Uninterruptible Durations

This study defines an uninterruptible duration as being from the moment when the estimated interruptibility turns to uninterruptible until the time when the estimation turns back to interruptible, as shown in Fig. 2. Since the employed system estimates the user’s interruptibility on three levels (1: low / 2: medium / 3: high) every 0.5 seconds, the onset of the uninterruptible duration is defined at the moment when the estimated interruptibility falls from high to medium. Similarly, its end is defined as the time when it changes to high. The uninterruptible durations are extracted from interruptibility estimates that are continuously logged while the user is engaged in their work.

2.3 Factors Affecting the Uninterruptible Duration

2.3.1 Size of the Task at Hand

The time for completing a task naturally depends on its size. In other words, easy tasks take a shorter time. Thus, we examined two presumable task-size-related features, i.e., type of application software and activity ratio between keystrokes and mouse clicks.

The size of a task is expected to have a statistical relationship with the sort of application software used.
instance, writing a manuscript, which is a typical complicated task, frequently takes a long time. In contrast, web browsing often ends (or someone takes a break) in a short time. Obviously, the actual duration, or task size, will have a natural variation even when using the same application software. Differences between individuals will also exist, but a general tendency may exist in a statistical sense.

Furthermore, one application software could be used for different tasks. For instance, although a word processor application can be used for writing documents as well as reading them, a worker would mainly use a keyboard for writing and a mouse for reading. Therefore, the activity ratio between keystrokes and mouse clicks would loosely reflect the size of the task and thereby relate to the uninterruptible duration as well.

2.3.2 Degree of Engagement in a Task

If a worker is highly motivated, the higher motivation will make his/her degree of engagement in the task higher and will encourage him/her to concentrate on the task longer. Thus, we examined two presumable engagement-related features, i.e., the amount of PC operation activities and the switching frequency of the application software.

The amount of PC operation activity, such as the number of keystrokes, has been reported to be a reverse index of user interruptibility [25], [29]. Moreover, the keystroke interval is known to be shorter under cognitive stress [30]. Thus, we expected that the amount of PC operation activity would reflect the worker’s engagement in the task at hand.

Now, let us suppose that a worker is engaged in a task that requires multiple software applications; e.g., she is writing a document with a word processor and using spreadsheet software to decide which data should be included in the document. In such case, she needs to switch software from time to time. Here, if the user is more concentrated on the task, she may complete the subtask using one piece of software in a shorter time; this in turn may increase the frequency of application switching (AS). Thus, the AS frequency could also be a feature reflecting the degree of engagement.

3. Analysis of Uninterruptible Duration

3.1 Experiment

To verify the relationship between the uninterruptible duration and PC operation features, we logged the PC operation activity of several workers in an actual working scenario.

3.1.1 Participants

Four faculty members and nine university students participated in the experiment. Table 1 summarizes the properties of the participants. All participants were engaged in research activities in information science and had touch typing skills. They spent most of their office hours at their own desk in the laboratory using their assigned PC. Before starting the experiment, they agreed to log their PC operation activity during work hours as volunteers without reward. The experiment was approved by the ethical review committee of the university. The faculty members, i.e., participants 1 to 4, worked on a fairly regular basis, while the students worked in a more flexible manner. In particular, participant 1 had a heavy schedule with many meetings and lectures. As for the students, participants 5 to 9 spent more time working on data analysis using a spreadsheet application, while 10 to 13 were mainly engaged in system development including programming.

3.1.2 Procedure and Tools

We explained the aim of the study, the function of the software, and the recorded information to the participants. Then, we instructed them to install the logging software on the PC they mainly use, and to set the software to be automatically started at PC activation. We did not put any restrictions on them or ask them to conduct any special activity. Since some participants joined and left during the data collection mainly due to their having been transferred, the length of the record naturally varied among the participants. The software worked in the background and counted mouse clicks, mouse wheel rotation, and keystrokes. It also recorded the name of the active window every 0.5 seconds. It estimated and recorded interruptibility in three levels at each sampling time. The features were detected by using the Application Programming Interface (API) functions of the Microsoft Windows OS.

3.1.3 Summary of Collected Data

The time length of the data set for each participant varied from 3 to 15 months because the beginning times of the recordings were different. Although imbalances in the amount of data among the participants might affect the accuracy of the developed model, we basically did not adjust it to avoid unintentionally influencing the data. Instead, we confirmed that similar tendencies could be obtained after equalizing the data amounts. The most frequently observed tasks in the collected data were document writing, data analysis, programming, web browsing, and e-mail. Since we did not pose any restriction on the activity, telephone calls, visitors, and other interruptions occurred in the same way as before the experiment.

We identified uninterrupted durations according to the definition described in Sect. 2.2. We targeted only durations longer than 180 s because of there being less necessity for predicting the durations of immediately ending tasks. As a result, a total of 23,931 uninterrupted durations was obtained. Here, to obtain a more general model, we needed to prevent our analysis from being influenced by the data from specific participants; we thus removed the data if the used software was uncommon (used by fewer than three participants), e.g., sound analyzing software. A total of 22,315
uninterruptible durations remained after 1,616 samples were removed. The average duration of the remaining data was 816.4 s, and the median value was 538.5 s.

In a study that analyzed the time spent editing Wikipedia[31], the average and median durations were found to be 900 s and 600 s. Our collected data appears consistent with this result, once we note that shorter tasks such as web browsing were included in our experiment.

3.2 Analyzed Features

As described in Sect. 2.3, we expected that uninterrupted durations can be predicted by using PC-operation features. Thus, we analyzed the relationship between them.

- Types of application software The application is defined as the software used the longest during the uninterrupted duration. We analyzed the uninterrupted durations for the seven most frequently observed types of application, i.e. Web browser (WB), PDF reader (PR), Word processor (WP), Spreadsheet (SS), Programming (PG), Mail (ML), and Presentation program (PP).
- Key-mouse ratio The activity ratio between keystrokes and mouse clicks (key-mouse ratio) is defined as the ratio between the number of samples when mouse clicks were detected and the number of samples when keystrokes were detected. If the ratio is close to 1.0, it indicates that keyboard and mouse are used equally in a task. A smaller ratio indicates that the work is keyboard-dominant, a larger ratio means mouse-dominant.
- PC operation rate The rate of PC operation activity (PC operation rate) is defined as the ratio of the number of samples in which any operation is detected (keystroke, mouse click, or mouse wheel operation) to the total number of samples during the uninterrupted duration.
- AS frequency The application switching (AS) frequency is the average number of application software switches per minute during the uninterrupted duration. We omitted switching to software which was used for less than 2 s, because such extremely short uses were obviously not for executing a task and were presumably part of a sequence of switchings.

3.3 Results of Analysis

Figures 3 (a) to (d) show the relationships between the uninterrupted duration and the four analyzed features. Except for application types, we divided all the data into six equal-sized bins in accordance with the feature values. The medians of the uninterrupted duration for each bin are shown.

3.3.1 Type of Application Software

Figure 3 (a) shows the variation in the uninterrupted duration for each type of application. To examine the significance of the difference, we applied a one-way Kruskal-Wallis test and detected any significant difference ($\chi^2(6; WB, n = 8217; PR, n = 581; WP, n = 4777; SS, n = 2731; PG, n = 1546; ML, n = 2130; PP, n = 2333) = 1525.8, p < 0.01$). In other words, Internet searches and reading documents take a shorter time on average than writing, data analysis, programming, etc. These results accord with our expectation that the used application software is related to the task size and consequently with the uninterrupted duration.

3.3.2 Key-Mouse Ratio

Figure 3 (b) shows the relationship between the key-mouse ratio and the uninterrupted duration. We applied a one-way Kruskal-Wallis test and detected any significant difference ($\chi^2(5; n = 3719) = 1244.5, p < 0.01$). The subsequent Steel-Dwass tests showed significant differences ($p < 0.01$) between adjacent pairs of bins except for the pair of the 2nd and 3rd left bins. As can be seen from the arched shape of the histogram, the uninterrupted duration tended to be shorter at both higher and lower sides of the key-mouse ratio.
ratio. In other words, the uninterruptible durations were longest in the cases that the workers operated the keyboard and the mouse roughly the same amount of time.

3.3.3 PC Operation Rate

Figure 3 (c) indicates the relationship between the PC operation rate and the uninterruptible duration. After confirming a significant difference ($\chi^2(5, n = 3719) = 1795.3, p < 0.01$) in the Kruskal-Wallis test, we applied Steel-Dwass tests between pairs of adjacent bins. Significant differences ($p < 0.01$) in all the pairs suggested that the uninterruptible duration tends to be longer in the tasks showing higher PC operation rates.

3.3.4 AS Frequency

Figure 3 (d) shows the variation of the uninterruptible durations in relation to the AS frequency. The Kruskal-Wallis test revealed a significant difference ($\chi^2(5, n = 3719) = 3579.9, p < 0.01$). The Steel-Dwass test between pairs of adjacent bins showed significant differences ($p < 0.01$). In the histogram, the uninterruptible duration tended to increase as the AS frequency increased, similar to the PC operation rate.

These results support our supposition that the PC operation rate and AS frequency reflect the degree of worker engagement in the task and have some relation to the uninterruptible duration.

3.4 Individual Variations

To verify the generality of the observed tendencies of the four features, we separately analyzed the data for each participant. Table 2 lists application types in ascending order of uninterruptible duration. The rows of PC operation rate and AS frequency show the correlation coefficients between those features and the uninterruptible durations. In the key-mouse ratio rows, the data is equally divided into three categories (Key-dominant, even, and Mouse-dominant) in terms of the key-mouse ratio (per two bins in Fig. 3 (b)), and the median value of the uninterruptible duration are shown for each bin.

Although there were some individual differences among the participants, we confirmed four common tendencies in the uninterruptible duration as follows. 1) The uninterruptible duration is relatively shorter in the tasks using the Web browser or PDF reader than in the tasks using other software. 2) The uninterruptible duration positively correlates with PC operation rate. 3) It also correlates with AS frequency. 4) The uninterruptible duration tends to get longer in when a worker evenly uses the keyboard and mouse.

We expected to find some sort of difference between
faculty members and students. However, no clear difference was observed between the two groups. Thus, we decided to develop a model on the basis of these four features for predicting the uninterruptible duration.

4. Work Continuance Model

4.1 Design Policy and Outline

Following the line of the previous studies [16]–[19], the current study attempted probabilistic prediction of uninterruptible durations based on their statistical tendency. Here, in the field of reliability engineering, the probability that a system continues operating for $N$ years is estimated to be the $N$th power of the non-failure rate per year of the system. Inspired by that idea, we here propose a model in which the duration of work is determined by the work continuance probability.

Figure 4 conceptually represents the time course of the cumulative work continuance probability. The uninterruptible duration is defined as the time it takes for the cumulative work continuance probability to decrease to 0.5.

In other words, the gradual decrease of the cumulative continuance probability stems from the nature of the probabilistic continuance itself. Although the actual work continuance probability may dynamically change depending on the concentration or fatigue of a worker, we assumed it to be constant for simplicity.

Similar to the non-failure rate in reliability engineering, the higher unit-continuance rate $Q$ reduces the probability that the current task will be discontinued; i.e., it makes the task last longer. We defined the predicted uninterruptible duration as the time until the cumulative work continuance probability $P(N)$ drops to 0.5.

As suggested in Sect. 3, the uninterruptible duration varies in accordance with four features. Thus, the model we propose here determines the work continuance probability (WCP) $Q$ by integrating the four individual continuance probabilities (ICPs), as shown in Fig. 5.

The problem with this model is the quantification of the ICP for each factor. To estimate WCP in a similar way.
as the system operation probability, we need to know the ICP for each affecting factor. However, the only observable probability is the integrated WCP, and we have no way to decompose it into the effects of individual factors from the observed data.

Thus, we calculated the ICP for each feature relative to the overall average probability calculated from the entire data set. This relativizing process allowed us to approximate the individual effect of each factor based on the logged data. The resulting ICPs were then integrated into one WCP, and the uninterruptible duration was calculated on the basis of this WCP.

4.2 Modeling the Effect of Each Factor

The first step in calculating the expected uninterruptible duration is the modeling of the variation in ICP in accordance with the values of the features. Therefore, we fitted the better of either an exponential function or a second-order polynomial function for representing the relationship between each feature and the observed work continuance probability. Figures 6 illustrate the results of the curve fitting to the data divided into 15 bins. The blue dots denote the work continuance probability calculated for each bin of data. The red lines represent the fitting functions, which equal to the continuance probability functions in Fig. 5.

As can be seen from the graphs, the effect of the features on the uninterruptible duration was successfully modeled by the fitting functions (PC operation rate: $R^2 = 0.953$, Key-mouse ratio: $R^2 = 0.876$, AS frequency: $R^2 = 0.890$). Since only the application type is a nominal scale, the median of the probability for each type was chosen for representing the effect.

4.3 Integrating the Effects of Four Factors

The second process is integration of the ICPs into one WCP. However, this poses three problems.

The first problem regards the separation of the effects of the individual features, as mentioned in Sect. 4.1. Since each continuance probability in Fig. 6 is calculated from actual work data, it is naturally affected by multiple factors. For example, the calculated continuance probability for a specific PC operation rate, e.g. 10%, is also affected by the type of application software in use, e.g. a word processor. Therefore, if the effects of the four factors are simply multiplied, the influence of each factor will be reflected more than once. To avoid this problem, we normalized each ICP by dividing by the average probability for the entire data set under the assumption that the number of the data is big enough for canceling the bias in the effect of the contaminant factors. This relative representation of ICP allowed us to model the effect of a feature without it being affected by the other factors.

The second problem is the exponential increase of the uninterruptible duration. Since the uninterruptible duration was calculated on the basis of the $N$th power of WCP, as the probability gets closer to 1, the uninterruptible duration excessively increases. Therefore, we applied a logarithmic transformation to each ICP before normalization.

The last problem is the accumulation of effects. In
other words, if the ICPs for two features both rise, the total WCP will become much higher. Thus, we empirically looked for an appropriate formulation and eventually chose a multiplicative integration method, which best reproduced the statistical distribution of WCP. Equation (2) represents the calculation for multiplicative integration.

\[
\log_{10} Q = \log_{10} p_{\text{app}} \times \log_{10} p_{\text{keymouse}} \\
\times \log_{10} p_{\text{pc}} \times \log_{10} p_{\text{mouse}}
\]

(2)

The predicted WCP value \( Q \) was calculated by multiplying the logarithms of the increasing/decreasing effects of the features. As described in the previous subsection, we transformed ICPs \( (p_{\text{app}}, p_{\text{keymouse}}, p_{\text{pc}}, p_{\text{mouse}}) \) into their logarithms and normalized with the average of the actual continuance probability for the entire training data set \( p_{\text{ave}} \).

Here, the continuance probability function for calculating each ICP is given by fitting the function to the training data set, as shown in Fig. 6. Thus, Eq. (2) allows calculating WCP from the values of four features. Consequently, the uninterruptible duration is predicted as the time when the \( N_{th} \) power of WCP drops to 0.5, as described in Sect. 4.1.

5. Prediction of Uninterruptible Durations

To examine the adequateness of the probabilistic work continuance model, we first tried to predict the uninterruptible durations by fitting it to the whole data set. Next, to examine the feasibility in predicting the uninterruptible durations for the independent data set, we conducted cross-validation across the subjects. Then, to simulate real-time prediction in an actual working scenario, we tried to predict the lengths of uninterrupted durations by using only the data available at a moment 3 minutes after their onset.

5.1 Prediction Using Whole Data Set

Figure 7 shows the distributions of the actual and predicted uninterruptible durations. The actual distribution appears to follow a log-normal distribution. Miyagi et al. reported that the time for arithmetic calculations followed a log-normal distribution[17], which is similar to the system available time in reliability engineering. Our collected data appears to be consistent with their results.

As demonstrated in the graph, the distribution of uninterrupted durations was successfully reproduced for the most part. On the other hand, uninterrupted durations longer than 30 minutes were not adequately predicted. In other words, the approximation of the synergistic effect among multiple features appeared insufficient for especially long-lasting tasks.

Table 3 shows the medians of the actual and predicted uninterrupted durations and their correlation coefficients. The median of the errors in predicting each uninterrupted duration was 228 s. Although the predicted durations have some errors, they showed positive correlation with the actual durations. These results suggest the feasibility of using the work continuance model for predicting uninterrupted durations.

5.2 Model Validation

To validate the prediction power of the proposed model for an independent data set, we conducted a leave-one-out cross-validation among the subjects. The entire data was divided into 13 subsets for each participant, and the uninterrupted durations in each subset were predicted using the model fitted to the data for the remaining 12 participants.

Table 4 shows the median of the actual uninterrupted duration for each participant, the median of the predicted duration by using the model fitted to the other 12 participants, and the correlation coefficient between the actual and predicted uninterrupted durations. Although the error levels varied from one participant to the next, the correlation coefficients were positive and greater than 0.3 for all participants. This suggests the existence of a common tendency across the participants between the uninterrupted duration and PC operation features. Furthermore, we conducted the cross-validation again after equalizing the data amount for each participant by truncating the latter part of the greater amount of the data subsets. Positive correlations between the actual and predicted durations remained after the equalization.

In the case of Participant 1, a professor who has a full work schedule, the actual durations were shorter than the predicted ones. Additionally, participants 12 and 13 had lower correlation coefficients than the other participants. The conceivable causes for these differences will be discussed in Sects. 6.1 and 6.2.

<table>
<thead>
<tr>
<th>Median of actual uninterrupted durations (s)</th>
<th>Median of predicted durations (s)</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>238.5</td>
<td>249.1</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Fig. 7 Histograms of actual uninterrupted durations and predicted durations by using multiplicative integration.
Table 4 Results of leave-one-out cross-validation. Each prediction was conducted using a model fitted to the data for other participants. Numbers in parentheses represent the correlation coefficients obtained after equalizing the amount of data for each participant. Each duration represents the median value for the condition.

<table>
<thead>
<tr>
<th>Participant No.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual duration</td>
<td>635.5</td>
<td>685</td>
<td>562</td>
<td>467.5</td>
<td>532</td>
<td>675.5</td>
<td>660</td>
<td>530</td>
<td>439.8</td>
<td>519</td>
<td>507.5</td>
<td>417</td>
<td>512</td>
</tr>
<tr>
<td>Predicted duration</td>
<td>807.3</td>
<td>638.0</td>
<td>467.0</td>
<td>494.9</td>
<td>601.0</td>
<td>552.3</td>
<td>482.3</td>
<td>484.8</td>
<td>450.5</td>
<td>405.4</td>
<td>418.1</td>
<td>383.1</td>
<td>367.9</td>
</tr>
<tr>
<td>Correlation coefficient</td>
<td>.33</td>
<td>.53</td>
<td>.44</td>
<td>.40</td>
<td>.48</td>
<td>.54</td>
<td>.57</td>
<td>.48</td>
<td>.51</td>
<td>.41</td>
<td>.57</td>
<td>.36</td>
<td>.34</td>
</tr>
<tr>
<td>(after equalization)</td>
<td>(.28)</td>
<td>(.51)</td>
<td>(.30)</td>
<td>(.40)</td>
<td>(.50)</td>
<td>(.54)</td>
<td>(.58)</td>
<td>(.46)</td>
<td>(.51)</td>
<td>(.40)</td>
<td>(.50)</td>
<td>(.26)</td>
<td>(.55)</td>
</tr>
</tbody>
</table>

Fig. 8 Histograms of actual uninterruptible durations and durations predicted by using data available 3 minutes after onset of uninterruptible duration.

Table 5 Actual durations, durations predicted by using the data available 3 minutes after onset of uninterruptible duration, and their correlation coefficient.

<table>
<thead>
<tr>
<th>Median of actual uninterruptible durations (s)</th>
<th>Median of predicted durations (s)</th>
<th>Correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>538.5</td>
<td>469.1</td>
<td>0.16</td>
</tr>
</tbody>
</table>

5.3 Prediction 3 Minutes after Onset of Uninterruptible Duration

Although the uninterruptible durations were successfully predicted, the expected time at which the user would become interruptible should be predicted in advance if a system based on our model is to be of practical use. Therefore, to simulate real-time prediction, we tried to predict the lengths of uninterruptible durations by using only data available 3 minutes after the onset of the uninterruptible duration.

Figure 8 shows the histograms of the actual and predicted uninterruptible durations, and Table 5 summarizes the medians of the actual and predicted uninterruptible durations and their correlation coefficients.

Similar to the results predicted using the entire data, the distribution patterns of the actual and predicted durations resembled each other. Compared with the results of using all the data, the difference between the medians of actual and predicted uninterruptible durations increased from 19.4 s to 69.4 s (Tables 3 and 5). The correlation coefficient with the actual uninterruptible duration also decreased from 0.39 to 0.16. Although the prediction error increased, the positive correlation demonstrated that the probabilistic work continuance model can predict the uninterruptible durations to some extent using only data available 3 minutes after the onset of the task.

6. Discussion

6.1 Affecting Factors on Uninterruptible Duration

As described in Sect. 2, we assumed that two factors affect the uninterruptible duration: the size of the task at hand and the degree of worker engagement. However, another factor exists, i.e. externally forced disruptions such as scheduled meetings and visitors. In reality, the proposed model over-predicted the uninterruptible durations of Participant 1, who is a laboratory supervisor having a busier schedule than those of students and assistant professors (Table 4). Frequent external disruptions might have forced him to switch his task regardless of the task state or degree of his engagement. Thus, this external factor should be taken into account. We speculate that integrating online calendar information into our model will improve the prediction accuracy because many business-related events are scheduled in advance.

In addition, it had been reported that the most prevalent distractions are surrounding noise, e.g. conversations occurring around the worker [1]. Although we have no way to assess how much these indirect factors affected our results, they are also worth considering in the future.

6.2 Approximation of Affecting Factors by Using PC-Related Features

The prediction results suggested that the proposed four features approximated the nature of PC tasks well for most of the data, but there were some cases in which the predictions were less effective. For example, Participant 13 occasionally opened two browser windows at the same time and repeatedly switched between them. However, our system did not count such switching as AS; thus, the system might have estimated the degree of the engagement as inadequately low. This case can be resolved by taking switching between the same applications into account.
Participant 12 used an additional PC other than his main unit to examine the program he was developing. However, this study predicted the uninterruptible duration only on the basis of activities detected on the main PC. Therefore, transitions to work on another PC were deemed task terminations or suspensions. To deal with this issue, activities across multiple PCs via a computer network should be integrated.

6.3 Individual Differences in Working Style

As shown in the cross-validation, our model successfully predicted the uninterruptible durations of a person not included in the training data set. However, Table 4 suggests that the prediction error varied greatly between the participants. This might be attributed to individual variations in working style. For instance, when using a word processor, the “copy” operation can be performed using keyboard-dominant operations, i.e. a shortcut key, or with a mouse-dominant operation. We could also divide workers into those who open more windows at a time and those who prefer to perform their tasks with a maximized single window. Their operation behavior might vary with their working style. Although in this study we aimed at obtaining a general model and conducted cross-validation without categorizing the participants, we believe that personalized models adapted to individual working styles would enable a better approximation.

6.4 Further Challenges and Applications of the Model

We attempted to make a prediction 3 minutes after the onset of the uninterruptible duration, as described in Sect. 5.3. Although this worked to some extent, the correlation coefficient decreased to 0.16 from 0.39 for the whole data set. Two major causes are conceivable. One is that the worker's engagement changed after the prediction. One is that the worker switched to a different task without taking a break after the prediction. We confirmed some cases supposed to be consecutive multiple tasks such as Web browsing followed by checking e-mail. However, it is not possible to predict future task transitions. Another cause might have been that the worker’s engagement changed after the prediction. Our model assumes that the cumulative work continuance probability monotonically decreases over time. In other words, our model does not allow ICPs to change during a task. Therefore, the prediction error could have increased when the degree of engagement changed after the prediction. One solution to deal with these issues may be to predict again when the system detects the change of task type or change in degree of engagement.

A potential application of our model is mediation between the interrupter and interruptee in remote-work environments. One presumable approach is to develop an intelligent agent, which accepts an interrupter’s request and arranges a time suitable for communication based on the prediction result for the interruptee. A more direct means would be to send the predicted time at which the interruptee becomes interruptible to the interrupter over a computer network. This would enable the interrupter to devote the time spent waiting to his/her own work, and thus, it would improve the total work efficiency of the distributed office.

However, prediction errors are unavoidable with probabilistic models; i.e., the median of the error in this study was 228 s even when using the whole data set. Thus, we gave a questionnaire to the participants asking their preferences as to how the prospective interruptible time is presented. As a result, 6 out of 7 respondees preferred to have a clue for judging whether to interrupt immediately or later, rather than a display of the predicted time on the second time scale. To support interrupter judgments, a probabilistic display such as one stating “The probability of terminating the task within the next 5 minutes is N%” may be more useful than displaying simply the predicted time.

6.5 Limitations

We defined the uninterruptible duration by using an interruptibility estimation method based on PC operation [25]. The reported precision of the method was 0.48 at three levels and 0.69 at two levels. Thus, this machine-estimated uninterruptible duration might be substantially different from the actual one the user felt. Explorations aimed at finding another index for a stricter determination of uninterruptible durations and validation of the proposed model for that index are needed.

Finally, in this study, we examined the feasibility of predicting uninterruptible durations of 13 participants working or studying at university. Although we selected features of PC operation, which supposedly have a general tendency across task types, the homogeneity of the participants prevents us from prospecting the applicable work scenarios for the proposed model. To support not-disruptive communication in the wider workplace, we need to verify the model in other types of work situations and need to improve it.

7. Conclusion

With the aim of predicting prospective interruptible times of PC workers, this study analyzed data collected from 13 participants and revealed the relationship between uninterrupted durations and four features. We proposed a probabilistic work continuance model and examined the feasibility of making predictions with it. The contribution of this study is to demonstrate the feasibility of predicting the uninterruptible durations and their statistical distribution in an actual working scenario.

Our future studies will verify the generality of the model, improve the prediction accuracy, and develop application systems for improving the productivity of office workers engaged in intellectual activities especially in remote environments.
Acknowledgments

We thank Mr. Hokuto Iga for his contribution in the early stage of this work. This work was partly supported by funds for smart space technology toward a sustainable society from the Ministry of Education, Culture, Sports, Science and Technology, Japan and funds from the Japan Society for the Promotion of Science (KAKENHI).

References

Shota Shiratori received his B.S. and M.S. degree from Tokyo University of Agriculture and Technology, Japan, in 2017 and 2019. His master's thesis was the development of a model for predicting uninterruptible durations of office workers. He is currently working as a research engineer at an automobile manufacturer.

Yuichiro Fujimoto received the B.S. degree in engineering from Osaka University, Osaka, Japan, in 2010 and the M.S. degree and Ph.D. in engineering from Nara Institute of Science and Technology, Nara, Japan, in 2012 and 2015. From 2015 to 2018, he was an Assistant Professor with the Engineering Department, Tokyo University of Agriculture and Technology in Japan. He is currently working as an Assistant Professor in the Division of Information Science, Nara Institute of Science and Technology in Japan. His research interests include analysis of humans from 3D information and spatial augmented reality.

Kinya Fujita received his Ph.D. degree from Keio University, Japan, in 1988. After working at Sagami Institute of Technology, Tohoku University, and Iwate University, he joined Tokyo University of Agriculture and Technology. He is currently working as a Professor in the Department of Computer and Information Sciences. His research interest is the development of smart human interface systems including interruption controls and remote communication systems. He is a member of IEEE, and IPSJ, and a board member of HIS and VRSJ.