Don't Skype & Type!

Acoustic Eavesdropping in Voice-Over-IP

Alberto Compagno, Mauro Conti, **Daniele Lain**, Gene Tsudik

Sapienza Univ. of Rome *University of Padua* *UC Irvine*

Abu Dhabi, 2017-04-06

ACM Asia Conference on Computer and Communications Security (ASIACCS) 2017

UNIVERSITÀ DEGLI STUDI DI PADOVA

SPRITZ SECURITY & PRIVACY RESEARCH GROUP

Outline

SPRITZ
Security & Privacy
Research Group

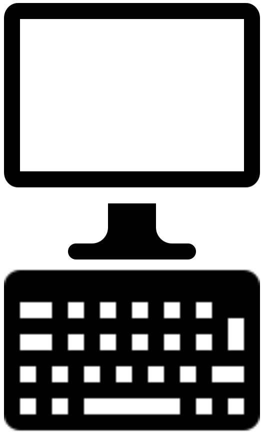
Università
degli Studi
di Padova

- Side Channels
  - *Eavesdropping physical emanations*
  - *Keyboard acoustic eavesdropping*

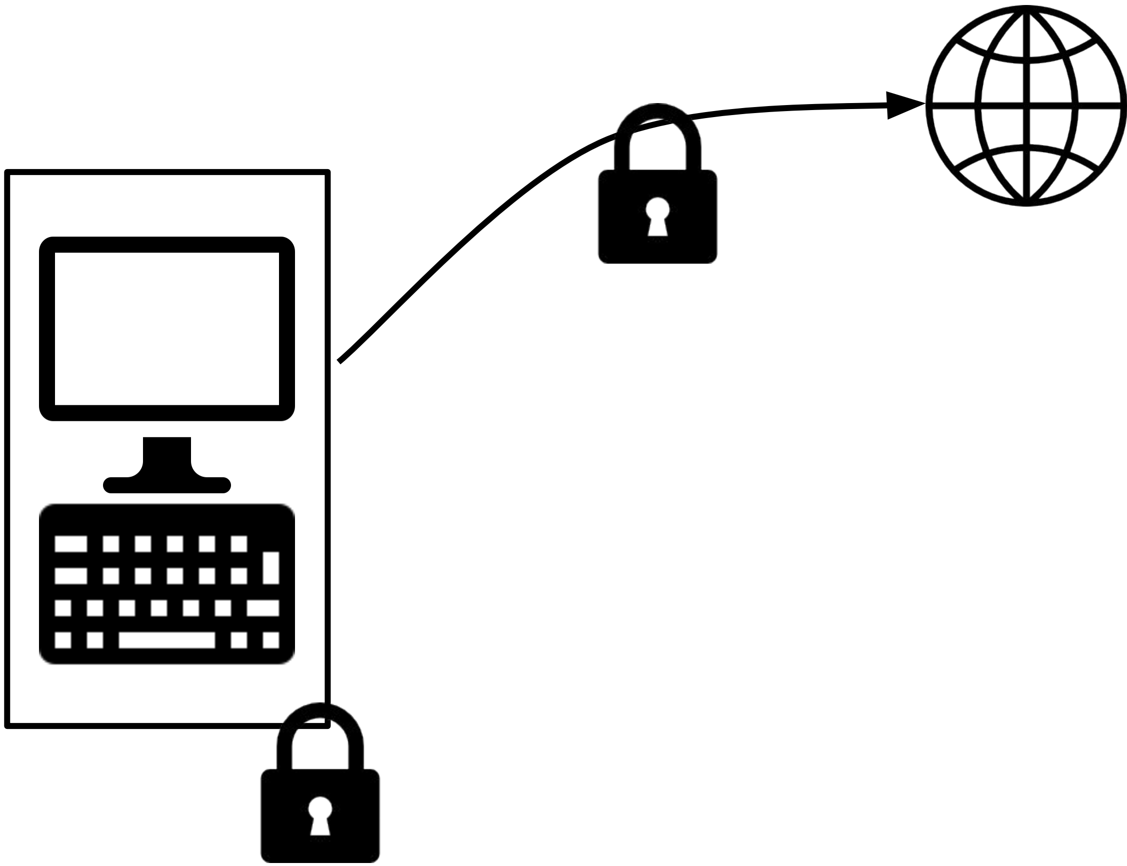
- Skype&Type Attack
  - *Design and setup*
  - *Evaluation*

- Conclusions and Future Work

Physical Emanations

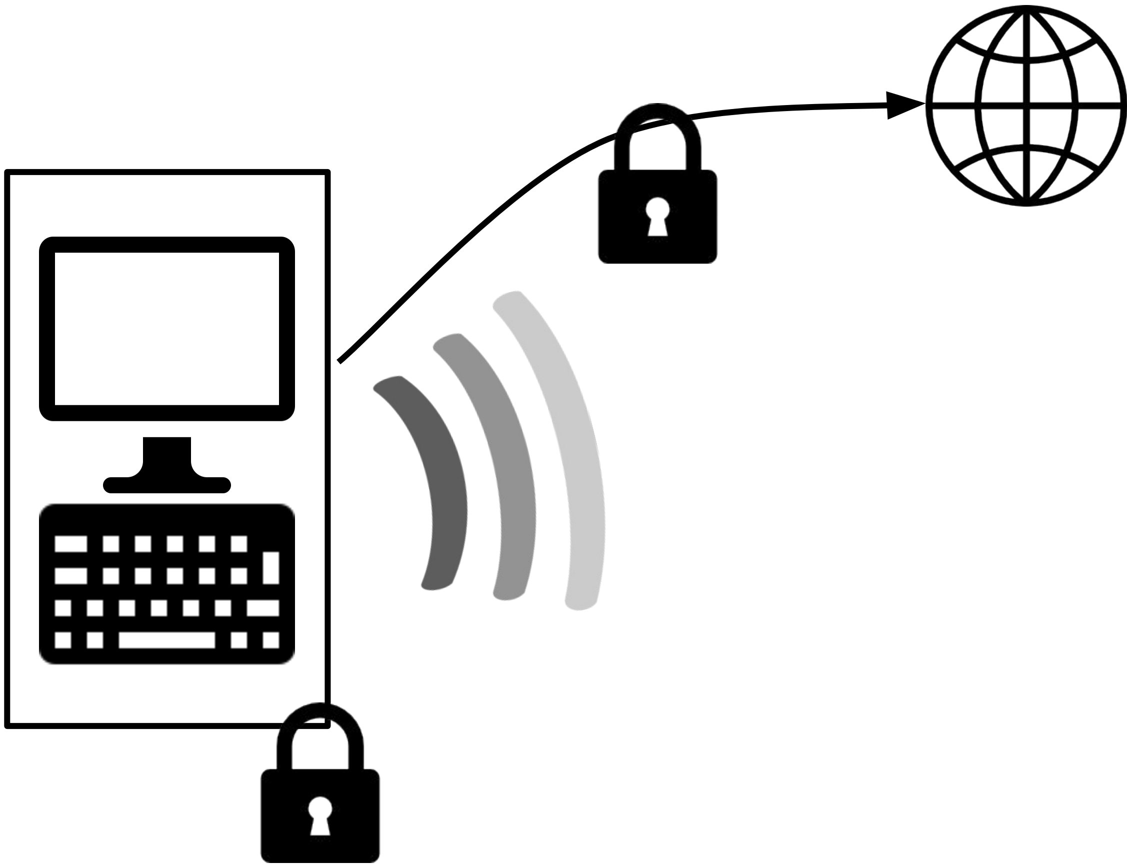


Physical Emanations



Physical Emanations

- Electromagnetic
  Data transmission
- Visual
  Videos, reflections
- Acoustic
  Hardware sounds
- Tactile
  Motion sensors



Physical Emanations Eavesdropping

Physical Emanations Eavesdropping

Eavesdropping

A convenient means to steal sensitive data

- Transmitting medium

  Network cables, wireless emanations, peripherals buses

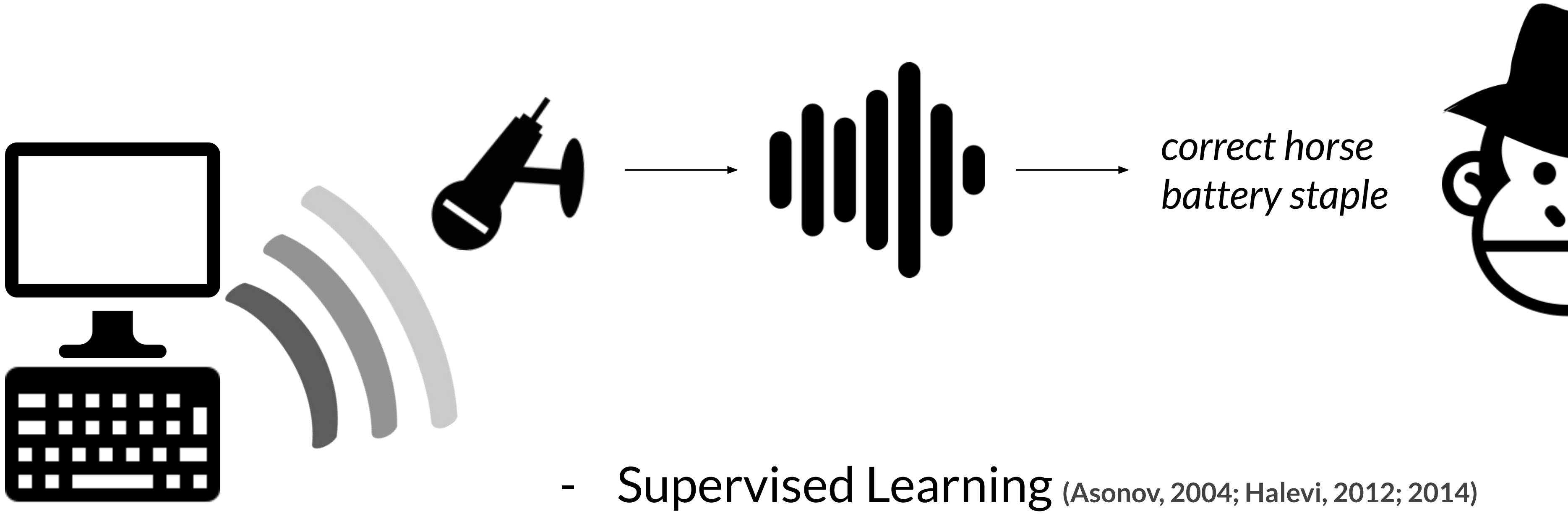
- Input/output devices

  Keyboards, touchscreens, monitors, printers

- Processing medium

  CPUs, hard drives, RAM

Keyboard Acoustic Eavesdropping

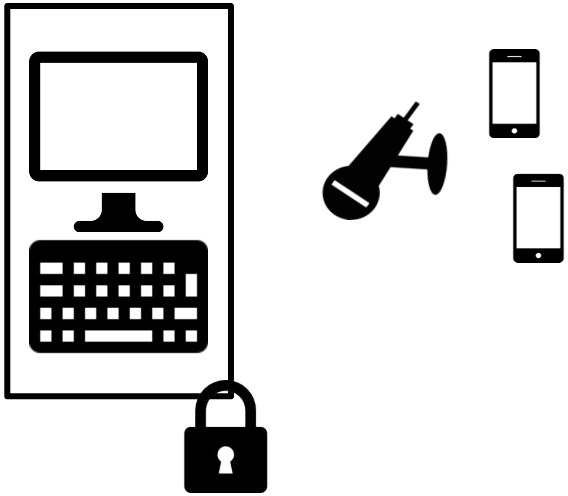


- **Supervised Learning** (Asonov, 2004; Halevi, 2012; 2014)
  Less input assumptions, more specific
- **Unsupervised Learning** (Berger, 2006; Zhuang, 2009)
  More input assumptions, more general

Problems

1) Precise training data - *how?*
   vs.
   Generic training data - *a lot, or in a known language (no passwords)*

2) Need ***physical proximity*** → unrealistic scenario
   To place microphones / mobile phones



Outline

- Side Channels
  - *Eavesdropping physical emanations*
  - *Keyboard acoustic eavesdropping*
- **Skype & Type Attack**
  - ***Design and setup***
  - *Evaluation*
- Conclusions and Future Work

Motivation

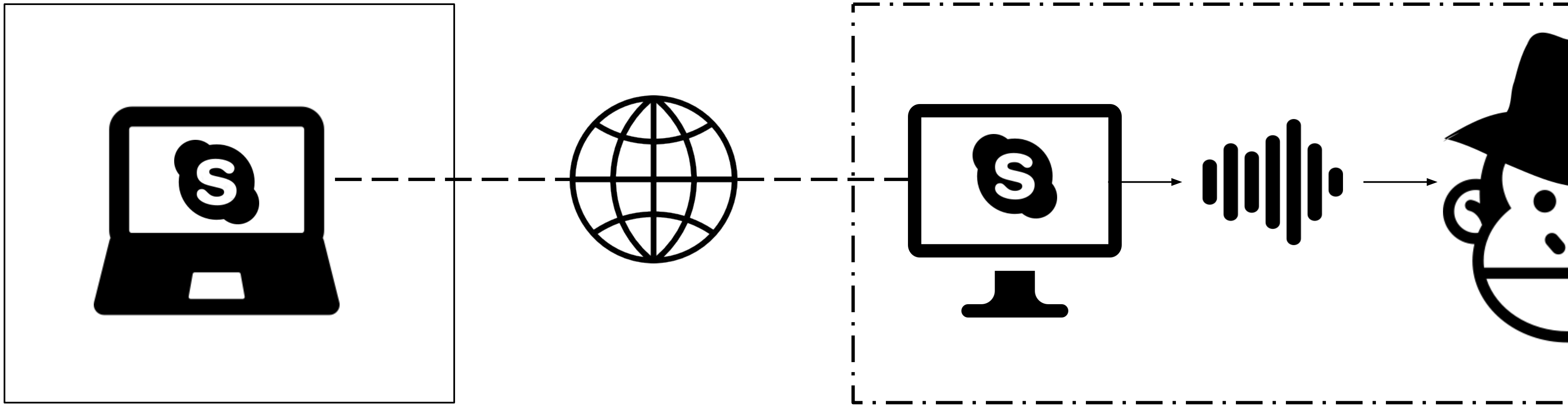
VoIP → one of the most used software: in academia, industry, at home

People type private stuff during Skype calls - it happens!

- *Login to websites*
- *Write a sensitive email*
- *Take notes*

We hear the keys' noise and use it to understand typed text

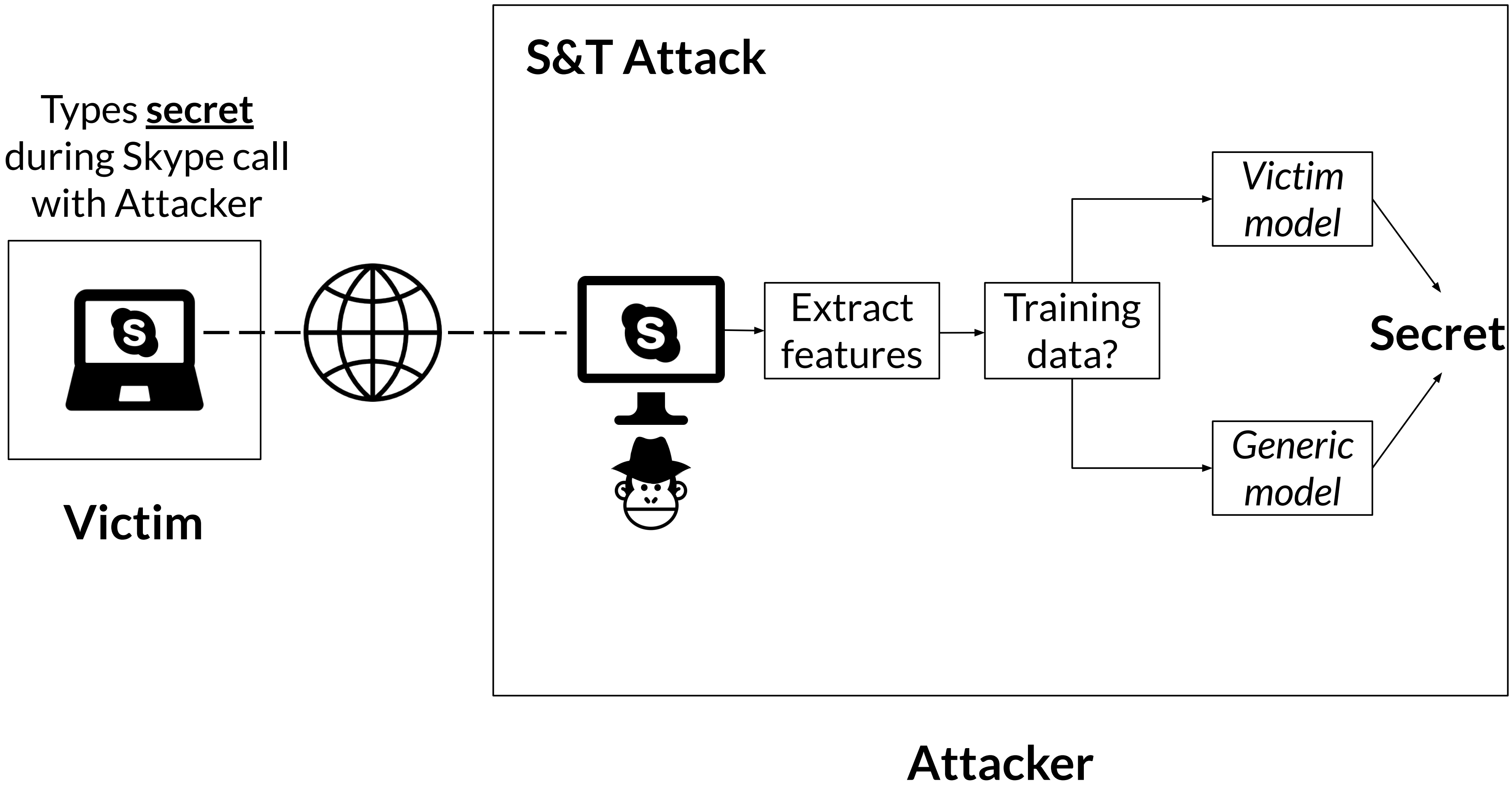
- *Victim is willingly giving us access to his microphone*



Skype&Type Attack

SPRITZ Security & Privacy Research Group

Università degli Studi di Padova



S&T - Tools

- **Data windowing and segmentation**

  To extract sound samples

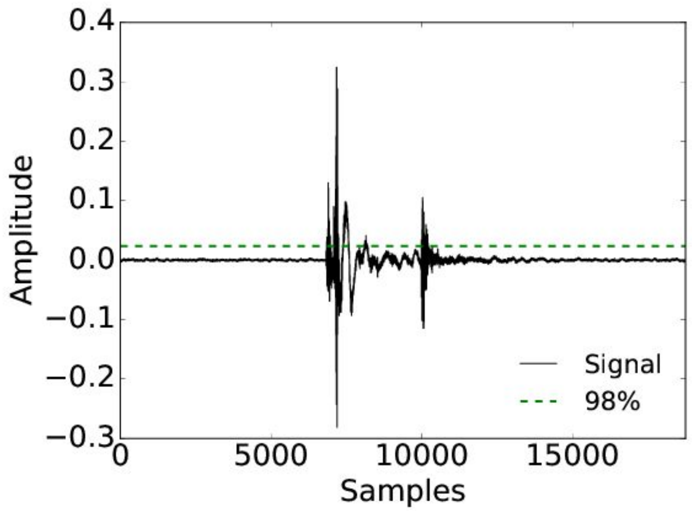
- **Mel frequency cepstral coefficients**

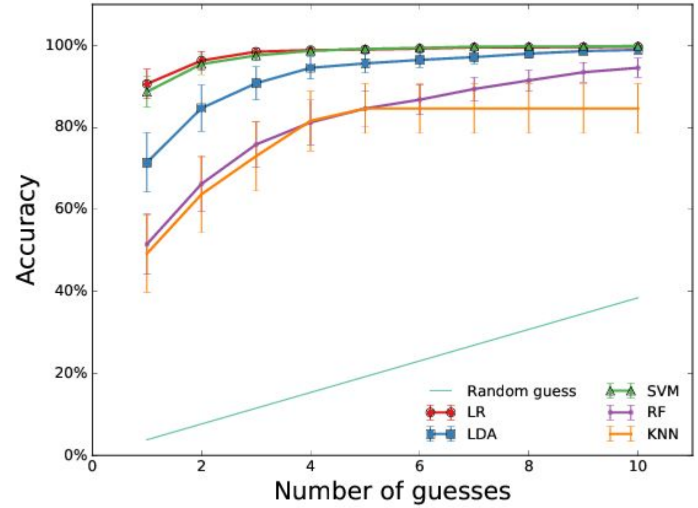
  Best performing and robust

- **Supervised learning paradigm**

  Target text can be possibly:
  - *Short (no clustering)*
  - *Random (no dictionary)*

- **Logistic Regression classifier**





Outline

- Side Channels
  - *Eavesdropping physical emanations*
  - *Keyboard acoustic eavesdropping*
- **Skype & Type Attack**
  - *Design and setup*
  - ***Evaluation***
- Conclusions and Future Work

Data Collection

- Try S&T in many scenarios
  - *With **5** different users over **Skype*** *(Google Hangouts also vulnerable)*
  - *Using **3** different common laptops: Macbook Pro, Lenovo, Toshiba*
  - *With **2** typing styles: single finger, and natural "touch" typing*

- Evaluate top-n accuracy of character recognition

  as a function of the number of guesses, focus on top-1 and top-5 accuracy

- Against a "dumb" random guess

  Might be a random password -- we can not use "smarter" approaches

Attack Scenarios

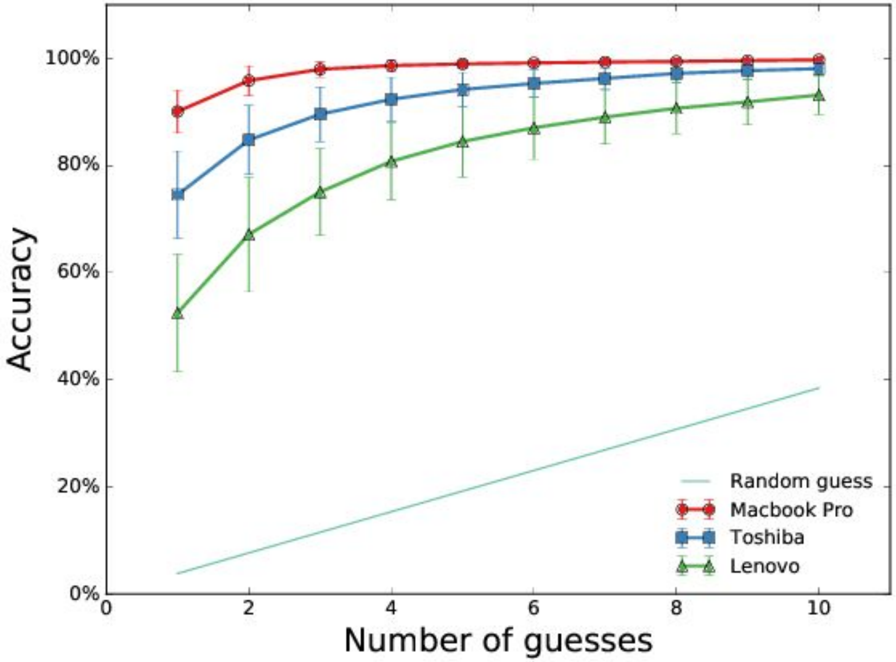
Evaluate the attack on two realistic scenarios

- Complete Profiling Scenario (Asonov, 2004; Halevi, 2012; 2014)
  - *Profiled the user on his laptop → specific training set*
  - *Ground truth disclosure, e.g., a short chat message*

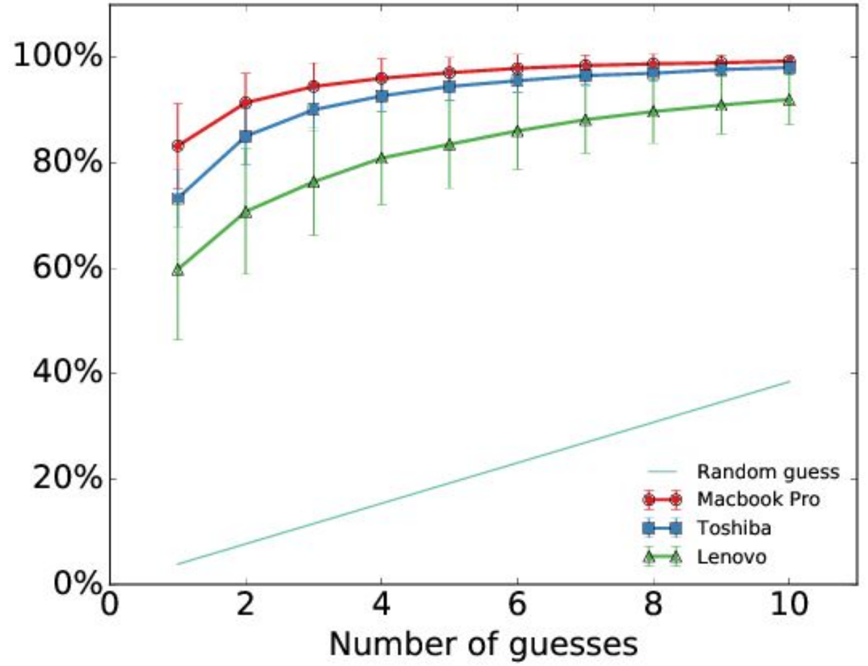
- Model Profiling Scenario
  - *Profiled a laptop of the same model on some users*
  - ***Victim is/can be unknown!***

Complete Profiling

Training set with the data the user disclosed



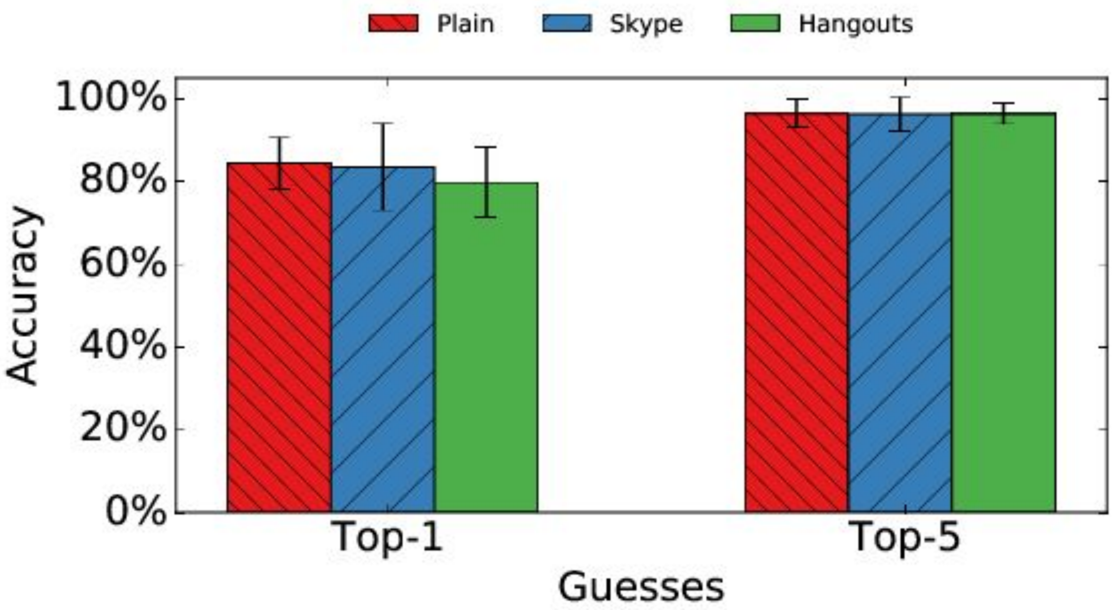
Hunt&Peck typing, unfiltered data



Touch typing, Skype filtered data

Complete Profiling

Is only Skype vulnerable to our attack?



No! It looks like a common problem for VoIP software

Model Profiling Setup

On the *Model Profiling* Scenario, the victim can be unknown

Someone the attacker does not know personally

First need to understand the laptop of the victim

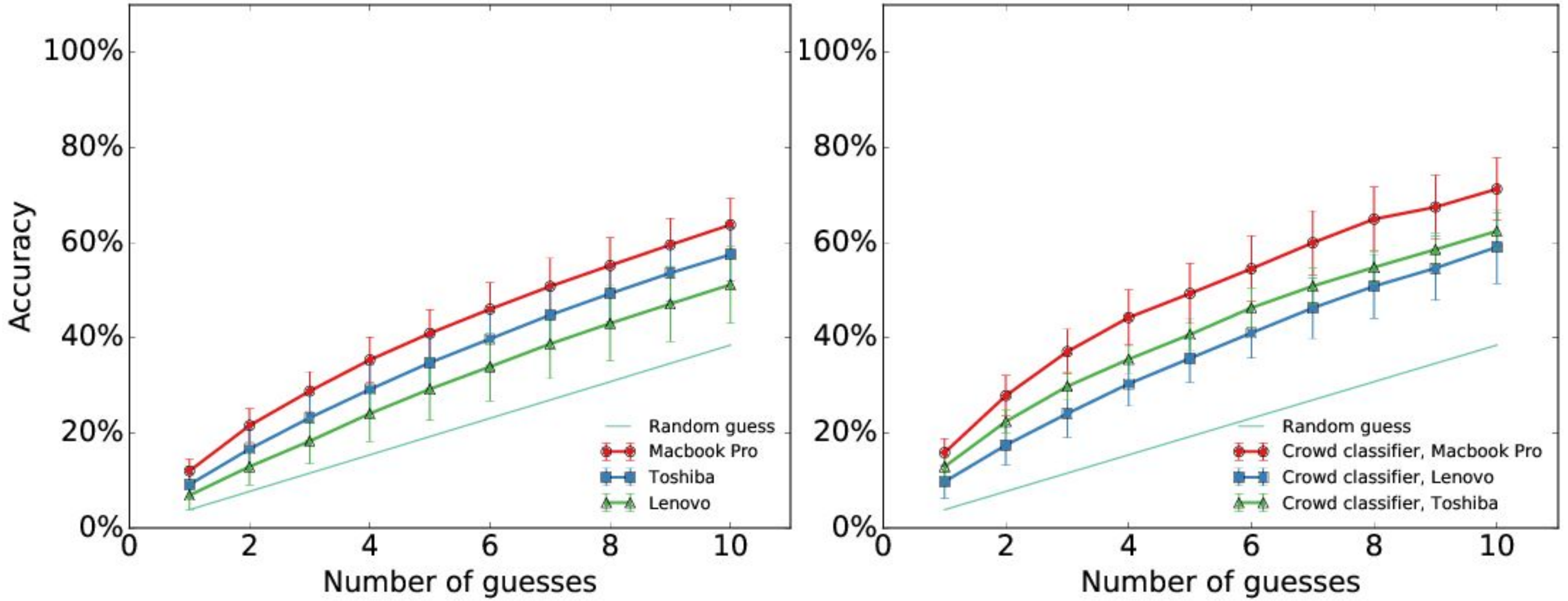
→ match it with a database of model signatures

- Guess correctly **93%** of the times if the model is known
- Statistical measures if the model is unknown

Model Profiling

One user

“Crowd” of multiple users



Summing Up Our Results

- Recognize a single character
  - *Complete Profiling: 90%+ accuracy*
  - *Model Profiling: 40%+ accuracy*

- Recognize a single word
  - *Complete Profiling: 98% correct letters*
  - *Model Profiling: 50% correct letters*

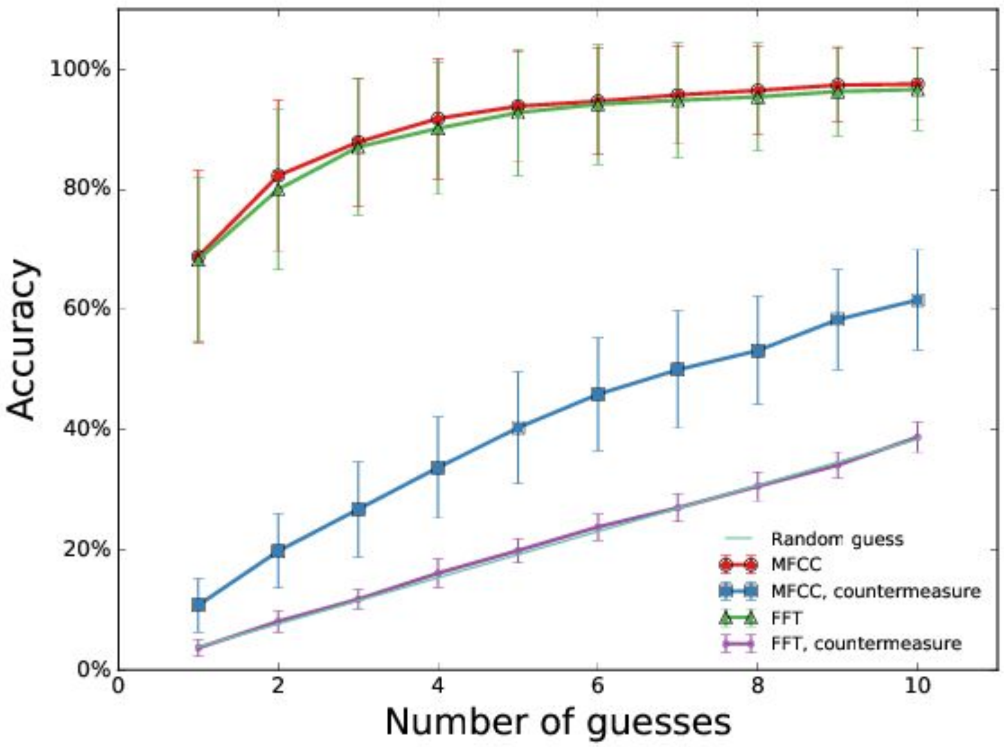
- Recognize a random password
  - *Improves 1-5 orders of magnitude time needed to guess the password*
  - *From 50 days to 42 seconds on a domestic PC*

Countermeasures

- <u>Don't Skype & Type</u>

- Remove volume when we detect a keypress sound
  - *Impacts voice, greatly degrades call quality*

- Disrupt spectral features with random equalization
  - *Assess impact on voice, real time feasibility*



Conclusions & Future Work

- VoIP Keyboard acoustic eavesdropping a serious threat

- Feasible and accurate:
  - *Realistic attack scenarios*
  - *91.71% on **Complete Profiling** scenario*
    - *Halevi (2012; 2014): 85.78%*
  - *41.89% on **Model Profiling** scenario*
    - *Novel attack vs. unknown victims*
  - *Robust to degradation and to voice*

- Future work:
  - *Try more users and different keyboards, and on more VoIP software*
  - *Try to attack another user in the same room*
  - *Analyze and improve the countermeasures*

References

Asonov, D., & Agrawal, R. (2004). Keyboard acoustic emanations. In *IEEE Symposium on Security and Privacy*.

Halevi, T., & Saxena, N. (2012). A closer look at keyboard acoustic emanations: random passwords, typing styles and decoding techniques. In *Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security*.

Halevi, T., & Saxena, N. (2014). Keyboard acoustic side channel attacks: exploring realistic and security-sensitive scenarios. *International Journal of Information Security*, Springer.

Berger, Y., Wool, A., & Yeredor, A. (2006). Dictionary attacks using keyboard acoustic emanations. In *Proceedings of the 13th ACM conference on Computer and communications security*.

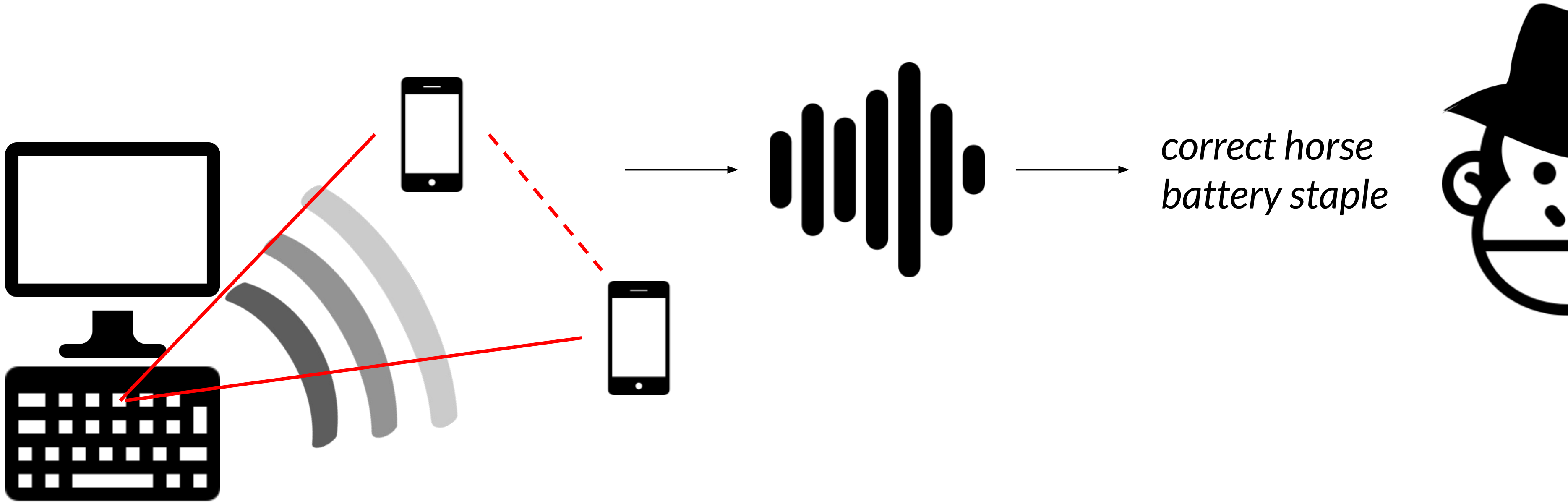
Zhuang, L., Zhou, F., & Tygar, J. D. (2009). Keyboard acoustic emanations revisited. *ACM Transactions on Information and System Security (TISSEC)*.

Liu, J., Wang, Y., Kar, G., Chen, Y., Yang, J., & Gruteser, M. (2015). Snooping keystrokes with mm-level audio ranging on a single phone. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*.

Zhu, T., Ma, Q., Zhang, S., & Liu, Y. (2014). Context-free attacks using keyboard acoustic emanations. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*.

Backup Slides

Keyboard Acoustic Eavesdropping



- Timing information (Liu, 2015; Zhu, 2014)
  Context-free, difficult to setup

Password Cracking

The goal was to crack the victim's random password

→ We need bruteforce techniques

Random password of 10 lowercase letters

- $\log_2(26^{10})$ = 47 bits of entropy

On the Complete Profiling Scenario (high accuracy)

- $\log_2(5^{10})$ = 23.22 bits of entropy

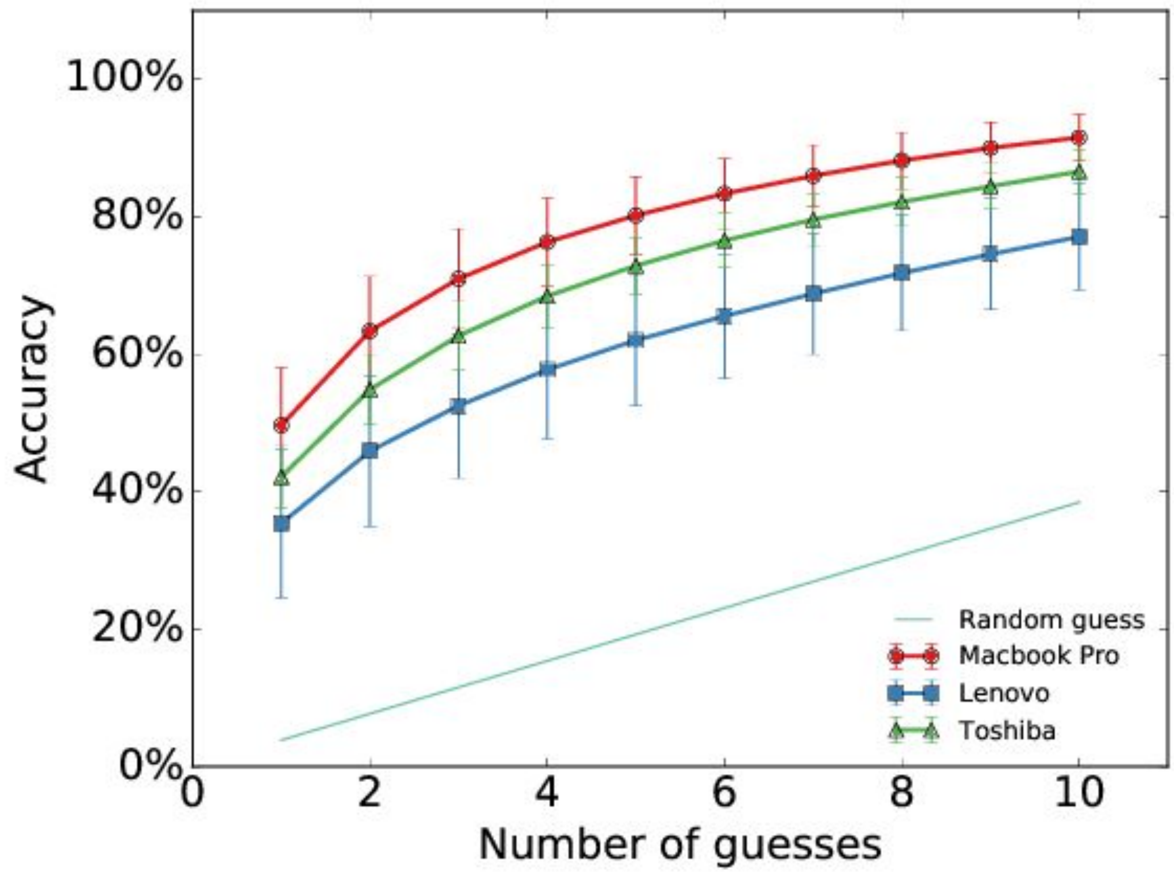
On the other scenarios - entropy is not meaningful

Evaluation - Small Training Set

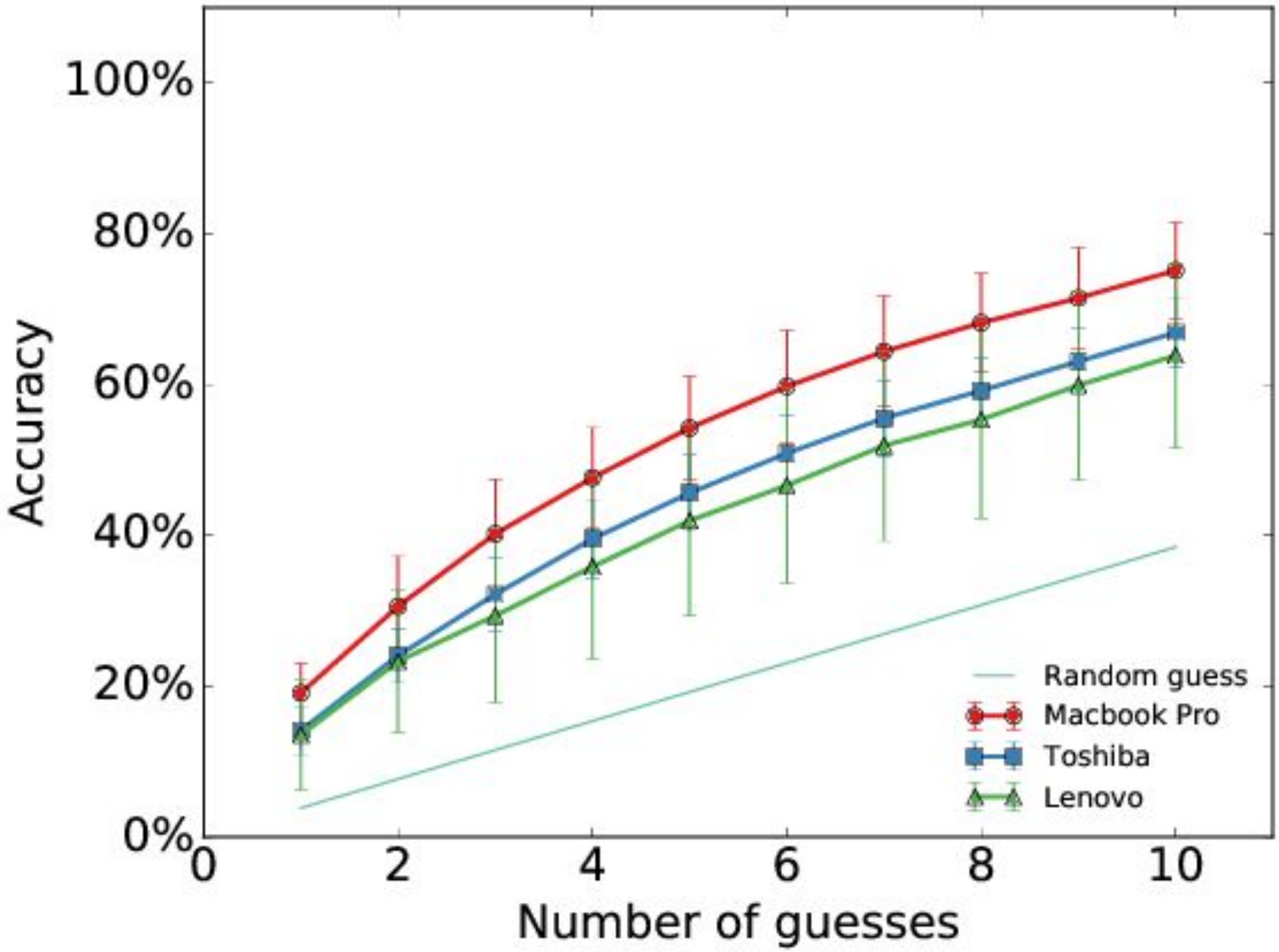
10 samples/character aren't your typical chat message

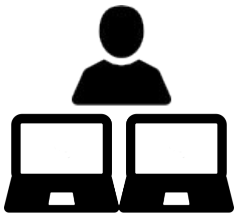
Training set *with realistic letter frequencies*
Test *against random password*

Character	*# Samples*
E	10
A	9
R	7
J	1
Z	1



Evaluation - User Profiling

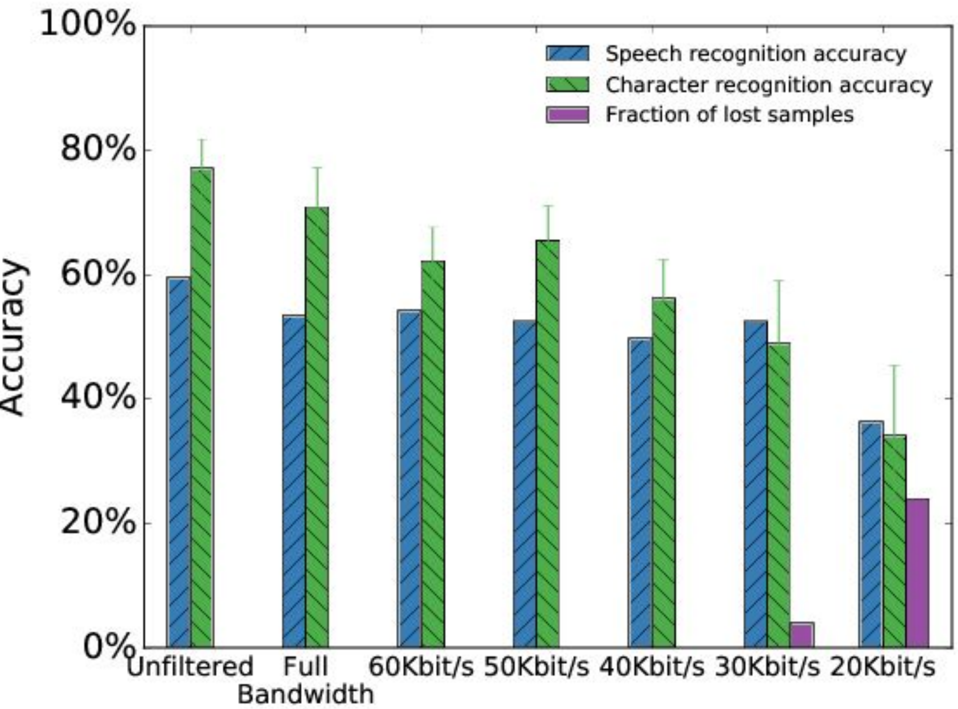




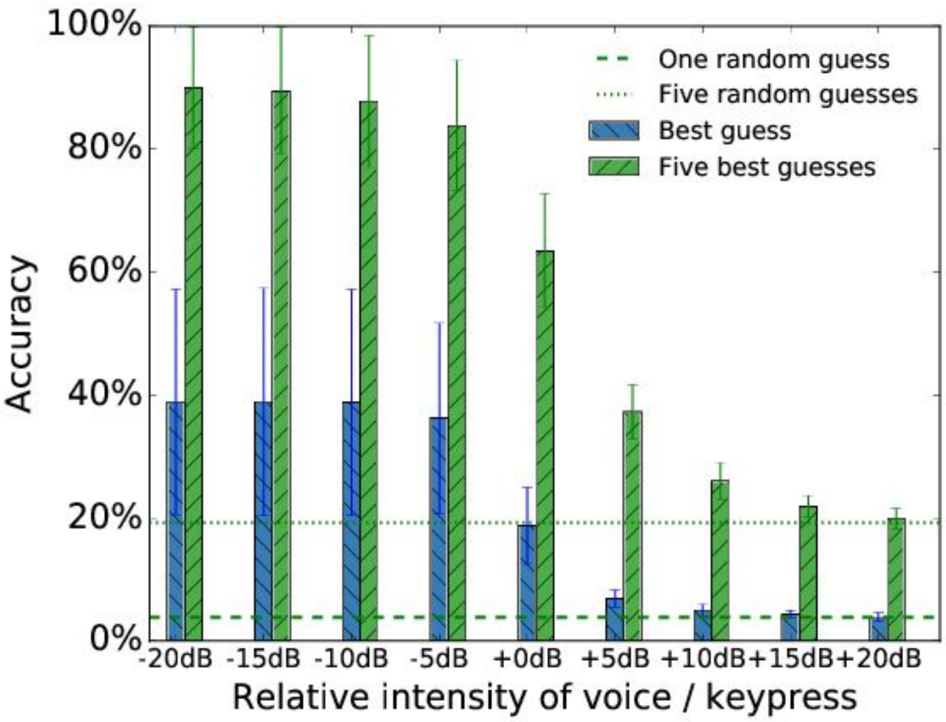
VoIP challenges

Using Skype poses additional challenges:

Slow Internet degrades call quality



Human voice on top of keypresses



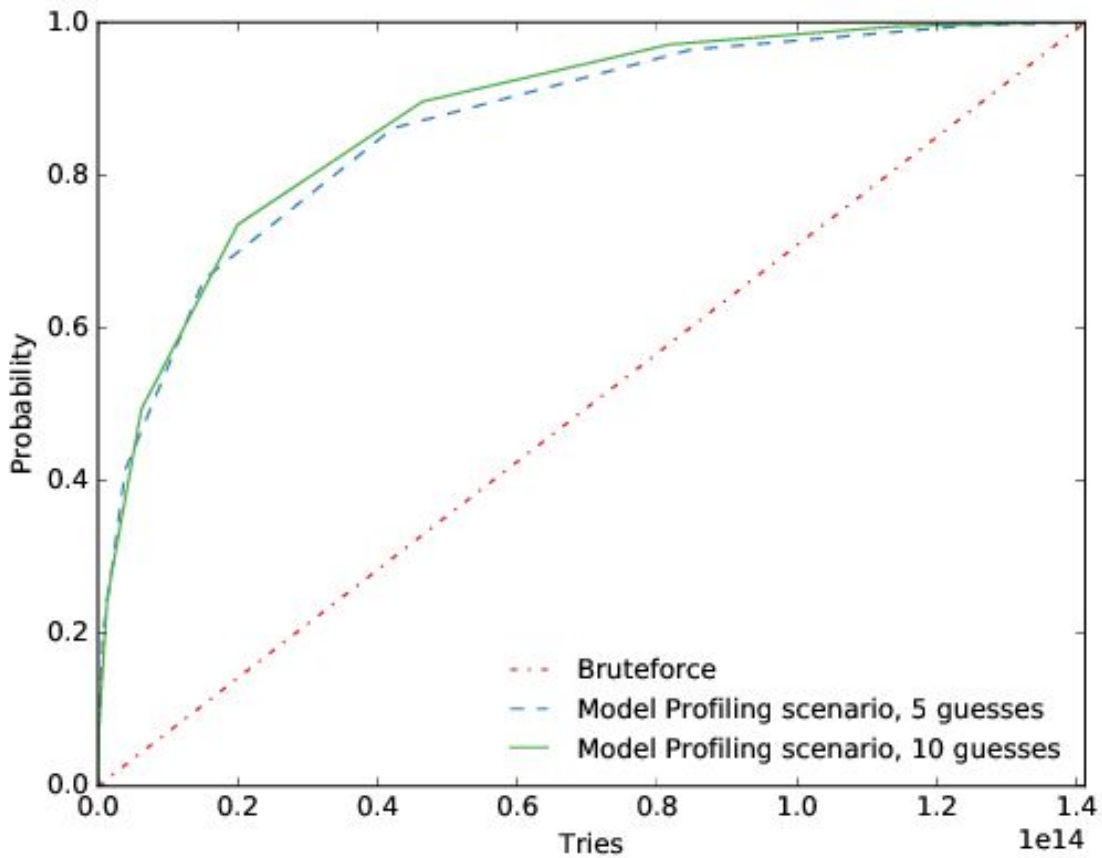
Password Cracking

SPRITZ Security & Privacy Research Group
Università degli Studi di Padova

Model Profiling Scenario → improved bruteforce

Take into account character probabilities

Evaluate the reduction of the average number of trials



Features

Fast Fourier Transform coefficients

$$S(f(t)) = 20\log_{10}\left(|\mathcal{F}(f(t))|\right)$$

$f(t)$ = signal
$\mathcal{F}$ = Discrete Fourier Transform function

Cepstrum coefficients

$$C(f(t)) = \left|\mathcal{F}^{-1}(S(f(t)))\right|^2$$

Mel frequency cepstral coefficients

$$MFC(f(t)) = DCT\left(\log_{10}\left(mel\{|\mathcal{F}(f(t))|\}\right)\right)$$

$$mel(f) = 2595\log_{10}\left(1 + \frac{f}{700}\right)$$

DCT = Discrete Cosine Transform