平仮名ランダム文字列を用いた

ワンタイムパスワードキーストローク認証に関する一検討

中村 優† 齋藤 詩織‡ 藤澤 匡哉‡

† 東京理科大学工学研究科経営工学専攻
162-8601 東京都新宿区神楽坂 1-3
yunaka@ms.kagu.tus.ac.jp

‡ 東京理科大学工学部経営工学科
162-8601 東京都新宿区神楽坂 1-3
fujisawa@ms.kagu.tus.ac.jp

あらまし キーストローク認証は打鍵時間等を利用して個人を識別する認証であるが，固定の文字列ではキーロガーにより個人の特徴，すなわち，打鍵時間等も併せて盗まれる恐れがある．ワンタイムパスワードを用いれば，この可能性を低くすることができるが，一方で個人の特徴を得るためのデータ量が多く必要となる．本稿では，日本人が日頃から打鍵し慣れている平仮名，すなわち「子音－母音」の組み合わせの打鍵時間のみを特徴として用いて認証を行い，Joyce-Gupta 法と順位相関係数を組み合わせることにより，個人の特徴を取得するために必要なデータ量の削減について検討する．

A note on a one-time password keystroke authentication using a random hiragana string

Yu Nakamura† Shiori Saito‡ Masaya Fujisawa‡

†Graduate School of Engineering, Tokyo University of Science,
1-3,Kagurazaka, Shinjuku-ku, Tokyo, 162-8601 Japan
yunaka@ms.kagu.tus.ac.jp

‡Faculty of Engineering, Tokyo University of Science,
1-3,Kagurazaka, Shinjuku-ku, Tokyo, 162-8601 Japan
fujisawa@ms.kagu.tus.ac.jp

Abstract In this paper, we propose a method of keystroke authentication using a one-time password which consists of random hiragana letters. This method uses a combination of the Joyce-Gupta method and rank correlation coefficients. Moreover, we show that it can reduce amount of input data during feature extraction in compare with the conventional methods.

1　はじめに

近年，通信環境の発展にともないネットワークを介したやり取りが増加しているが，同時になりすましなどの不正アクセスも急増している．このなりすましを防ぐために重要となるのが本人認証である．本人認証方法として，パスワード等の個人の知識を用いたものやICカード等の所有物を用いたものがある．知識や所有物は忘却，盗難の恐れがあるという問題から，個人の生体的な特徴を用いた認証方法が注目されている．生体認証方法には，指紋や虹彩など個人の身体的特徴を利用したものと，筆跡や歩行など個人の行動的特徴を利用したものがある．身体的特徴を利用したものは人工的に作られた指や虹彩などによるなりすましの問題があることから，他人が模倣することが困難である行動的特徴を利用した認証方法に関心が高まりつつある．

行動的特徴を利用した生体認証方法の一つに，キーボードを操作するときのスピードなど，タイピング時の固有の特徴を解析して識別を行うキーストローク認証がある．しかし，特定の文字列の入力によるキーストローク認証[1,2]は，リズムを記憶するキーロガーにより行動的特徴を盗まれる危険性がある．ワンタイムパスワードを用いることで行動的特徴が盗まれる危険性を低くすることが可能であるが，個人の特徴を取得するために大量のデータが必要となり，ユーザの負担が大きくなることが問題となる．

日本人が認証を利用する場合，英字打鍵による認証では，被験者が英字打鍵の慣習がないため，打鍵が安定せず特徴抽出に必要なデータ量も多くなる．打鍵し慣れている日本語文を用いることにより，英字打鍵と比較し，認証率が向上することが知られている．

文献[3]では，日本語の特徴である「子音－母音」の中からとくに出現頻度が高い組み合わせに着目することで，個人の特徴を抽出できることを確認している．また，文献[4]では長文である日本語文入力において，日本語入力に特化した特徴量を用いることにより高い認証率を実現している．

本研究では，日本語打鍵に着目し，日頃から打鍵し慣れている連続した英字二文字（以下，二連字）のみ，すなわち，「子音－母音」間の打鍵時間を特徴として用いる．日本語打鍵を用いたワンタイムパスワード方式は特徴として取得する二連字の総組み合わせ数から，特徴抽出に必要なデータ数は少なくなるが，十分に小さくなるわけではなく，ユーザの負担は依然として大きい．

ワンタイムパスワード方式では，認証毎に毎回異なるランダムな文字列を認証に用いる．文字列がランダムではなく通常の文章を用いた認証において，認証毎に任意の文章（非定型文）を用いる認証は，固定の文章（定型文）を用いる認証より，共通する二連字のデータ数が少ないため認証精度が悪化することが知られている．ランダム文字列の場合も同様のことがいえる．文献[5]では，各二連字の順位の並び方の向きに着目したKendallの順位相関係数を用いることで非定型文を用いた場合の認証精度を改善しており，利用できるデータ数が少ない場合でも有効な手法であるといえる．

本研究では，パスワード等の固定文字列を認証文字列とするキーストローク認証方式の一つであるJoyce-Gupta法[1]と，データの分布の歪みや外れ値の影響を受けにくいSpearmanの順位相関係数を組み合わせて用いることで，認証精度の改善を示す．さらに，Joyce-Gupta法に順位相関係数を組み入れた提案法に対し，特徴抽出のための取得データ数の削減の可能性について検討をおこなう．

2　ワンタイムパスワードキーストローク認証

2.1　特徴の抽出

キーストローク認証は，キーボード入力操作時における打鍵時間，打鍵の圧力，あるいは，ミスタイプ等に表れる各個人の特性を利用することにより認証を行う．計測のしやすさから，特徴としてキーの打鍵時間が用いられることが多い．打鍵時間は 1 つのキーを押してから離すまでの時間とあるキーを離して次のキーを押すまでの時間で構成される．連続する 2 つのキーの間の時間としては，押す－押す，押す－離す，離す－押す，離す－離す の 4 つの時間間隔があるが，本研究では押した時刻と次のキーを押した時刻の時間間隔を用いる．

ランダムな文字列を用いるためには，全二連字の打鍵間隔のデータを取得する必要がある．英字による全二連字の時間間隔数は 26×25 であるのに対し，1 文字からなる平仮名を用いる場合 60 程度である．また，ローマ字入力で日本語の通常の文章を入力する場合，文章や単語をひとかたまりとして入力するため，同じ二連字でも前後の関係により打鍵の時間間隔に差が出ると考えられる．一方，ランダムな文字列を入力する場合には，予測しながら打鍵することは困難なため，平仮名一文字ずつ打鍵する形となり，「子音－母音」の組み合わせの打鍵時間は安定すると考えられる．したがって，本研究で用いる平仮名は「子音－母音」で構成されるものを取り扱うことにする．母音「あ」，「い」，「う」，「え」，「お」や「ん」は含めず，また，「し」(shi ，si) や「つ」(tsu，tu)等の 2 通りの入力が存在するものも除いている．以降では，この平仮名(または，二連字)の集合を H と表記する．

2.2　ワンタイムパスワードキーストローク認証の概要

固定文字列をパスワードとして用いる場合，キーボードからの入力を監視して記録するソフトであるキーロガーにより打鍵リズムやパスワードを盗まれる危険性がある．ワンタイムパスワードを用いることで打鍵リズムを盗まれにくくし，第三者の侵入を防ぐことができる．

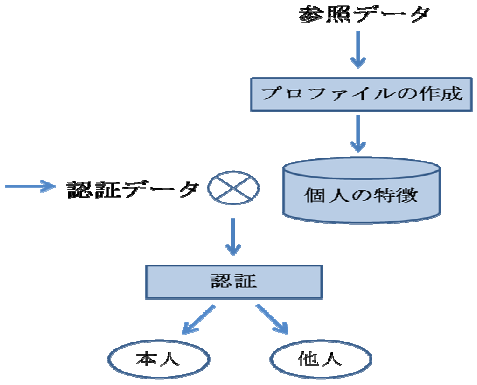

図 1 ワンタイムパスワードキーストローク認証

ワンタイムパスワードキーストローク認証は，特徴を抽出するプロファイル作成ステップと認証ステップの 2 つのステップにより構成される（図 1 に概要を示す）．以下に，詳細な手順を示す．

Step 1（プロファイル作成）

1) 打鍵データの取得
　ランダムに文字 $\alpha \in H$ を n 文字表示し，打鍵させる．これを m 回繰返す．全体打鍵量は $N(=n$ 文字 $\times\ m$回$)$ 文字になる．取得した打鍵データにおいて，各二連字 α の平均 μ_α，標準偏差 σ_α を計算する．
2) 外れ値の除去
　平均 μ_α，標準偏差 σ_α に対して，平均から $3\sigma_\alpha$ 以上離れた打鍵データを稀なデータとして除外し，再度平均 μ_α，標準偏差 σ_α を計算する．
3) プロファイルの作成
　対応するデータの組 $(\alpha, \mu_\alpha, \sigma_\alpha)$ を各個人の特徴のプロファイルとして保持する（表 1）．

Step 2（認証）

1) 認証データ取得

ランダムに文字$\alpha \in H$をn　文字表示し，打鍵させ，認証データを取得する．認証データの中で同じ文字が複数回出た場合は平均をとり，この打鍵時間$\tilde{t}$　を認証データとして保持する．

2) 判定

ユーザuのプロファイルから，入力文字に対応するデータを取得する（これを参照データと呼ぶ）．参照データと認証データを比較し，適当な尺度を用いて設定したしきい値θより小さいか否かで本人であるかどうかの判定を行なう．

表１　プロファイル

文字α	二連字	平均	標準偏差
か	ka	μ_{ka}	σ_{ka}
き	ki	μ_{ki}	σ_{ki}
く	ku	μ_{ku}	σ_{ku}
け	ke	μ_{ke}	σ_{ke}
⋮	⋮	⋮	⋮

Joyce-Gupta 法では，プロファイルの参照データと認証データの比較の尺度として以下の距離Dを用いる[3]．ここで，$B \subset H$は認証時に打鍵させた二連字のうち入力したデータが有効な二連字の集合とし，t_βは各二連字$\beta \in B$に対応する打鍵時間（認証データ）とする．また，$b = |B|$　はBの要素数である．

$$D = \frac{1}{b}\sum_{\beta \in B}\frac{|t_\beta - \mu_\beta|}{\sigma_\beta} \qquad (1)$$

次に，順位相関係数を用いる場合には，認証方法の判定手順を以下の手順に置き換える．

2)’ 判定

認証データとして入力したデータが有効な二連字$\beta \in B$に対し，参照データと認証データの値別に順位をつける．各二連字$\beta \in B$の参照データでの順位と認証データでの順位の差をd_βとする．比較尺度として，以下の Spearman 順位相関係数ρを用いる．

$$\rho = 1 - \frac{6\sum_{\beta \in B} d_\beta^2}{b(b^2-1)} \qquad (2)$$

Spearman 順位相関係数は参照データと認証データの相関の強さを見る指標であり，順位のばらつき度合いを表している．Kendall の順位相関係数は順位の一致度合いにより順位の並びの類似性を表している．打鍵時間がほぼ同じである二連字がいくつかある場合にはそれらの順位が大きく変わる可能性があるため，相関係数の値も大きく変わる可能性がある．本研究では Spearman の順位相関係数を用いる．

2.3　提案法

Joyce-Gupta 法[1]を利用した認証手法では，ランダムな平仮名文字列に対して，打鍵時間の安定した二連字のみを用いており，使用するデータは取得データのごく一部である．

そこで全二連字を利用することのできる Spearman の順位相関係数を組み合わせた認証方式を提案する．

認証手順，概要図を以下に示す．

認証

Hの元をランダムにn字表示し，入力させる．

1. 入力データに対応した文字だけで順位をつけ直した後，Spearman の順位相関係数ρを求め，しきい値θ_1より小さければ$(\rho < \theta_1)$，次のステップに進み，そうでない場合には他人と判定する．

2. 1. において正規ユーザのデータとして識別された入力データに対し，Joyce-Gupta 法を利用して，認証データの平均値 $\mu_\beta(u)$，標準偏差 $\sigma_\beta(u)$ から尺度距離を求めしきい値 θ_2 より小さいか否かにより本人か否かを判定する．

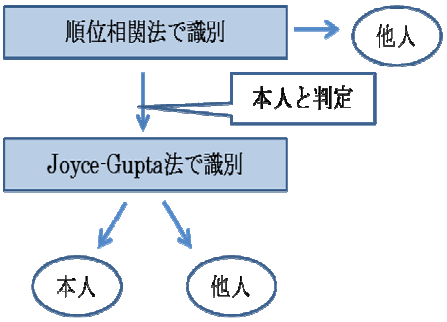

図 2　提案手法の概要図

3　評価実験

3.1　実験概要

実験に使用するデータとして，14 人の被験者対し，ランダムな平仮名文字 $n=20$ 文字を表示し，ローマ字入力で $m=22$ 回×12 セット入力させ，被験者 1 人あたり平仮名文字列 $N=2{,}640$ 字(ローマ字 5,280 字)のデータを得た．

また，認証実験においては，14 人の被験者対し，ランダムな平仮名文字列 $n=20$ 文字を $m=14$ 回(本人の認証文字列 1＋他人の認証文字列 13)×5 セット入力させ，$N=1400$ 字のデータを得た．

3.2　結果と考察

提案手法を用いた認証結果を表 2 に，Spearman の順位相関法のみを用いた認証結果を表 3 に，Joyce-Gupta 法[1]を利用した手法のみ用いた認証結果を表 4 に示す．Joyce-Gupta 法を用いるにあたって，本人と判定するための閾値には，本人拒否率(FRR)と他人受け入れ率(FAR)が一致する値を用いる．

図 3 に Spearman の順位相関法を用いた場合のしきい値に対する FRR と FAR を示す．図より，FRR と FAR が等しい EER(等誤り率)はしきい値が 0.44 のときである．これより今回順位相関法において用いるしきい値は 0.44 とした．また，提案手法での順位相関法におけるしきい値は 0.36 とした．

表 2 に示すように，提案手法の FRR は 7.14%，FAR は 5.49%となり，表 3 の順位相関のみを用いた認証結果と比較し，FRR，FAR ともに大幅に改善することができた．また，表 4 の Joyce-Gupta 法を利用した手法をのみ用いた認証結果と比較しても FAR を改善することができた．

これにより，Joyce-Gupta 法と，Spearman の順位相関係数を組み合わせた手法を用いることで，Joyce-Gupta 法のみ，Spearman の順位相関法のみを用いたものに対して認証精度の改善を示すことができた．

表 2　提案手法

	本人	本人以外
本人と判定	92.86%	5.49%
本人以外と判定	7.14%	94.51%

表 3　順位相関法

	本人	本人以外
本人と判定	85.71%	12.09%
本人以外と判定	14.29%	87.91%

表 4　Joyce-Gupta 法

	本人	本人以外
本人と判定	92.86%	8.38%
本人以外と判定	7.14%	91.62%

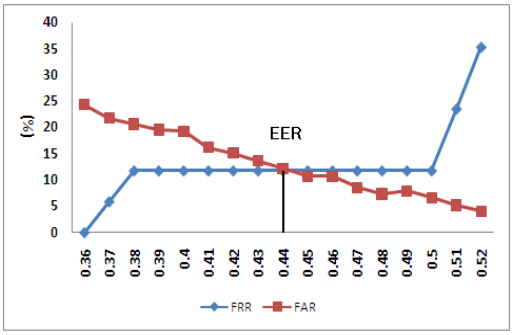

図 3 認証精度の評価

4　データ量削減の検討

4.1　取得データ数の削減

本実験では, 参照データ数 2,640 字を得た. しかし，参照データを取得する度に打鍵するには，ユーザにとって大きな負担となる量である.

順位相関を加えた提案手法を用いることにより，個人の特徴をより少ないデータ数から抽出できるようになると考える.

参照データ取得の際の個人への負担の軽減を目的とし，取得したデータ数を 1,500 字まで減少させ同提案手法で行った.

4.2　結果と考察

表 5，6 の認証結果より，データ数 1,500 で提案手法を用いた場合，FRR7.14%，FAR6.59%となり，2,640 字を用いた認証結果と同等の結果を得ることができた．また，順位相関法のみを利用した結果と比べても，FAR を大幅に改善できていることがわかる.

このことから，提案手法がデータ量の削減において有効であることがわかった.

表 5　提案手法(1,500 字)

	本人	本人以外
本人と判定	92.86%	6.59%
本人以外と判定	7.14%	93.41%

表 6　順位相関法(1,500 字)

	本人	本人以外
本人と判定	92.86%	14.29%
本人以外と判定	7.14%	85.71%

5　まとめ

本研究では，ランダムな平仮名文字列を認証文字列とするワンタイムパスワードキーストローク認証において，従来の Joyce-Gupta 法に加えて Spearman の順位相関係数を組み合わせた方法を提案した.

提案手法は, FRR は 7.14%, FAR は 5.49%と，順位相関法のみを用いた結果である FRR14.29%，FAR12.09%と比べ認証精度を改善することができた．また，Joyce-Gupta 法のみを用いた認証結果と比較しても，FAR を改善することができた.

さらに, 提案手法に対して, 個人の特徴を取得するために必要なデータ量の検討をおこなった結果, データ数を 1,500 文字まで削減した場合においても, 削減前のデータ数で得られた認証結果と同等の結果を得ることが確認できた.

今後の課題は被験者数を増やした認証実験の実施や, より多くのキーストローク特徴を取り入れた解析方法の検討である.

参考文献

[1]　R. Joyce and G. Gupta, “Identity Authentication Based on Keystroke Latencies,” Comm. ACM, vol.33, no.2, pp.168-176, 1990.

[2] 粕川正充, 森裕子, 小松賢嗣, 赤池英夫, 角田博保, “打鍵データに基づく個人認証システムの評価と改良,”情報処理学会論文誌, vol.33, no.5, pp.728-735, May.1992.

[3] 倉橋勇気, 横山和也, 小松尚久, “キーストロークダイナミクスの特徴と個人照合アルゴリズムの提案,”信学技法, vol.105, no.40, pp.7-12, May.2005.

[4] 佐村敏治, 西村治彦, “キーストロークダイナミックスによる日本語文での個人識別,”システム制御情報学会研究発表講演会, 5F3-6, pp.707-708, May.2007.

[5] 片岡祥啓, 宮本貴朗, 青木茂樹, 泉正夫, 福永邦雄, “キーストロークの統計情報を利用した個人認証手法の提案,”情報処理学会, vol.2007, no.71, pp.23-30, 2007.