Privacy-Preserving Implicit Authentication

Nashad Safa  Rei Safavi-Naini  Siamak Shahandashti
Outline

• Device, Implicit Authentication
  • Usage patterns, authentication decision making
  • Cost: privacy!

• Our Basic Protocol
  • Preserves privacy against carrier, benign illegitimate users

• Our Improved Protocol
  • Preserves privacy against malicious illegitimate users as well

• Privacy Guarantees, Computation & Communication Cost

• Concluding Remarks
Implicit Authentication

- Idea: authentication by device usage pattern
  - Implicit: does not need user interaction, runs in the background
- Usage pattern is compared with history
  - If conforming: no action
  - If not conforming: user asked to provide the first factor for authentication
- Result: legitimate user not burdened much, illegitimate user caught
Example Scenario

4. Auth. Response

1. Service Request

2. Auth. Request

3. Authentication Protocol

5. Service Response
Storage of Usage Pattern History

**Usage pattern history needs to be stored on the carrier side!**

- Otherwise, loss of device = loss of usage pattern history
  - = ability to mimic (physically or artificially) the usage pattern
  - = loss of authentication security!
  - = loss of privacy!
Usage Pattern Data

• 3 categories of usage pattern data:
  • 3rd party (App server / cloud) data: app usage pattern, app data, ...
  • Carrier data: call, sms, data usage patterns, location pattern, ...
  • Device data: WiFi usage pattern, sensor data, device usage pattern, ...

• Device (3rd party) data needs to be shared with carrier for effective implicit authentication

• We claim this is unnecessary!

• and propose “privacy-preserving implicit authentication”

• Idea: store encrypted usage pattern data
User Profiles & Authentication

• User profile: vector of features
• Each feature belongs to a user-specific distribution
• Feature distributions are approximated by feature history
• On a new reading, a decision is made if it belongs to the distribution

• Observation: often the distribution is a collection of clusters
e.g. based on time of day
A Simple Decision Maker

• For a distribution $D$, calculate a measure of dispersion $d$
  • E.g. standard deviation, average absolute deviation (AAD)
• On a new reading $x$, calculate the area under the distribution curve between $x - d$ and $x + d$
  • This ‘similarity measure’ is between 0 and 1
  • Can be approximated by the number of points recorded in the history
• Only needs comparison, addition, calculation of dispersion $d$
Calculation in the Ciphertext Space

• **Homomorphic Encryption (HE):** enables addition in ciphertext space
  • \( H.\text{Enc}(a + b) = H.\text{Enc}(a) \oplus H.\text{Enc}(b) \)
  • Hence, \( H.\text{Enc}(c \cdot a) = c \odot H.\text{Enc}(a) \)

• Comparison in the ciphertext space
  • Possible using homomorphic encryption, but needs interaction
  • **Order-Preserving Symmetric Encryption (OPSE)**
    • \( a > b \iff OP.\text{Enc}(a) > OP.\text{Enc}(b) \)

Boldyreva et al. EuroCrypt’09
Our Protocol: Idea, Pre-computation

Basic idea:
• Device sends *encrypted* readings to carrier periodically, which are stored on the carrier side as history:

\[ H.\text{Enc}(v(t_i)), \ OP.\text{Enc}(v(t_i)) \]

Pre-computation:
• Carrier finds order in history using order-preserving encryptions, finds encrypted median, calculates average absolute deviation (AAD):

\[ H.\text{Enc}(\text{AAD}(v)) \]
Our Protocol: Authentication, Update

**Authentication:**

- Carrier calculates, sends them to device:
  
  $$H.\text{Enc}(v(t_i) - AAD(v)), \quad H.\text{Enc}(v(t_i) + AAD(v))$$

- Device decrypts, calculates OP encryptions, sends back:
  
  $$OP.\text{Enc}(v(t_i) - AAD(v)), \quad OP.\text{Enc}(v(t_i) + AAD(v))$$

- Carrier locates values, counts no. of ciphertexts within the range

**Update:**

- If authentication succeeds (either implicit or explicit), update AAD
  
  - Only needs a few calculations to account for the difference
Privacy of our Protocol

- Definition based on secure two-party computation guarantees:
  - Device only learns AAD of history
  - Carrier only learns order of current reading compared to history

- Proven our protocol secure against an *honest-but-curious* device, an *honest-but-curious* carrier
  - User privacy is preserved against carrier
  - If device stolen or lost, user privacy preserved against illegitimate users, as long as the device is not ‘hacked’
  - For ‘hacked’ devices, need to consider privacy against *malicious* devices
Improving Security

• To achieve security against malicious devices:
  • Device required to send a *proof of knowledge* of plaintext with the ciphertext $H.\text{Enc}(v(t_i))$
  • Order-preserving encryption replaced by interaction with device to compare ciphertexts
    • Compare $OP.\text{Enc}(v(t_i) \pm \text{AAD}(v))$ with history records via binary tree search
    • $\log \ell$ rounds of interaction for a history of size $\ell$
  
• Proven our protocol secure against a *malicious* device
  • If device stolen or lost, user privacy preserved, even if device ‘hacked’
Comparing Homomorphic Ciphertexts

• Goal: compare $a, b$ given $H. Enc(a), H. Enc(b)$, device has key
• Naïve: send to device, get response, but device learns $a, b$, might cheat
• Equivalent: Calculate $H. Enc(a - b)$, compare with zero
• Randomise: $H. Enc(r(a - b))$, so device does not learn $a - b$, but still might cheat
• Mix with $k - 1$ other values $H. Enc(c_i)$ for known $c_i$, now device might still cheat, but will be caught with high probability
Computation & Communication Cost

Cost of privacy for device: encryption

• Basic protocol:
  • 3 homomorphic, 3 order-preserving encryptions
  • Authentication: 300ms on 2.66 GHz single-core processor
  • Only 2 rounds of communication

• Improved protocol:
  • $k \log \ell$ homomorphic encryptions for security parameter $k$
  • Authentication failure discovered 4 seconds with $k = 2, \ell = 100$
  • $\log \ell$ rounds of communication
Final Remarks

• Implicit authentication improves security without degrading usability
• However it requires giving up on privacy! Is this necessary?
• We proposed privacy-preserving implicit authentication
• Guarantees privacy against carrier, also illegitimate users in case of loss of device
• Does not incur prohibitive extra computation, communication cost
• A step towards showing that

  the trade-off between privacy & security is a false one!
Thank you!

Full version:  eprint.iacr.org/2014/203

Contact me:  siamak.shahandashti@ncl.ac.uk

www.esperez.com