

## 話者認識と音声認識を統合した携帯電話セキュリティシステムの開発\*

平野 邦彦 (法政大学情報科学部), 伊藤克亘 (法政大学情報科学部)

## 1 まえがき

現在携帯電話は目覚ましい発展を遂げている。それに伴い扱える情報量も増加している。ID や SUICA などの電子マネーなども扱うようになり、さらに機密資料や個人情報など外部に漏れてしまてはいけない情報も多く含まれている。

現在多くの携帯電話に搭載されているセキュリティシステムは数字 4 桁によるダイヤルロックである。しかしダイヤルロック方式では堅牢性が低く、現在新たに注目されているのが生体認証である。

生体認証ならユーザーを生体を使用したものをパスワードとしているの紛失する恐れも無い。現在生体認証には指紋、静脈、声紋、網膜認証など様々な方法がある。文献 [3] などの先行研究がある。しかしいずれのシステムも特殊な機器を必要としそれを携帯端末上で実装しようとする、端末に負荷がかかり効率的ではない。しかし携帯端末にはマイクだけは当初から実装されており声紋認証だけは端末に負荷がかからず実現できるといえる。

本研究で作成するシステムは話者照合を使用したものである。さらに音声認識のシステムと統合することでより堅牢性の高いシステムの作成を行う。その実現のためにシステムに最適なシステムの作成・評価実験を行った。

## 2 音声認識と話者照合を統合した認証システム提案

現在携帯電話で主流に使用されているシステムはダイヤルロック方式のものである。しかしこれには欠点が存在する 1 つ目パスワードの紛失や盗難・流出である。これにより使用者以外のものがパスワードを入手してしまった場合、他者によるアクセスが簡単に可能になってしまう。、2 つ目総当たり攻撃による不正アクセスされてしまうことである。10 個の数字を 4 つ組み合わせているこのシステムは 10000 通りの組み合わせが存在する。しかしこれは 10000 回アクセスを行ってしまえば簡単にセキュリティを破られてしまうのである本研究では携帯端末に話者認識による話者認証のセキュリティシステムを実装することにより、携帯端末の堅牢性を向上させる。話者認識による生体パスにすることでユーザーに余計な操作を強いることも、端末に特殊な機器を付ける必要も無い。音声認識を使いパスワードを随時変化するワンタイムパスに設定することでパスを認証できるものをユーザーのみに限定し盗聴による不正アクセスも防止する。話者認識と音声認識の両者とも的一致が行われて初めてユーザーは携帯端末にアクセスすることができる。このようにして携帯端末の堅牢性を高めると共に操作性も悪くすることは無い。

## 3 システム概要

本研究で目指すシステムは、話者認識と音声認識の両立したセキュリティシステムである。ランダムで生成された 3 つのアルファベットをユーザーが読み上げ携帯端末に音声データとして入力する。話者認識ではユーザーを判断し、音声認識では読み上げた語句を文字として認識しアルファベットと比較する。お互いが一致することのみでユーザーが先に進むことができる強固なシステムを作成する。(図 1)

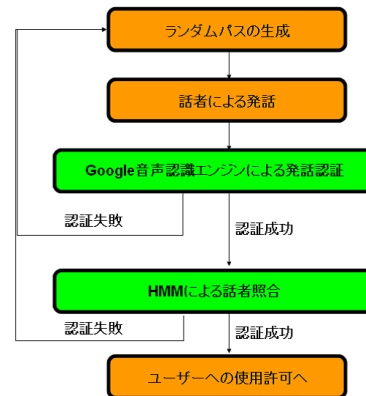


図 1. システム概要

## 3.1 システム実装により向上する点

本研究では音声認識を導入することによりセキュリティシステムの向上、携帯端末の堅牢性の向上を目指してきた。しかし堅牢性の向上以外にも音声認識を使用するうえで以下のような利点がある。

代表的な認証方法である指紋認証やキーロックではタイピングや指紋読み取りなど手を用いた操作が必要になる。しかし、提案するシステムでは機器に触れる必要性がなく、操作性が高められる。

語句認証を併用することで話者認識の欠点である音声の盗聴などの不正アクセスを防げる。

パスワードはランダムで提示される単語であるため、ユーザー以外の音声（盗聴による音声データ）では対応することが難しく、さらにランダムに提示される単語をキーとするためにパスワードの紛失や盗難を抑えることができる。

認証時にマイクが必要となるが、携帯電話にはマイクが常備されており、新たな機器の追加は不要である。

現在主流の暗証番号 4 ケタの物より組み合わせ数が多いため、端末の堅牢性をより向上させることができる。

## 4 各システムの性能実験

## 4.1 音声認識エンジン性能実験

音声認識を行う上でさまざまな音声認識ソフトが存在する。先行研究ではフリーソフトである julian などの

\*security system of cellular phone that integrates speaker recognition and phrase recognition by Kuni-hiko Hirano. (HOSEI university) et. al.

音声認識ソフトを使用することでユーザーからの音声認識を行っていた。しかし、認識精度とモバイル端末への実装の両方を考慮した結果、実現化が極めてむずかしいとされた。そこで本研究では音声認識ソフトの代用として Google のモバイル用音声認識エンジンを使用した。他のシステムとの性能評価実験を行った。結果として google の誤り率は 35%、julius は 92.5%、julian では 85% となった。

当システムでは話者照合の盗聴による不正アクセスを禁止するために音声認識システムを組み合わせた。音声データが流出したと仮定し、総当たり攻撃による認証実験を行った。ランダムに提示されるパスワードに対して固定したアルファベット 3 文字 (abx) で 100 回の認証実験を行った。結果としては認証率は 0% となり、結果として盗聴に対しての堅牢性を向上した。

#### 4.2 話者照合性能実験

今回話者照合は HTK により隠れマルコフモデル (HMM) を使用し単語を単位とした確立モデルを作成した。ユーザーの学習データは 10 種類 \* 4 回、他者の学習データは男女 2 ずつ 10 種類 \* 1 回をおこなう。どちらも総数 40 個の音声データを使用し個人モデルの作成を行った。今回、使用する個人特徴量としては MFCC (メル周波数ケプストラム係数) を使用した。モデル構造としてはユーザーの A~Z のモデル、と他者の A~Z のモデルを作成し比較を行うことでユーザーか他者かの判断を行うものである。男性 3 人と女性 2 で実験を行い性能評価を行った。結果として本人棄却率 3.4%、他人受率率 2% となった。

認識率の評価としては本人棄却率と他人受率率の 2 つが存在する。本人棄却率は本人を他人と認識してしまうことであり、他人受率率は他人を本人として認識してしまうことである。当システムでは堅牢性の向上を目標にシステム作成を行っている。そのため本人棄却率よりも他人受率率を下げることを優先して行う必要がある。

表 1. HMM 文字単位モデルによる認識結果

	誤認識数 (個)	認識総数 (個)	誤り率 (%)
本人棄却	5	160	3.12
他人受率	8	240	2

#### 4.3 本研究の評価方法

本研究で目標とすることは話者認識に音声認識を統合することで盗聴による不正アクセスの防止。それに伴うシステムの堅牢性の向上である。音声認識では盗聴データからのアクセスを話者照合単体の堅牢性の評価方法としては、上記に書いたように本人棄却率と他人受率率を使用して行う。本人が他人として認証される事は堅牢性の低下には影響しない。他人受率率の低さが堅牢性と直結しているため評価は他人受率率の低さで行う。話者モデルに使用していない評価用のデータを使用し他人受率率を求める。システムの総合的な評価は 2

つのシステムを統合させた認証率の論理和による認証精度を評価尺度とする。

#### 4.4 システム統合による実験結果

評価用データにより認証実験を行ったところ、盗聴により不正アクセスが行われる可能性は 0% である。この危険性を音声認識を組み合わせることで 0% になった。このことは堅牢性の向上に大きく繋がる。さらに話者照合単体の認識率は他人受率率が 2% となった。これは先行研究文献 [6] と比べると 1% 低い結果になってしまった。全体の認証精度の平均としては 23% となった。

表 2. システムの総合認識率 (%)

被験者	男性 A	男性 B	男性 C	女性 A	女性 B
話者照合	1	20	0	0	0
音声認識	20	30	20	10	20
認識・論理積	21	44	20	10	20

#### 4.5 考察

今回の実験は 2 つのシステムの統合により行った。システムの認識率を向上させるためには両システムの認識率の向上が必用に感じた。特に音声認識システムの認識率を下けているものはある限られた単語が複数誤答を行っているため認識率下げる結果となった。理由として考えられるものは D・B など音素が繋がり発音が似ている物の誤認識が多い。その他にあげられるものは KQ などの 2 音が繋がることで日本語として意味を成してしまうものが google の言語データベースにより修正されてしまいご認識になってのだと考えられた。以上の理由から音声認識エンジンの認識率の降下に繋がるものと考えられる。

#### 5 あとがき

二つのシステムを統合し認証システムを作成した。今後の方針としては閾値の設定による話者照合の認証制度の向上や、google の音声認識システムの欠点であるネットワークのない環境下での認証を可能にすることである。

#### 参考文献

- [1] 住吉貴志, 李晃伸, 河原達也, “音声認識エンジン Julius/Julian の API 実装” 情報処理学会研究報告. SLP, 音声言語情報処理, 09196072, 社団法人情報処理学会, 2001, 2001, 68, 91-96
- [2] 長内隆, 尾関和彦, 谷本益巳, “単独発声母音を用いた話者照合における特徴量変換” 日本音響学会誌 03694232, 社団法人日本音響学会, 2006, 62 巻 12 号 pp.848-855
- [3] 小林光, 田中章浩, 岸田悟, 渡部徹, 長谷川弘, “指紋と声紋によるハイブリッド認証システムの構築” 電子情報通信学会技術研究報告, 社団法人電子情報通信学会, 2008008-03-05, P 421-426
- [4] 松井知子, 古井 貞熙, “音韻・話者独立モデルによる話者照合尤度の正規化” 電子情報通信学会技術研究報告, 社団法人情報処理学会, 1994-06-16, P 61-66
- [5] 荒木 雅弘, “フリーソフトでつくる音声認識システム - パターン認識・機械学習の初歩から対話システムまで” 森北出版
- [6] 松井 知子, “MM による話者認識” 電子情報通信学会技術研究報告, 社団法人情報処理学会, 1996, 1996-01-18, 467, 17-24