

キーボード打鍵音による入力推定攻撃とその対策

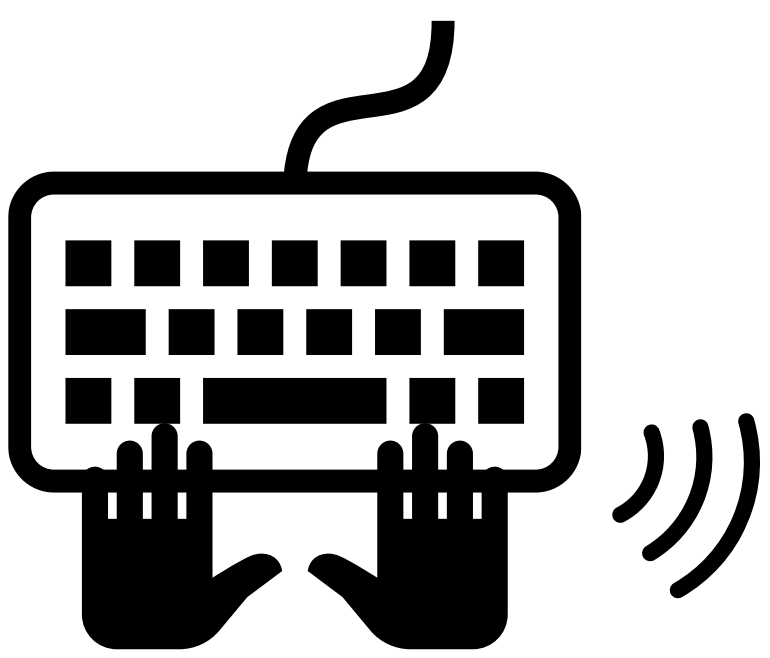
研究駆動コース 飯田雅裕

身近に潜む攻撃のリスク

新型コロナウイルスの蔓延によって、我々の働き方は大きく変化し、オンライン会議システムやバーチャルオフィスツール等の利用機会が急増 → **キーボードの打鍵音**を耳にするように

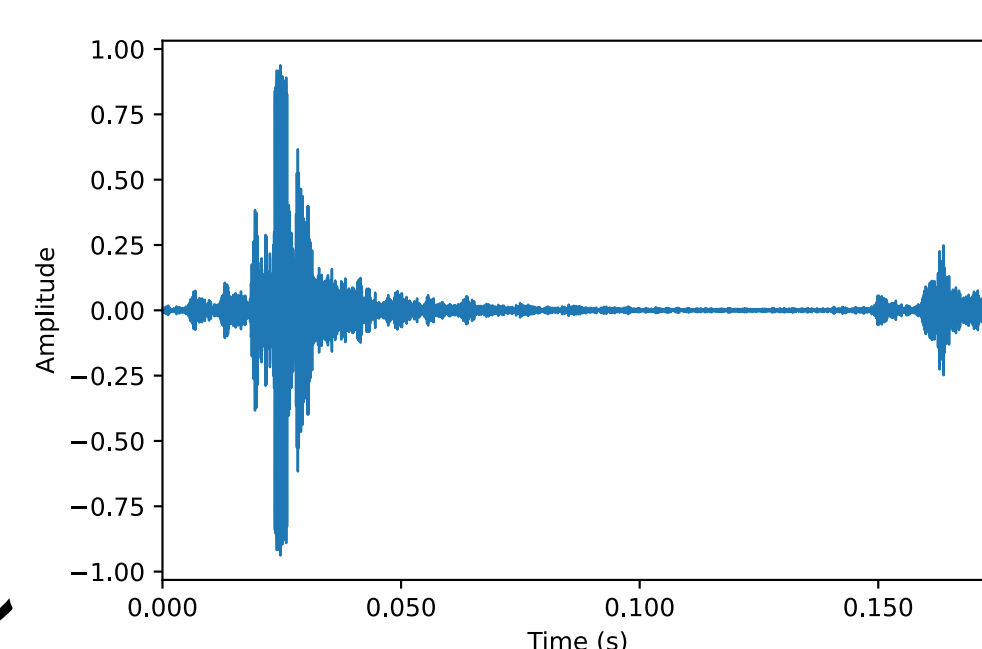
もし打鍵音から入力内容が推定できると...

- ・オンライン会議の音声から入力内容の推定
- ・ブラウザ経由でのキーロギング
- ・盗聴器によって作成した機密文書が流出
- ・生放送中に配信者のプライベートな情報が流出



これまでの研究

- ・キーボードの打鍵音を利用しユーザが入力したキーを推定する攻撃として **Keyboard Acoustic Emanations** が知られている[1-3]
- ・打鍵時に発生する音から特徴から機械学習を用いてキーの候補を推定
- ・辞書情報等から単語を推定
- ・先行研究では英語が攻撃対象として推定モデルの構築や辞書情報の利用
- ・**英語以外の言語に対する有効性は明らかでない**

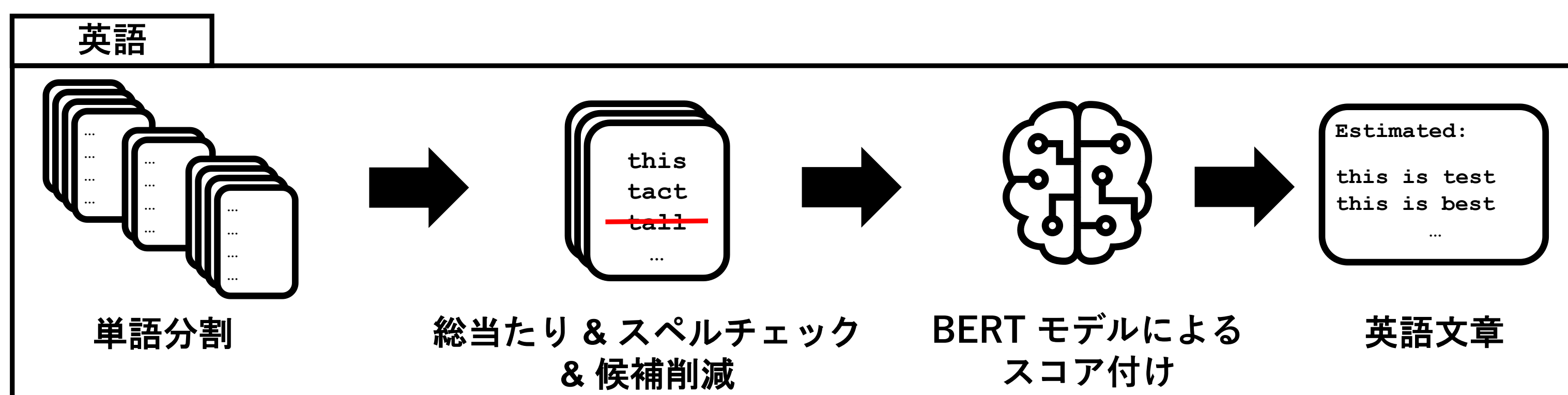
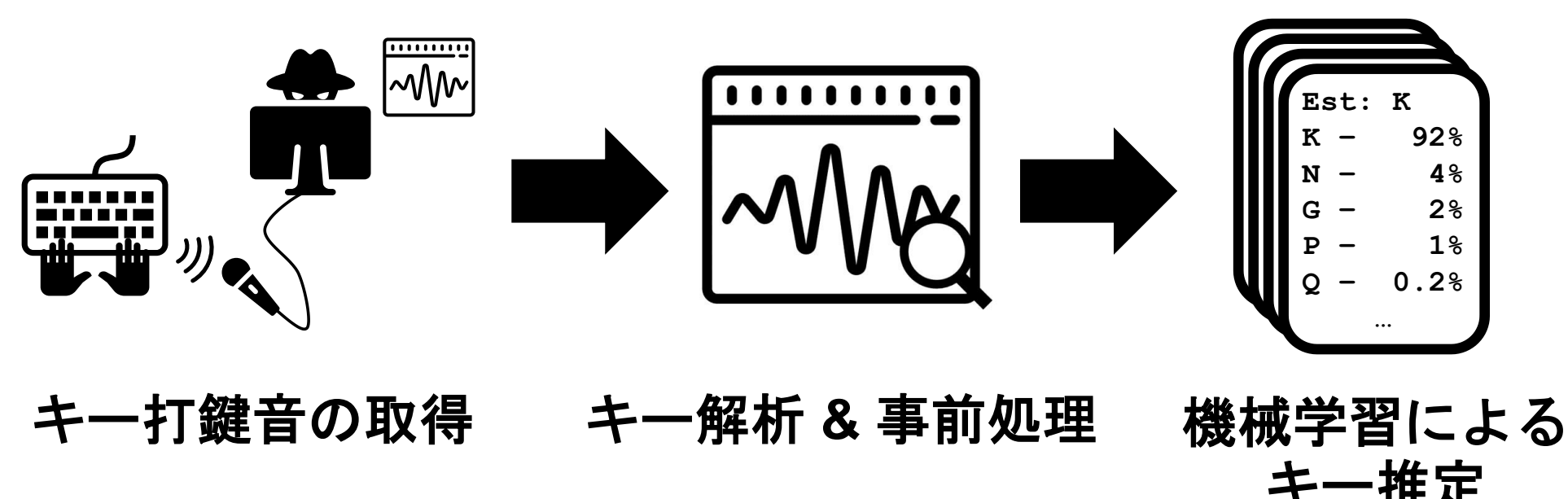


キーボード打鍵音の例

キーの推定にBERTモデルを適用した自然言語テキストに対する攻撃手法の提案

Keyboard Acoustic Emanations に対して自然言語処理モデルであるBERTモデル (Bidirectional Encoder Representations from Transformers) を適用し、**単語間の繋がりを考慮**することで自然言語テキストに対する推定精度の向上を図り、日本語に適した処理を加えることで、従来の英語に対する攻撃だけでなく**日本語に対しても攻撃可能**

1. キー打鍵音の取得から機械学習(ロジスティック回帰)によるキーの候補推定までは各言語共通で、英語の場合は推定候補の最上位がスペースキーの場合に当該箇所を区切りとして単語に分割、日本語の場合は日本語変換の確定時に押されるエンターキーを区切りとして文節に分割
2. 分割後、英語文章の場合は各キーの打鍵音から得られた推定候補の組み合わせを総当たりし、スペルチェックを行い候補となる単語を抽出、日本語文章の場合は英語同様に組み合わせを総当たりし、ローマ字変換可能な候補を抽出
3. どちらも候補を抽出後、各言語に対応するBERTモデルを用いて PLL (Pseudo-log-likelihood score)[4] でスコア付け



制約

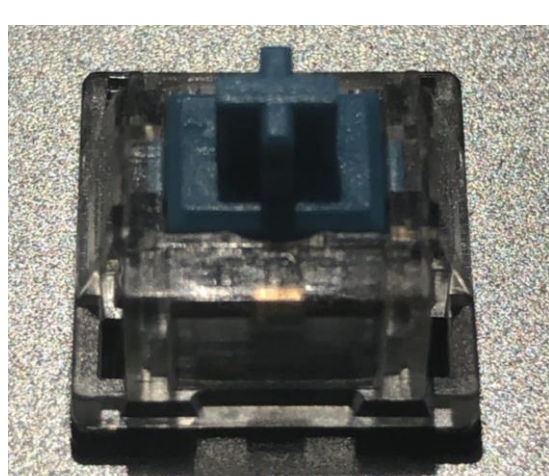
- ・攻撃対象となるユーザーが使用するキーボードの種類や入力テキストの言語が事前に特定できている必要がある
- ・教師あり機械学習を用いるため、少なくとも各キーで5回程度のラベル付き打鍵音が必要
- ・日本語文章の推定においては、単文節変換でのテキスト入力を想定しており、連文節変換の場合には分割されたキーの文字列が長くなることで、総当たりによる候補の生成が現実的な時間内に終わらない可能性がある

提案手法の検証

- ・学習に用いる各キーのデータ数と推定精度の関係は**キーボードによって異なる結果を得た**(Top-N は推定結果の上位N個の候補に正しいキーが含まれる確率を表す)
- ・学習に使用する各キーのデータ数を20個とし、英語・日本語の文章を各5文、計611回分の打鍵音を用いて推定精度を検証
- ・推定候補上位10個の中に正解が含まれていれば成功とみなす場合の確率として 10-best WCR、10-best CCR を併せて計算
- ・比較的多くの単語・文節で構成される自然言語テキストの場合、PLL を用いることで**高い精度で推定することに成功**



メカニカル (青軸) 方式



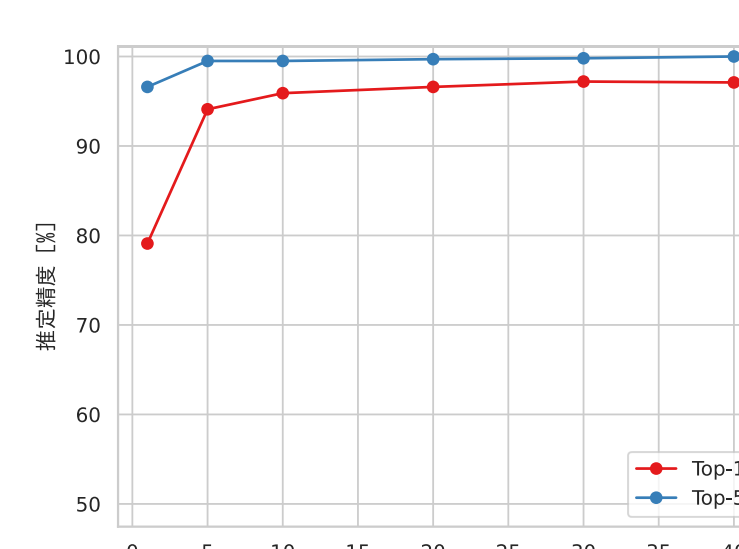
静電容量無接点方式



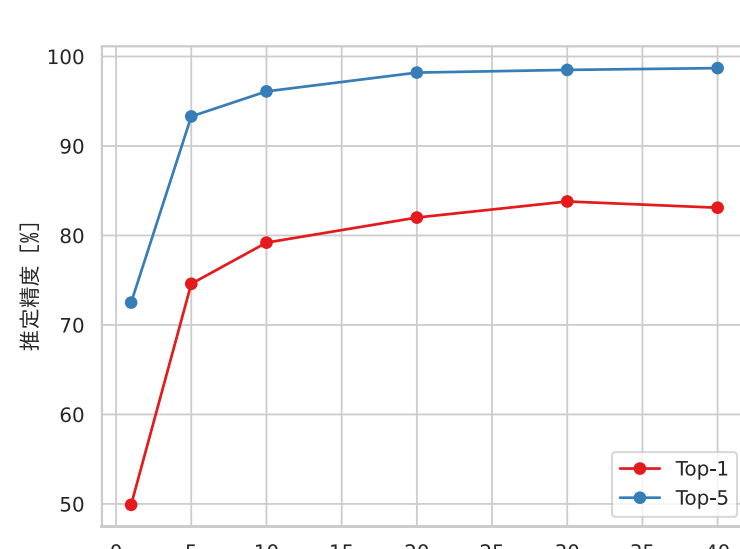
静電容量無接点方式

実験環境	
キーボード	メカニカル (青軸) 方式 静電容量無接点方式
被害者側	録音マイク Web カメラ内蔵モノラルマイク
攻撃者側	CPU AMD Ryzen 9 3900X GPU NVIDIA GeForce RTX 3080 メモリ 64GB

注) WCR (Word Correct Rate): 単語単位の推定成功率
CCR (Character Correct Rate): 入力キー単位の推定成功率



(a) メカニカル (青軸) 方式



(b) 静電容量無接点方式
学習に使用する各キーのデータ数と推定精度の関係

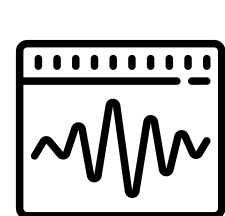
自然言語テキストに対する推定結果 (メカニカル(青軸)方式の結果: 静電容量無接点の結果)

#	推定対象となる入力テキスト	WCR[%]	CCR[%]	10-best WCR[%]	10-best CCR[%]	処理時間 [秒]
1.	this is a test message	80:80	91:96	100:100	100:100	48:40
2.	the employers have not made an official offer on cutting the working week	100:100	100:100	100:100	100:100	123:296
3.	the united states has entered a swamp from which it can in no way get out safely	100:71	100:86	100:77	100:90	90:569
4.	they managed to warn off the crew and hit something that was the least nuisance to everybody	94:82	98:92	100:100	100:100	164:385
5.	huge oil platforms dot the gulf like beacons usually lit up like christmas trees at night	100:100	100:100	100:100	100:100	170:342
6.	これはテスト文章です	100:60	100:80	100:100	100:100	36:33
7.	今日はとてもよい天気である	71:86	77:92	86:86	92:92	32:36
8.	私は辞書を開いて単語を調べた	100:80	100:86	100:80	100:86	55:59
9.	私はコンビニに行き大好きなおにぎりを購入しました	92:92	92:96	92:92	92:96	104:91
10.	今日は朝から天気がよく洗濯物を干すには絶好の日である	94:59	96:73	94:65	96:73	127:142

キー推定への対策

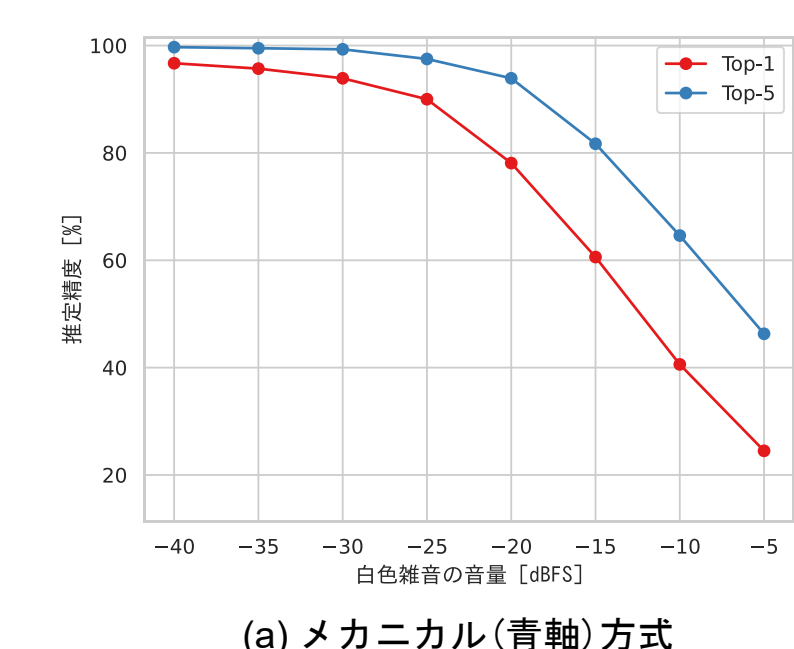
ノイズの挿入

打鍵音にノイズが含まれていると推定精度が下がるため音楽等を流すことなどが**有効**

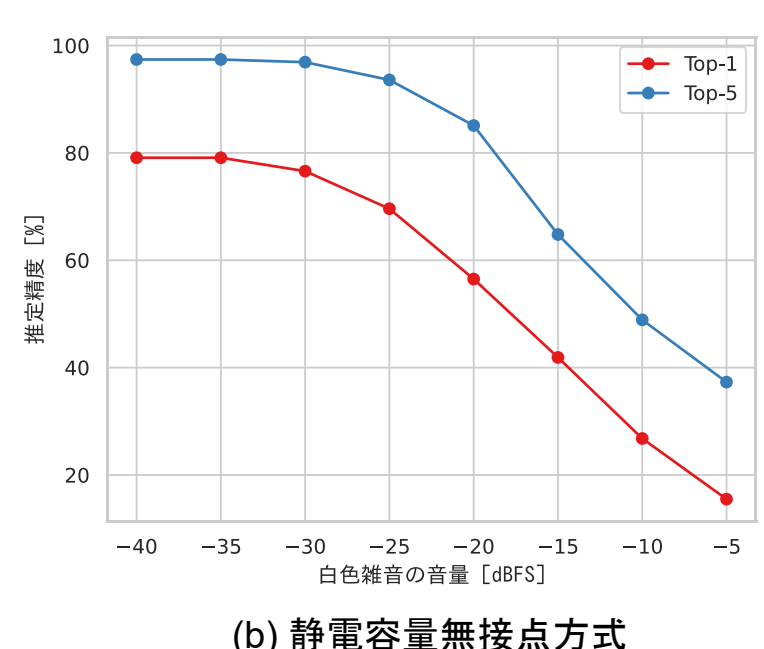


攻撃を受けにくいキーボードの使用

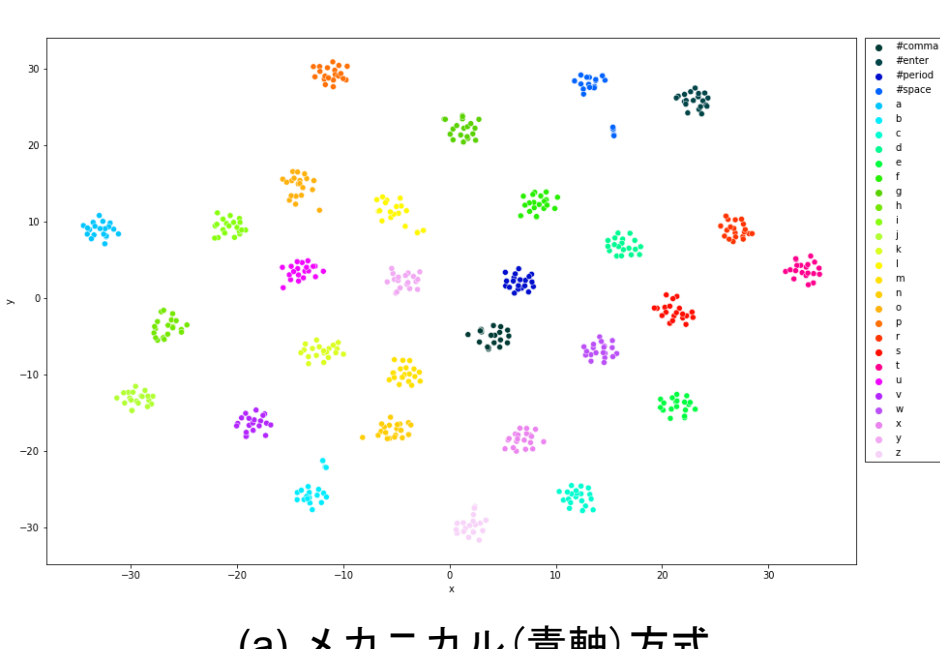
打鍵音が発生しづらいものを使用することで**推定精度を下げる**ことができる



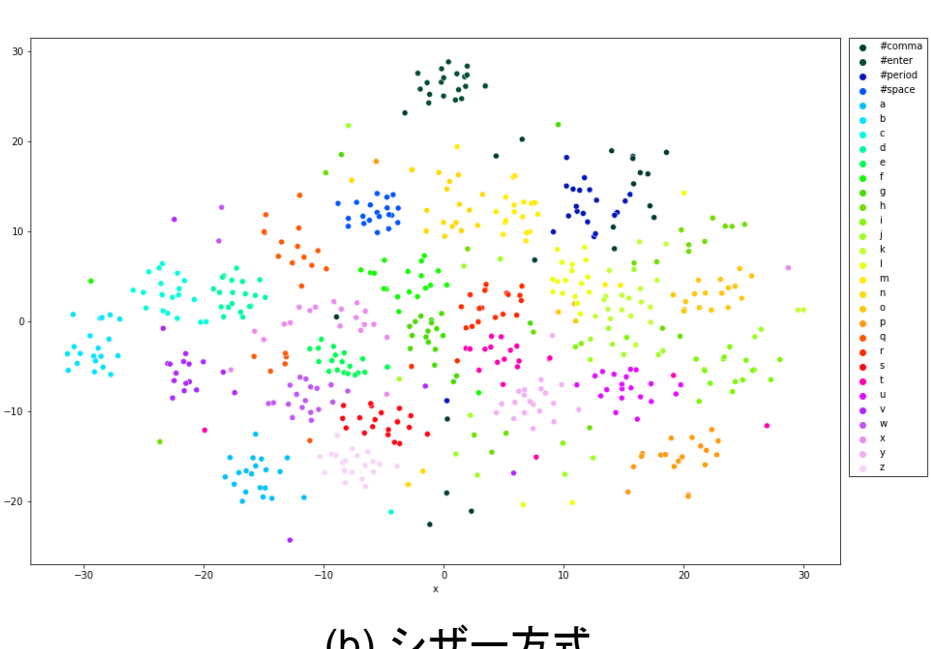
(a) メカニカル (青軸) 方式
白色雑音挿入時の推定精度への影響



(b) 静電容量無接点方式



(a) メカニカル (青軸) 方式
打鍵音の特徴量可視化



(b) シンザー方式

今後の展望

- ・ラベル付き打鍵音が不要となる教師なし機械学習での**自然言語テキストの推定**
- ・他言語への適用
- 様々な言語のBERTモデルが公開されているため、これらを用いて容易に適用できる可能性
- ・オンライン会議システム等のインターネット経由での攻撃可能性の検証
- ・攻撃を防ぐための対策手法についてまとめる

文献

- [1] D. Asanov and R. Agrawal, "Keyboard acoustic emanations," Proceedings of the 2004 IEEE Symposium on Security and Privacy, pp.3-11, 2004.
- [2] L. Zhuang, F. Zhou, and J.D. Tygar, "Keyboard acoustic emanations revisited," ACM Transactions on Information and System Security, vol.13, no.1, p.1-26, 2009.
- [3] S. Cecconello, A. Compagno, M. Conti, D. Lain, and G. Tsudik, "Skype & type: Keyboard eavesdropping in voice-over-ip," ACM Transactions on Privacy and Security, vol.22, no.4, p.1-34, 2019.
- [4] J. Salazar, D. Liang, T.Q. Nguyen, and K. Kirchhoff, "Masked language model scoring," Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, p.2699-2712, 2020.

外部発表

ICSS研究会 BERTモデルを用いたキーボード打鍵音による入力推定攻撃とその対策
飯田雅裕・秋山満昭・神菌雅紀・笠間貴弘・服部祐一・井上博之・猪俣敦夫