様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　２ 年　 ６ 月 １５ 日現在

機関番号：３２６８９
研究種目：挑戦的研究（萌芽）
研究期間：2018～2019
課題番号：１８Ｋ１９７８９

研究課題名（和文）音声セキュリティ研究の開拓

研究課題名（英文）Pioneering Acoustic Security Research

研究代表者
森　達哉（Mori, Tatsuya）

早稲田大学・理工学術院・教授

研究者番号：６０７０８５５１

交付決定額（研究期間全体）：（直接経費）　　4,800,000円

研究成果の概要（和文）：本研究課題は、音声を入出力として利用する音声アシスタントシステムに固有なセキュリティ脅威の解明と、その対策技術の開発に取り組んだ。テーマの方向性として、(1) 音声アシスタントシステムに対する攻撃と(2) 音声アシスタントによる攻撃の2つがあり、その両方に取り組んだ。前者のテーマでは、　音声アシスタントシステムへのコマンドインジェクション攻撃を指向性スピーカを用いることで実現する攻撃の評価とその対策技術の開発を実施した。後者のテーマでは音声アシスタントシステム上で動作するアプリ(以下、VAアプリ)に着目し、多数のVAアプリを収集・解析しその挙動やあるべき設定を明らかにした。

研究成果の学術的意義や社会的意義
本研究で取り組んだ2つのテーマは、より広くはアナログ信号を扱うIoT機器の入出力にかかわる研究テーマの創出、およびユーザからは見えない、音声アシスタントシステムの裏側のロジックであるVAアプリに着目した研究で、ユーザから見える音声インタフェースの背後にある、クラウド側で動作するサービスに内在するセキュリティにかかる、新たな研究テーマの創出につながった。本研究課題は「音声セキュリティ研究の開拓」と銘打ち、音声を使ったシステムにかかる新たなセキュリティ脅威の同定と、その対策方法を開発するためのフレームワークづくりが大きな目標であったが、その目標は十分に達成できたと考えられる。

研究成果の概要（英文）：This research project aims to understand the security threats inherent in voice assistant systems. We also developed technologies that aim to mitigate such threats. There are two main themes: (1) attacks on voice assistant systems and (2) attacks by voice assistant systems. We tackled both themes. In the former theme, we studied the feasibility of command injection attacks on voice assistant systems using directed sound beams. We also developed an effective countermeasure against the threat. In the latter theme, we collected and analyzed a large number of voice assistant applications and clarified their behavior and settings.

研究分野：情報セキュリティ

キーワード：IoT　セキュリティ　アプリ

様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９（共通）

１．研究開始当初の背景

システムに対する入力がセキュリティ脅威になるケースは広く知られている．例えば Web セキュリティにおいては，SQL インジェクションと呼ばれる攻撃が有名である．これは Web システムが受け取る入力中に，秘密裏に SQL コマンドを紛れ込ませることにより，意図しないデータベースへのアクセスを許す攻撃である．SQL インジェクションが成立するとデータベース中に格納された個人情報の漏えいや，データベースの改竄，破壊などが可能になる．このような攻撃に対する対策としては，データの検査|すなわち入力値を適切にエスケープ処理するエスケープ関数やバインド機構の利用が有効である．もうひとつの例はバッファオーバーフロー攻撃である．攻撃者が悪意をもってプログラムに対して値を入力することにより，プログラム開発者が意図していないようなメモリ領域の破壊と上書きが生じ，結果として機器の制御権を奪うようなセキュリティホールが生じる．このような不正なメモリアクセスが起きないよう，プログラミング言語，ライブラリ，CPU による実行保護など様々なレベルで対策技術が開発されている．上述の例はいずれもデジタルデータがシステムあるいはプログラムへの入力となるケースであった．一方，様々なセンサー技術の進展とともに，コンピューターは様々なアナログデータを受け取るようになった．その一例は音声であり，音声認識・音声合成技術の飛躍的な性能向上に伴い，社会において利用される場面が増えている．本研究の目的において示したように，音声情報は必ずしも信頼できる相手から発信されるとは限らない．したがって偽の音声情報がもたらす脅威への対策として，　音声情報を何らかの方法で検査する手段の確立が必要不可欠であるとの考えに至った．

２．研究の目的

背景で示したような、悪意のある音声情報を人間の耳から隠蔽した状態でシステムに送信したり，偽の音声情報を人間に送信することで生じる脅威を本研究では「音声セキュリティ」と呼称する．本研究の目的は，　音声セキュリティにおいて起こりうる脅威の解明と有効な対策手段の開発を行うことにある．

３．研究の方法

音声セキュリティの課題は(1) 機械に対して秘密裏に音声入力を行う脅威と(2) 人間に対して偽の音声情報を送られる脅威に分類することができる．本研究ではこれらの脅威を個別の研究課題として取り組む．(1) については指向性スピーカを用いることにより、周囲の人間には聞こえない音声コマンドを AI スピーカに印加する攻撃の実現性評価と対策技術の開発に取り組んだ。(2) に関しては音声アシスタントシステムで動作するアプリに着目し、アプリの挙動を解析する手法の開発と、その手法を用いた大規模なアプリ解析に取り組んだ。(1) については、機械への音声入力を騙す攻撃として，機械の近くにいる周囲の人間に気が付かれることなく，機械に音声認識を成功させる脅威を考える．そのような脅威を実現する具体的な方法として，攻撃者が音波に指向性を持たせることによって，対象となる機械にのみ音波を送信し，周りの人間に対しては音波が届かなくする方法の実現可能性を検証した．(2)については，音声アシスタントシステム上で動作するアプリとの対話を自動化し，考えうる入力に対する出力を収集し，それらのデータを解析することで，音声アシスタントアプリがどのような情報を収集しているかを詳細に分析した．

４．研究成果

本研究を通じ，以下の成果を得た．

査読あり論文

- Ryo Iijima, Shota Minami, Zhou Yunano, Tatsuya Takehisa, Takashi Takahashi, Yasuhiro Oikawa, Tatsuya Mori, “**Audio Hotspot Attack: An Attack on Voice Assistance Systems Using Directional Sound Beams and its Feasibility**,” IEEE Transactions on Emerging Topics in Computing PP(99):1-1 · November 2019 (online early access: https://ieeexplore.ieee.org/document/8906174)

- Ryo Iijima, Shota Minami, Zhou Yunano, Tatsuya Takehisa, Takashi Takahashi, Yasuhiro Oikawa, Tatsuya Mori, “**Audio Hotspot Attack: An Attack on Voice Assistance Systems Using Directional Sound Beams**,” Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security 2222-2224 (ポスター発表)

- Atsuko Natatsuka, Ryo Iijima, Takuya Watanabe, Mitsuaki Akiyama, Tetsuya Sakai, and Tatsuya Mori, “**A First Look at the Privacy Risks of Voice Assistant Apps,**” (ポスター発表), Proc. of ACM Conference on Computer and Communications Security

査読なし発表

- 飯島涼，南翔汰，シュウ・インゴウ，竹久達也，高橋健志，及川靖広，森達哉，“**超音波の分離放射による音声認識機器への攻撃:ユーザスタディ評価と対策技術の提案**” コンピュータセキュリティシンポジウム 2018 論文集，Vol．2，pp．17-24，2018 年 10 月
- 刀塚敦子，飯島涼，渡邉卓弥，秋山満昭，酒井哲也，森達哉，”**Voice Assistant アプリの大規模実態調査**“ コンピュータセキュリティシンポジウム 2019 論文集，Vol．2019，pp．618-625，2019 年 10 月
-

表彰

1. **SCIS2018 論文賞** 飯島涼，南翔汰，シュウインゴウ，及川靖広，森達哉，”**パラメトリックスピーカーを利用した音声認識機器への攻撃と評価**“
2. **サイバーセキュリティシンポジウム道後 2018 最優秀学生研究賞** 飯島涼，”**パラメトリックスピーカーを利用した音声認識機器への攻撃と評価**“
3. **CSS2018 最優秀論文賞** (受賞率: 2/183) 飯島涼,南翔汰,シュウインゴウ,竹久達也,高橋健志,及川靖広,森達哉，“**超音波の分離放射による音声認識機器への攻撃:ユーザスタディ評価と対策技術の提案**” コンピュータセキュリティシンポジウム 2018 論文集
4. **CSS2019 最優秀論文賞** 刀塚敦子,飯島涼,渡邉卓弥,秋山満昭,酒井哲也,森達哉,”**Voice Assistant アプリの大規模実態調査**“ コンピュータセキュリティシンポジウム 2019 論文集，Vol．2019
5. **2019 年度（令和元年度）山下記念研究賞** 飯島涼， 超音波の分離放射による音声認識機器への攻撃:ユーザスタディ評価と対策技術の提案 (CSS2018)

報道

- 2019/12/16 How to Silently Hack a Smart Speaker By Michelle Hampson (IEEE Spectrum) https://spectrum.ieee.org/tech-talk/consumer-electronics/audiovideo/how-to-silently-hack-a-voice-assistance-system
- 2018/7/27 ＩｏＴクライシス サイバー攻撃があなたの暮らしを破壊する 「ＡＩスピーカーも乗っ取り可能」 (NHK スペシャル取材班)
- 2018/4/6 声で家電を操作「注目集める AI スピーカー」家電量販店、今後に期待 (苫小牧民報)

５．主な発表論文等

〔雑誌論文〕　計3件（うち査読付論文　3件／うち国際共著　0件／うちオープンアクセス　3件）

１．著者名 Iijima Ryo、Minami Shota、Yunao Zhou、Takehisa Tatsuya、Takahashi Takeshi、Oikawa Yasuhiro、Mori Tatsuya	４．巻 1
２．論文標題 Audio Hotspot Attack: An Attack on Voice Assistance Systems Using Directional Sound Beams	５．発行年 2018年
３．雑誌名 CCS '18 Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security	６．最初と最後の頁 2222-2224
掲載論文のDOI（デジタルオブジェクト識別子） 10.1145/3243734.3278497	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 －

１．著者名 Iijima Ryo、Minami Shota、Zhou Yunao、Takehisa Tatsuya、Takahashi Takeshi、Oikawa Yasuhiro、Mori Tatsuya	４．巻 1
２．論文標題 Audio Hotspot Attack: An Attack on Voice Assistance Systems Using Directional Sound Beams and its Feasibility	５．発行年 2019年
３．雑誌名 IEEE Transactions on Emerging Topics in Computing	６．最初と最後の頁 1-15
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/TETC.2019.2953041	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 －

１．著者名 Natatsuka Atsuko、Iijima Ryo、Watanabe Takuya、Akiyama Mitsuaki、Sakai Tetsuya、Mori Tatsuya	４．巻 1
２．論文標題 A First Look at the Privacy Risks of Voice Assistant Apps	５．発行年 2019年
３．雑誌名 CCS '19 Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security	６．最初と最後の頁 2633-2635
掲載論文のDOI（デジタルオブジェクト識別子） https://doi.org/10.1145/3319535.3363274	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 －

〔学会発表〕　計7件（うち招待講演　5件／うち国際学会　0件）

１．発表者名 飯島涼、南翔汰、シュウインゴウ、竹久達也、高橋健志、及川靖広、森達哉
２．発表標題 音波の分離放射による音声認識機器への攻撃:ユーザスタディ評価と対策技術の提案
３．学会等名 コンピュータセキュリティシンポジウム2018
４．発表年 2018年

１．発表者名
飯島涼

２．発表標題
音声認識×セキュリティ研究の最新動向 〜 超音波の分離放射による音声認識機器への 攻撃と対策手法の提案 〜

３．学会等名
ハードウェアセキュリティフォーラム2018（HWS2018）（招待講演）

４．発表年
2018年

１．発表者名
森達哉

２．発表標題
アナログ信号のセキュリティ

３．学会等名
電子情報通信学会総合大会計測セキュリティと今後の方向性（招待講演）

４．発表年
2019年

１．発表者名
森達哉

２．発表標題
機械学習のセキュリティ

３．学会等名
第41回情報理論とその応用シンポジウム(SITA)（招待講演）

４．発表年
2018年

１．発表者名
刀塚敦子，飯島涼，渡邉卓弥，秋山満昭，酒井哲也，森達哉

２．発表標題
Voice Assistant アプリの大規模実態調査

３．学会等名
コンピュータセキュリティシンポジウム2019

４．発表年
2019年

１．発表者名 森達哉
２．発表標題 音声アシスタントシステムのセキュリティ・プライバシー
３．学会等名 SecurityDay 2019（招待講演）
４．発表年 2019年

１．発表者名 Iijima Ryo、Minami Shota、Yunao Zhou、Takehisa Tatsuya、Takahashi Takeshi、Oikawa Yasuhiro、Mori Tatsuya
２．発表標題 Latest Trends in Voice Assistance Systems and Security Research: An Attack on Voice Assistance Systems Using Directional Sound Beams
３．学会等名 The 14th International Workshop on Security (IWSEC)（招待講演）
４．発表年 2019年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

―

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考
研究協力者	飯島　涼 (Iijima Ryo)	早稲田大学・理工学術院・大学院生 (32689)	
研究協力者	刀塚　敦子 (Atsuko Natatsuka)	早稲田大学・理工学術院・大学院生 (32689)	