ライトコマンド：レーザーを用いて音声コマンドを挿入する攻撃

菅原 健[1,a)]　ベンジャミン シア[2]　サラ ランパッジ[2]　ダニエル ゲンキン[2]　ケビン フー[2]

概要：本稿は，国際会議 USENIX Security Symposium 2020 [10] で発表済みの内容に基づく．光が音に変換される物理現象を用いて，マイクに信号を挿入する新しいクラスの攻撃を提案する．まず，振幅変調した光をマイクの開口部に当てることで，対象のマイクに任意の音声を挿入できることを示す．また，この方法を用いると，音声制御システムに対して，遠隔から音声コマンドを挿入できることを示す．Amazon Alexa, Apple Siri, Facebook Portal, そして Google Assistant の評価を行い，110 メートルの遠方や，隣の建物から，それらの制御を奪うことができることを示す．コマンド挿入により引き起こされる被害は，音声制御システムに接続された機器による．本稿では特に，次のような攻撃ができることを示す: (i) スマートロックを解錠する，(ii) ガレージドアを開ける，(iii) ウェブサイトで買い物をする，(iv) 対象の Google アカウントに結びついた車を解錠してエンジンを始動させる（例: Tesla と Ford）．最後に，この攻撃に対するソフトウェア・ハードウェア的な対策法について述べる．

キーワード：センサのセキュリティ，シグナルインジェクション攻撃，音声コマンド，レーザー

Voice Command Injection Using Laser

Takeshi Sugawara[1,a)]　Benjamin Cyr[2]　Sara Rampazzi[2]　Daniel Genkin[2]　Kevin Fu[2]

Abstract: This paper is based on the study by the same authors published in USENIX Security Symposium 2020 [10]. We propose a new signal injection attack on microphone using light; we can inject arbitrary signal to a microphone by shining light modulated with the target signal. An attacker can exploit this technique to inject commands to voice controllable systems silently. We verify the attack on various products based on Amazon Alexa, Apple Siri, Facebook Portal, and Google Assistant, and demonstrate that the attack is feasible from 110 meters away and through a glass window across buildings. The consequence of a command injection depends on the devices associated with the voice controllable systems, which includes (i) unlocking a smart lock, (ii) opening a smart garage door, (iii) buying things on e-commerce websites, and (iv) locating, unlocking, and starting vehicles. Finally, we discuss the possible countermeasures against the proposed attack.

Keywords: Security of Sensors, Signal Injection Attack, Voice Command, Laser

1. はじめに

情報技術の発展により，コンピュータが人間の言葉を認識して理解できるようになりつつある．そこで，人間が発した言葉をコマンドとして処理する音声制御システム (VCS: Voice Controllable System) を，Apple, Google, Facebook, そして Amazon などの IT 企業がこぞって市場に投入している．その結果，Alexa [1], Siri [2], Portal, Google Assistant [6] を搭載した何千万台もの機器がすでに販売されている．これを受け，VCS との連携機能を有する Internet of Things (IoT) 製品が，多くのサードパーティ企業より発売されている．以上により，VCS のユー

[1] 電気通信大学
The University of Electro-Communications
[2] ミシガン大学
University of Michigan
[a)] sugawara@uec.ac.jp

ザは言葉を発するだけで— 物理的なキーボード，マウス，タッチスクリーンに触れることなく—製品から情報を得たり，その制御を行うことができるようになった．

VCS とその対応製品に大きな期待が寄せられるのに対して，ソフトウェア的・ハードウェア的な攻撃への安全性については分かっていないことが多い．実際，先行研究 [5], [7] は VCS の重大な課題として，適切な認証方法が欠けているため，潜在的な攻撃者によるコマンドをコマンドを実行できてしまうことを指摘している．初期の攻撃は，大声で叫ぶといったものであったため，近くにいる正規ユーザが容易に気づくことができた．しかし，より新しい攻撃は，ユーザに気づかれづらいこと（ステルス性）に主眼が置かれており，場合によっては無音でコマンドを入力することができる [3], [4], [8], [9], [11], [12], [13].

音声だけで認証を行うのは難しいため，VCS の安全性は物理的な距離に頼っている．すなわち，攻撃者は物理的な壁，鍵のかかったドア，および窓などにより阻まれるため，機器の近くにいるユーザは正規ユーザと見なせるはずである，という想定である．実際，利用者にとってありがたいことに，最新の攻撃であっても，従来の攻撃法の射程は 7〜8 メートルに留まっている [9]．この射程は，開けた空間における値であり，窓などの物理的な障壁がある時はさらに短くなる．

それに対して私たちは，以下の問いを立てて研究を行った: 「VCS に対し，遠隔から気づかれることなくコマンドを入力することはできるだろうか？もしできるならば，攻撃者に課せられる現実的な制約の元でも実行可能だろうか？また，そのような遠隔からのコマンド入力が VCS に対応したサードパーティ製品に与える影響範囲はどれほどであろうか？」

2. 貢献

遠隔から，気づかれることなく VCS にコマンドを入力できる Light Commands （ライトコマンド）と呼ばれる攻撃法を提案した [10].

レーザーを用いた音声信号のインジェクション

マイクの仕様と物理の間に横たわるセマンティックギャップとして，マイクが光に反応し，あたかも音を聞いたかのように認識してしまう現象を発見した．この現象を利用すると，音声信号で振幅変調したレーザービームをマイクに照射することで，マイクに音声信号を受信させることができる．

音声制御システムの安全性評価

最も普及している VCS（Alexa, Siri, Portal および Google Assistant）について，レーザーを用いた音声信号のインジェクションの安全性評価を行った．その結果，(i) 多くのスマートホーム機器に音声コマンドを挿入するには 5 mW (米国におけるレーザーポインタの規制値) という微弱な出力のレーザーで十分であること，(ii) その 10 倍にあたる 50 mW のレーザーを用いればスマートフォンやタブレットでも攻撃が成功することを明らかにした．

攻撃距離の延伸

カメラ用望遠レンズを用いてレーザーを集光することで，これまでよりも遥かに長い射程からの攻撃が可能であることを明らかにした．いくつかのスマートスピーカー製品では，110 メートル（用意した実験環境の限界値）遠方からでも攻撃が成功することを示した（図 1）．また，より現実的な環境として，ある建物から数十メートル離れた別の建物へ向けて，ガラス窓を貫通して攻撃が可能であることも示した（図 2）．従来法による攻撃距離には，音波を使うことに起因する限界があったのに対し，提案法はレーザーパワー，光学的な集光能力，および照準の精度にしか制限されていない．

不十分な認証の発見

上記の遠隔からの攻撃が，VCS のセキュリティに与える影響について調査した．VCS では認証が全く無い場合や，もしあったとしても不適切な実装がされている（例: PIN を総当りできる）場合があることを明らかにした．その結果，レーザーによる音声信号のインジェクションにより (i) スマートロックの解錠，(ii) ガレージドアの開放，(iii) E コマースサイトによる買い物，および (iv) VCS に接続された特定車種 (Tesla および Ford) の位置特定，解錠，およびエンジン始動が行えることを示した．

攻撃の最適化: 低コストな攻撃と目立たない攻撃

上記の攻撃が，（科学用の実験機器でなく）安価に販売されているレーザーポインターとレーザードライバを組み合わせることで実現できることを明らかにした．また，不可視な赤外レーザーや，VCS のボリューム制御機能（例: Alexa のささやきモード）を悪用することで，正規ユーザに気づかれづらい目立たない攻撃を行うことができることを示した．

対策法

上記攻撃への対策法について，ソフトウェアおよびハードウェアの両面から検討した．ソフトウェアによる対策法としては，認証の層を追加することは攻撃の緩和に役立つ．また，ユーザと機器の間で簡単なインタラクションをするようにすれば，対象機器のレスポンスを観測できない攻撃者による攻撃を防ぐことができる．ハードウェアの対策法として，複数マイクを用いて異常検知を行うようにすれば，（攻撃に用いるレーザーの本数という意味で）攻撃の難易

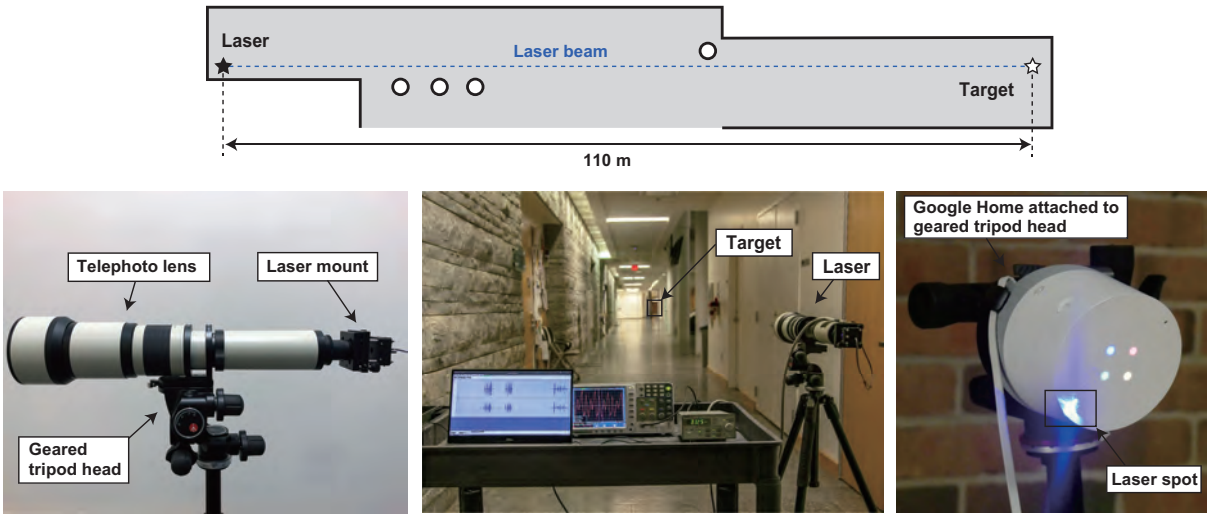

図 1 文献 [10] より引用．攻撃可能距離を調べるための実験セットアップ．（上）110 メートルの廊下のフロアプラン．（左）レーザーを取り付けた望遠レンズを，照準合わせのためのギア雲台に搭載した様子．（中央）110 メートルの廊下を介したレーザーの照準合わせ．（右）三脚に搭載した対象デバイスにレーザーが照射された様子．

度の難易度を挙げることができる．また，機器をカバーで覆ったりすることで，攻撃者からマイクまでの視線を遮ることも対策になる可能性がある．

3. 安全と脆弱性開示

レーザーの安全性

高出力のレーザーは火災や，目・皮膚への怪我を引き起こすため，実験には特別な安全措置を講じる必要がある．本研究の再現を試みる研究者は，正式な安全トレーニングを受講するとともに，管理者から実験計画の承認を受けるべきである．本研究で行った実験は全て，ミシガン大学の安全委員会で承認を受けた標準作業手順に基づいて実施した．

脆弱性開示プロセス

責任ある脆弱性開示 (Responsible Disclosure) の実践に従い，本研究で得た知見を，Google, Amazon, Apple, August (スマートロック製造者), Ford, Tesla, および Analog Devices (MEMS マイクの製造者) と共有した．各企業のセキュリティチームに加え，ICS-CERT およびアメリカ食品医薬品局 (FDA) と連絡を取り合っている．各機関と積極的に協力して，発見した攻撃の理解と緩和法について取り組んでいる．本稿，および元論文 [10] に記載した脆弱性は，各企業・機関と合意した情報解禁日（2019 年 11 月 4 日）に開示済みである．

4. まとめと今後の課題

本稿は，光を照射することで音声制御システムに対してコマンドを挿入する攻撃であるライトコマンドを提案した．音声コマンドで振幅変調した光を照射すると，マイクの内部で光から元の音声コマンドに復元されることを攻撃の原理とする．Siri, Portal, Google Assistant, および Alexa を搭載するたくさんの市販機器の安全性評価を行い，100 メートル以上の遠方からや，ガラス窓を貫通した攻撃が可能であることを示した．また，提案法が音声制御システムのセキュリティ上の問題と結びつくことで，スマートロックや車などのサードパーティ製品に攻撃が波及することを示した．

光が音に変わる物理現象を詳しく解明することで，新たな攻撃や対策法が得られると考えている．特に，同じ原理を用いれば，音波を用いる既存の攻撃法（例: モーションセンサに超音波を照射する攻撃）を，光を用いて実現できる可能性がある．また，レーザーによる加熱も，シグナルインジェクション攻撃を行う新たな手段になりうると考えている．

謝辞 本研究において，菅原は JSPS 科研費 JP18K18047 と JP18KK0312 の助成を受けた．

参考文献

[1] Amazon, “What is Alexa?” Amazon Alexa official site, `https://developer.amazon.com/alexa` (Accessed: 2019-08-20).

[2] Apple, “Siri,” `https://www.apple.com/siri/` (Accessed: 2019-08-20).

[3] Nicholas Carlini, Pratyush Mishra, Tavish Vaidya, Yuankai Zhang, Micah Sherr, Clay Shields, David Wagner, and Wenchao Zhou, “Hidden voice commands,” In *USENIX Security Symposium*, pages 513–530, 2016.

[4] Moustapha M Cisse, Yossi Adi, Natalia Neverova, and Joseph Keshet, “Houdini: Fooling deep structured visual and speech recognition models with adversarial examples,” In *Advances in neural information processing systems*, pages 6977–6987, 2017.

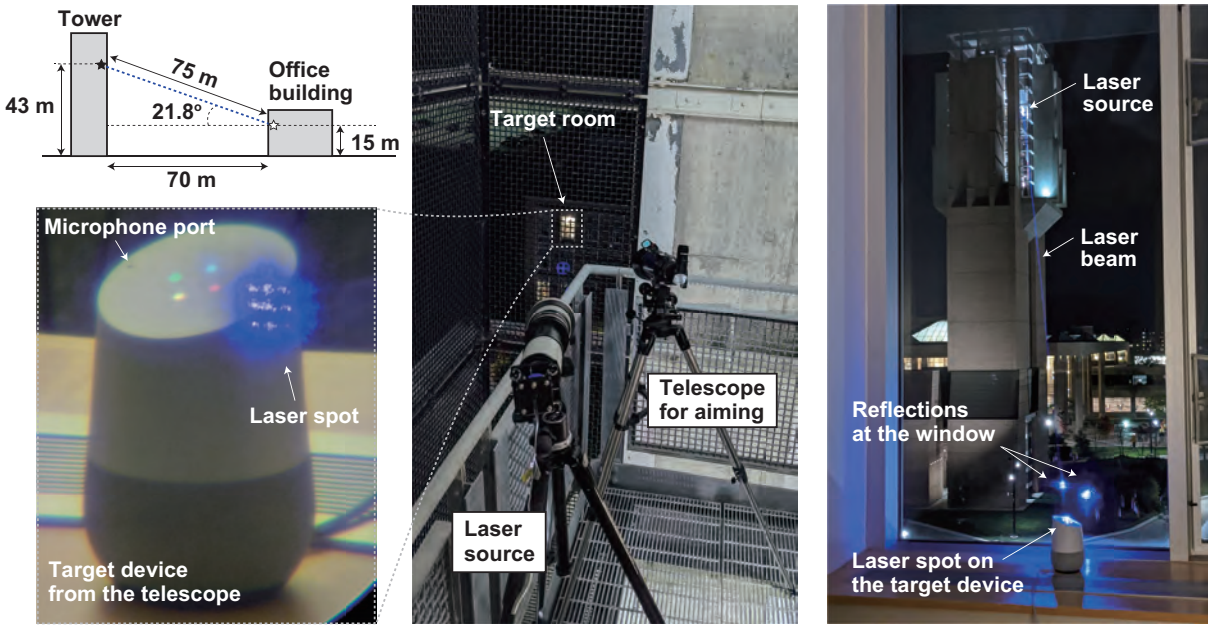

図 2 文献 [10] より引用．低出力レーザーで建物をまたいだ攻撃を行うセットアップ．（左上）レーザーと対象の位置的な関係．（左下）攻撃者が望遠鏡で見た対象デバイス．（中央）攻撃者側の建物の様子：望遠レンズに取り付けたレーザーで対象に狙いをつけている．（右）被害者側の建物の様子：対象のデバイスにレーザーが照射されている．

[5] Wenrui Diao, Xiangyu Liu, Zhe Zhou, and Kehuan Zhang. “Your voice assistant is mine: How to abuse speakers to steal information and control your phone,” In *Proceedings of the 4th ACM Workshop on Security and Privacy in Smartphones & Mobile Devices*, pages 63–74. ACM, 2014.

[6] Google, “Google Assistant, your own personal Google,” `https://assistant.google.com` (Accessed: 2019-08-20).

[7] Yeongjin Jang, Chengyu Song, Simon P Chung, Tielei Wang, and Wenke Lee. “A11y attacks: Exploiting accessibility in operating systems,” In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pages 103–115. ACM, 2014.

[8] Nirupam Roy, Haitham Hassanieh, and Romit Roy Choudhury, “Backdoor: Making microphones hear inaudible sounds,” In *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*, pages 2–14. ACM, 2017.

[9] Nirupam Roy, Sheng Shen, Haitham Hassanieh, and Romit Roy Choudhury. “Inaudible voice commands: The long-range attack and defense,” In *15th USENIX Symposium on Networked Systems Design and Implementation (NSDI 18)*, pages 547–560, 2018.

[10] Takeshi Sugawara, Benjamin Cyr, Sara Rampazzi, Daniel Genkin, and Kevin Fu, “Light commands: Laser-based audio injection attacks on voice-controllable systems,” In *29th USENIX Security Symposium (USENIX Security 20)*, 2020.

[11] Tavish Vaidya, Yuankai Zhang, Micah Sherr, and Clay Shields, “Cocaine noodles: exploiting the gap between human and machine speech recognition,” *Presented at WOOT*, 15:10–11, 2015.

[12] Xuejing Yuan, Yuxuan Chen, Yue Zhao, Yunhui Long, Xiaokang Liu, Kai Chen, Shengzhi Zhang, Heqing Huang, XiaoFeng Wang, and Carl A Gunter, “CommanderSong: A systematic approach for practical adversarial voice recognition,” In *27th USENIX Security Symposium (USENIX Security 18)*, pages 49–64, 2018.

[13] Guoming Zhang, Chen Yan, Xiaoyu Ji, Tianchen Zhang, Taimin Zhang, and Wenyuan Xu, “DolphinAttack: Inaudible voice commands,” In *ACM Conference on Computer and Communications Security*, pages 103–117. ACM, 2017.