# User Identification Based on Mouse Operation Toward Automatic Home Appliance Configuration

Kyohei Suda*, Shigemi Ishida*, Hiroshi Inamura*
*Graduate School/School of Systems Information Science, Future University Hakodate, Japan
Email: {g2122037, ish, inamura}@fun.ac.jp

***Abstract*—Smart home appliances such as coffee machines with precise taste control are prevalent nowadays. These home appliances are often shared with family members, while the individual has a preferred configuration. In this paper, we propose an automatic home-appliance configurator that identifies a user and automatically configures the home appliance based on the user's preference. As a first step of the proposed system, we present a user identification system based on mouse operations to wake a computer from a sleep state, which is an example of daily home appliance control operations. The key idea is to extract the features of user-specific mouse operations and identify the user by supervised learning. We collected mouse operation logs from 24 subjects and confirmed that the proposed system identified users with an accuracy of 93.5%.**

***Index Terms*—user identification, home appliance operation, mouse operation, supervised learning.**

## I. Introduction

Recently, smart home appliances, such as coffee machines with precise taste control and toaster with baking quality control, are becoming prevalent due to the advances in information and communication technologies. These home appliances are shared among family members and allow users to configure settings based on individual's preference. We usually reconfigure the home appliance with the custom configuration when we use the home appliance. This research aims to realize an automatic home-appliance configurator that identifies the user and automatically configure the home appliance.

To identify users, gesture-based authentication methods have been proposed. The gesture-based authentication methods, however, rely on additional devices or on specific gestures that are apart from daily gestures to achieve high accuracy [1], [2]. These are also true for mouse operations [3]–[5].

We are developing a user identification system relying on home-appliance operations. As the first step for realizing the user identification system, this paper focuses on mouse operations to wake a computer from a sleep state, where user-specific movements can be observed. We collected mouse operation logs from 24 subjects and confirmed that our user identification system successfully identified users with a mean accuracy of 93.5%. For seven of the 24 subjects, data were collected for 10 days to evaluate the performance degradation over time passage. The results revealed that we need to retrain the learning model using the latest two to five days to maintain high accuracy.

## II. User Identification System Based on Mouse Operation

We assume that users wake a computer from a sleep state by (1) moving a computer mouse, (2) entering or selecting the user name, and (3) entering a password. When a user moves a computer mouse, our user identification system identifies the user and automatically select or enter the user name.

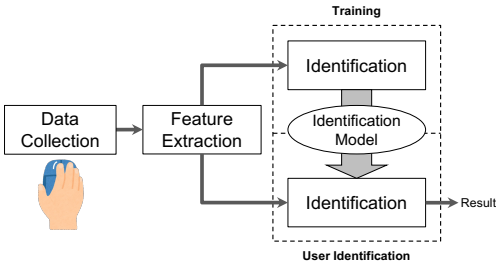


Fig. 1. System overview

TABLE I
Features

| Feature | Description |
|---|---|
| Operation time length | Time from the beginning to the end of the mouse operation |
| Cursor speed statistics | Mean, median, and standard deviation of the mouse cursor speed |
| Trajectory range | Cursor trajectory ranges in $x$-axis and $y$-axis |
| Total distance traveled | Total length of trajectory |
| Number of data samples | The number of OS notifications of the mouse cursor movements |

Figure 1 shows the overview of our user identification system. The user identification system consists of the data collection, feature extraction, and identification blocks. We use supervised learning that requires model training before the actual use. During the training phase, mouse operation logs are collected to train an identification model. When a sufficient amount of data is collected, the system switches to the user identification phase. We use the authentication mechanism as it is because our system aims user identification.

### A. Data Collection

The data collection block collects trajectory of a mouse cursor on a computer display during the user's mouse operation. The trajectory data samples are collected at the timing when the cursor movement is notified by the operating system (OS).

### B. Feature Extraction

The feature extraction block extracts five kinds of features shown in Table I to calculate a feature vector for each trial, which results in eight-dimensional feature vectors.


doi: 10.1109/GCCE56475.2022.10014053

### C. Identification

The identification block identifies the user as a multi-class classification problem using a supervised learning model with the feature vectors calculated in the feature extraction block. Because the dimension of the feature vector is small, we used a Support Vector Machine (SVM) model with a linear kernel in this paper. We used default parameters of the scikit-learn library in an actual implementation. There is room to discuss the classification model.

## III. Evaluation

### A. Experiment Environment

We collected actual mouse operation logs to evaluate the performance of our user identification system. We developed a computer-wake-up emulator and asked 24 subjects to move a mouse such that they wake a computer to collect mouse trajectory data. The data collection was repeated 100 times for each subject. For seven of the 24 subjects, data were collected for 10 days separated by up to 178 days. The experiments were conducted under permission from the ethics committee of Future University Hakodate (permission #2021004).

Although our goal is identifying a user rather than authenticating the user, successive failures spoil the usability. We therefore defined a target accuracy based on the success rate within a specific number of trials. Let $a$ be the user identification accuracy. The success rate $p$ of user identification within $n$ trials is calculated as

$$p = 1 - (1 - a)^n. \tag{1}$$

We derive the lower bound of accuracy as

$$a \geq 1 - \sqrt[n]{1 - p}. \tag{2}$$

In this paper, we assumed the number of allowable retries is two, substituted $n = 3$ and $p = 0.99$ to Eq. (2) deriving $a \geq 0.78$. The target accuracy was defined as 0.80.

### B. User Identification Accuracy

Considering the use case of our system, we limit the number of users to identify in the evaluation. We randomly selected four of the 24 subjects and performed a 10-fold cross-validation to calculate the user identification accuracy. The user identification with the random subjects was repeated 100 times and the mean accuracy was calculated.

The mean user identification accuracy was 93.5%. We also calculated the user identification accuracy while changing the number of data from 10 to 100 used in the cross-validation and derived accuracies greater than 90% for all the cases.

### C. Performance Degradation Over Time

We evaluated the performance degradation over time without the identification model update. In the training phase, 10 trials on the first day for each subject were selected to train the user identification model. We performed user identification with the remaining data and averaged the user identification accuracy for each day.

Figure 2 shows the user identification accuracy as a function of the day on which evaluation data were derived. The result shows that time passage from the data collection for identification model training decreased the identification accuracy.

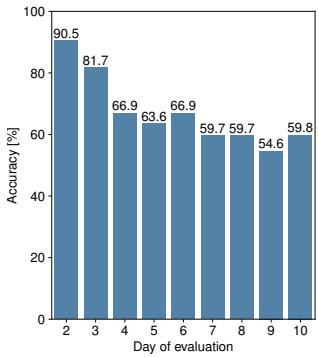


Fig. 2. User identification accuracy as a function of the day on which evaluation data were derived

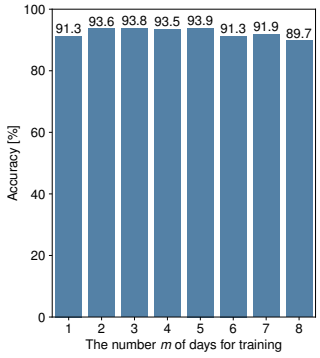


Fig. 3. User identification accuracy as a function of the number $m$ of days for training

To maintain high accuracy, the identification model update is mandatory. We evaluated the influence of the data amount used in the model update. When we evaluate the identification accuracy on the $n$-th day, we trained the model with the first 10-trial data derived from each of the $m$ days immediately before the $n$-th day. For each of $m$ $(1 \leq m \leq 8)$, the mean identification accuracy was calculated. Note that the range of $n$ is $m + 1 \leq n \leq 10$.

Figure 3 shows the user identification accuracy as a function of the number $m$ of days for training. The highest accuracy was derived when $m = 5$, though the accuracy was almost the same when $2 \leq m \leq 5$.

## IV. Summary

In this paper, we proposed the user identification method based on home-appliance control operations. Specifically, we focus on mouse operation to wake a computer from a sleep state in this paper. We collected actual mouse operation logs from 24 subjects and evaluated the user identification performance. The results show that our user identification system successfully identified the user with an accuracy of 93.5% for randomly-selected four subjects. We plan to extend this method to support various operations on various home appliances such as a coffee maker and toaster.

## Acknowledgment

This work was supported in part by JSPS KAKENHI Grant Numbers JP20KK0258 and JP21K11847 as well as the Cooperative Research Project of RIEC, Tohoku University.

## References

[1] P. Bours and C. J. Fullu, "A login system using mouse dynamics," *Int. Conf. Intell. Inf. Hiding Multimedia Signal Process.*, pp. 1072–1077, Sep. 2009.

[2] S. Mare, R. Rawassizadeh, R. Peterson *et al.*, "Continuous smartphone authentication using wristbands," *Workshop Usable Security*, 2019.

[3] J. Zhao and J. Tanaka, "Hand gesture authentication using depth camera," *Adv. Inf. Commun. Netw.*, Springer, vol. 887, pp. 641–654, 2019.

[4] D. Qin, S. Fu, G. Amariucai *et al.*, "MAUSPAD: mouse-based authentication using segmentation-based, progress-adjusted DTW," *IEEE TrustCom*, pp. 425–433, Dec. 2020.

[5] S.-M. Shin and M. Kim, "PC user authentication using hand gesture recognition and challenge-response," *J. Adv. Inf. Technol. Convergence*, vol. 8, no. 2, pp. 79–87, Dec. 2018.