

## 調査研究活動成果報告書

### 調査研究課題名 「迷惑電話防止に有効な通話システムの構築と評価」

調査研究代表者：弓削商船高等専門学校・長尾和彦

調査研究期間：平成29年7月26日 ～ 平成30年3月31日

### 概要

#### 1. 研究の目的

過疎化高齢化により増加した独居老人にとって、電話は離れた家族と連絡を取り、近況を把握するために大切な役割を果たしている。一方、高齢者を狙った迷惑電話・特殊詐欺が多発し、問題となっている。警察庁による特殊詐欺の認知・検挙状況によれば、平成21、22年の一斉検挙により一時的に減少したものの、年々増加を続けており、26年には過去12年で最悪の565億円の被害が発生している。

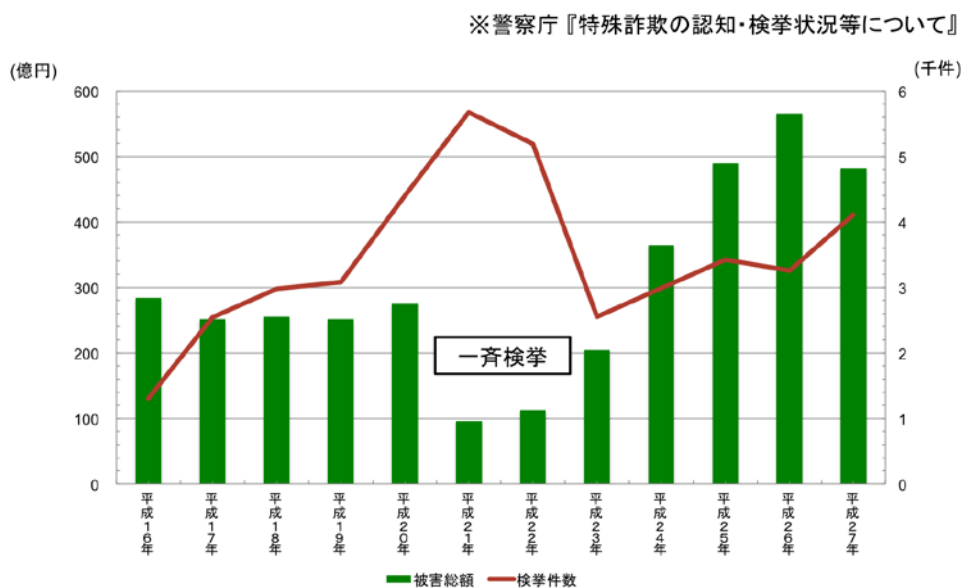


図1 特殊詐欺被害・認知状況

インターネットの普及により、IP電話が各自治体を中心に普及しつつある。IP電話では、通話内容をコンピュータ処理することが比較的容易である。我々は特殊詐欺被害を減少させることを目的とし、IP電話による話者認識・音声認識技術を用いた迷惑電話防止判定機能を有するシステムの開発に取り組んでいる。本取組では、音声認識技術の動向調査、IP電話網構築による検証実験環境の構築、話者認識・音声認識に対応したIP電話アプリケーションの開発、固定電話に対応したシステムの設計についての検討を行った。

## 2. 研究によって得られた成果

### ● 音声認識技術の動向

スマートフォンや家電製品を操作するための音声認識を組み込んだ製品が発表されている。これらのほとんどはサーバに接続して音声認識を行う方式である。IP電話を対象とした場合は、音声データがサーバを介するため、比較的容易にシステム構築が可能となる。固定電話の場合はインターネットに接続された専用のアタッチメントが必要と考えられる。

### ● 迷惑電話防止の取り組みの動向

最近の電話機には、ナンバー通知機能が標準となっており、誰からの電話かを容易に確認することができるようになってきている。これらの番号を用いて、着信拒否する機能なども提供されている。また、防犯上の目的から、自動応答、自動録音する機器も提供されている。徳島県警では、これらの機器を貸し出すことで、迷惑電話防止の効果を上げている。

### ● 検証実験用 IP 電話網の構築

研究室内にサーバを構築し、電話交換機能、話者認識エンジン、音声認識エンジンを実装した。端末間で相互に通話できることを確認した。

- サーバ PC: GPU 搭載 (GTX1080)
  - ・ OS: Ubuntu Linux 16.04
  - ・ 電話交換機 (IP-PBX) サーバ: asterisk13.1.0
  - ・ 話者認識サーバ: ALIZE
  - ・ 音声認識サーバ: Julius
- 端末
  - ・ SIP アプリ
    - ・ PC (Mac): X-Lite
    - ・ iPhone: Zoiper
    - ・ SIP 電話機: KX-UT123N

### ● IP 電話を用いた迷惑電話防止システムのコンセプトとプロトタイプの実装

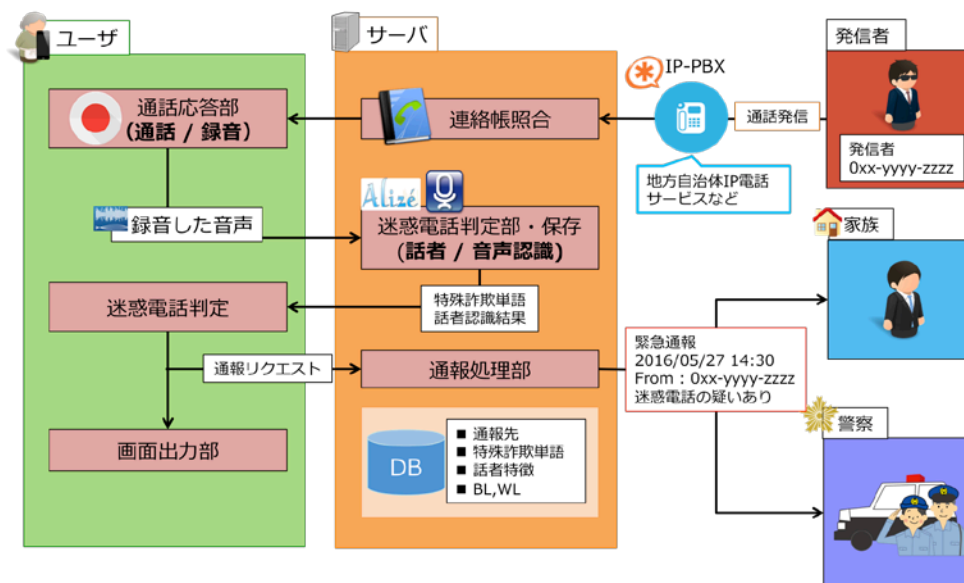


図2 システムコンセプト

我々が開発した迷惑電話防止システムについて概要を説明する。ユーザは専用アプリを用いて通話を行い、通話中の音声をサーバに送信する。サーバでは迷惑電話判定を行い、その結果を危険度としてユーザに返す。

- ログイン

ユーザはログインをすることで、接続サーバ・ユーザ認証を行う。

- 電話帳照合

本システムでは、信頼できるユーザ（電話帳：ホワイトリスト）と迷惑であると認定されたユーザ（ブラックリスト）の2つを持つ。

ホワイトリストからの通話は接続を行うが、ブラックリストは着信拒否を原則とする。いずれにも含まれない番号の場合は、警告を表示する。



図3 着信時の画面表示

- 通話中のモニタリング

警告が行われた通話は自動的に迷惑電話判定モードになり、端末側で通話録音を行う。話者認識の場合は、明示的ユーザを選択しなければならない。音声データは定期的（10秒単位）にサーバに転送され、音声・話者認識が行われる。判定結果は端末側に表示される。



図4 判定結果の表示

- プロトタイプの問題点

端末側にSIP対応の専用アプリを開発したが、標準の電話機能に比べ操作性に劣る。また、話者認識はなりすましを対象としているため、電話帳に登録されているユーザであり、一定時間の通話に基づいて音声モデルが作成されている必要がある。これらの自動化はまだ実現されていない。自分の音声と通話相手の音声はデータの的に分離することは比較的容易ではあるが、ハウリングなどの問題を完全に解消できるものでない。

● 話者認識率の向上に関する実験

学生（男 10 女 10）を対象として IP 電話を用いた音声録音を行い、話者認識率に関する評価を行った。通話端末は iPad(専用アプリ), iPhone5s(zoiper) である。得られた音声は 10 秒ごとの音声に分割し、モデル作成用音声、比較用音声を作成した。それぞれの音声ファイルからパラメータを変更し特徴ベクトルファイルを抽出、比較を行った。モデル生成に要する時間（UBM 学習時間）は数十秒以上要するが、照合時間は 1 秒未満である。認識率は 82 ~ 85% である。

● 固定電話アタッチメントの設計

既存の自動録音応答装置は電話回線と電話機の間接続される。この方法により音声出力を取得することができる。ただし、アタッチメントをどのようにサーバに接続するかはコスト面での課題となる。相互の音声を分割する方法についても課題が残る。

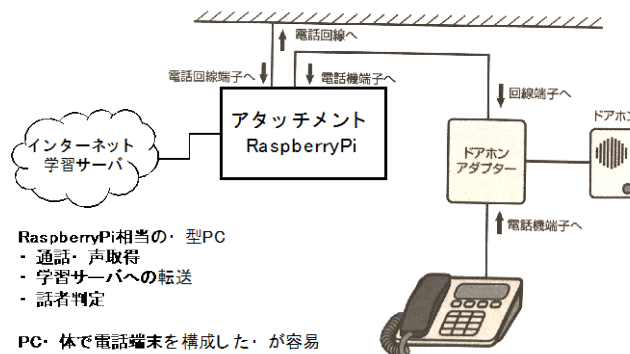


図5 固定電話アタッチメントの設計

● CATV に対する調査

四国地区 CATV 事業者（40 組織）に対し、IP 電話サービスに関するアンケート調査を行なった。回収率は 35% (14 件) であった。ほとんどの事業者で IP 電話サービスを外部委託しており、詳細について不明であった。

迷惑電話対策を行なっている：2 件

本システムは有用：非常に思う：5 件 少し思う：5 件

導入を検討したい：非常に思う：2 件 少し思う：5 件

3. 今後の課題

● 固定電話アタッチメントの開発

固定電話用アタッチメントについてシステムの実装を行う。インターネットへの接続が課題となるが、WiFi などによる接続環境を前提として検討を進め

たい。

- 認識率の向上とオフラインでの実装  
実験環境において話者認識率は80%以上を達成したが、実用にはまだ十分でない。モデル学習の効率化を含め、学習方法の改善を進め、RaspberryPiなどの小型マイコンで実装ができるようにする。
- 汎用 SIP アプリの開発  
汎用的に利用できる SIP アプリを開発し、ユーザへの普及を進める。
- 通信事業者への啓蒙  
電話番号管理による着信拒否、通話録音は迷惑電話防止に効果的であることがわかっている。通話内容の解析による迷惑電話対策は通信傍受に抵触するため、利用者双方の同意が必要となるが、犯人特定や被害防止に役立つことが期待される。これらのサービスを標準機能として提供できるように働きかけていきたい。

#### 4. 添付資料

- システムリーフレット（高専プロコン）
- 学会原稿（情報処理学会）

# ミ・マモーレ

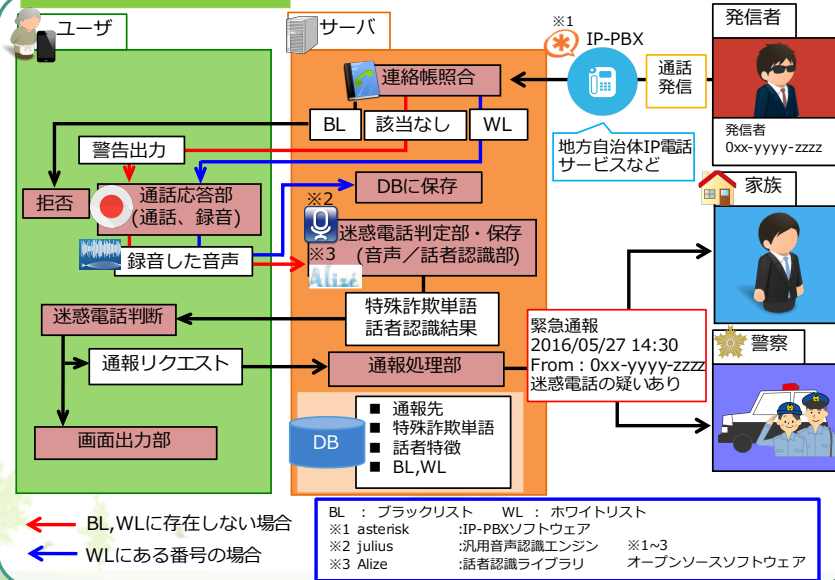
第27回全国高専プログラミングコンテスト  
自由部門04 弓削商船高等専門学校

## ～迷惑電話防止システム～

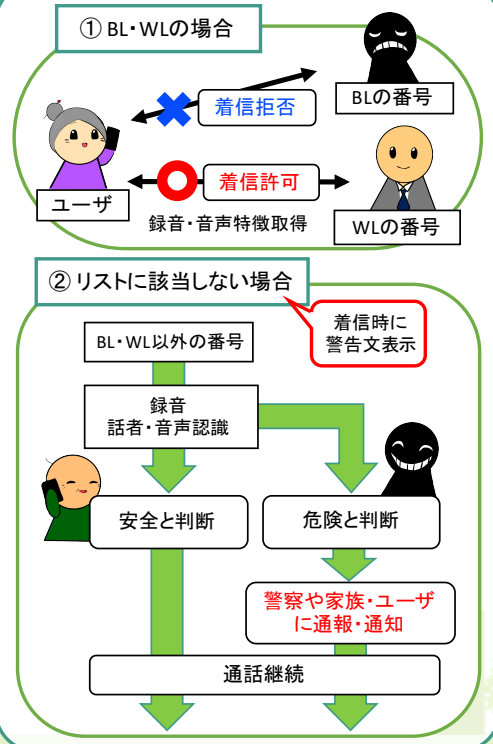
### はじめに

近年、様々な通信機器が開発されていますが、**高齢者にとって電話は最適な通信手段**となっています。それに伴い、高齢者を狙った迷惑電話や特殊詐欺が多発しています。そこで私たちは低コストで利用が容易な**IP電話を対象とした、迷惑電話対策システム**を提案します。

### システム構成図



### 利用イメージ



### 話者・音声認識



#### 話者認識

家族との会話を録音した音声を元に声の特徴をとります。それを通話相手の音声特徴と比較することで、相手が家族かどうかの判定をします。

#### 音声認識

詐欺で使用する主な単語を検出し、詐欺の危険度判定を行います。

#### バイブレーション

危険度が一定以上に達するとバイブレーションで通知をします。

### 連絡帳



### 通知



### 録音

通話開始時に「通話内容を録音すること」を伝え、通話中に音声録音を行います。その音声をサーバへ転送し、話者・音声認識を行います。

### まとめ

私たちは、**高齢者を対象とした手軽な迷惑電話対策システムを提供**します。現状ではIP電話を利用するシステムとなっていますが、固定電話用の専用機器を開発することにより対応を進めたいと考えています。

# IP 電話を用いた迷惑電話防止システムの開発

Calvin Janitra Halim<sup>†</sup> 吾藤 秀亮<sup>†</sup> 青井 佑太<sup>†</sup> 伊藤 宏紀<sup>†</sup> 長尾 和彦<sup>†</sup>

弓削商船高等専門学校<sup>‡</sup>

## 1. はじめに

過疎化・高齢化により増加した独居老人に電話は大切な通信手段であるが、一方で、高齢者を狙った迷惑電話や特殊詐欺が多発している。警察や企業で特殊詐欺対策システムを提案し対策を行っているが、手口の巧妙化により詐欺を完全に防ぐことは困難である。新たな通信の形として、各自治体では IP 電話が普及しており、平成 26 年度末時点で 3564 万件に導入されている。そこで我々は IP 電話を対象とした迷惑電話防止システムを開発した。

## 2. 迷惑電話防止システムの開発

### 2.1. 研究の目的

本研究は、年々増加傾向にある特殊詐欺の被害を減少させることを目的とし、IP 電話による迷惑電話防止システムの開発を行い、電話口からの音声に音声認識などを用いて特殊詐欺の抑制ができるか検証する。

通常の電話機能に加え、電話帳照合・通話録音機能、音声・話者認識処理といった機能を使用し、迷惑電話の対策が可能である。IP 電話網は、オープンソースの IP-PBX ソフトウェアである Asterisk<sup>[1]</sup>を用いて構築しており、8kHz のサンプリング周波数でやりとりが行われる。図 1 にシステム構成図を示す。

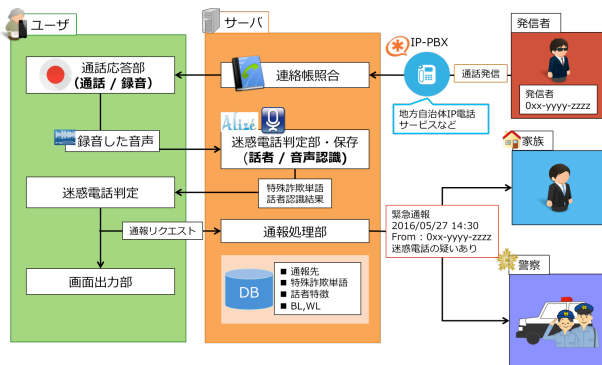


図 1 システム構成図

### 2.2. 電話帳照合

電話帳機能を備えたホワイトリスト(WL)に家族や知人の電話番号、名前などを登録する。WLにある番号とは制限なく通話が行える。怪しいと感じた電話番号はブラックリスト(BL)に登録することで、その番号からの通話を拒否できる。

### 2.3. 迷惑電話判定

WL, BL のどちらにも該当がない場合、迷惑電話判定機能が実行される。音声認識と話者認識を行い、その結果を「安全」「要注意」「危険」の三段階のレベルで確認できる。判定には音声を通話端末で録音し、リアルタイムに行う。

### 2.4. 音声認識

警察庁ホームページ<sup>[2]</sup>や地元警察から提供を受けた「特殊詐欺に使用される主な単語」から作成した詐欺の種類別のリストと、通話中に検出された単語と比較される。検出には汎用音声認識エンジンである Julius<sup>[3]</sup>を利用する。

### 2.5. 話者認識

あらかじめ家族との会話を録音し、声の特徴を抽出してデータベースに登録する。通話相手の音声特徴と、家族の音声特徴を比較することで家族かどうか判定する。話者認識には ALIZE<sup>[4]</sup>ライブラリを利用し、アルゴリズムとして GMM-UBM<sup>[6]</sup>を使用する。

### 2.6. 緊急通知

迷惑電話判定において「危険」と判断された場合、システム利用者にバイブレーションで通知を行う。事前に登録されている家族などの連絡先には状況を記載したメールを自動送信する。

### 2.7. 管理ツール

WEB 上にて提供し、管理者によりユーザごとのログ情報の表示を行う。将来、詐欺電話のトレンド解析などの活用が期待できる。

## 3. 通話音声からの話者認識率の調査

本システムは IP 電話を利用するため、通話端末におけるサンプリング周波数が 8kHz と通常の 16kHz よりも低い。また、回線状況により認識率の低下が生じると考えられる。そこで、男女 10 名ずつを対象に認識率を算出した。

発声側から通話を発信し、受信側が通話に应答した時点から音声の録音を開始する。また、GMM-UBM アルゴリズムで関係するバックグラウン

ドモデル(UBM)の生成回数と混合ガウス分布数を上昇させ、認識率の変動を確認した。

### 3.1. 録音環境

発声側は、騒音などの外乱が録音に影響を及ぼさない部屋(騒音の平均 39.82db 程度)で行い、通話端末に iPad mini 1 を使用し、クライアントには Zoiper を使用した。受信側では、通話端末として iPhone5s を使用した。クライアントは本システムを使用し、マイクをミュートにして外乱の発生を防止した。

### 3.2. 録音音声

録音音声は本校国語科の講義に使用されていた『高専の国語』に掲載されている現代文 2 作品とした。実際に音声を録音したところ、登録する音声の長さは 02:01~02:50, 比較する音声の長さは 01:11~01:41 となった。

### 3.3. パラメータごとの認識率の調査

UBM の生成回数と混合ガウス分布数を上昇させたときの処理時間と認識率の結果を表 1 に示す。混合ガウス分布数を 32, 128, 1024 に、UBM 生成回数を 5, 10, 15 に設定した。評価値として EER(等価エラー率)を使用する。また、16kHz, 英語音声での実験結果<sup>[5]</sup>と値変化の比較を行った。

表 1 パラメータごとの認識率

混合ガウス分布数	UBM 生成回数	UBM学習時間(秒)	1分11秒のモデル作成+判定時間(秒)	EER (等価エラー率)	認識率	EER (16kHz, 英語) <sup>[5]</sup>
32	5	21.99	0.38	17.82%	82.18%	6.02%
	10	35.64	0.35	17.40%	82.55%	-
	15	48.27	0.35	17.43%	82.49%	5.30%
128	5	67.86	1.21	16.09%	83.90%	2.40%
	10	112.00	1.22	14.94%	85.05%	-
	15	155.62	1.21	16.20%	83.79%	2.22%
1024	5	496.73	8.55	14.91%	85.11%	1.46%
	10	842.43	8.47	13.88%	86.12%	-
	15	1173.62	8.63	14.94%	85.05%	1.43%

次に、UBM の生成回数 5 とするときの DET カーブを図 2 に示す。

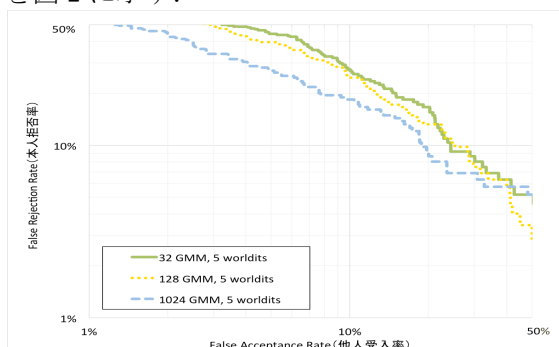


図 2 DET カーブ

16kHz の音声と値変化を比較したとき、8kHz の音声の EER はいずれも減少傾向にある。さらに、今回の実験で設定した UBM 生成回数のうち、生成回数 10 のときに認識率が最大となり、生成回数 15 で低下する傾向にあることがわかった。また、混合ガウス分布数を上昇させた際には、認識率の向上が見られた。

## 4. 考察

本研究では、IP 電話による迷惑電話防止システムを提案、開発した。IP 電話上での話者認識実験を行い、話者一致度の精度が 8 割以上であることを確認した。

混合ガウス分布数を上昇させることにより、EER が減少し、処理に必要な時間が増えた。しかし、学習は通話中の時間以外で行われるため判定時間に影響を及ぼさない。また、1 人あたりの比較には 1 分間の音声に対して応答が 10 秒未満で返ってきている。これらのことから、認識率の向上に有効であり、システムを運用する上での有用性を確かめることができた。その一方、UBM 生成回数を増加させた際に、逆効果を及ぼし認識率が低下してしまった。これは、過学習によるものではないかと考えている。

## 5. 今後の課題

今回の実験は IP 電話上での音声を対象としたが、今後サンプリング周波数の低下に影響されない普通の音声での認識率を算出し、比較を行い、どの程度認識率が低下するかを確認したい。

また、作成したシステムは、iPhone 上で動作するシステムであり、システムの対象者である高齢者は自宅にある固定電話の利用が大半を占めている。そのため、固定電話に対応させるための機器の開発が求められる。

## 6. おわりに

卑劣な手口によって 1 日あたり数百万円の被害が発生している特殊詐欺は非常に重要な問題である。各組織や機関と連携して本システムの実用化を進め、被害を未然に防ぐ一手段としていきたい。

## 7. 参考文献

- [1] Asterisk : <http://www.asterisk.org/>
- [2] 警察庁(2016.12.01) : [https://www.npa.go.jp/safetylife/seianki31/1\\_hurikome.htm](https://www.npa.go.jp/safetylife/seianki31/1_hurikome.htm)
- [3] Julius : <http://julius.osdn.jp/>
- [4] ALIZE : <http://mistrall.univ-avignon.fr/>
- [5] Hernandez, L et al. (2015). Evaluation of the ALIZE/LIA\_RAL Speaker Verification Toolkit on an Embedded System. Technische Universität Wien
- [6] Bimbot, F et al. (2004). A Tutorial on Text-Independent Speaker Verification. EURASIP Journal on Advances in Signal Processing, 430-451.