

# Implicit Authentication for Mobile Devices

Markus Jakobsson

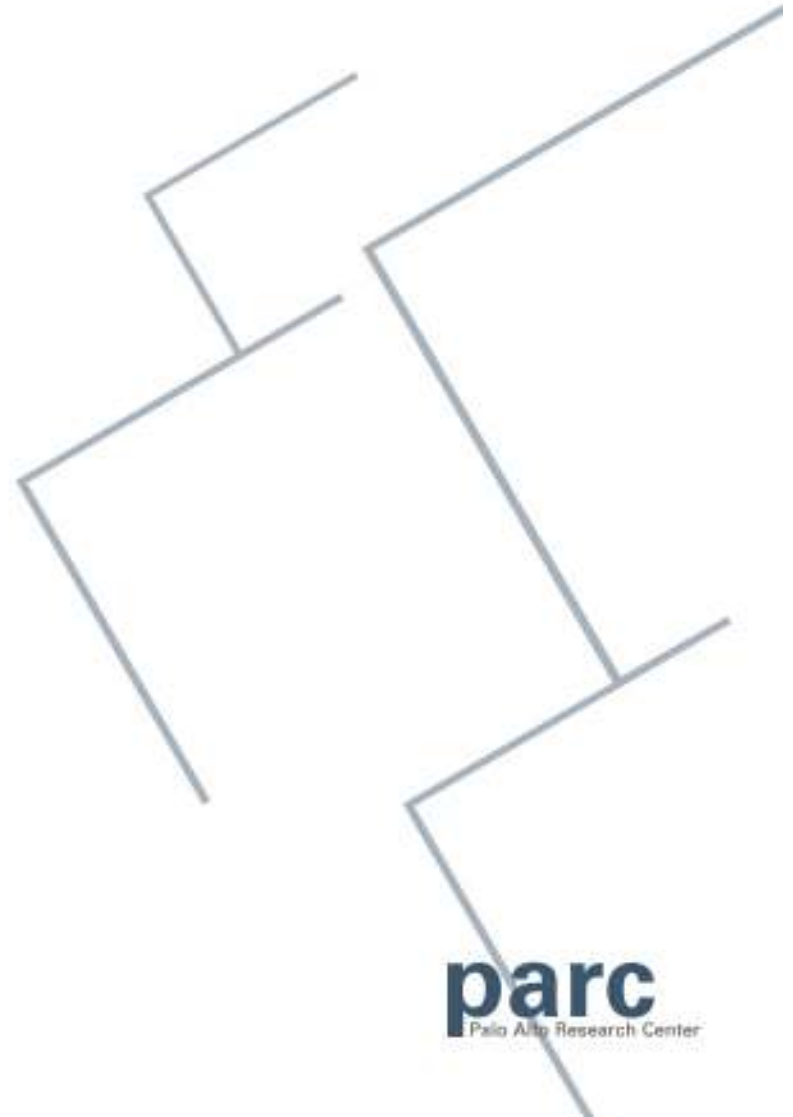
Elaine Shi

**Philippe Golle**

Richard Chow

(Palo Alto Research Center)

Thanks to Yuan Niu (UC Davis)



# Authentication trends

- Increased demand for authentication
  - Hosting of applications and services on the Web
  - Rapid growth of mobile commerce
  - Need to authenticate both users and devices
- Need for higher-assurance authentication
  - Limits of password authentication
    - » Passwords are weak, re-used, shared, lost, ...
  - Mandates for two factor authentication
  - HIPAA legislation
- Growth of mobile Internet devices (MID)
  - Used to access personal, financial, medical data
  - Privacy and liability concerns if the device is lost or stolen
  - Password hard to type (limited input interface)
  - Need for authentication with no/limited user involvement

# Conflicting requirements

We want authentication to be

More secure

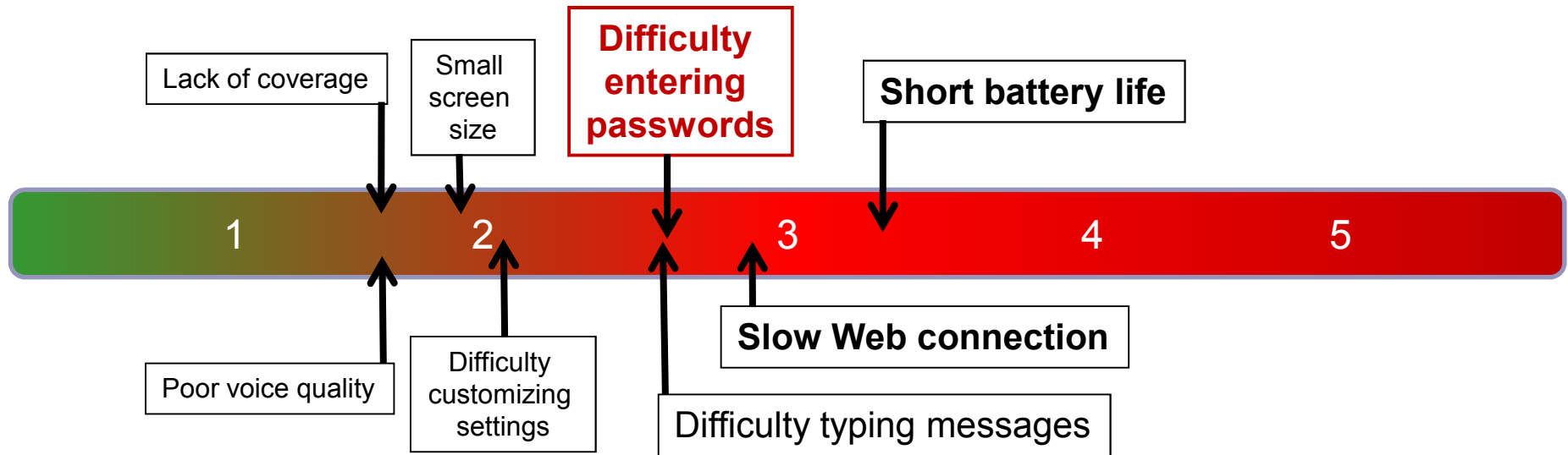
More usable

# Pain of authentication on MID

- Survey of 50 iPhone, BlackBerry and Android users recruited on mTurk
- Password use on mobile devices is *common*
  - 30% need a password to unlock device
  - 46% enter a password once or more / day
  - 24% enter a password 5 or more times / day
- Mobile device passwords are *weak*
  - 44% contain 4 characters or fewer
  - 88% contain only digits
- Mobile passwords are a *pain point*
  - 56% mistype a password 1 in 10 times or more
  - Harder to type passwords on a mobile device (5.0 on scale 1-7,  $\delta=1.1$ )



# Pain of authentication on MID



“It isn’t difficult as much as annoying”

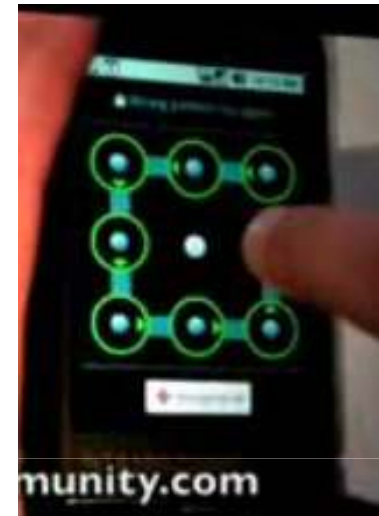
“I really don’t like it when I have to put my passwords in”

# Authentication technologies

- Two factor authentication
  - E.g.: password + token
  - Standard in the enterprise
  - Begins to penetrate high-value consumer markets
  - Concerns with usability and cost
- Proxy solutions
  - Examples
    - » Browser-based password manager
    - » Server-side “remember me” functionality
    - » Single Sign-On (SSO)
  - Address problem of too-frequent authentication requirements
  - Vouch only for identity of device, not user
    - » Do not defend against theft or compromise of device
    - » Do not address voluntary account sharing
  - Poor implementations are insecure

# Authentication technologies

- Graphical passwords
  - Higher entropy
  - Better retention
  - But not in widespread use
- Biometrics
  - Fingerprints, typing patterns, voice prints
- Heuristics to authenticate
  - Transactions (credit card companies, telcos)
  - Machines: OS, browser version, etc (the41.com)



# Implicit Authentication: Our habits authenticate us!

## ■ Opportunity: Rich I/O on mobile devices

- Phone calls (date, time, duration)
- Location
- Calendar events
- SMS in and out
- New email detection
- Opening/closing email messages
- Adding/removing email messages to/from folders
- Creation and sending of new email message
- Types of email attachments
- Accelerometer data
- Adding, removing, editing contacts
- Task list items and memo pad entries
- Holster in/out
- Alerts started/stopped
- Battery level (high, medium, low)
- Etc.



# The case for implicit authentication

- Vision: authenticate users implicitly based on observed behavior
- According to [Furnell et al, 2008]
  - Users want a solution that “authenticates the user continuously/periodically throughout the day in order to maintain confidence in the identity of the user”
  - Receptive to biometrics and behavioral indicators
  - Not receptive to security tokens
- Greendstadt and Beale called for a multi-modal approach “in which many different low-fidelity streams of biometric information are combined to produce an ongoing positive recognition of a user.”

# Implicit authentication on MID

- Data sources
- System architecture
- Learning framework
- Experiments
- Usage scenarios

# Data for authentication

## ■ Types of data

- Location and co-location
  - » GPS coordinates
  - » WiFi, Bluetooth, USB connections
- Application usage
  - » Call, SMS and Web browsing patterns
  - » Software installation
- Biometric measurements
  - » Typing patterns, voice,
  - » Pulse, temperature, blood pressure
- Contextual data
  - » Calendar entries

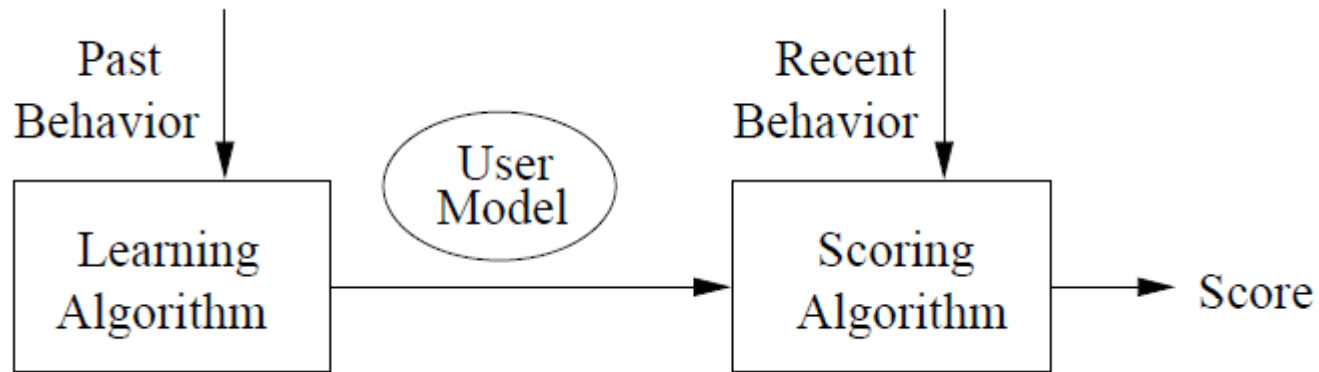
## ■ Data sources

- Mobile device
- Carrier
- Cloud

# System architecture

- Authentication score computed by device, network provider or other third party
- Score computed on device
  - Protects user privacy
  - Does not defend against theft or corruption of device
- Score computed by network provider
  - Established trust relationship with device
  - Natural ability to communicate with device
  - But privacy concerns!
    - » Delete identifying information
    - » Report pseudonymous information
    - » Report coarse-grained or aggregated data
- Authentication score consumed by
  - Mobile device (e.g. to grant access to some resource)
  - A service provider (e.g. a bank)

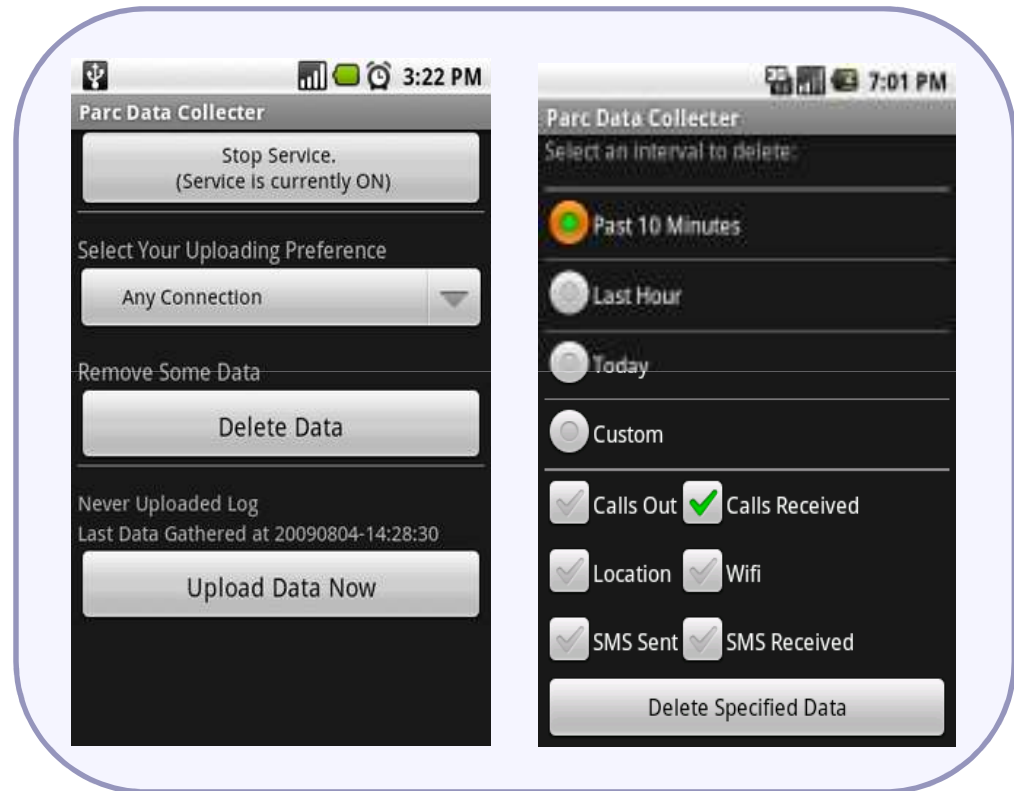
# Learning framework



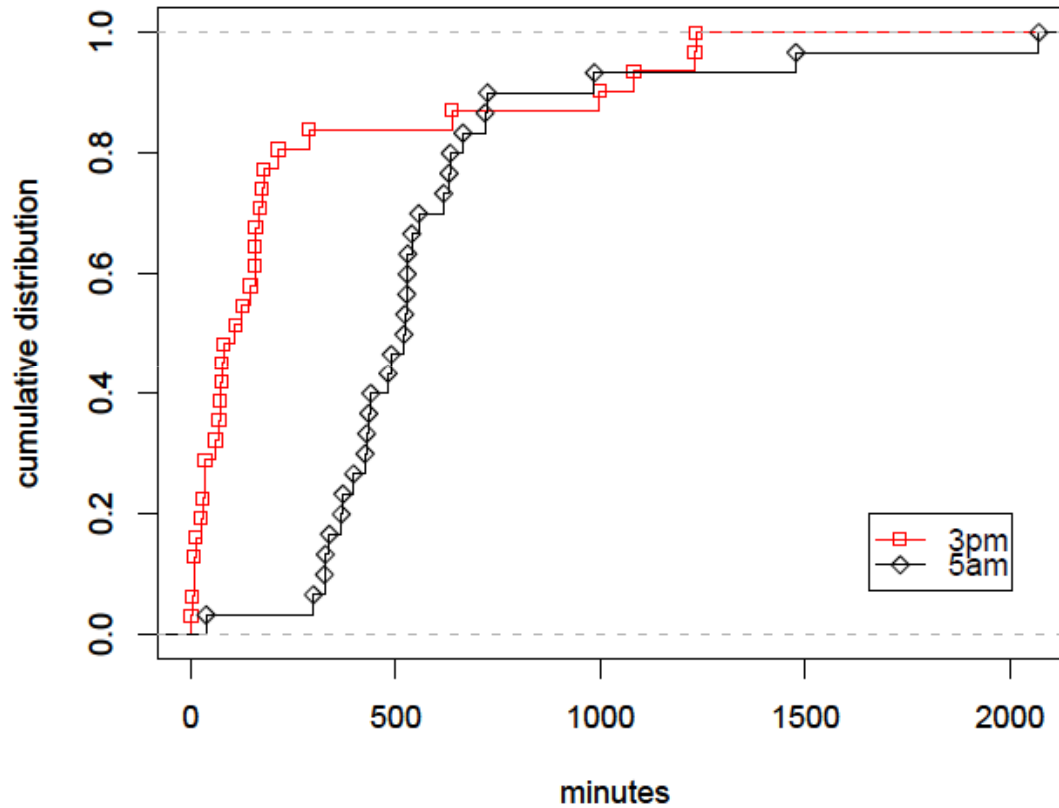
- Independent feature model
  - User model is product of  $k$  probability density functions conditioned on time
- Scoring independent features
  - Score based on observed feature value and pdf
- Authentication score
  - Combination of feature scores
  - Learn combination weights with ML

# Experiment

- Emails
- Calls
- SMSs
- Location
- Contacts
- Calendar
- Tasks
- Memos
- Alerts
- Battery level
- (Un)holstering
- USB connections
- Power on/off
- SD card removal/insertions



# Experiment: phone calls



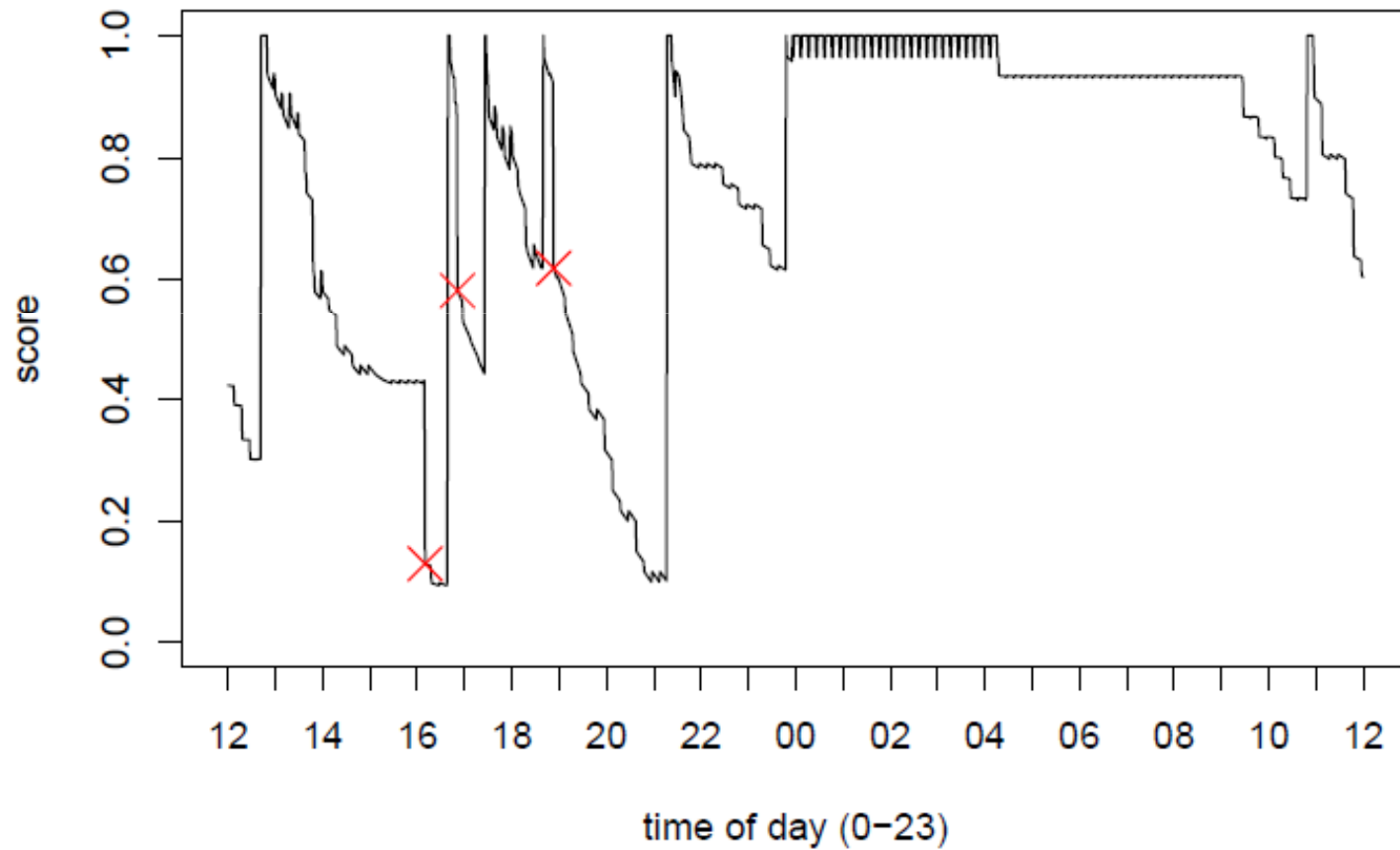
Cumulative distribution of time elapsed since last phone call

# Experiment: location





# Experiment: authentication score



# Usage scenarios

- IA as a password replacement (better usability)
  - Access online email or calendar
  - Log on to online services (Facebook, etc)
  - Unlock phone / medical device
  - Small online purchases
- IA as a second factor (richer backend decisions)
  - Larger online purchases
  - Access of patient records
- Offline uses of IA
  - Paying for subway tolls, pay a vending machine
  - Banking (ATM use, credit card use)

# Future work and conclusion

- Ongoing large scale experiment
  - Train scoring function
  - Model dependencies between features
  - Model adversarial behavior
  - Estimate false positive and false negative rates
- Download our app from the Android marketplace!
- Questions?