4ZB-05

IP 電話を用いた迷惑電話防止システムの開発

Calvin Janitra Halim† 吾藤 秀亮† 青井 佑太† 伊藤 宏紀† 長尾 和彦†
弓削商船高等専門学校‡

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[1]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．

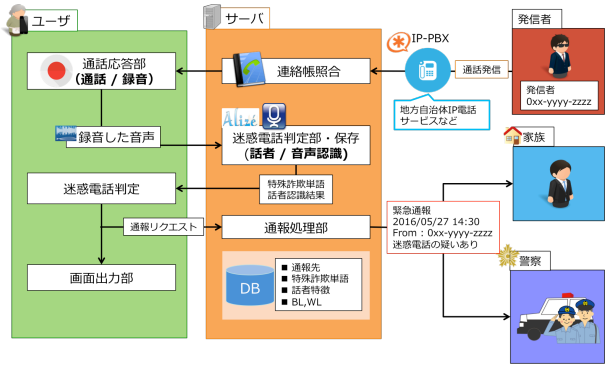

図 1 システム構成図

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[2]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[3]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[4]ライブラリを利用し，アルゴリズムとして GMM-UBM[6]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，受信側が通話に応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

Development of VoIP-based Unwanted Call Prevention System
† Calvin Janitra Halim, Hideaki Goto, Yuta Aoi, Hironori Ito, Kazuhiko Nagao
‡ National Institute of Technology, Yuge College

Copyright ©2017 Information Processing Society of Japan.
All Rights Reserved.

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[5]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)[5]
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．

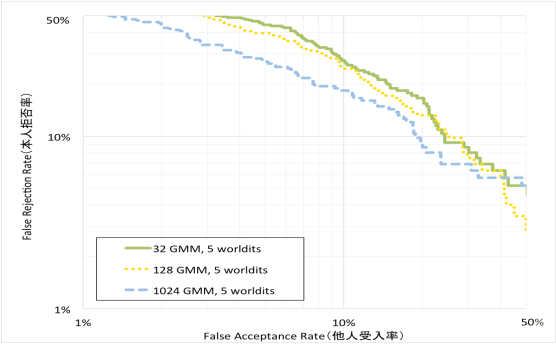

図 2　DET カーブ

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段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
[3]Julius：http://julius.osdn.jp/
[4]ALIZE：http://mistral.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.

Copyright ©2017 Information Processing Society of Japan.
All Rights Reserved.