

# キーログを用いた英単語タイピングの確信判定

丸市 賢功<sup>†</sup> 坂本 周司<sup>†</sup> Olivier Augereau<sup>††</sup> 黄瀬 浩一<sup>†††</sup>

<sup>†</sup> 大阪府立大学工学域 〒599-8531 大阪府堺市中区学園町 1-1

<sup>††</sup> 大阪府立大学大学院工学研究科 〒599-8531 大阪府堺市中区学園町 1-1

E-mail: <sup>†</sup>{maruichi,sakamoto}@m.cs.osakafu-u.ac.jp, <sup>††</sup>augereau.o@gmail.com, <sup>†††</sup>kise@cs.osakafu-u.ac.jp

あらまし 英単語の学習において、復習は重要なプロセスである。一般的に復習は間違えた英単語に対して行われる。しかし、正解はしたものの正解であると確信が持てなかった単語を見落としては十分復習したとは言いきれない。そこで本研究では、英単語タイピング問題解答時の確信の有無を判定し、それを解答の正誤とともにユーザに提示することで、より効果的な復習の実現を目指す。確信の判定にはキーログから取得したタイピングに関する情報を利用する。実験参加者 12 名に英単語タイピング問題 120 問に解答してもらい、その際の解答に対する確信の有無を判定した。その結果、ユーザ依存で学習を行った場合は 91.3%、ユーザ非依存で学習を行った場合には 89.8% の平均判定精度が得られた。

キーワード 英語学習, 語彙学習, キーログ, 確信判定, SVM

## Confidence Estimation for English Word Typing by Using a Key-Logger

Takanori MARUICHI<sup>†</sup>, Shuji SAKAMOTO<sup>†</sup>, Olivier AUGEREAU<sup>††</sup>, and Koichi KISE<sup>†††</sup>

<sup>†</sup> College of Engineering, Osaka Prefecture University 1-1 Gakuen-cho, Naka-ku, Sakai, Osaka 599-8531, Japan

<sup>††</sup> Graduate School of Engineering, Osaka Prefecture University 1-1 Gakuen-cho, Naka-ku, Sakai, Osaka 599-8531, Japan

E-mail: <sup>†</sup>{maruichi,sakamoto}@m.cs.osakafu-u.ac.jp, <sup>††</sup>augereau.o@gmail.com, <sup>†††</sup>kise@cs.osakafu-u.ac.jp

**Abstract** Reviewing plays an important role in English vocabulary acquisition. However, we tend to overlook the words that were correctly answered without confident. In this research, we estimate confidence in answering English word typing questions to let learners reviewing more effectively. We use keystroke information acquired by a key-logger to estimate the confidence. In order to evaluate the proposed method, we employed the dataset obtained from 12 participants answering 120 questions. From the experimental results on estimating the confidence, the proposed method has achieved the average accuracy of 91.3% in the case of user-dependent learning, and 89.8% in the case of user-independent learning.

**Key words** English learning, vocabulary acquisition, keystroke analysis, confidence estimation, SVM

### 1. ま え が き

国際化の影響により、英語学習への関心が高まっている。英語の習得には英単語の暗記、英文法の理解が必要である。その中でも、英単語の暗記は、英文法の理解に比べて、暗記の確認を反復して行う必要があるため、非常に手間のかかる地道な作業である。

地道な努力を要する英単語の暗記を、ゲーム感覚で効率的に行えるツールとしてスマートフォンなどで利用できる英単語アプリが注目を集めている。一般的に英単語アプリでは、英単語を暗記し、英単語の定着度合いをテスト形式で確認できる。そ

の中から、間違えた英単語を復習リストに登録して、復習リストを復習し、テストに再挑戦する学習方法を採用している。

しかし、間違えた英単語だけを復習するだけでは、まだ不十分である。なぜなら、正解した英単語にも復習が必要な場合があるからである。具体的には、確信が持てずに正解した英単語を復習せず放置していると、再びテストを行った際には間違えてしまう恐れがある。そのため、確信が持てずに正解した英単語も復習リストに加える必要がある。また、正解であると確信をもって解答したにも関わらず間違えてしまった場合にも、英単語の意味や綴りを勘違いしている可能性がある。この場合についても、注意して復習に取り組みなければ再び間違えてしま

う。したがって、確信をもって解答したにも関わらず間違えてしまった英単語については、復習リストでも注意を促す必要がある。

そこで、本研究では、英単語アプリにおける問題解答時の確信を判定し、問題の正誤と合わせて提示することで、復習をより効果的に進めることのできる英単語学習システムの構築を目指す。確信を判定する手法として、タイピングに関する情報に着目する。タイピングに関する情報には人の心理状態に関する情報が含まれていることが先行研究から明らかになっている [4-7]。具体的には、解答に確信が持てない場合、迷いが生じるため、タイピング速度が遅くなり、タイプする文字数も多くなる。したがって、確信の判定には、タイピングに関する情報が利用できるを考える。

タイピングに関する情報を取得するためにキーログを用いる。キーログとは、入力したキーやキーを入力した時間などを記録することができるソフトウェアのことである。キーログは、現在のスマートフォンに搭載されている機能のみで実装可能であり、問題解答時の妨げにもならないという利点がある。

本稿では、キーログを用いて、英単語タイピング問題解答時の確信を判定する手法を提案する。提案手法では、キーログから取得したタイピングに関する情報から特徴量を抽出し、SVMを用いて確信を2クラスに分類した。実験参加者12名に対し、実験を行った結果、提案手法の平均判定精度は、各実験参加者のデータセットについて特徴量の選択と分類を行ったユーザ依存の場合には、91.2%、データセット全体について特徴量の選択と分類を行ったユーザ非依存の場合では89.8%であった。なお、本研究は、大阪府立大学工学研究科倫理委員会の承認を得ている事を付記しておく。

## 2. 関連研究

キーログから取得したタイピングに関する情報が人の心理状態の推定に有効であるであることがこれまでの研究から分かっている。本節では、キーログを用いた研究や確信判定に関する研究について述べる。

まず、キーログを用いた研究について述べる。キーログから取得したタイピングに関する情報は個人認証の分野で数多く取り上げられてきた [1, 2]。しかし、人の心理状態の判定に利用している研究も数は少ないが存在している。中村ら [3] は、e-learning 教材において設問に解答する際の顔画像、マウス操作、キー入力といった情報からユーザの教材に対する主観的難易度を判定している。キー入力に関する特徴としては、キー入力の平均発生間隔のみを使用している。Epp ら [4] は、キーログから取得したタイピングに関する情報から、ユーザの心理状態を判定する手法を提案している。ただし、対象としているのは、用意された文章をそのままタイプ入力するというタスクで、問題に解答するような形式はとっていない。Nahin ら [5] は、Epp らの同様のタスクで、Epp らの用いた特徴量に加えて、文書情報を用いることで、心理状態を判定している。Allen ら [6] は、タイピングでエッセイを作成している際にキーログから取得した情報と文書情報から退屈度、集中度を判定する手

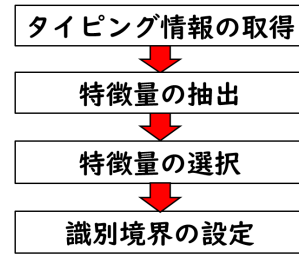


図1 処理の流れ

法を提案している。退屈度の判定ではタイピングの停止時間とタイプした文字数が、集中度の判定ではそれらに加えタイピング速度とミスタイプの回数が有効であることが述べられている。Ghosh ら [7] は、スマートフォンでタイピングを行った際にキーログから取得した情報を用いて、ユーザのストレス度合いなどを判定する手法を提案している。Ghosh らの手法においても、タイピング速度、タイプした文字数、そしてタイプミスの回数が有効であることが示されている。

次に、問題解答時の確信を取り上げた研究について述べる。浅井ら [8] は、タブレットで数学の問題に解答している際の筆圧や筆記速度、筆記間隔などの情報を用いて、つまづいた箇所を検出する手法を提案している。三木ら [9] は、英単語の並べ替え問題を解答している際のマウスの軌跡情報や解答時間を用いて迷いの判定を行っている。山田ら [10] は、多肢選択問題を解答している際の視点情報を用いて、解答している問題に対する確信度を判定する手法を提案している。山田らの手法では、確信度判定のために、視点情報から抽出した特徴量や解答時間を使用している。

しかし、キーログから取得したタイピングに関する情報と問題解答時の確信についての関係性について触れている研究は今のところ確認されていない。

## 3. 提案手法

そこで我々は、キーログから取得したタイピングに関する情報を用いて、英単語タイピング問題解答時の確信を判定する手法を提案する。

処理の流れを図1に示す。まず、キーログからタイピング情報を取得し、特徴量を抽出する。次に、抽出した特徴量から有効なものを選択し、SVMを用いて学習・識別を行う。

### 3.1 タイピング情報の取得

スマートフォンアプリを用いてタイピング情報の取得する。取得した情報とは、具体的には、タップした箇所の中心座標といった位置に関する情報や図2のように、タップしてから指を離すまでの時間、指を離してから次にタップするまでの時間といった時間に関する情報である。

次に、取得した情報に含まれる座標から、どのキーをタップしたのかを特定し、キーに対応する文字情報を含んだタイピング情報に変換する。キーが配置されていない座標については、どのキーにも対応していないので、タイピング情報には含めないように設定する。

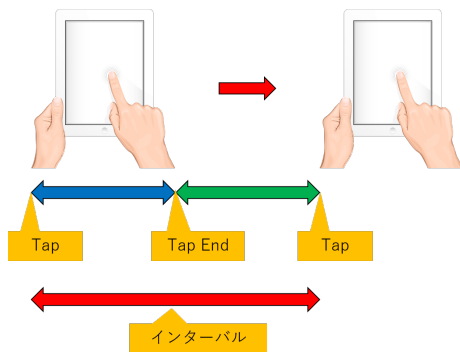


図2 インターバルの設定

### 3.2 特徴量の抽出

取得したタイピング情報から、問題ごとに特徴量を抽出する。抽出する特徴量を表1に示す。特徴量は、上から順に、時間に関連するものが8個、キー情報に関連するものが2個、正解の単語に関するものが2個、キー情報と正解の単語どちらにも関連するものが1個の合計13個である。解答時間とは、エンターキーをタップしてから次にエンターキーをタップするまでの時間、インターバルとは、図2に示す通り、キーをタップしてから次のキーをタップするまでの時間とする。最初のインターバルとは、問題が提示されてからキーを押すまでにかかった時間、最終のインターバルとは、最後の文字をタイプしてから、エンターキーを押すまでの時間のこととする。正解単語の使用頻度には、Word Count<sup>(注1)</sup>というWebサイトのデータを利用した。また、編集距離とは、タイプした文字列と正解単語がどれ程異なるのかを示す指標で、具体的には、文字の挿入、削除、置換といった操作の数のことである[11]。

### 3.3 特徴量の選択

抽出された特徴量について、有効な特徴量の選択を行う。本手法では全ての特徴量の組合せについて探索を行う。その上で、選択される特徴量の数が最も少なく、かつ判定精度が最も高い特徴量の組合せを選択する。

### 3.4 識別境界の設定

選択された特徴量を基に、SVMを用いて識別境界の設定を行い、表2の2クラスに分類する。

## 4. 実験

提案手法が確信判定に有効であるかを検証するための実験について述べ、実験によって判定された確信判定の精度について検証を行う。

#### 4.1 実験条件・実験手順

実験参加者の日本人大学生12名(男性9名、女性3名)に図3のような形式の問題に解答するよう指示し、その際のタイピング情報を記録した。問題にはリクルート社の英単語アプリ・スタディサプリ英単語<sup>(注2)</sup>のWriting問題の中から12セクション、合計120問を使用した。このとき単語の難易度には偏りが出ないように注意して選択した。

表1 特徴量

No	特徴量
1	解答時間 [ $\mu$ s]
2	インターバルの平均 [ $\mu$ s]
3	インターバルの標準偏差
4	インターバルの最大値 [ $\mu$ s]
5	インターバルの最小値 [ $\mu$ s]
6	インターバルの中央値 [ $\mu$ s]
7	最初のインターバル [ $\mu$ s]
8	最終のインターバル [ $\mu$ s]
9	タイプした文字数
10	デリートキーをタイプした回数
11	正解単語の文字数
12	正解単語の使用頻度
13	編集距離

表2 クラス分類

ラベル	クラス
0	解答が正解であるという確信が持てた
1	解答が正解であるという確信が持てなかった

実験手順について述べる。まず、Auto Touch<sup>(注3)</sup>というスマートフォンアプリを起動して問題を1セクション解答するよう指示した。解答が終わった後に各問題について確信が持てたかどうかのアンケートを行う。この手順を12セクション全てに対して行った。なお、解答するセクションの順番は実験参加者ごとにランダムになるように設定した。

#### 4.2 評価方法

評価方法について述べる。本稿では、2つの場合について検証を行う。

1つ目はユーザ依存での検証である。ユーザ依存とは、識別器の学習に際して、特定のユーザのデータのみを用い、同一ユーザに対して確信判定を行う場合をいう。ユーザ依存で学習された識別器を他のユーザに用いることはできない。ユーザ依存での学習を行うことで、そのユーザが解答する際に確信がある場合あるいは確信がない場合にどのような振舞いを行うのかを学習することができる。その振舞いの差が識別できれば、確信判定がより容易になる。また、全ての実験参加者について、各特徴量の選択回数を計算することで、ユーザの個人差に依らず、どの特徴量が判定に寄与しているのかを求めることができる。

評価方法については、図4のように、各実験参加者について、1セクション分のデータ10問をテストデータ、残りのセクションを学習データとして分割し、全てのセクションについて評価を行うLeave-one-section-out Cross-validationを用いた。

2つ目はユーザ非依存での検証である。ユーザ非依存とは、学習データを取得したユーザとは異なるユーザに対して識別器を適用する場合である。ユーザ非依存で確信の判定が可能であれば、まだ一度も問題を解いたことのない新しいユーザの確信を判定する場合でも、本実験で作成したデータセットを用いて判定を行うことができる。したがって、実用上はユーザ非依存

(注1): <http://www.wordcount.org/>

(注2): <https://eitangosapuri.jp/>

(注3): <https://autotouch.net/>



図 3 問題のフォーマット

表 3 事前確率

実験参加者	事前確率 [%]
A	65.8
B	63.3
C	43.3
D	32.5
E	38.3
F	56.7
G	67
H	50
I	41.7
J	65
K	35.8
L	35.8
平均	49.6

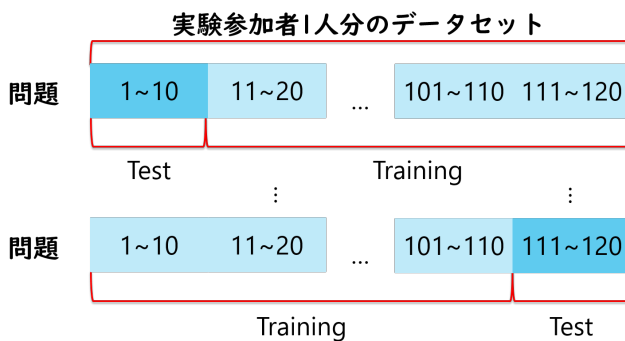


図 4 Leave-one-section-out Cross-validation

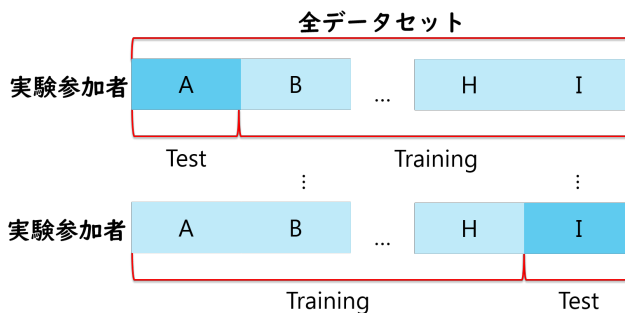


図 5 Leave-one-user-out Cross-validation

での判定精度も高ければよりよい手法であるといえる。したがって、ユーザ非依存場合についても検証を行っておく必要がある。

評価方法については、図 5 のように、ユーザ 1 人分のデータ 120 問をテストデータ、残りのユーザのデータを学習データとして、全てのユーザについて評価を行う Leave-one-user-out Cross-validation を用いた。

これら 2 つの評価方法を適用した提案手法を正誤情報のみを用いた手法と比較する。ここで、正誤情報のみを用いた手法とは、正解した問題は確信があると判定し、間違えた問題は確信がないと判定する手法のことである。

#### 4.3 実験結果・考察

各実験参加者の事前確率を表 5 に示す。本実験での事前確率とは、アンケートにおいて各実験参加者が確信があると回答した割合のことである。事前確率の平均値が 50%前後であることから、2 つのクラスのデータ数には偏りがないことが確認された。

##### 4.3.1 ユーザ依存での検証

提案手法のユーザ依存の場合と正誤情報のみの場合の判定精度を比較した結果を図 6 に示す。平均判定精度は正誤情報のみの場合では 86.7%、提案手法のユーザ依存の場合では 91.2% であった。実験参加者全体としては、問題の正誤と確信度が一致している傾向が見られた。具体的には、正解した問題については確信が持てたと回答し、間違えた問題については確信が持てなかったと回答したケースが多かった。そのため、正誤情報のみの場合でもほとんどの実験参加者で高い判定精度が得られた。一方、提案手法は全ての実験参加者について正誤情報のみでの高い判定精度を上回ることに成功した。

実験参加者ごとで見られた特徴としては、実験参加者 F, G, H, K では、提案手法と正誤情報のみの場合ともに高い判定精度で判定できることが確認された。これらの実験参加者では、少し考えて分からない問題はスキップする傾向が認められた。またその場合には、スキップしていない問題については確信あり、スキップした問題については確信なしとする場合が多いことも確認できた。一方で、実験参加者 A, B, E, J, L では、提案手法が正誤情報のみの場合を大きく上回った。これらの実験参加者では、分からない問題は時間をかけて思い出そうとする傾向が確認された。またその場合には、解答時間はもちろん、タイピング速度も遅くなることが確認できた。

実験参加者 D の判定精度が低い理由について述べる。この実験参加者は、ある 1 セクションにおいて、確信が持てず正解した問題と、確信を持ちつつ間違えた問題の数が他のセクションと比べて極端に多かったということが確認された。Leave-one-section-out Cross-validation を行った場合、これらのパターンが学習データには少なかったために誤認識が起きやすくなったと考える。

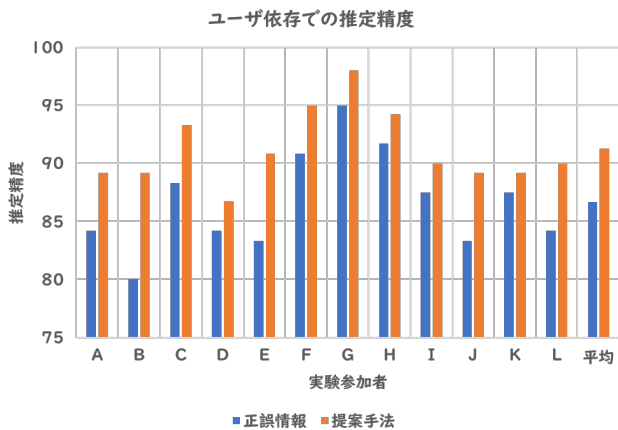


図 6 ユーザ依存の場合の推定精度

表 4 特徴量の選択回数

選択回数	特徴量
11	編集距離
6	解答時間 正解単語の文字数
5	インターバルの平均 最終のインターバル タイプした文字数
4	インターバルの標準偏差 インターバルの最大値
3	インターバルの最小値 最初のインターバル 正解単語の使用頻度
2	インターバルの中央値
1	デリートキーをタイプした回数

各特徴量の選択回数を表 4 に示す。今回の実験では、実験参加者は 12 人なので、選択回数の最大値は 12 回である。最も多く選択された特徴量は編集距離であった。次に多く選択されたのは解答時間、正解単語の文字数であった。3 番目に多く選択されたのはインターバルの平均、最終のインターバル、タイプした文字数であった。編集距離は問題の正誤情報も含んでいるため、選択回数が多くなったと考える。正解単語の文字数は単語の難易度に関連する情報であり、文字数が多い単語ほど暗記するのが難しいという傾向がある。解答時間、インターバルの平均、最終のインターバル、タイプした文字数は、解答に確信が持てない場合には、値が大きくなるため、選択回数が多くなったと考えられる。特に、最終のインターバルは、自身の解答への迷いを顕著に表しているため、多く選択されたと考えられる。

提案手法と事前確率、提案手法と正誤情報のみの場合について符号検定 [12] を行った結果を表 5 に示す。なお、表中の\*\*は有意水準 1% で、\*は有意水準 5% で提案手法と各手法との間に十分な差がみられたことを示す。

表 5 から提案手法と事前確率との間には十分な差がみられた。しかし、提案手法と正誤情報のみの場合との間には全体としては十分な差があるものの、各実験参加者については実験参

表 5 ユーザごとの検定結果

実験参加者	事前確率	正誤情報
A	**	
B	**	**
C	**	
D	*	
E	**	
F	**	
G	**	
H	**	
I	*	
J	**	
K	*	
L	**	
全体	**	**

加者 B 以外での十分な差が認められなかった。このことから、12 セクション分のデータでは差を検出するには不十分なのではないかと推測する。したがって、各実験参加者に解答してもらうセクション数を増やすことを検討する必要がある。これについては、今後の課題としたい。

#### 4.3.2 ユーザ非依存での検証

ユーザ非依存の場合の結果を図 7 に示す。提案手法の平均判定精度は 89.8% であった。表 6 に示す通り、これはタイピング情報を用いなかった他の 4 つの手法を上回る精度であった。選択された特徴量は表 7 の通りである。

一般的に、ユーザ非依存の場合には、ユーザ固有の振舞いは反映されないため、判定精度が下がる。本実験でも、多くのユーザに対してそのような傾向がみられた。特に、確信がない場合の振舞いは、学習が難しく、ユーザ依存の場合と比べて誤認識が起きやすくなることが分かった。なぜなら、確信がある場合は、頭に浮かんでいる文字列をそのままタイプしていただくだけで、その振舞いには大きな個人差はないが、確信がない場合の振舞いは人それぞれで、具体的には、キーを押さずに頭の中でじっくり考えて解答する人、キーを押して自身のタイプした文字列を見ながら解答を探す人、あるいはスキップする人など、様々であったからである。

その中でも、実験参加者 A, B, C については、ユーザ依存の場合と同等の判定精度を得ることができた。この理由については、A, B, C においてユーザ依存の場合で選択された特徴量の組合せが似通っていることから、この 3 者のタイピング作業における振舞いが似ていたということが推測される。

また、実験参加者 G についても、ユーザ依存の場合と同等の判定精度を得ることができた。これは、A, B, C の場合とは異なり、この実験参加者は確信が持てない問題はほとんどスキップしていたので、得られたタイピングに関する情報は確信が持てた問題に限られていた。確信が持てた問題はユーザ非依存の場合でも判定が容易であったので、判定が上手くいったと考えられる。

実験参加者 E について、ユーザ依存の場合と比較して判定精度が著しく下がった理由について考察する。この実験参加

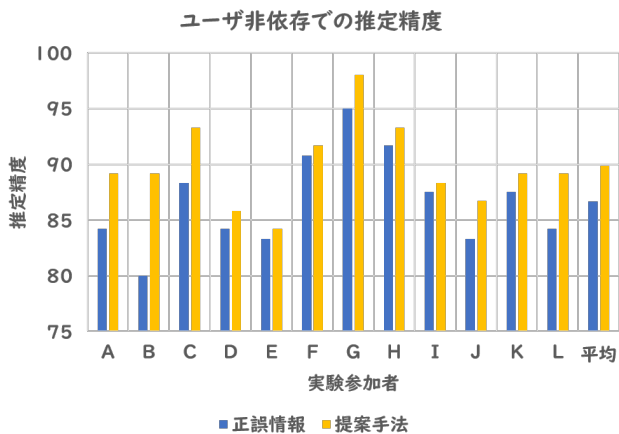


図 7 ユーザ非依存の場合の推定精度

表 6 タイピングに関する情報を用いない手法との比較

手法	判定精度
提案手法	89.8%
正誤情報	86.7%
正誤情報 + 解答時間	86.6%
正誤情報 + 正解単語の情報	86.5%
正誤情報 + 解答時間 + 正解単語の情報	86.7%

表 7 ユーザ非依存の場合に選択された特徴量

特徴量
解答時間
インターバルの平均
インターバルの最大値
インターバルの中央値
最初のインターバル
最終のインターバル
正解単語の文字数
正解単語の使用頻度
編集距離

者については、デリートキーをタイプした回数が確信判定に寄与していることが確認されたが、ユーザ非依存の場合にはこの特徴量は選択されなかった。したがって、ユーザ非依存で学習を行った際に、ユーザ依存で学習を行った場合に比べて確信がない場合の振舞いを学習しきれていないことが確認された。このことから、この実験参加者は確信がない場合の振舞いが特殊であるということが推測される。Leave-one-user-out Cross-validation を行った場合、確信がない場合の振舞いが他の実験参加者の振舞いと合致しなかったため、誤判定が起きやすくなったと考える。

以上のことから、実験参加者数を増やすことで、確信がない場合の振舞いのパターンを増やす必要があると考える。これについても、今後の課題としたい。

## 5. まとめ

本稿では、確信を考慮した英単語学習システムを実現するために、キーログを用いて、タイピング問題解答時の確信度を判定する手法を提案した。実験の結果、提案手法の判定精度はユー

ザ依存の場合は 91.2%、ユーザ非依存の場合は 89.8%であった。

しかし、本稿で行った実験はあくまで英単語タイピングに特化したものである。一般的な問題に解答する際には、長時間思考する場合もあるため、本稿で述べた実験結果とは異なる結果となる可能性がある。より一般的な問題についても提案手法が適用できるのかを検証する必要がある。

また、本稿で得られた実験結果は編集距離に大きく依存した結果となった。英作文などでは、解答の多様性から、明確な解答が提示されていない場合が多いため、編集距離や正誤情報を用いずに確信を判定する手法を考案する必要がある。具体的には、まず、タップしてから指を離すまでの時間、指を離してから次にタップするまでの時間を加える。それに加えてキーをタップした際の圧力も特徴量として導入する。確信がない場合には、迷いによりキーをタップする力が弱くなることが推測されるからである。しかし、Auto Touch には現状、そのような機能は実装されていないため、機能を追加する必要がある。さらに、タイピングに関する情報以外の新たな情報を追加することで、判定精度が向上するのではないかと考える。

謝辞 本研究の一部は、JST CREST (Grant No. JP-MJCR16E1)、日本学術振興会科学研究費補助金挑戦的萌芽研究 (15K12172)、ならびに大阪府立大学キーププロジェクトの補助による。

## 文 献

- [1] F. Monrose and A. Rubin "Keystroke dynamics as biometric for authentication," *Future Generation Computer Systems* 16, 2000, pp.351-359
- [2] F. Bergadano, D. Gunetti, and C. Picardi, "Identity verification through dynamic keystroke analysis," *Intelligent Data Analysis*, 2003, pp. 469-496.
- [3] 中村 和晃, 角所 考, 美濃 導彦 "e-learning 環境における学習者の観測に基づく主観的難易度の判定," *MIRU*, Jul. 2007.
- [4] C. Epp, M. Lippold, and R. L. Mandryk "Identifying Emotional States using Keystroke Dynamics," *CHI*, May. 2011.
- [5] N. Nahin, J. Alam, H. Mahmud and K. Hasan "Identifying emotion by keystroke dynamics and text pattern analysis," *Behaviour & Information Technology*, Jul. 2014, pp.987-996.
- [6] L. Allen, C. Mills, M. Jacovina, S. Crossley, S. D'Mello, and D. McNamara "Investigating Boredom and Engagement during Writing Using Multiple Sources of Information: The Essay, The Writer, and Keystrokes," *Learning Analytics and Knowledge*, Apr. 2016, pp.114-123.
- [7] S. Ghosh, N. Ganguly, B. Mitra, and P. De "Evaluating Effectiveness of Smartphone Typing as an Indicator of User Emotion," *ACII*, 2017.
- [8] 浅井 洋樹, 野澤 明里, 苑田 翔呉, 山名 早人 "オンライン手書きデータを用いた学習者のつまづき検出," *DEIM*, 2012.
- [9] 三木 良介, 宮崎 佳典, 厨子 光政, 法月 健 "英単語並べ替え問題における迷い抽出機能の実装," 情報処理学会第 76 回全国大会, 2014.
- [10] 山田 健斗, 大社 綾乃, 藤好 宏樹, 星加 健介, Olivier Augereau, 黄瀬 浩一 "英語多肢選択問題解答時の視線に基づく確信度判定," *PRMU2016-192*, Feb. 2017.
- [11] 中村 光貴, 荒見 篤郎, 上久保 竜輝, 福嶋 政期, 苗村 健 "英単語暗記システムのための綴りの類似性に関する基礎検討," *PRMU2017-116*, Jan. 2018.
- [12] 高木 英行, "使える! 統計検定・機械学習: I: 2 群間の有意差検定," *システム/制御/情報: システム制御情報学会誌* vol.58 No.8, pp.345-351, 2014.