PinDr0p: Single-ended Audio Features To Determine the Call Provenance

Balasubramaniyan et al. CCS'10

Wajih Ul Hassan

11-17-2106

UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

Big Picture

Given the **audio** of a phone call, it's possible to determine, using **audio analysis**, where the call is actually **originating** from

Why this issue exists?

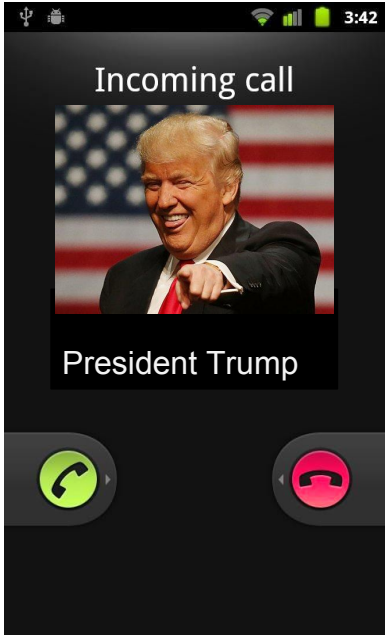
- Caller-ID information being transmitted over networks without verification
- Attackers can manipulate this data and make it appear like an incoming call is coming from a different source
- Services like caller-id spoofing are widely available

Caller ID Spoofing

- A caller deliberately falsifies the information transmitted to your caller ID display to ***disguise*** their identity

Caller ID Spoofing

- A caller deliberately falsifies the information transmitted to your caller ID display to ***disguise*** their identity
- Used By
  - Fraudsters
  - Scammers
- Legitimate Use



Who Cares About it?

- Banks
- E-tailers
- Call centers

Using Caller-id Spoofing to Craft Call Center Attacks

- Call centers have moved on to stronger authentication
  - Knowledge-based authentication
- Social engineering or weak KBA leads to password resets via the phone channel
- New password is used to attack the web channel
  - Funds transfer from online accounts

What we need

*A **fool-proof** way to determine the **origin** of a call could be the way to provide a much-needed layer of security on the phone channel, where the **caller ID system**, which was never designed with security in mind, is completely **broken**.*

Solution: *Call Provenance*

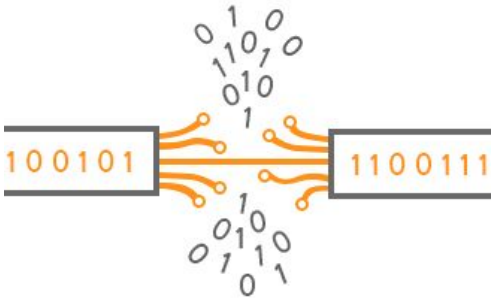
The provenance of a call describes the characteristics(features) of the source and traversed networks. (***Phoneprinting***)

Help distinguish and compare different calls in the absence of verifiable end to end metadata

PinDr0p is an infrastructure to help determine provenance of a call

Call Features (Artifacts)

An example of **feature**(artifact) is when a call is on VoIP network it experiences packet loss and packet loss results in tiny breaks in the call audio

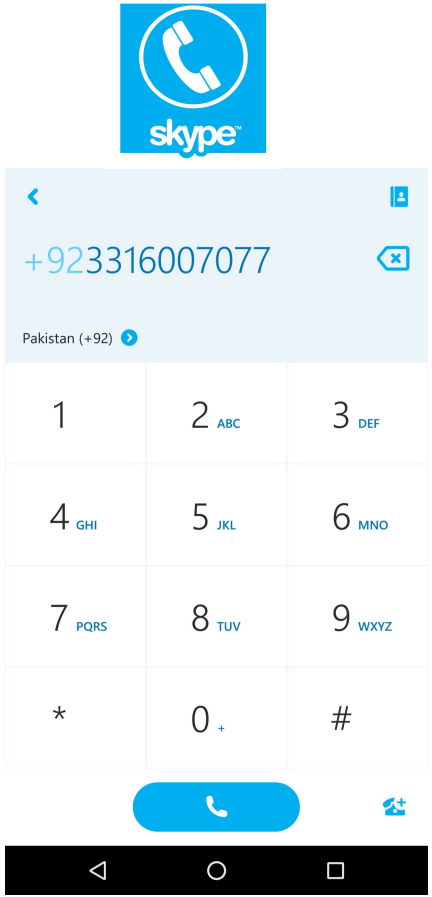


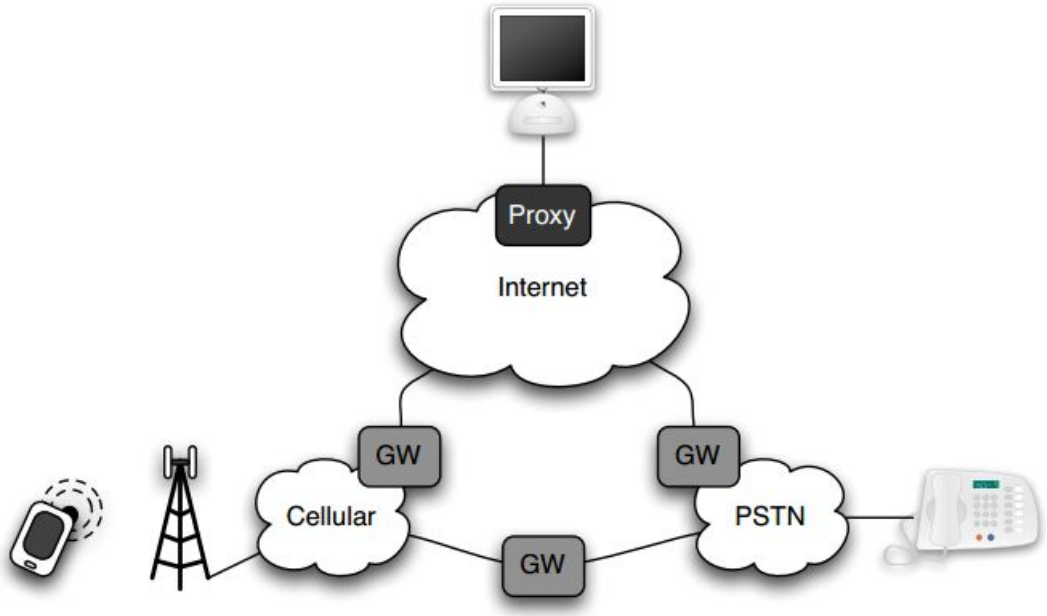
BACKGROUND
Existing Telephony Infrastructures

- Public Switched Telephone Networks (PSTN)
  - Traditional circuit switched
  - Lossless connections with high fidelity audio
- Codecs Used: For encoding and Decoding audio
  - G.711 (capture speech without any compression and require much higher bandwidth (64 kbps) than most other codecs)

- Cell Phones
  - Circuit switched core with some portions replaced by IP links
- Codecs Used:
  - GSM FR

- Voice over IP (VoIP)
  - Run on top of IP links and share Internet-based traffic paths
  - Almost always experience packet loss
- Codecs Used:
  - iLBC
  - Speex
  - G.729




Proxy
Internet
GW
GW
Cellular
PSTN
GW

How Phoneprinting works

- Use the different networks the call traverse through to identify call provenance

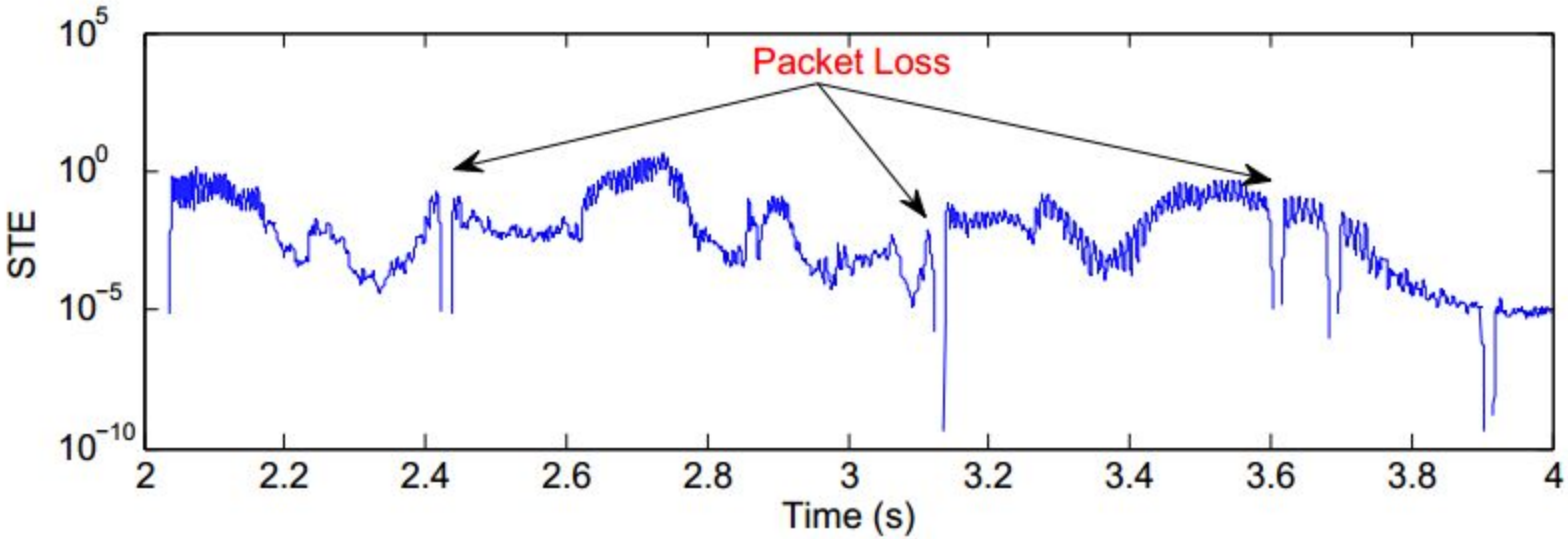
- Packet loss, bit errors and noise are hard for an adversary to control
  - an adversary bounded by a lossy connection, many miles away, cannot spoof a lossless, dedicated PSTN line to a bank

How it works

VOIP Network Detection

- Use Packet Losses
- Relate Packet loss to short time energy drop
- Amount of energy drop related to codec used

Therefore, when a call traverses a potentially lossy VoIP network, the packet loss rate and the codec used in that network can be extracted from the received audio.


Packet Loss
STE
10^{5}
10^{0}
10^{-5}
10^{-10}
2
2.2
2.4
2.6
2.8
3
3.2
3.4
3.6
3.8
4
Time (s)

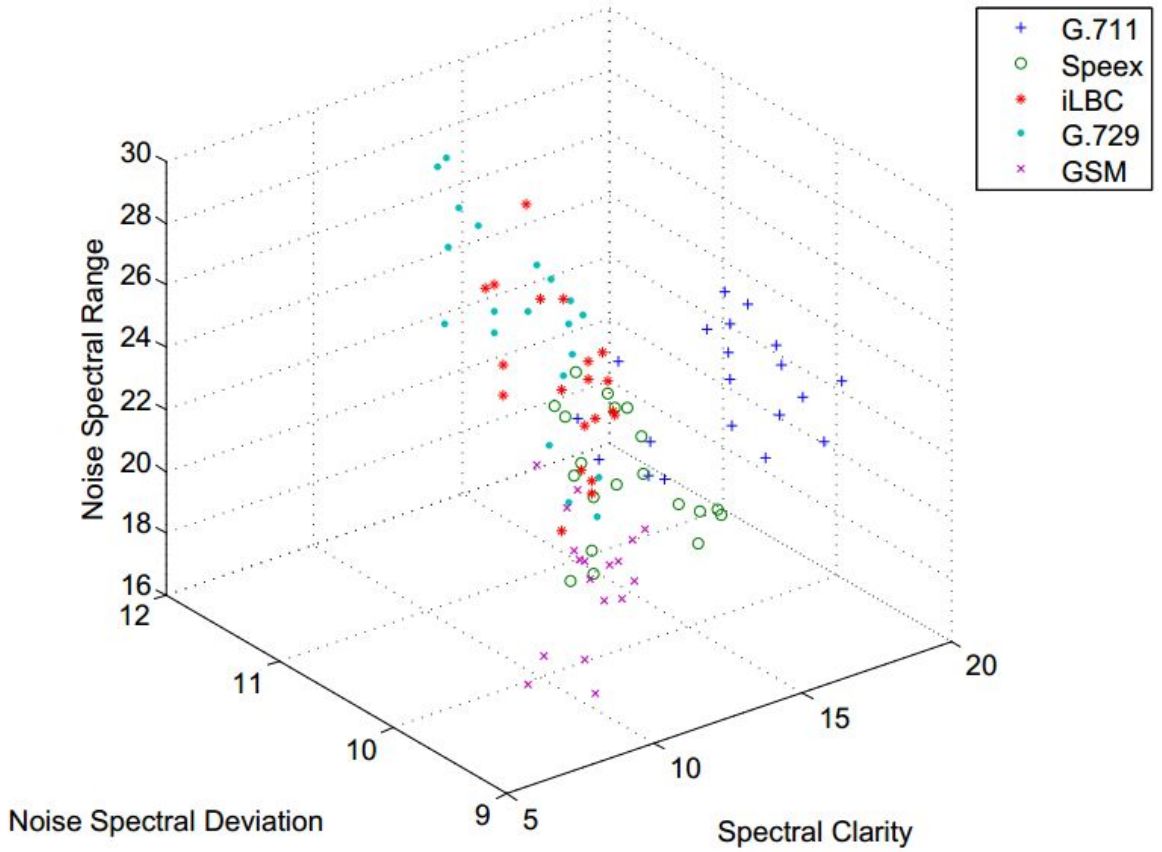
PSTN and Cellular Networks Detection

- PSTN and cellular networks can be identified and characterized due to their vastly different noise characteristics.

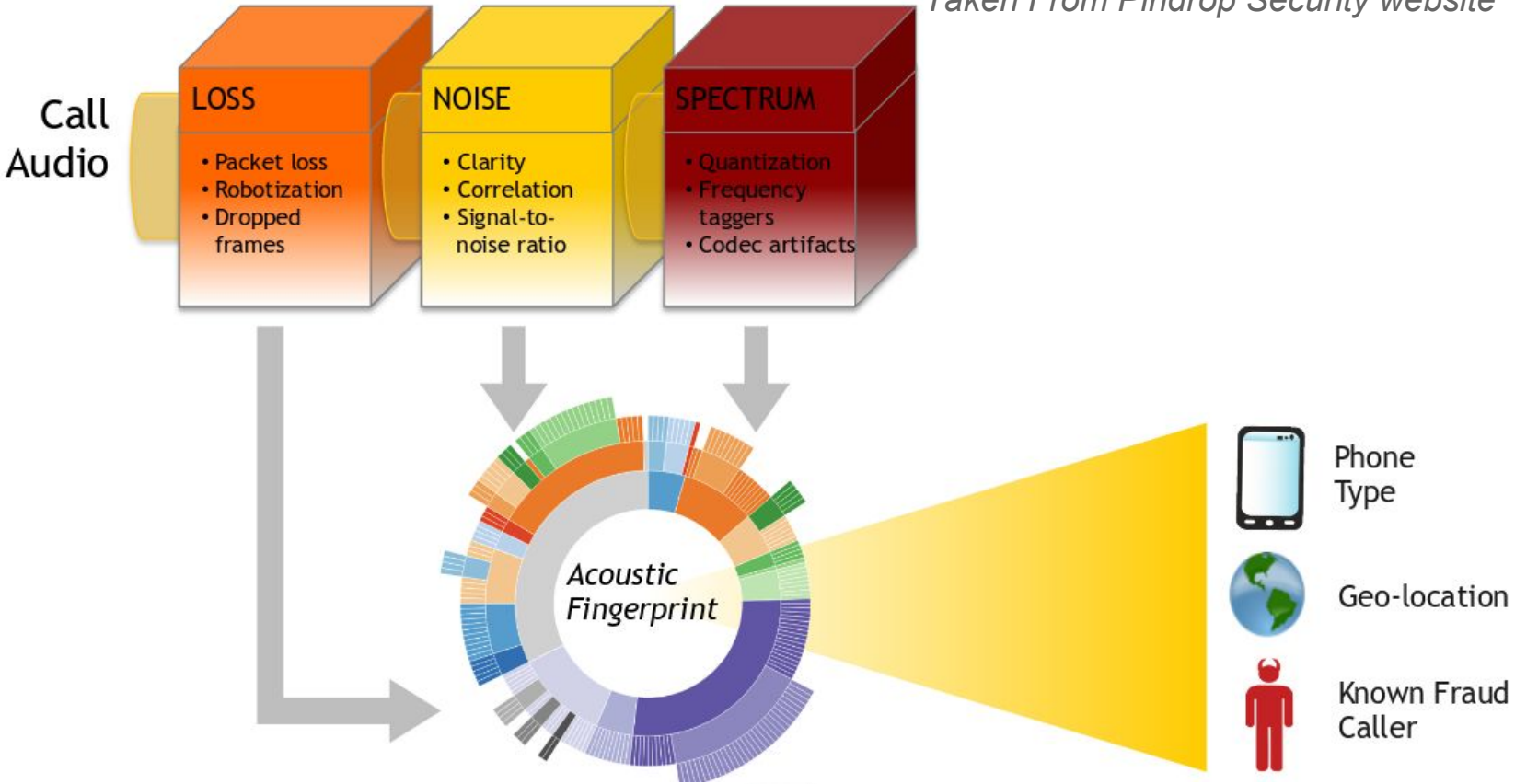
Spectral clarity quantifies the perceptible difference in call quality that we experience when talking on a landline versus a mobile phone

1. PSTN uses G.711:
    a. Without any compression and require much higher bandwidth (64 kbps)
    b. The spectral clarity for such a codec, or the measured crispness of the audio, is very high

2. Cellular Networks use GSM-FR:
    a. High compression codecs like with lower bandwidth (13 kbps)
    b. spectral clarity of such codecs suffer due to the significant compression


G.711
Speex
iLBC
G.729
GSM
Noise Spectral Range
30
28
26
24
22
20
18
16
Noise Spectral Deviation
12
11
10
9
Spectral Clarity
5
10
15
20

Taken From Pindrop Security website



In A Nutshell

The complete provenance fingerprint of a call consists of the path traversal signature, and profiles for packet loss, concealment, noise and quality.

Evaluations

Evaluated based on:

- Accuracy of multi-label classifier in predicting the correct network traversal signature of a call
- Ability of provenance fingerprint to consistently identify a call source

Predicting network travel signatures

Experiments are conducted by taking speech samples from the Open Speech Repository and encoding it with the appropriate codec using PJSIP.

Each sample is subjected to codec transformations and network degradations depending on the networks it traverses

Classification of a Call

- A feature vector consist of:
  - packet loss,
  - noise and quality measurements
- A sample has five labels, each indicating the presence or absence of a codec
- Multi-label classifier is trained on each sample’s feature vector and label

Multi-Label Classifier

Multi-label classifiers can use a variety of reduction techniques to convert the multi-label into a single label.

- Random k-Labelsets (RAkEL)

We use C4.5 decision trees as the underlying single-label classifier

The results show that we are able to predict which networks a call traversed with high accuracy

Provenance fingerprint to consistently identify a call source

If this fingerprint remains consistent for a call source, it can be used to identify and distinguish different calls

Asked different users to make a set of 10 live calls to our testbed in Atlanta, GA from 16 different locations around the world,

Provenance fingerprint to consistently identify a call source

Extract features from the received audio and then label all calls from a call source with the same unique label.

Then, trrain a neural network classifier

The results show that even if a single set of 16 calls is labeled, the remaining sets of calls from the 16 different locations are identified with the correct call source label with 90% accuracy.

Limitations

The majority of misclassifications occur for samples that traversed a VoIP network with 0% packet loss rate.

Plan to study when there is no degradation

Couple other limitations in the paper

Take Aways

Identified robust source and network path artifacts extracted purely from the received call audio
Developed call provenance classifier architecture
Demonstrated our robustness in identifying call provenance for live calls
PinDr0p makes VoIP-based phishing attacks harder and provides an important first step towards a Caller-ID alternative

Discussion

- Criticisms / limitations of the paper ?
- Would this work in a real world with a moving source?
- Any other feature or artifact we can use to identify caller?

Backup Slides

Codecs

Voice is encoded and decoded in each telephony network using a variety of codecs
Different networks use different codecs
Depends on sound quality, robustness to noise, and bandwidth requirements

Codec	Networks	Applications
G.711	PSTN, VoIP	Standard Telephony
GSM-FR	Cellular	Cellular Telephony
iLBC	VoIP	VoIP over Cable
Speex	VoIP	XBox Live
G.729	VoIP	SkypeOut/SkypeIn

- Provenance detection
  - Check packet loss
  - Use correlation algorithm to detect packet loss concealment
  - Extract noise profile and add to feature vector

Our packet loss and packet loss concealment detection algorithms identify three aspects about the provenance of a call: (1) Whether the call traversed a VoIP network, (2) the packet loss rate in that network and (3) the codec used in that network. (1) identifies if there are VoIP networks in the path of a call and (2) and (3) characterize the VoIP network.

Call Traversal Scenarios

Configuration	Scenario	# Simulated Samples
Single Network Traversal		
PSTN - PSTN	Plain old telephone call	20
Mobile - Mobile	Short distance call b/w cell phones	20
VoIP - VoIP	Unfederated call b/w VoIP clients e.g., Google Talk	60
Two Network Traversal		
PSTN - Mobile	Call b/w PSTN landline and cell phone	320
PSTN - VoIP	Call b/w PSTN landline and VoIP client e.g., SkypeOut	360
Mobile - VoIP	Call b/w cell phone and VoIP client	560
Three Network Traversal		
PSTN - VoIP - Mobile	International call using calling cards	1200
PSTN - VoIP - PSTN	Same as above	240
Mobile - VoIP - Mobile	VoIP call bridging b/w two mobile phones e.g., Google Talk	960
Mobile - PSTN - VoIP	Call b/w mobile using a PSTN core network and a VoIP client	400
Mobile - PSTN - Mobile	Similar as above	80
VoIP - PSTN - VoIP	Call b/w two commercial VoIP clients e.g., typical Vonage call	720
		Total = 4940