Practical Acoustic Eavesdropping On Typed Passphrases

April 8, 2025

Darren Fürst and Andreas Aßmuth

Agenda

Ostbayerische Technische Hochschule Amberg-Weiden

- Side-Channel Attacks
  - First discovered side-channel attack
  - Historical Meaning and Implications
- Side-Channel Attacks on User Input
- Acoustic Attack Experiment
  - Audio Side-Channel
  - Segmentation
  - Clustering
  - Passphrase Recovery
  - Results

Side-Channel Attacks

- Uses unintended effects to gather information
- Doesn't attack the algorithm itself
- Examples: Power Consumption, Time Differences, Noises

Discovery

- From a now declassified NSA document.
- An engineer at Bell Labs found a vulnerability in one of their typewriters.
- There are correlating spikes to be seen on an oscilloscope.
- With an antenna and amplifier, these pulses could be decoded from almost 25 meters away.
- These are seen whenever the Bell Labs 131-B2 typewriter encrypts a letter.

Sigaba m134c, similar type of telewriter
Photograph: Mark Pellegrini

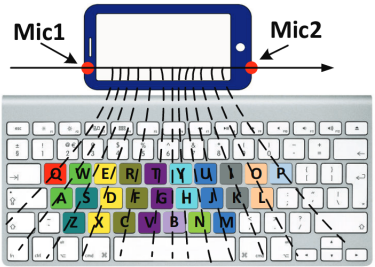
Historical Meaning and Implications

- The Bell Labs 131-B2 typewriter was used during the second world war.
- It was used by the army and the marine to transmit confidential messages.
- German and japanese forces were not able to break the algorithm used to encrypt letters.
- This shows, that theoretically safe encryption can have unwanted effects when implemented.
- These effects can potentially be used to gather information about the encryption process.

Side-Channel Attacks on User Input

Acoustic Inference

- Uses acoustic signals gathered during the typing on a keyboard
- Examples: Measuring timing differences in the arrival with two microphones [2], Frequency differences of different keys

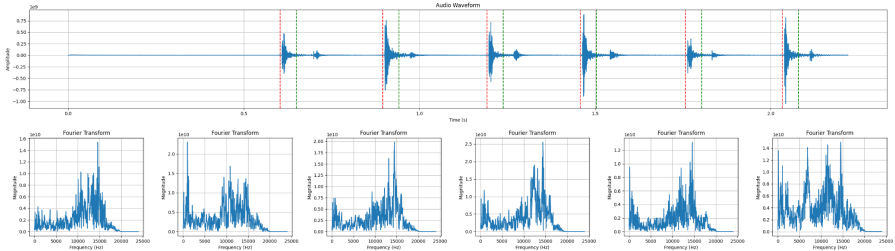


(b) Theoretical key groups and corresponding half hyperbolas.

[0][2]Snooping Keystrokes with mm-level Audio Ranging on a Single Phone

Acoustic Inference

- Uses acoustic signals gathered during the typing on a keyboard
- Examples: Measuring timing differences in the arrival with two microphones [2], Frequency differences of different keys



Movement-based Inference

- Evaluation of movement patterns of hands or fingers during typing
- Example: Analysis through smartwatches, other wearables, or typing on a mobile phone

[0][3]When Good Becomes Evil: Keystroke Inference with Smartwatch

Vibration-based Inference

- Analysis of the mechanical vibrations generated by keyboard inputs
- Example: Placement of a mobile phone on the desk near the keyboard

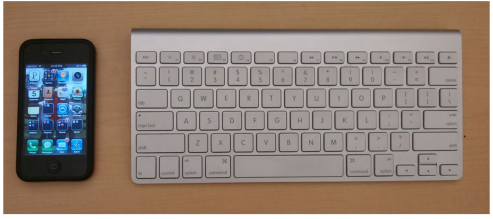


Figure 1: Our experimental placement of a mobile phone running a malicious application attempting to recover text entered using the nearby keyboard.

[0][4](sp)iPhone: Decoding vibrations from nearby keyboards using mobile phone accelerometers

Visual-based Inference

- Evaluation of visual signals
- Example: Capturing movements using a camera

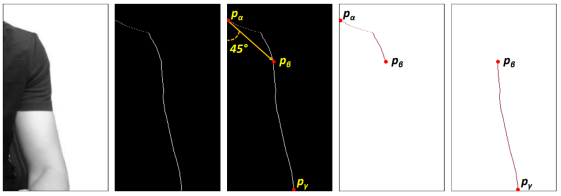
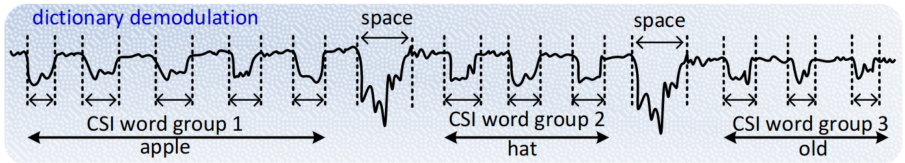


Fig. 4: *(a)* A keystroke frame segment, *(b)* Outer contour (OC), *(c)* 45° projection from p_α that intersects OC at p_β, *(d)* Shoulder contour (SC), and *(e)* Arm contour (AC).

[0][5]Zoom on the Keystrokes: Exploiting Video Calls for Keystroke Inference Attacks

Wi-Fi Channel-State-Information Based Inference

- Analysis of the channel state information influenced by keyboard inputs
- Example: Setting up Wi-Fi travelling through a keyboard

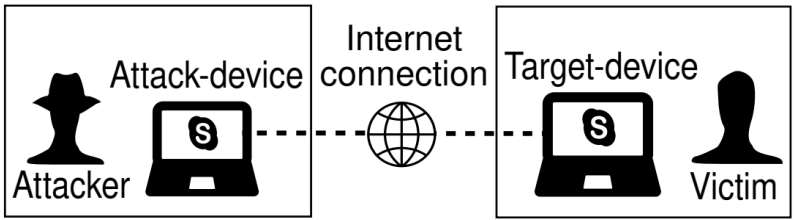


[0][7]Wireless Training-free Keystroke Inference Attack and Defense

Acoustic Attack Experiment

Audio Side-Channel

- Placed microphone
- Video conferences [1]



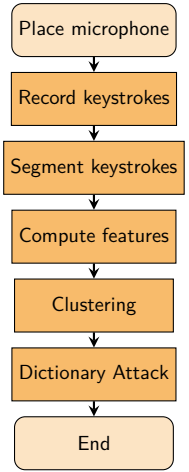
[0][1]Don't skype & type!

Audio Side-Channel

My Attack

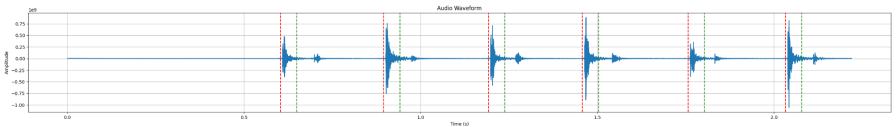
- Assumption: Microphone placed or laptop itself used as a microphone
- Supervised learning should not be used
  - Pre-trained models per user do not make practical sense
  - Unsupervised clustering for the identification of identical keys pressed
- Dictionary attack
  - Analysis of the length and character frequency and position in sample

Attack


Place microphone
Record keystrokes
Segment keystrokes
Compute features
Clustering
Dictionary Attack
End

Segmentation

Adaptive thresholding is used for segmentation. The average volume of the environment is calculated and sounds above a certain threshold are clipped out.



Clustering

Ostbayerische Technische Hochschule Amberg-Weiden

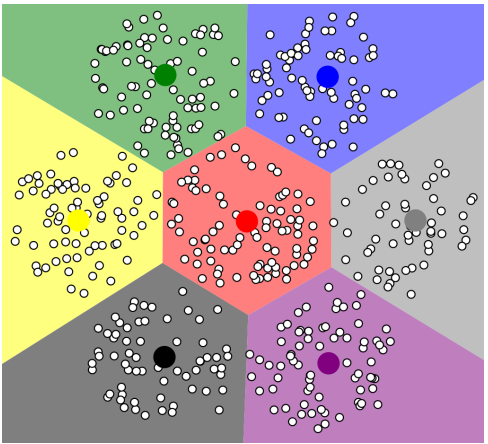
- Allows for automatic grouping of keystrokes based on features
- With the right choice of features, we can group keys by sound

Possible Features:

- MFCC (Mel Frequency Cepstral Coefficients)
- FFT
- Cross Correlation

Clustering Models

- K-Means (Hard Clustering)
  - Distance-based



Space partitioned into Voronoi cells. Created using [6].

Assumptions:

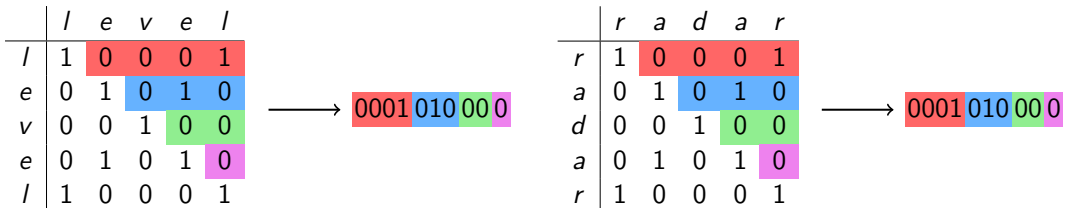
- Target language is know (English)
- Natural language text is used
- Only small letters and space is used

Features of the clusters:

- Word length
- Amount and positions of same characters in each word

Passphrase Recovery

To perform a dictionary attack, we need to find suitable word candidates. For this we need a feature-based encoding:



These same length words with the same features can be encoded like so: "0001010000". In this encoding we can tell the length of a given word, as well as the overlapping character places.

Ostbayerische Technische Hochschule Amberg-Weiden

Cluster: 1234_5166_789

acid (1234) ➔ 000000	jazz (5166) ➔ 000001	elk (789) ➔ 000
taco	doll	jar
myth	huff	elk
acid	jazz	chi

The candidates were chosen based on their matching relationship matrix. If we start concatenating words and computing the relationship matrix of the cluster concatenation and the dictionary words we can remove candidates.

Passphrase Recovery
Example

Text	Relationship Encoding
1234	000000
taco	000000
myth	000000
acid	000000

Passphrase Recovery

Example

Text	Relationship Encoding
12345166	00001000000000000000000000001
tacodoll	00000000000000000001000000001
mythdoll	00000000000000000000000000001
aciddoll	00000000000000000001000000001
tacohuff	00000000000000000000000000001
mythhuff	00000000000000000001000000001
acidhuff	00000000000000000000000000001
tacojazz	00000000001000000000000000001
mythjazz	00000000000000000000000000001
acidjazz	00001000000000000000000000001

Passphrase Recovery

Example

Text	Relationship Encoding
12345166789	0000100000000000000000000000000000000000000000001000000000
acidjazzjar	0000100010000000000000000000000000000010000010100000000
acidjazzelk	0000100000000000000000000000000000000000000001000000000
acidjazzchi	0000100000000000100000000010000000000000000001000000000

Passphrase Recovery

Example

acid (1234) ➔ 000000	jazz (5166) ➔ 000001	elk (789) ➔ 000
acid	jazz	elk

Passwords generated

Ostbayerische Technische Hochschule Amberg-Weiden

1) peroxide hacking arena
2) goldfish augmented yoyo
3) nugget iguana nylon
4) finalist caviar cufflink
5) ipad decal uptown
6) lukewarm pedometer litter wreckage
7) juggle gibberish hacking luxurious
8) unmarked vaseline aluminum jasmine
9) poison amendment sizable angelfish
10) taco ferret circle deliverer
11) velcro jelly duplex magazine silicon
12) hefty frosting acid zookeeper patio
13) daughter pyramid onyx pogo palm
14) cahoots arena cement statue mutation
15) blade banana awhile elsewhere tadpole
16) oxygen remote diffuser engine lettuce acid
17) oncoming feline glucose sushi abdomen judiciary
18) nullify scarf deepness modify euphemism grumbling
19) apple unnoticed bullfrog datebook vicinity glove
20) unhinge zodiac movie tadpole tapestry waffle
21) habitat gullible jingling mule envoy device erratic
22) licorice breath thumb navigate saddlebag yahoo voucher
23) festival yearbook fountain underwear nastiness dedicate licorice
24) scooter urchin albatross sneezing itunes gumdrop cubical
25) bagpipe earlobe aerosol aliens ivory clubhouse pantyhose
26) couch crawfish mundane goggles rupture florist rancidity degree
27) hefty tree riverboat sculpture junkyard awhile isotope unveiled
28) sled dyslexia jelly clergyman fruit family blade rancidity
29) payphone rupture awoke virus tuesday upbeat knapsack amnesty
30) afloat ardently fox emission exquisite dagger jersey lubricant

Participants Samples

Participant	Passphrases
1	30
2	30
3	16
4	30
5	19
6	27
7	22
8	30
9	19

Samples per participant

Participant	Keyboard Model	Microphone Model	Mechanical
1	Tecurs	IdeaPad 5 Pro 14ACN6	✓
2	Laptop	Laptop Webcam	✗
3	Apple Magic (2014)	iMac 2014	✗
4	HIGROUND Base 65	Auna CM 900B	✓
5	Keychron K8 Pro	MacBook Air M1	✓
6	Redragon	Macbook Pro 14	✓
7	Cherry	Laptop	✓
8	Corsair K55 Gaming	Lenovo ThinkPad T14s	✗
9	Cherry	DELL Notebook	✓

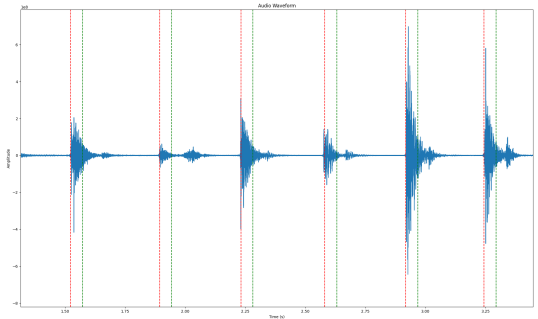
Hardware used by participants.

Hyperparameter Search

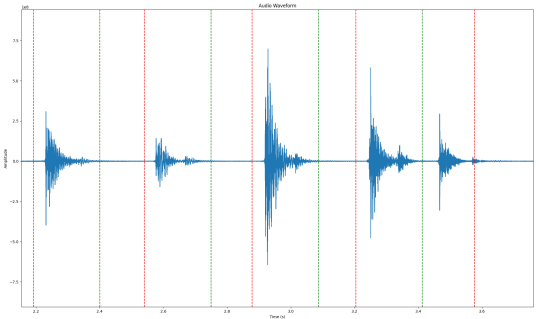
Hyperparameter	K-Means	Cross-Correlation
Feature	FFT, MFCC, FFT+MFCC	Raw Audio, FFT, MFCC
Smoothing	true, false	
Smoothing Window	5 to 300	
Scaling	true, false	
PCA	true, false	
PCA Components	1 to 20	1 to 12
Keystroke Span	P, PR	

Hyperparameters for K-Means and Cross-Correlation-based Models.

Keystroke Segmentation



a) Keystroke Press Segmentation.



b) Keystroke Press-Release Segmentation.

Segmentation Comparison.

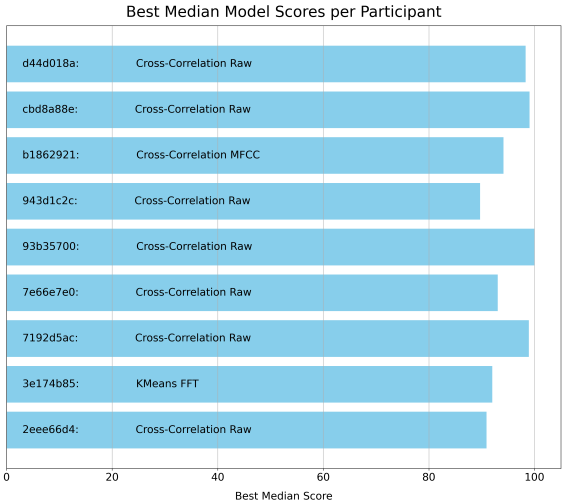
Hyperparameter Search

Hyperparameter	K-Means	Cross-Correlation
Feature	MFCC	Raw
MFCC Components	180	
PCA	True	False
PCA Components	1	
Smoothing	False	False
Scaling	True	False
Keystroke span	PR	P
Median Score	90.27	93.12
Mean Score	88.95	93.21
Max Score	91.77	98.91
Min Score	83.07	85.44

Best Model Scores and Their Hyperparameters.

Best Model per Participant

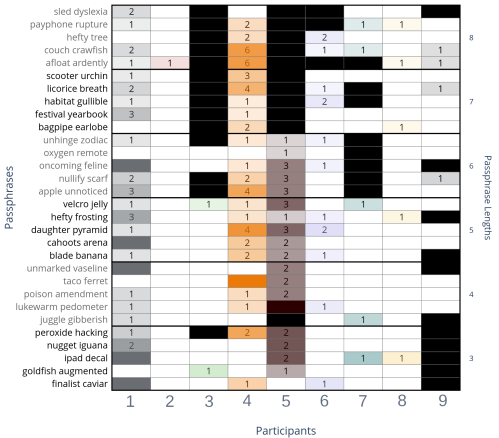
- Using specific hyperparameters per target can improve recovery results
- After initial recovery of a passhprase or message the recovery can be used to conduct hyperparameter tuning



Best Median Model per Participant.

Passphrase Recovery Results

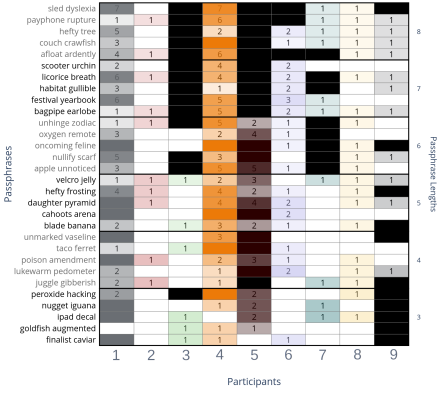
Not exactly great...



Recovery results using one set of clusters from the best model.

Passphrase Recovery Results

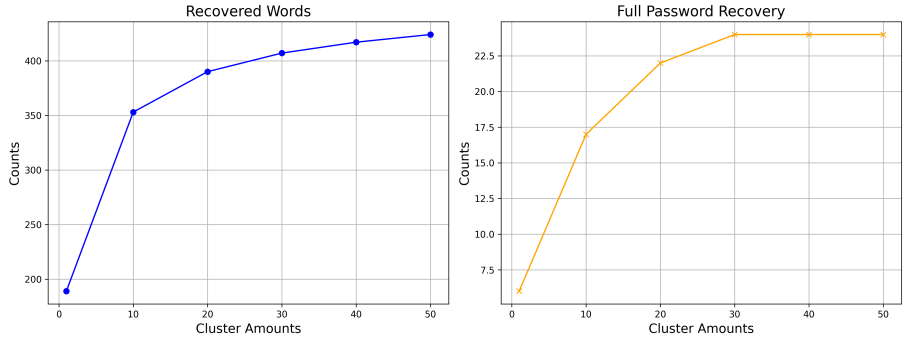
The Machine Learning Thing to Do



Recovery results using ten clusters.

Passphrase Recovery Results

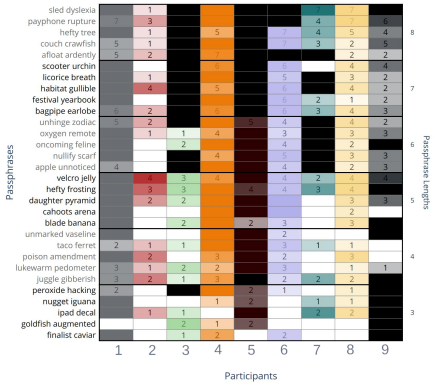
So let's just do more clusters?



Recovery as a function of clusters.

Passphrase Recovery Results

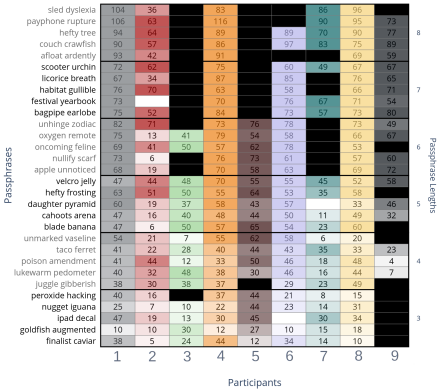
Alright. Then let's just Brute-Force from here



Recoveries brute-forcing combinations of partial recoveries from ten clusters.

Passphrase Recovery Results

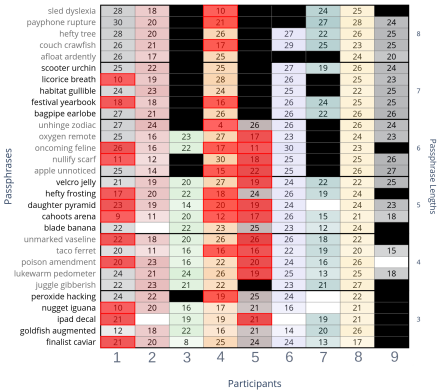
Why is my cloud bill so high?!



Amount of combinations of demodulated words from ten cluster results. The table shows the exponents to the base of 2.

Passphrase Recovery Results

AI final



Brute-Force attempts needed, when starting with most likely candidates from ten cluster results. The table shows the exponents to the base of 2. Red marked cells are full recoveries.

Email Recovery

Example

Message

the vision is coming into focus and it has a new name over the past several months the it help desks at enron net works have incorporated new processes and methodologies to optimise customer service delivery as part of this effort we have implemented a new name resolution center the following it help desks are included in this

Recovered Message

the vision is coming into focus and it has a new name over the past several months the it help des_s at enron net wor_s have incorporated new processes and methodologies to optimi_e customer service delivery as part of this effort we have implemented a new name resolution center the following it help des_s are included in this

Concatenated Words
Skipped Words

Literature

Compagno, A., Conti, M., Lain, D., and Tsudik, G.

Don't skype and type! acoustic eavesdropping in voice-over-ip, 2017.

Liu, J., Wang, Y., Kar, G., Chen, Y., Yang, J., and Gruteser, M.

Snooping Keystrokes with mm-level Audio Ranging on a Single Phone.
In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking* (Paris France, Sept. 2015), ACM, pp. 142–154.

Liu, X., Zhou, Z., Diao, W., Li, Z., and Zhang, K.

When good becomes evil: Keystroke inference with smartwatch.
In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security* (New York, NY, USA, 2015), CCS '15, Association for Computing Machinery, p. 1273–1285.

Marquardt, P., Verma, A., Carter, H., and Traynor, P.

*sp*iphone: Decoding vibrations from nearby keyboards using mobile phone accelerometers.
pp. 551–562.

Sabra, M., Maiti, A., and Jadliwala, M.

Zoom on the keystrokes: Exploiting video calls for keystroke inference attacks, 2020.

Harris, N.

Visualizing K-Means Clustering.
`https://www.naftaliharris.com/blog/visualizing-k-means-clustering/`, Jan. 2014. Accessed: Apr. 11, 2025.

Yang, E., Fang, S., Markwood, I., Liu, Y., Zhao, S., Lu, Z., and Zhu, H.

Wireless training-free keystroke inference attack and defense.
IEEE/ACM Transactions on Networking 30, 4 (2022), 1733–1748.