AuthentiCall: Efficient Identity and Content Authentication for

Brad Reaves‡, *Logan Blue*†, *Hadi Abdullah*†, *Luis Vargas*†, *Patrick Traynor*†, *and Tom Shrimpton*†

‡North Carolina State University

†University of Florida

Problems with Phone Calls…

YouMail:

2.4 billion robocalls were placed **nationwide in November 2016**,
equaling roughly **7.3** calls per person affected.

The Perfect Weapon: How Russian Cyberpower Invaded the U.S.

Читать статью по-русски

By **ERIC LIPTON**, **DAVID E. SANGER** and **SCOTT SHANE** DEC. 13, 2016

Yared Tamene, the tech-support contractor at the D.N.C. who fielded the call, was no expert in cyberattacks. His first moves were to check Google for "the Dukes" and conduct a cursory search of the D.N.C. computer system logs to look for hints of such a cyberintrusion. By his own account, he did not look too hard even after Special Agent Hawkins called back repeatedly over the next several weeks — in part because he wasn't certain the caller was a real F.B.I. agent and not an impostor.

"I had no way of differentiating the call I just received from a prank call," Mr. Tamene wrote in an internal memo, obtained by The New York Times, that detailed his contact with the F.B.I.

OCT 6, 2016 @ 09:30 AM 92,738 VIEWS

Dozens Arrested In IRS Phone Scam Call Center Raids



Kelly Phillips Erb, FORBES STAFF

I cover tax: paying tax is painful but reading about it shouldn't be. FULL BIO

theguardian

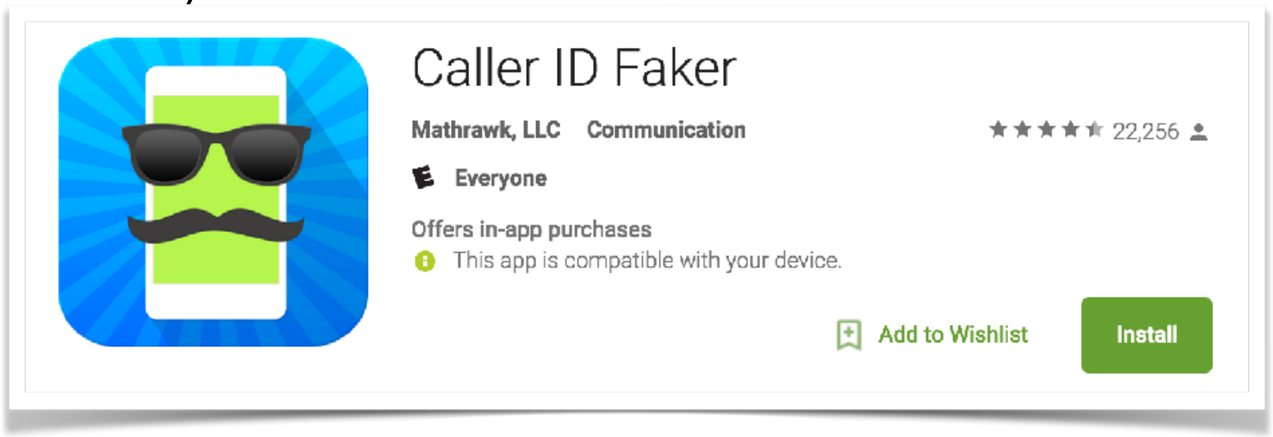
The terror of swatting: how the law is tracking down high-tech prank callers

All of these problems have a single root cause:

Users cannot authenticate the calls they place or receive

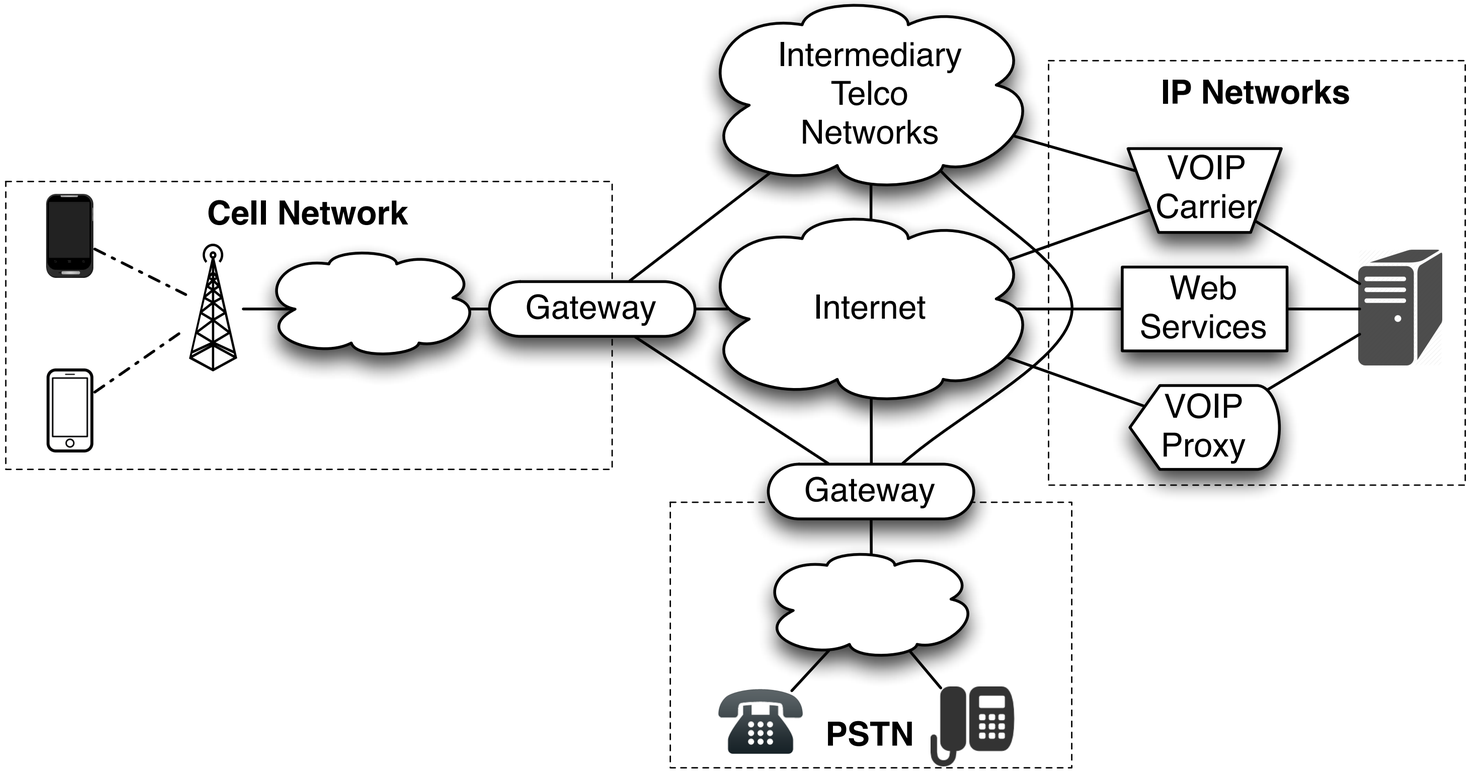
The root of the problem

- In general, neither **carriers** nor **end users** can provide guarantees about who is on a call or where the call is coming from
- Faking a Caller ID is easy

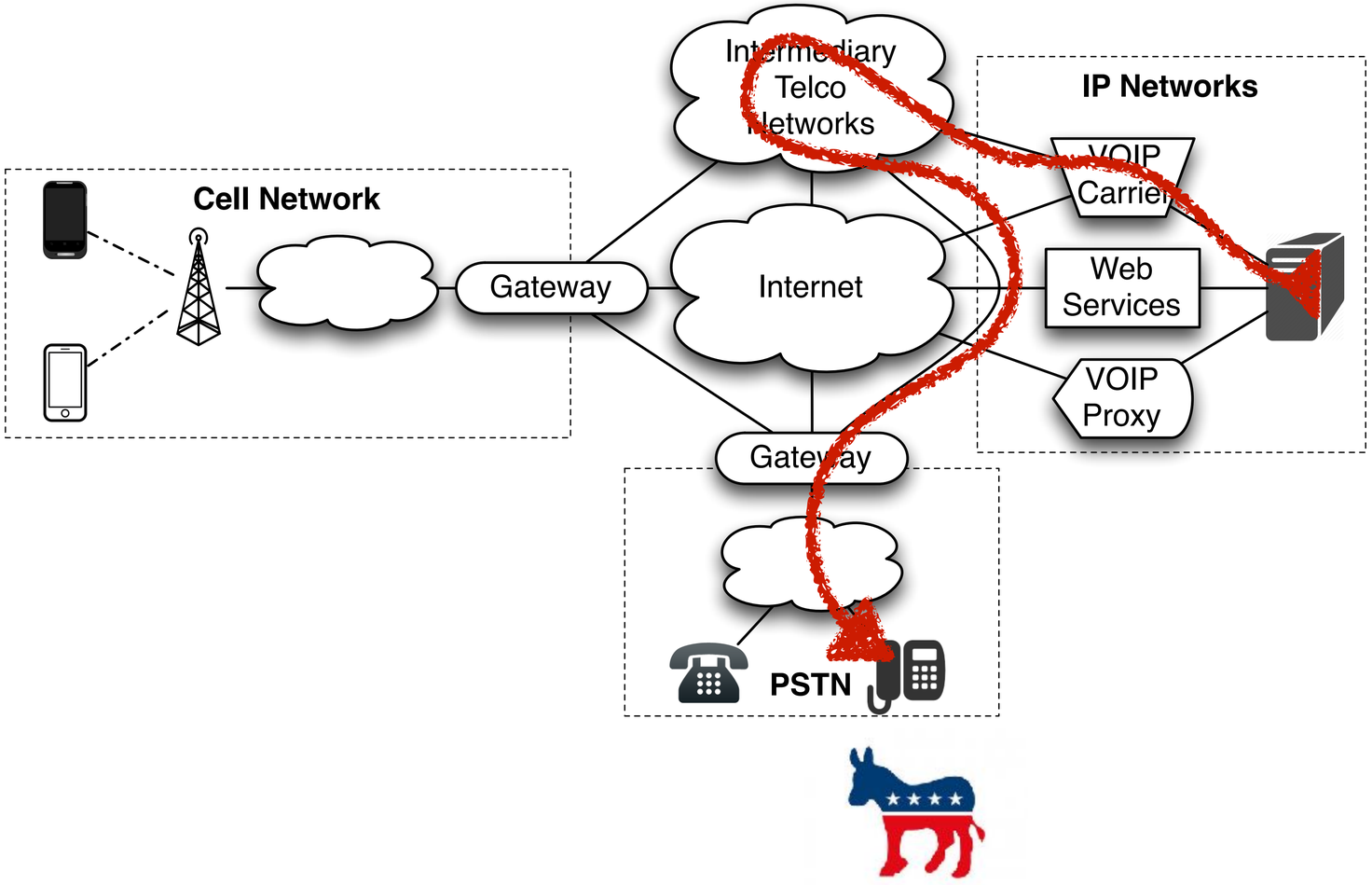


- What is at the root of this problem?

Modern Telephone Systems



Modern Telephone Systems

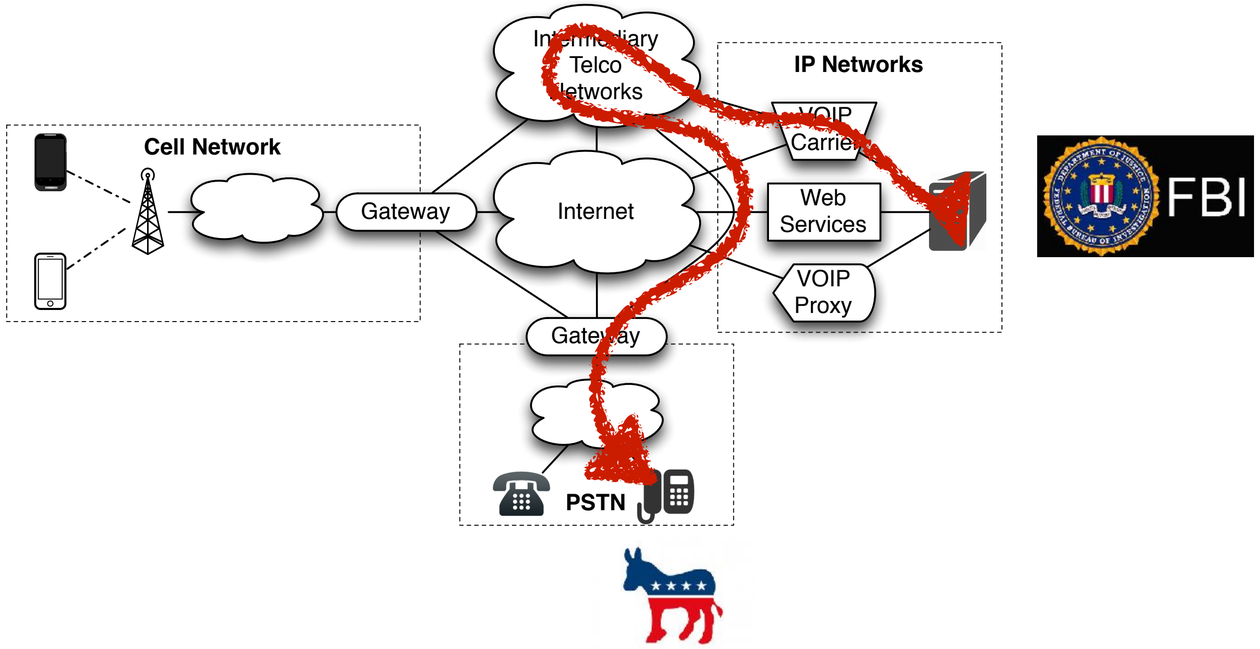




What happens to the call

Along the way:

1. Identity is **asserted**, not **attested**
2. Signaling protocols change at gateways
   - No guaranteed shared data path between endpoints
3. Audio is modified as it flows through the network



The heterogeneity of the network complicates authentication

Insight: Use data channels

- Most phones now have access to an (often low-bitrate) Internet connection



Cellular: 2G, 3G, 4G data



VoIP: Inherent data



Landline: Ubiquitous WiFi

- Data channel may not support bandwidth or quality of the phone network
- **How can this data channel be used to authenticate phone calls?**

AuthentiCall

- **AuthentiCall cryptographically authenticates both call parties and call content end-to-end for *regular phone calls* through an auxiliary data channel**
- **Before the Call:** AuthentiCall verifies identity before the call, making call experience similar to what users already do
- **Fast:** AuthentiCall adds 1-1.4 seconds to call setup
- **Mutual authentication:** Both caller and callee are identified
- **Protects content:** Call content can be verified as authentic
- **Protects against abuse**: Protects callee privacy and can prevent abuse by users

AuthentiCall Mechanisms

- **Protocols** to *enroll* phones and *authenticate* incoming calls using a data channel



- **Audio Authentication Techniques** to protect call content and ensure call integrity
    - Existing content authentication techniques will not solve this problem

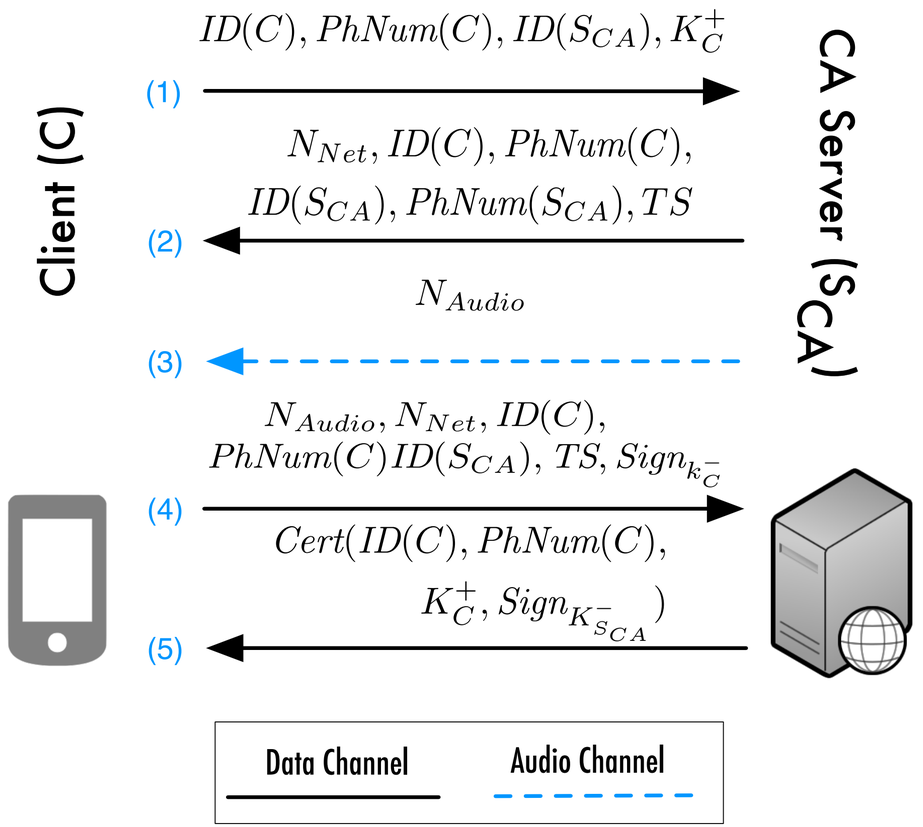


AuthentiCall Protocols

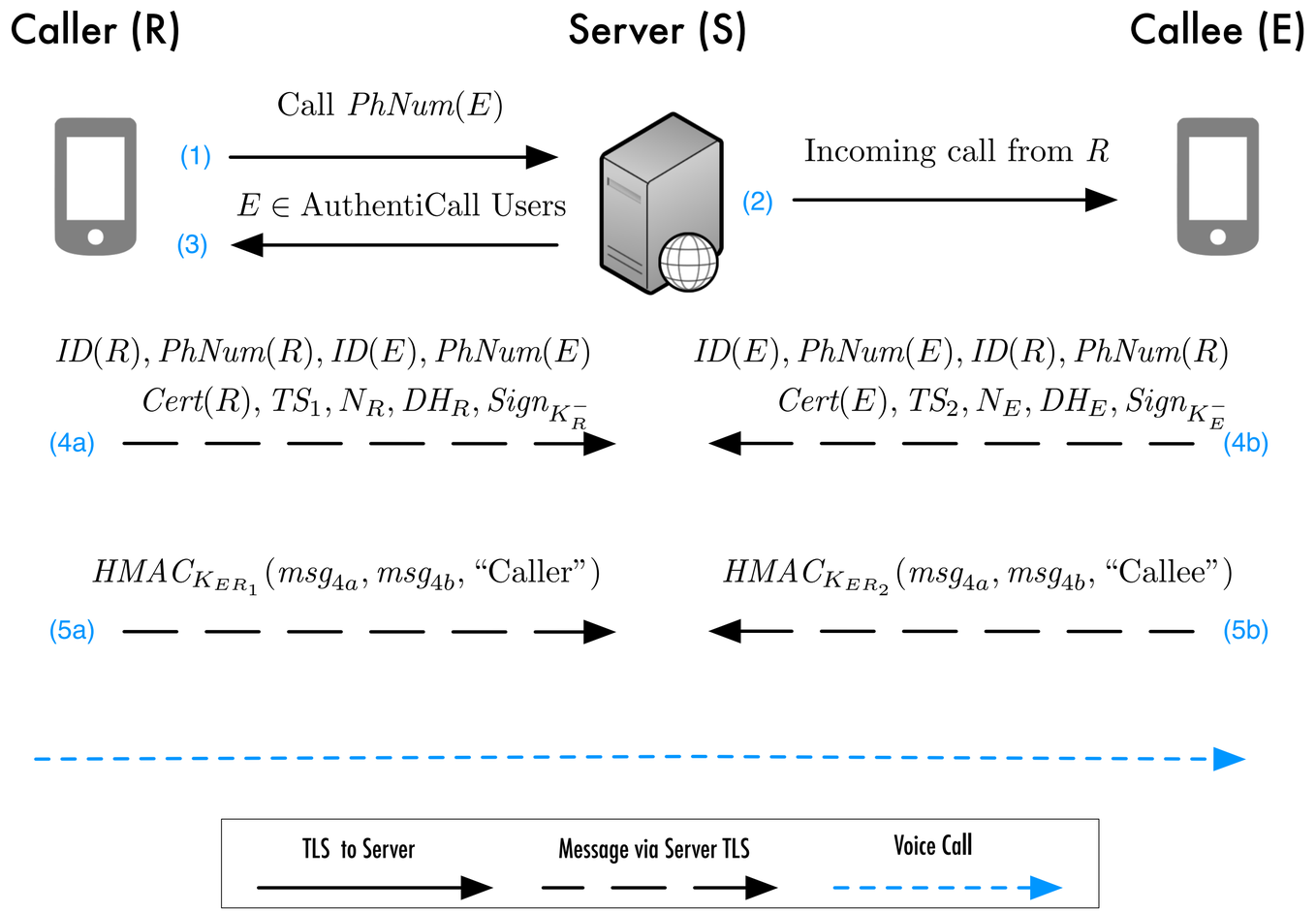
- We need to provide authentication…
- In a setting where neither party is likely to have a public-facing IP…
- And in which adversaries have a strong incentive to flood the system with connection requests
- And may lie about owning a particular phone
- A trusted intermediary can solve these issues…
  - … But at a risk of compromise of the intermediary itself
- *We need custom protocols to address these issues*

AuthentiCall Enrollment

- AuthentiCall issues certificates to authenticate users
- Need protocol to ensure the client actually owns the phone number
- No human in the loop!
- Limitations: similar to the Internet CA model



AuthentiCall Handshake



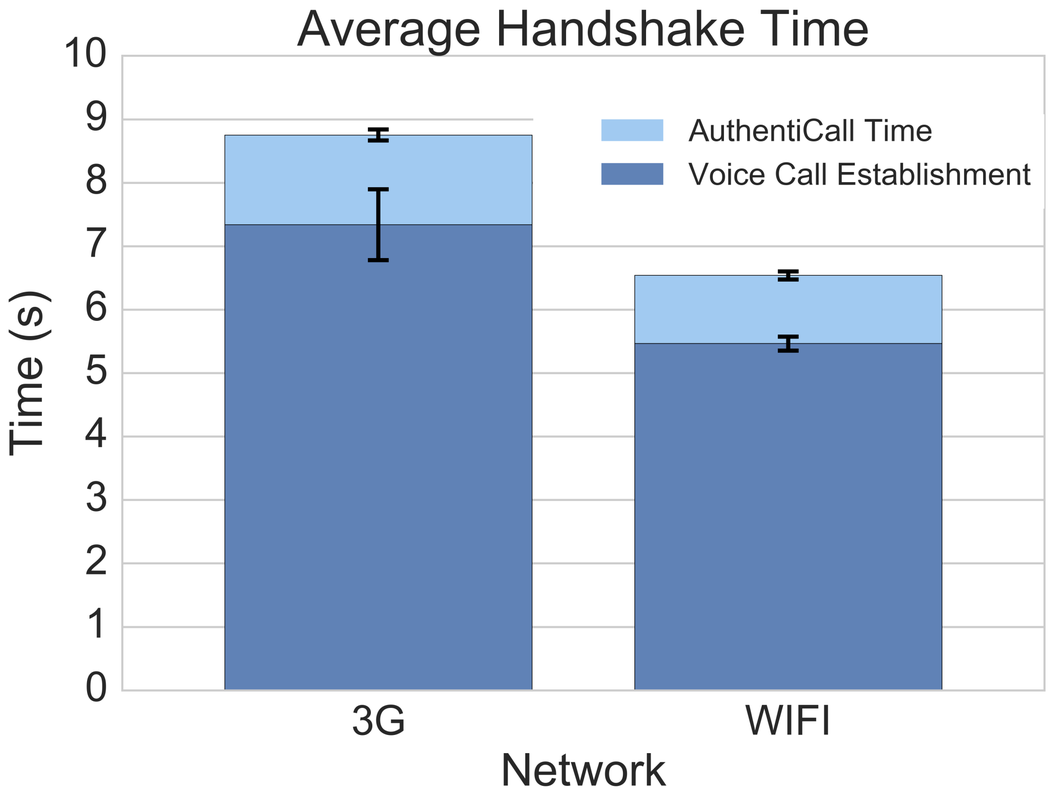
Information to establish a key end-to-end securely in presence of adversary

Handshake complete
Both sides authenticated

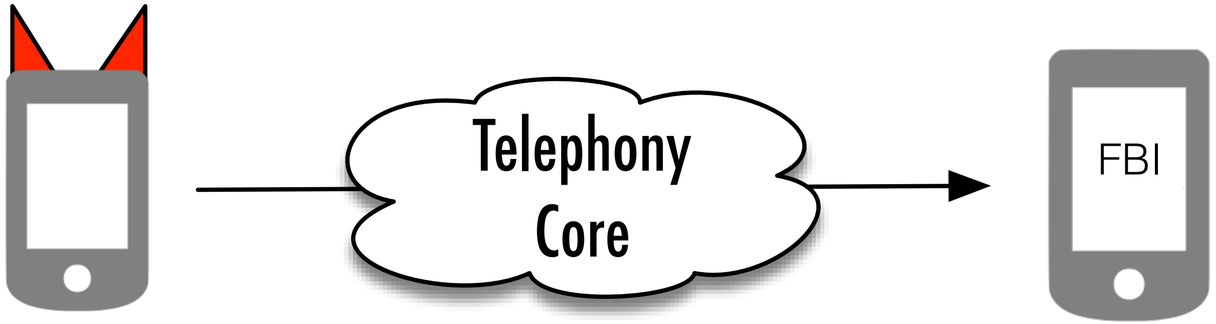
Normal Voice Call Can Proceed

Handshake Performance

- Performance = "How long to authenticate a phone call"
- Regular calls take many seconds to set up with high variance
- AuthentiCall calls experience a negligible 1-1.4 additional seconds for call establishment

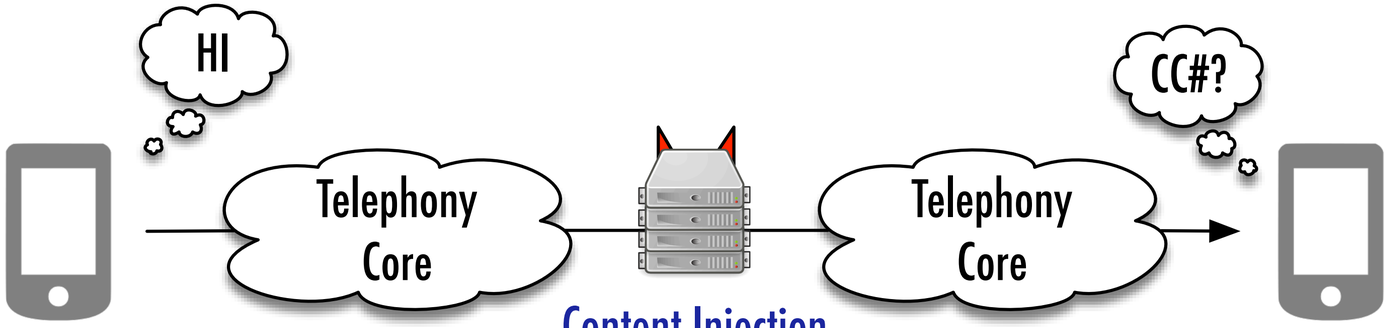


The Handshake deals with this attack:



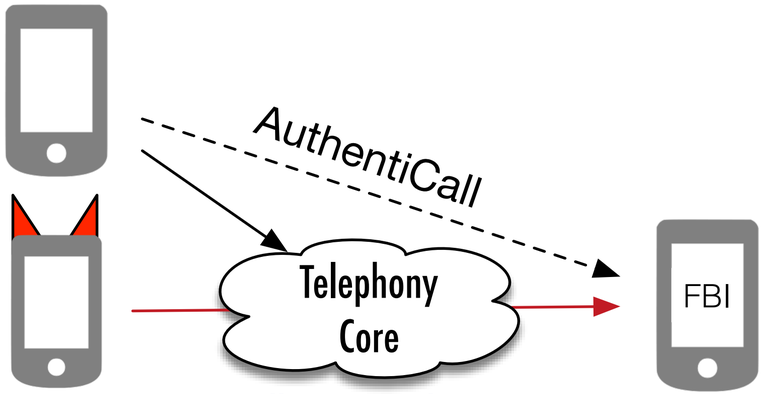
Caller ID Spoofing

But what about this one?



Content Injection

Or this one?



Call Race Condition

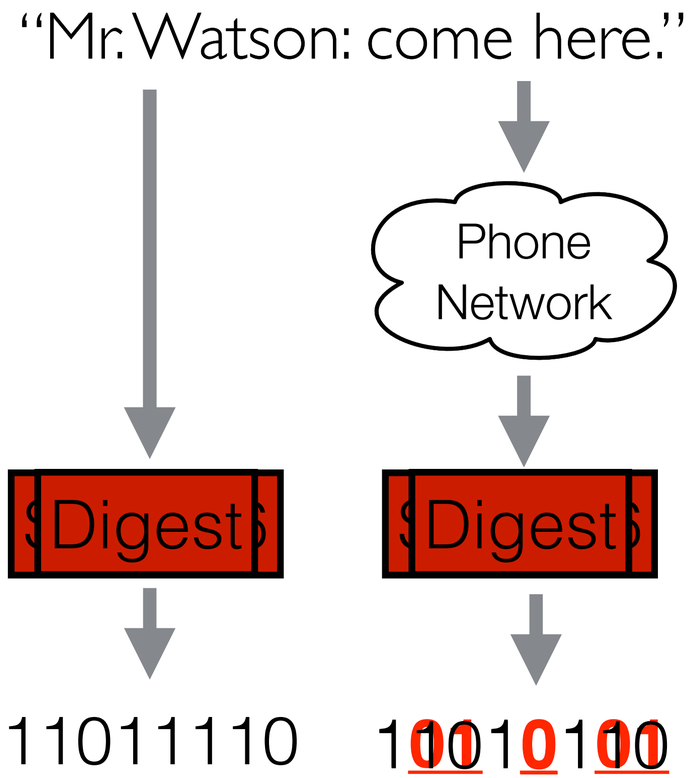
Call Content Protection

- We need to **bind the voice and data channel** and assure **liveness** and **content integrity**
- How do we use a low bandwidth side channel to authenticate call audio?

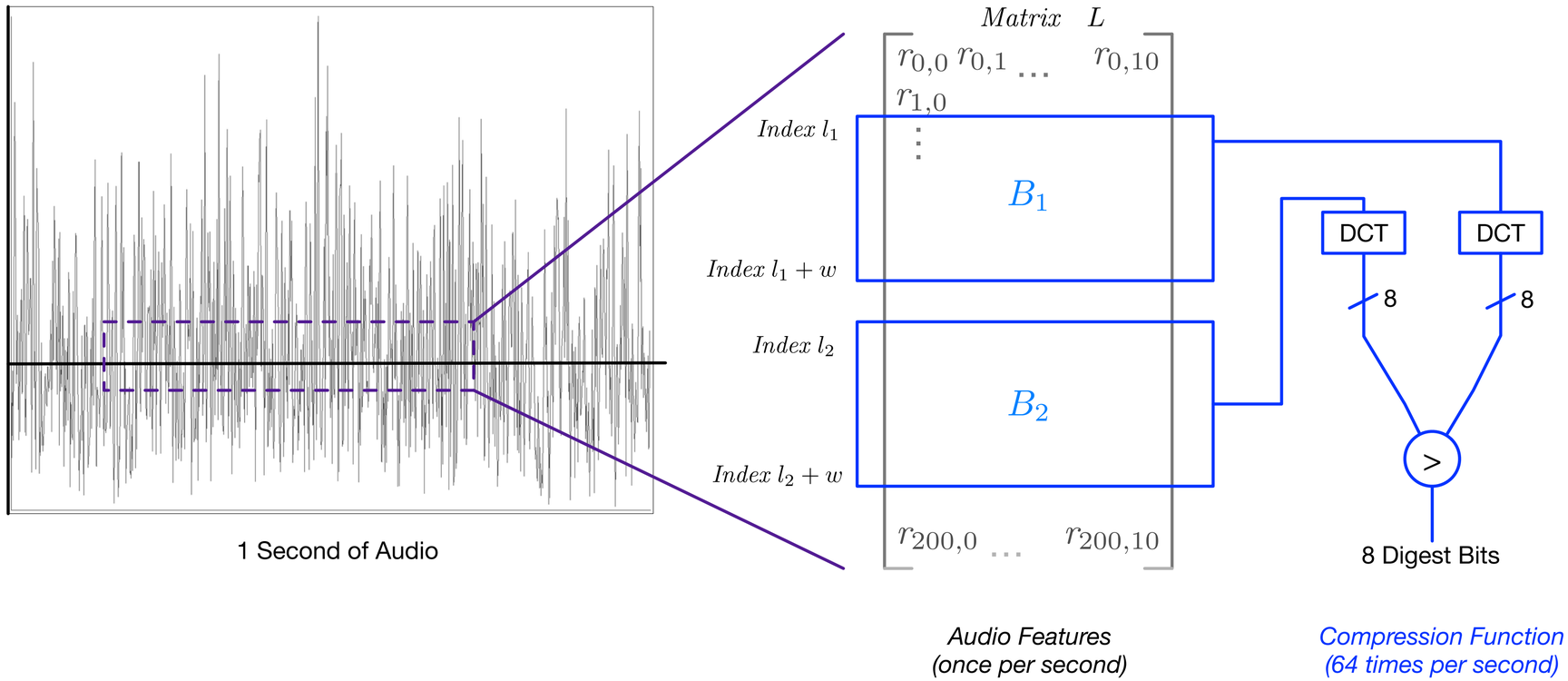


Solution: Robust Digests

- Solution: send "digests" of the call audio
- Digesting is not simple because call audio is legitimately modified in transit
  - Cryptographic hashes won't work here!
- We need a "robust digest" that preserves audio semantics while ignoring legitimate modifications
- Cryptographically authenticating data that can change in unknown ways is **hard**



RSH: The Robust Hash



RSH compresses 1 second of audio into 512 bits

Audio differences can be measured with bit error

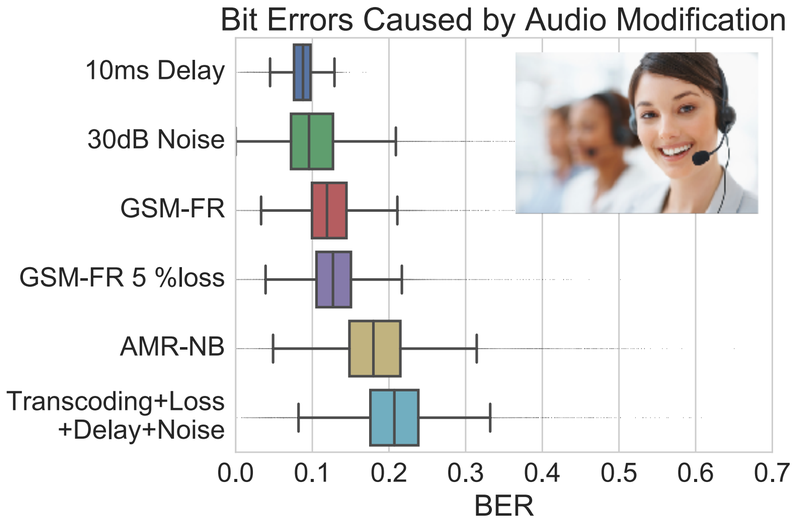
Y. Jiao, L. Ji, and X. Niu, "Robust Speech Hashing for Content Authentication," IEEE Signal Processing Letters, vol. 16, no. 9, pp. 818–821, Sep. 2009.

RSH: Our Contribution

- RSH was not intended for this domain and only briefly analyzed in the literature

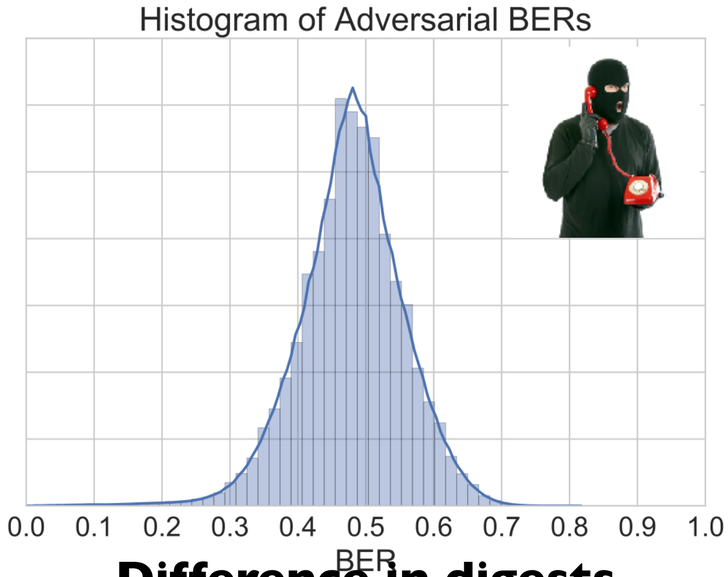
- We need to show:
  - That the randomized construction indeed protects all of the audio
  - That the algorithm will work in real systems
  - That the algorithm will work in an **adversarial setting**

Audio Digest Performance



Change in digests before and after network transmission

Average case: 10-20% of bits



Difference in digests between different sentences of audio

Average case: 48% error

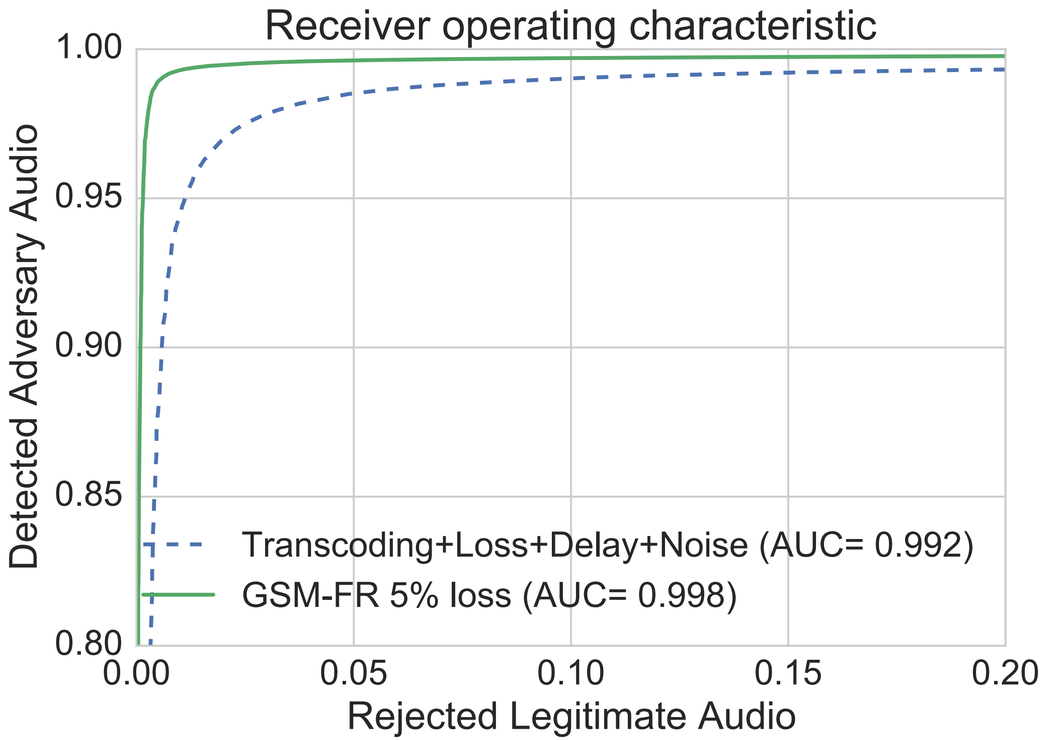
Legitimate changes to audio result in small changes to digests while substituting content results in large changes

Making adversaries easy to detect

Audio Digest Performance

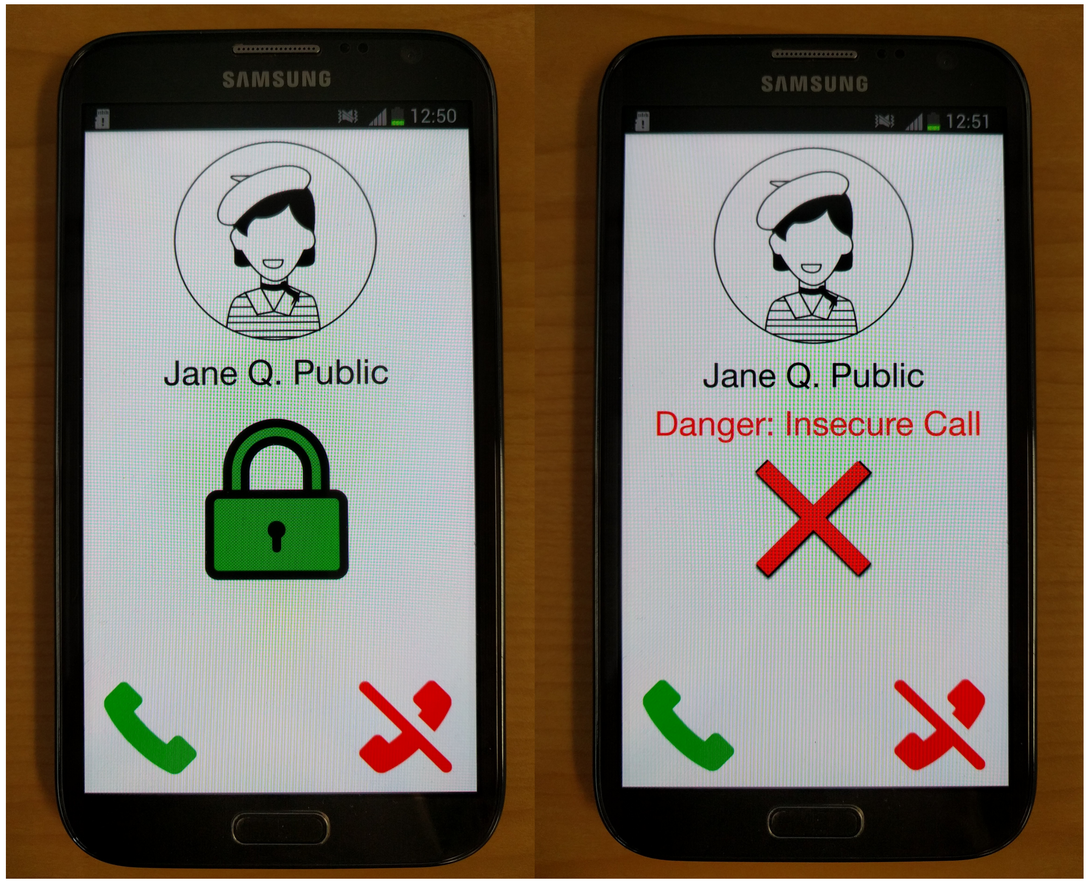
- Individual seconds of audio: 90% detection (false positives: 0.58%, or 1 every 3 minutes)
- 3-out-of-5 seconds: 99%(with a single FP roughly every 6 years)

This ensures both users are on the correct call (channel binding) and that call audio is unmodified



Take Away

- AuthentiCall can authenticate calls before answering — mitigating fraud and unwanted robocalls while increasing trust in the phone system
- We can protect call content from modification
- All at negligible cost to call experience!



(a) (b)

Questions?

Takeaway: AuthentiCall cryptographically authenticates both call parties and call content end-to-end for *regular phone calls* through an auxiliary data channel



Brad Reaves
BradReaves.net