打鍵認証におけるなりすまし防止に最適なデータ解析手法に関する研究

1140289　赤澤 健人　【 清水研究室 】

1　はじめに

近年，インターネットの普及に伴い，個人認証が重視されいる．その際に使用されるのが本人の知識を用いたパスワード認証や生体的特徴を使用した生体認証である．しかし，これらの認証方式はアクセス時のみしか認証を行わず，認証後のなりすましに対処することができない．これに対応した技術として個人の打鍵特徴を使用した打鍵認証があげられる．打鍵認証は特殊な機器を必要とせずに継続的に認証を行うことが可能である．打鍵認証はキーボードから特徴量を取得する部分と特徴量を解析する部分の 2 つに分けられる．特徴量を取得する部分はどのような特徴量が有効かといった研究がなされているが，データ解析の部分はどの手法が最適かという研究がなされていない．本研究では，既存研究で使用されているデータ解析手法をサンプルデータを用いて比較し，打鍵認証に最適な解析手法の検証を行う．

2　解析手法の比較

本項では，比較する各データ解析手法について，実験手順，評価について解説する．

2.1　データ解析手法

本研究ではユーグリッド距離，WED 法，AD 法，WED + AD 法，自己組織化マップの比較を行う．ユーグリッド距離は 2 点間の絶対的な距離を示す．

WED 法はある定めれた距離尺度に基づく 2 点間の距離をもとに評価を行う手法である．

AD 法は各特徴量の不揃度を数値化することで評価を行う手法である [1]．

WED + AD 法は 2 点間の絶対的値である距離と相対的値である不揃度を組み合わせ認証率の向上を図った手法である．

自己組織化マップは多次元データの解析に用いられる競合学習型ニューラルネットワークである [2]．これは，多次元データを解析し，類似した性質を持つものが接近するように低次元化を行う．打鍵認証では特徴量をマッピングし，プロファイルデータと認証データのユーグリッド距離を用いて認証を行う．

2.2　実験手順

被験者のプロファイルデータを作成するため a ～ z が 5 回以上出現する文字列を打鍵してもらい，打鍵の際のプレス-リリース間隔 (以下 p-r 間隔) を取得する．次に各アルファベットの平均 p-r 間隔を計算し，各個人のプロファイルデータとする．これを 5 回繰り返し A1～A5 の 5 つのプロファイルデータを作成する．

次にランダムに選定された非定型文を 200 文字分打鍵してらい，取得した特徴量の中から 5 回以上出現したアルファベットの平均 p-r 間隔を計算し，認証用データ UK とする．

A1～A5 と UK を比較する 5 つのデータ解析手法を用いて解析し，適切な閾値を設定して認証を行う．

2.3　評価

データ解析手法の評価は，正規のユーザが不正ユーザとして認証される本人拒否率と，不正ユーザが正規ユーザとして認証される他人受容率の 2 つを比較することで行う．本研究では，本人拒否率と他人受容率が交差する時のそれらを比較対象とし優位性を示す．

3　おわりに

5 つのデータ解析手法比較し打鍵認証に最適な手法の検証を行った．以下は各手法を比較した結果をまとめた表である．

	平均本人拒否率	平均他人受容率
ユーグリッド距離	22.4 %	15.9 %
WED 法	15.8 %	8.2 %
AD 法	17.1 %	12.3 %
WED + AD 法	12.3 %	6.7 %
自己組織化マップ	10.2 %	10.0 %

その結果，本人拒否率では自己組織化マップが 10.2%，他人受容率では WED + AD 法が 6.7%でそれぞれ最小値になることが分かった．本研究では，認証後のなりすましを防ぐことを目的としているため，本人がサービスを利用できなくなることよりも，他人の不正利用を防止することを優先すべきであると考えた．そこで，本研究では他人受容率の低い WED + AD 法が打鍵認証には最適であると結論付ける．

今後は，解析するサンプル数を増やし，本人拒否率，他人受容率がどのように変化するかや閾値の設定，特徴量を複数にした際の最適な解析手法などを検証する必要がある．

参考文献

[1] D. Gunetti and C.Picardi, “ Keystroke Analysis of Free Text ”, ACM Trans. Inf.Syst.Secur., vo.8, no.3, pp.312-347, 2005

[2] 徳高 平蔵, 藤村 喜久郎,“ 自己組織化マップ (SOM) とその応用 ”, 日本ファジィ学会誌, Vol.13, No.4, pp.345-355, 2001