Privacy in Mobile Communications

Valtteri Niemi
University of Turku, Finland

FRISC Winter school
FINSE, 24th April, 2013

Talk Outline

- Bad news: Mobility exacerbates the privacy problem

- Good news: Mobile devices are better positioned to deploy privacy enhancing technologies (PETs)

- Surprising news: Mobility can in fact help privacy

Mobility Exacerbates the Privacy Problem

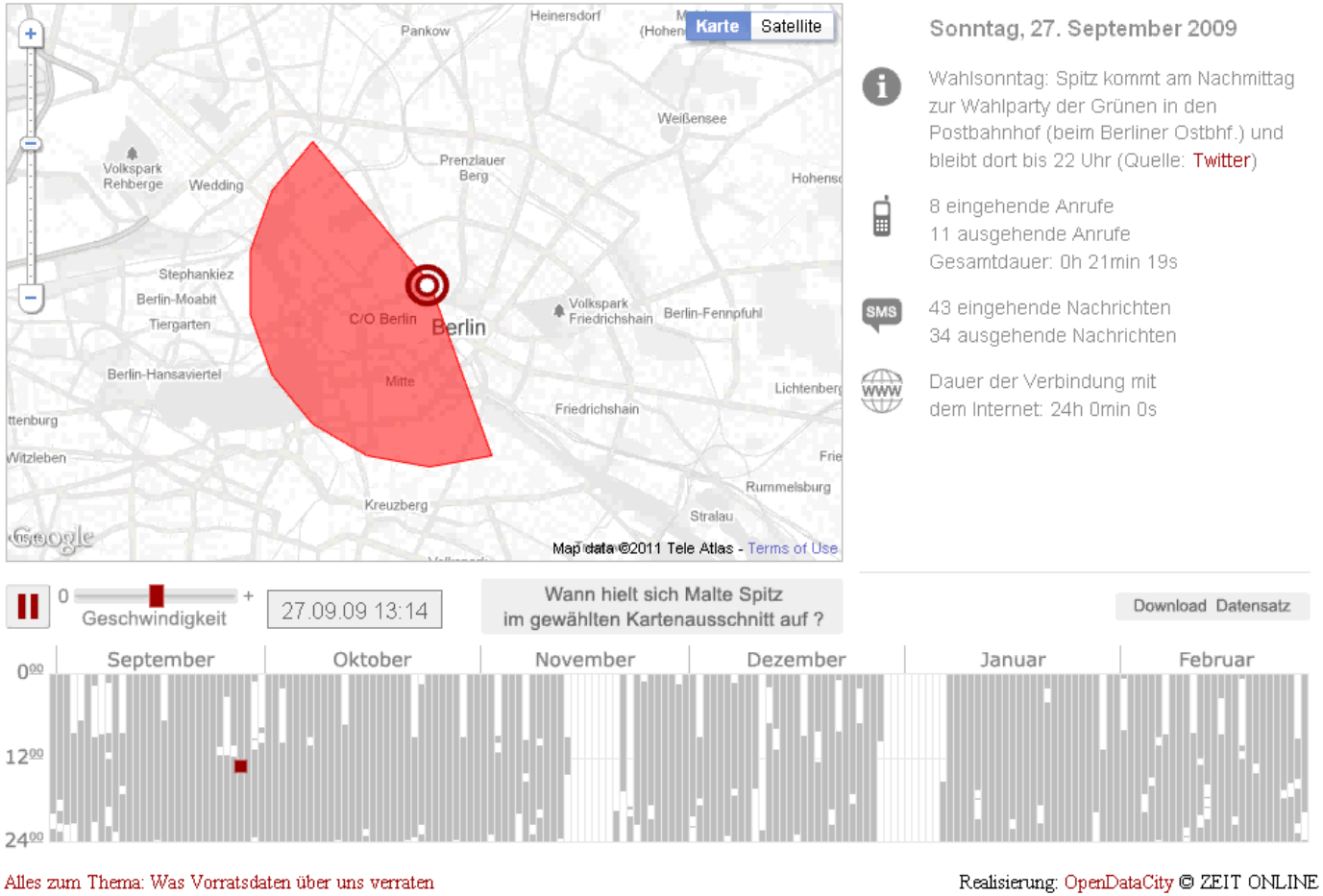
- Operators
- Service Providers on the cloud
- Your device

Mobility Exacerbates the Privacy Problem

- Operators
  - know rough location, contacts, communication patterns

"Six Months of My Life"

Blog entry by Malte Spitz Zeit.de animation



Mobility Exacerbates the Privacy Problem

- Operators
- Service Providers on the cloud
  - want to improve service quality
  - ask for location, know content consumption patterns...
  - …will know financial transactions, health/wellness data…

Mobility Exacerbates the Privacy Problem

- Operators
- Service Providers on the cloud
- Your device
  - Sees all your activities
  - Combining different types of user data is valuable
    - Hence the push to collect data and upload to Service Providers
  - …but risky, at least from a privacy perspective

Threats against identity privacy

- Identifiers in many layers / interfaces
  - MAC address (for WLAN, Bluetooth, cellular, ….)
  - IP address
  - Application layer identifiers (email address, SIP address, usernames, …)
- Correlation between different identifiers
- Long-standing pseudonym = permanent identity
- Active attacks typically easy to do: “please identify yourself”
- Tracking person’s device = tracking person
- Radio fingerprinting

Nokia Instant Community – and a trial

[Bindschaedler, Jadliwala, Bilogrevic, Aad, Ginzboorg, Niemi, Hubaux, ’12]

[Bilogrevic et al. ’12]

Nokia Instant Community

Main Features:

- Create location and context-based networks between people and places
- Automatic – no user actions needed
- Know who/what are close to you

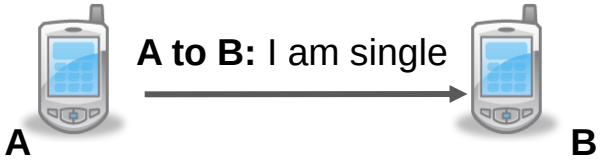
Application examples:

- Find how many people are close to certain point, and, if necessary, who they are
- Connect these people to each other, share information
- Give navigational help and information for places close to you: Point-of-Interest contacts you, not vice versa
- Points of sales give information and advertisements to their customers
  - Loyalty card programs; locating customers in the shop; targeted ads, guidance,…

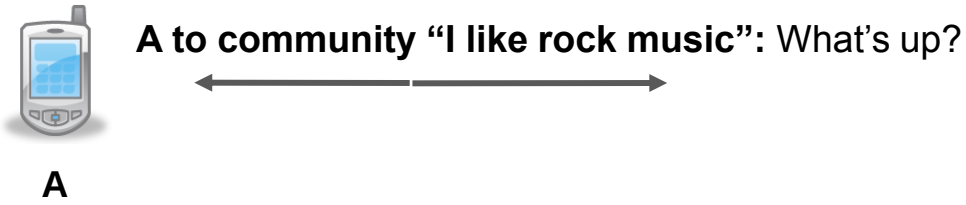
- Start group communication automatically (chat, content sharing etc…)
- Engage the crowd in big events
  - Group messages sent for the event participants
  - Social games and collaborative content creation and sharing

Nokia Instant Community: privacy and identity issues

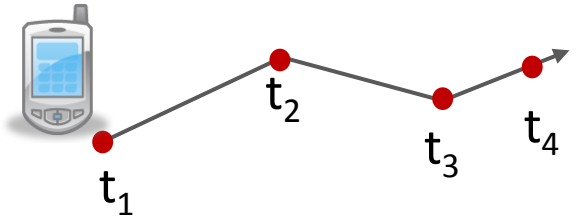
- Privacy (User, Community) supported



- Community anonymity, unlinkability



- Pseudonyms, traceability (location privacy)



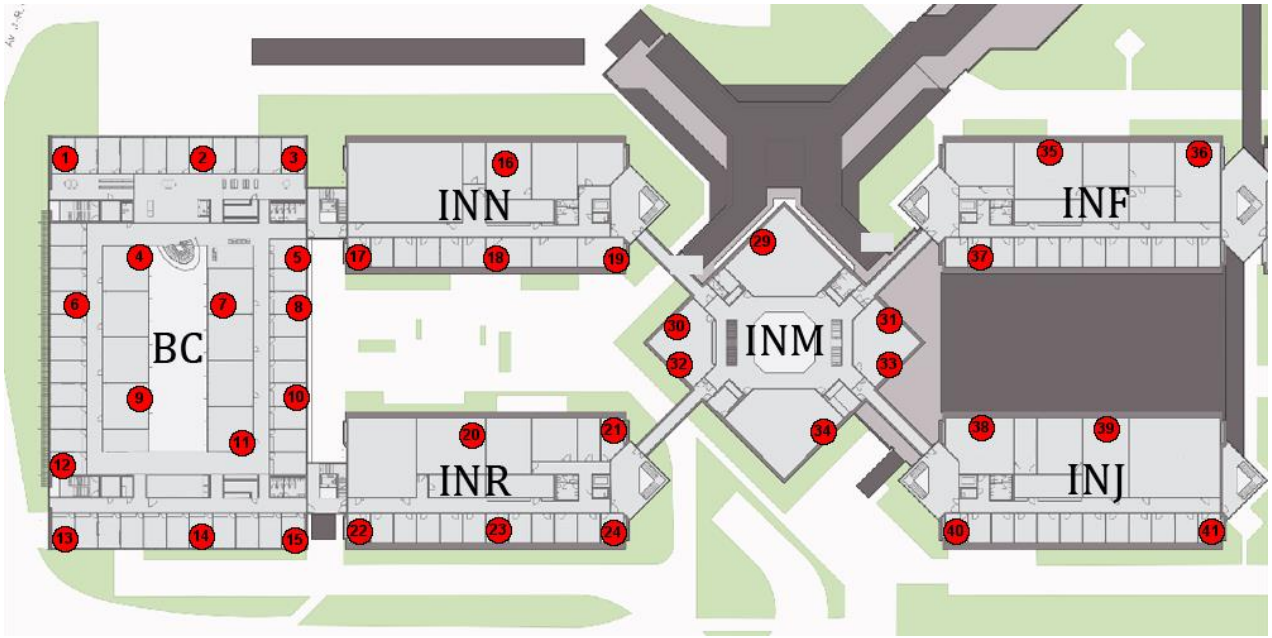
- Message encryption, signing

Nokia Instant Community privacy trial

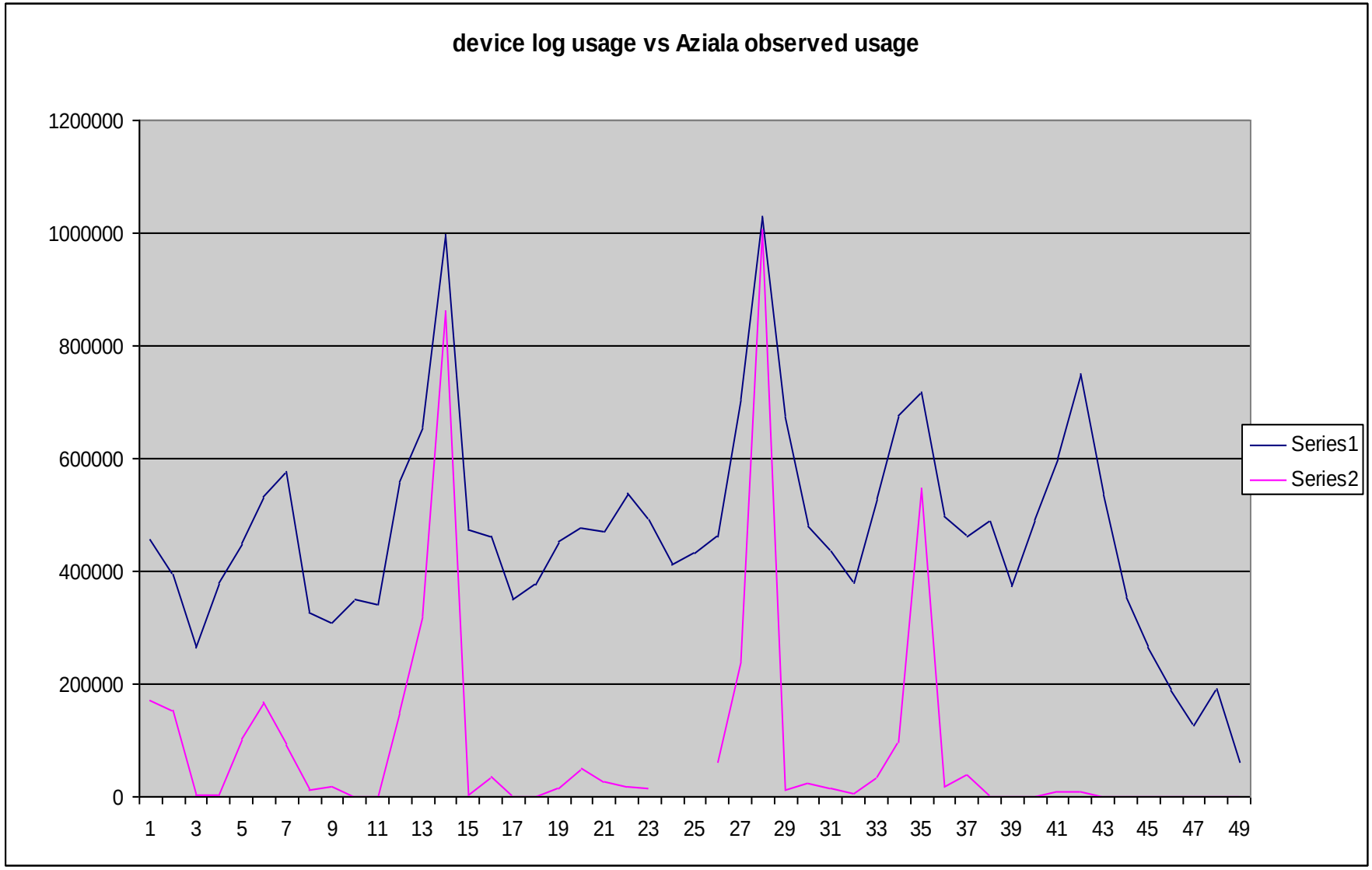
- March-May 2011
- 80 participants in EPFL campus, Lausanne, Switzerland
- Seven applications, some developed by EPFL students
- Several privacy features tested
  - Pseudonyms with pseudonym change algorithm
  - Encryption of messages (with shared community key)
  - Privacy-triggered networking
- Participants carried Nokia N900 device but typically not as their primary phone
- Logs collected in devices and independently with Aziala network (see next slide)

Attacker model

- Passive
- Eavesdrops on communications
- Mesh network of 37/41 sniffing stations (Aziala)
- Coverage 200 m x 100 m



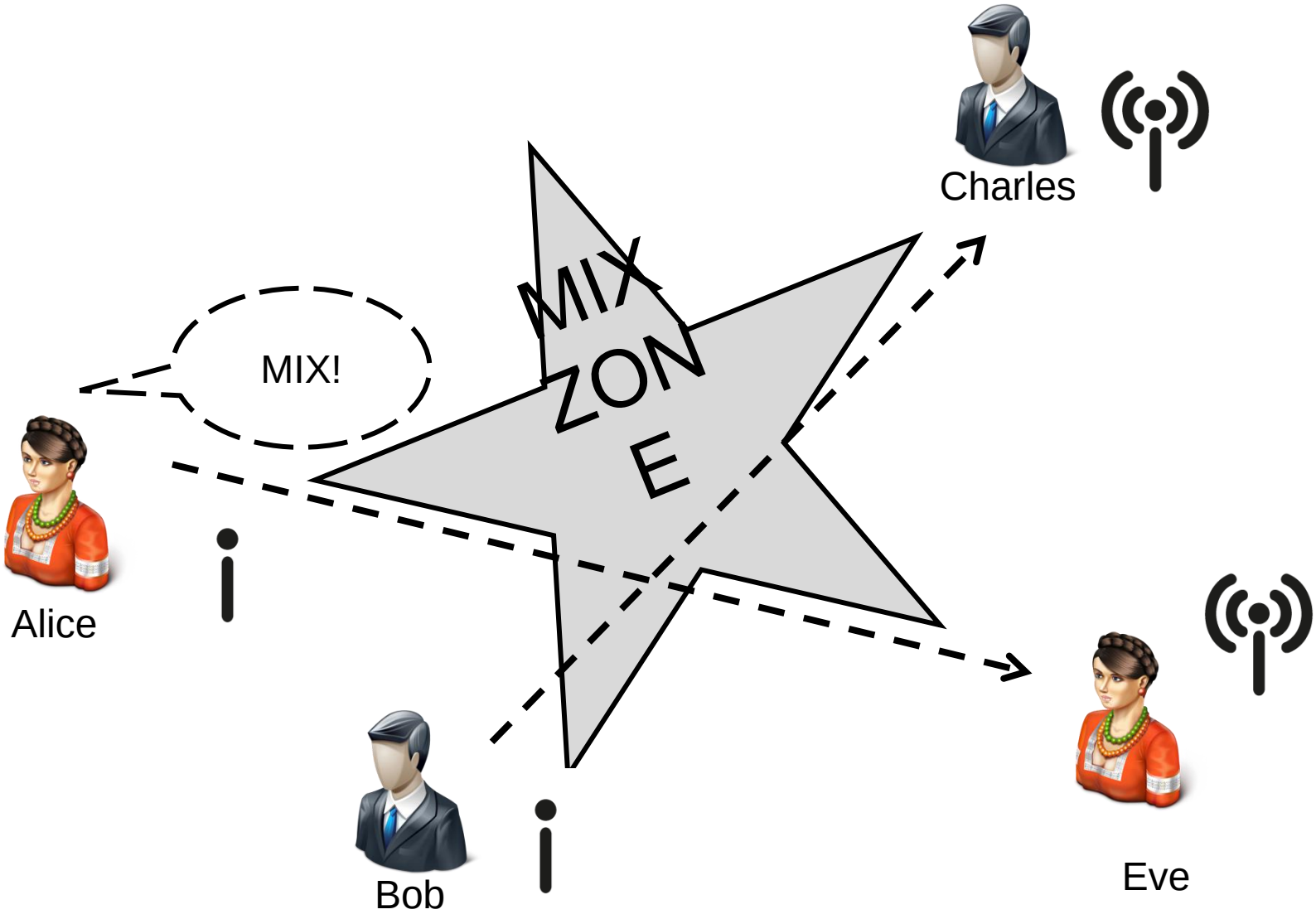
Usage patterns



Attacker Goals

- Track users by de-anonymizing pseudonym-user mapping
- Reconstruction attack
  - Produce a sequence of pseudonyms believed to have been used successively by a given user
- Simple identification attack
  - Seeing who sent one message with one pseudonym is enough to identify the user associated to the reconstruction sequence

Pseudonym Change Algorithm (PCA)



PCA Principles

- Change pseudonym:
  - at fixed times
  - based on context
  - on reception of mix request
  - at midnight

- When a pseudonym change decision is made, the device broadcasts a mix request

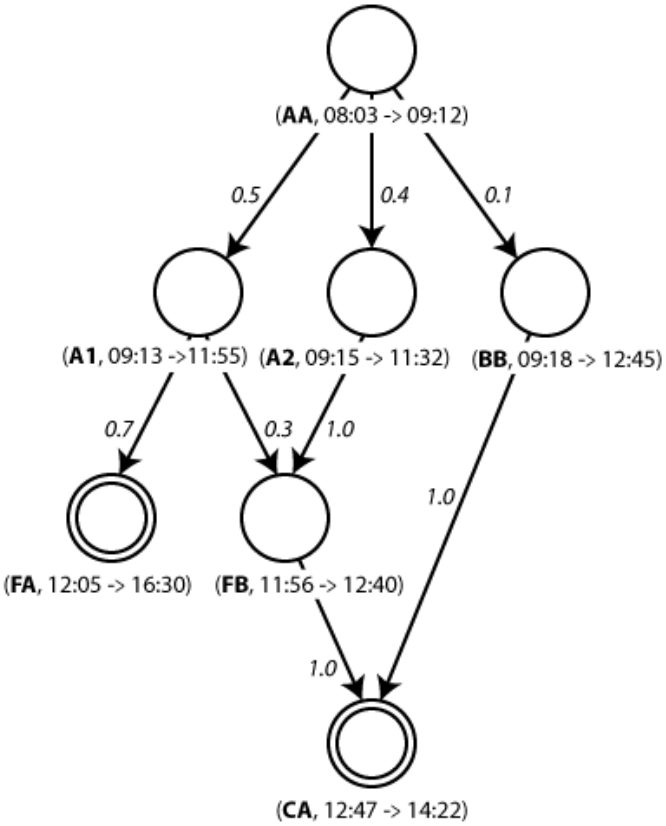
- A quota is placed on the number of allowed pseudonym changes to prevent network performance collapse

PCA Parameters

	Group 1	Group 2	Group 3
Forced timer	14400s	7200s	3600s
Context timer	3600s	1200s	300s
Change threshold	7200s	1800s	600s
Neighbor threshold	1	2	3
Daily change quota	5	20	50

Tracking Model

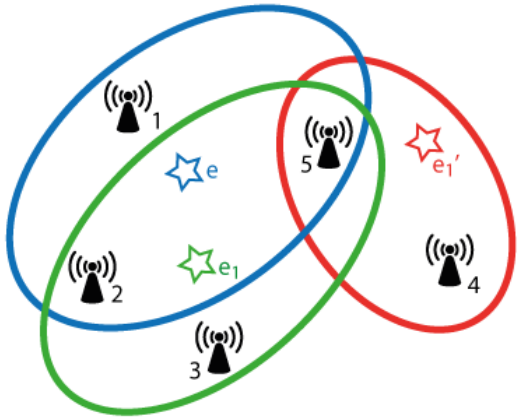
- States
- (pseudonym, first event → last event) -- everything about the use of a given pseudonym
- Transition matrix associates a probability to each pair of states
- In each state, the possible next candidates are those states in the matrix with nonzero probability
- State space is a directed acyclic graph



Heuristics For The Transition Probabilities

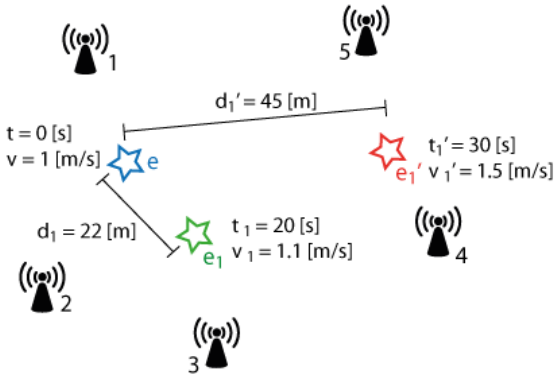
Common sniffing stations

«The more sniffing stations in common between the current state and the next state candidate, the more likely the candidate»



Speed matching

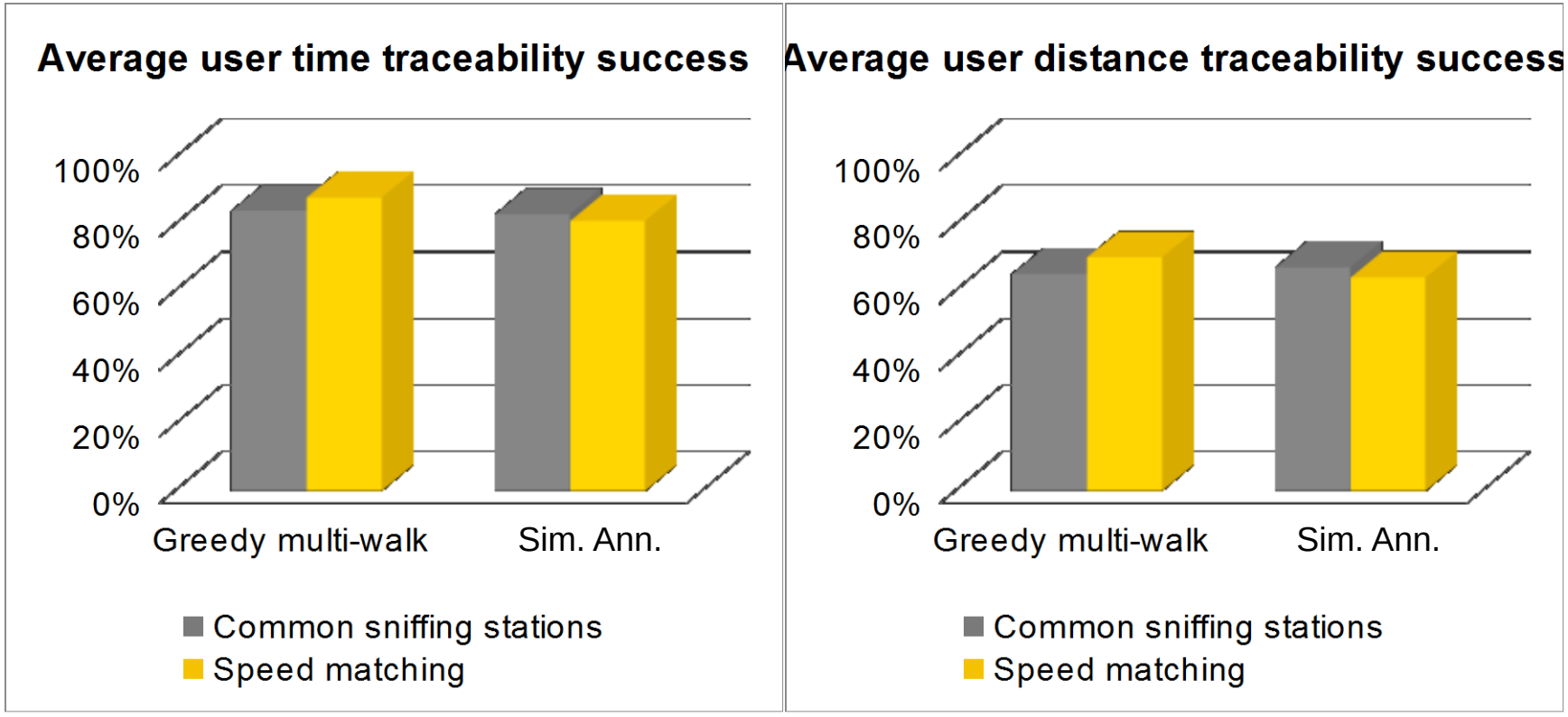
«The closer the user speeds between the current state and the next state candidate, the more likely the candidate»



Results in a Glimpse for single target

- Tracking success > 60% of time, >50% of space
- The two heuristics work almost equally well
- The more reference points, the better the tracking

Tracking Results for multi-target case



Multi-target tracking average traceability (τ)
(6 days, up to 48 users per day)

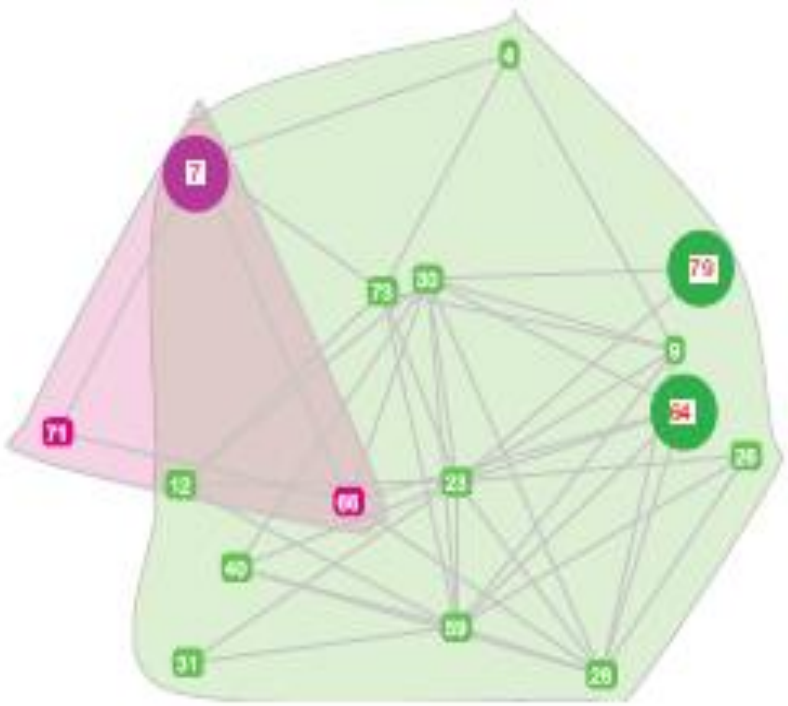
Multi-target Results in a Glimpse

- Tracking success > 80% of time, > 60% of space
- Both algorithms have a similar success rate
- As expected, tracking multiple users simultaneously improves success by removing “collisions”
- MTT can handle users leaving monitored area and coming back during the same day

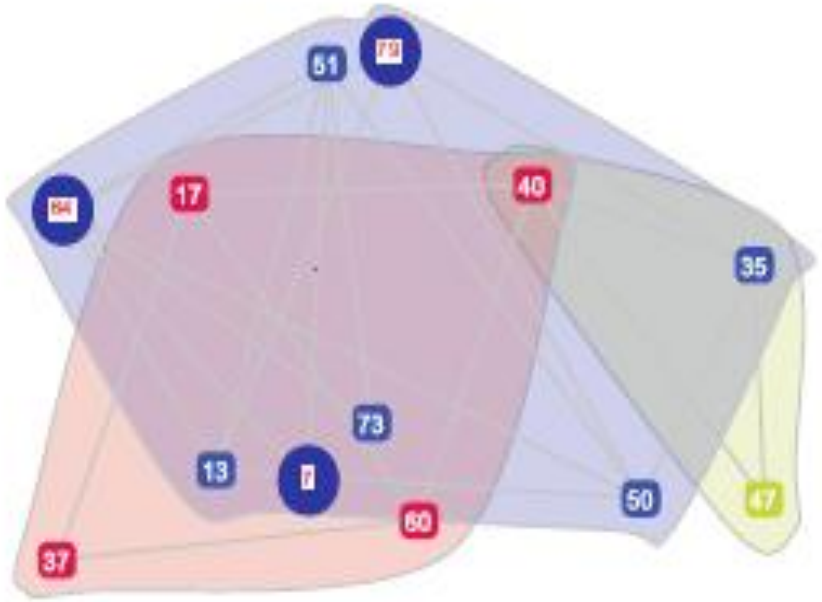
Another attack: Community detection based on proximity

- Attacker finds proximity between nodes by observed RSSI values and trilateration (similarly as in the tracking attack)
- Tries to deduce social contacts based on proximity data (distance, length, frequency)

Community construction

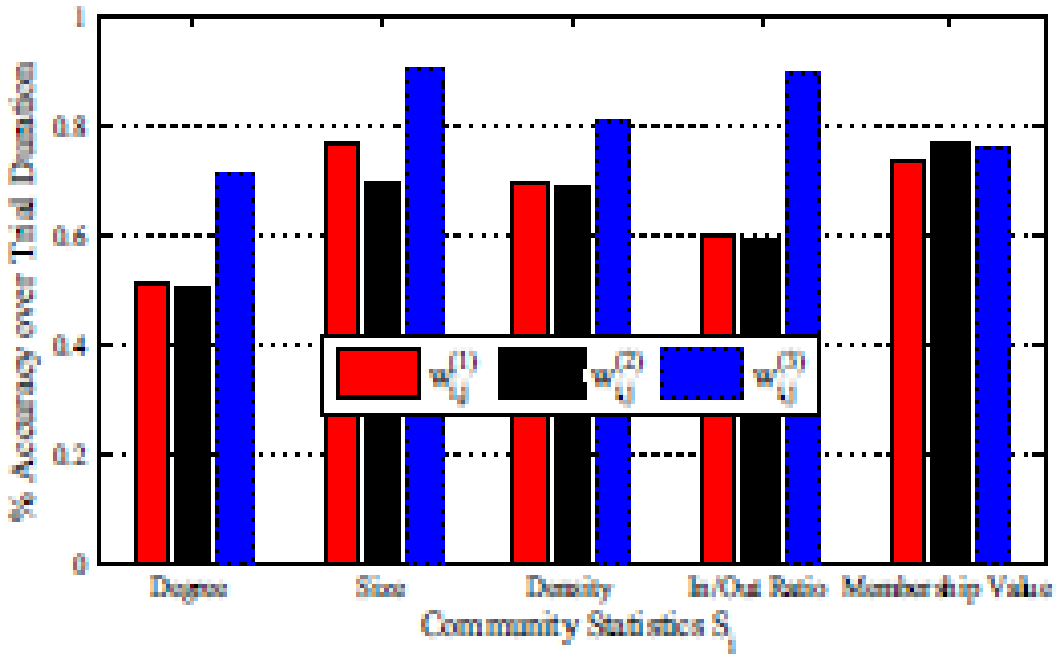


(a) Communities inferred by using internal data.



(b) Reconstructed communities by the adversary.

Adversary has a good success in detecting communities



(b) Community statistics accuracy.

Mobile Devices are PET-friendly

- PETs have significant computational needs
- PETs can benefit from platform security
- Examples of PETs

(PETs = Privacy Enhancing Technologies)

Mobile Devices are PET-friendly

- PETs have significant computational needs
  - Mobile devices are no longer anemic

Mobile Devices are PET-friendly

- PETs have significant computational needs
- PETs can benefit from platform security

Platform Security Widely Deployed

Hardware and software mechanisms

Both IMSI and IMEI require physical protection.

Physical protection means that manufacturers shall take necessary and sufficient measures to ensure the programming and mechanical security of the IMEI. The manufacturer shall also ensure th
(where applicable) remains secur

GSM 02.09, 1993

The IMSI is stored securely within the SIM.

The IMEI shall not be changed after the ME's final production process. It shall resist tampering, i.e. manipulation and change, by any means (e.g. physical, electrical and software).

NOTE: This requirement is valid for new GSM Phase 2 and Release 96, 97, 98 and 99 MEs type approved after 1st June 2002.

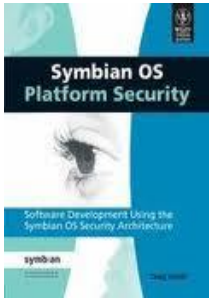
~2001

TrustZone®
Security Foundation by ARM®

~2002

J2ME

~2005



Different starting points:
widespread use of hardware
and software platform security

On-board credentials: what and why

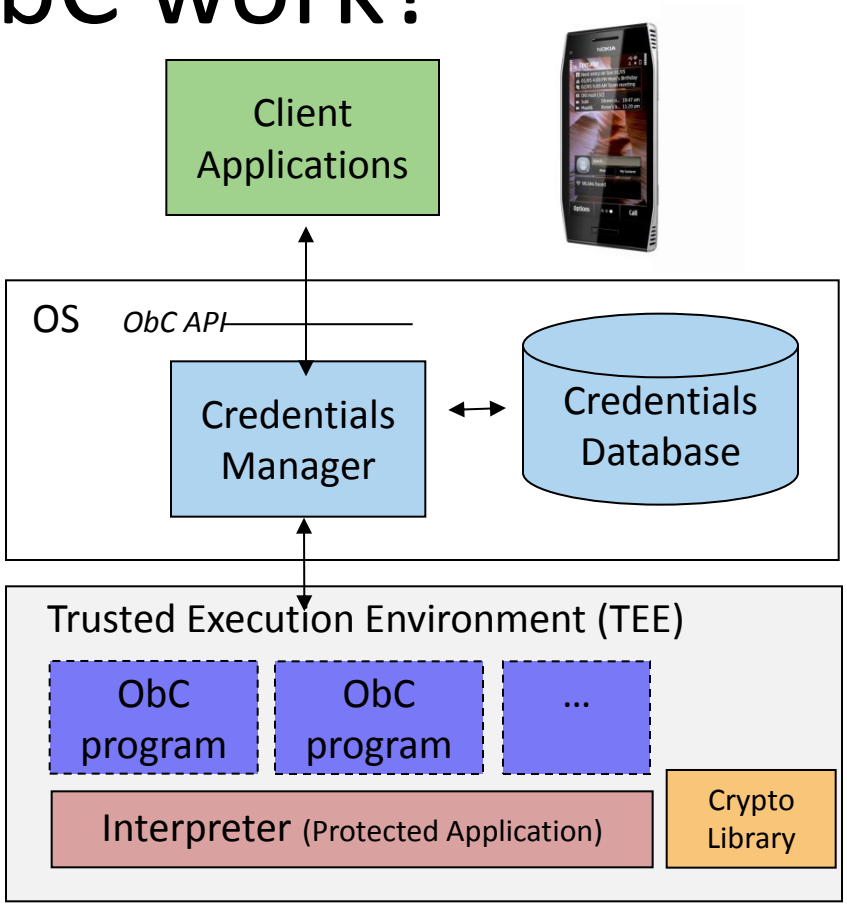
n open

A credential platform that leverages on-board trusted execution environments

Secure yet inexpensive

How does ObC work?

- Apps use ObC API to provision/enroll/use from TEE
- New types of credential logic as ObC programs (bytecode)
- Credentials protected by TEE: never disclosed outside it
- Processing done inside TEE



Paper: ACM ASIACCS '09

Mobile Devices are PET-friendly

- PETs have significant computational needs
- PETs can benefit from platform security
- Examples of PETs
  - Usage Control

Usage Control

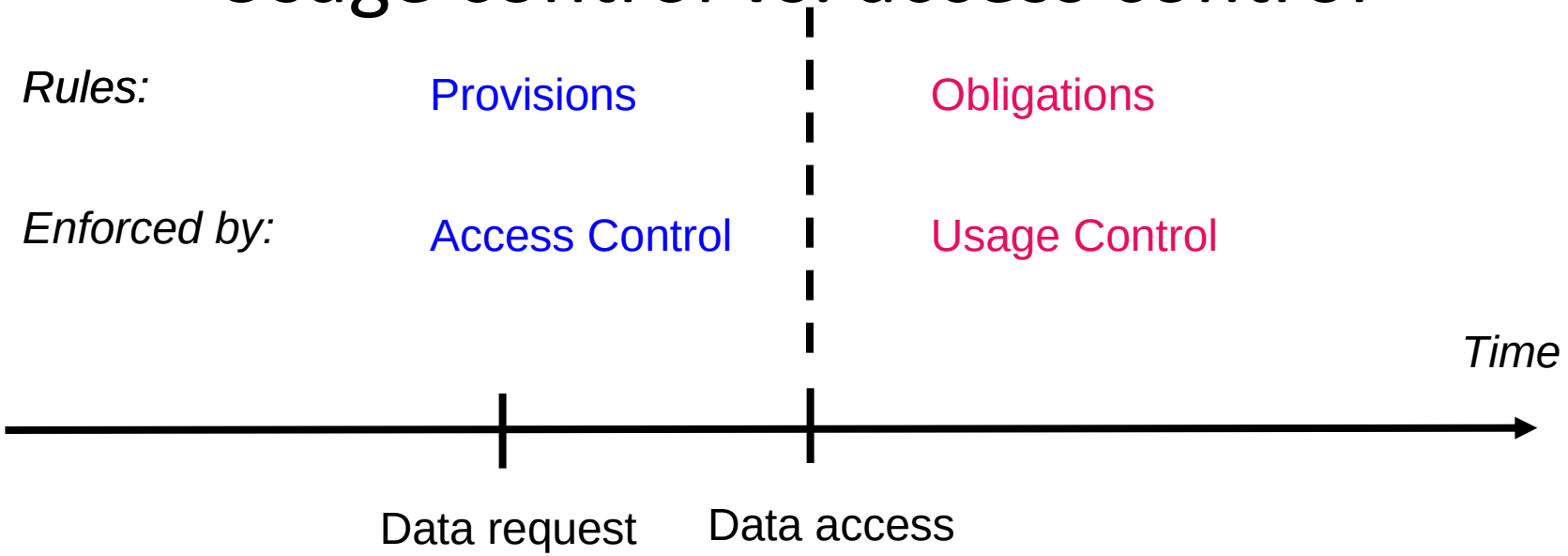
[Basin, Klaedtke, Müller, Pfitzmann,'08]

[Biswas, Nefedov, Niemi, '11]

Usage control

- Control usage of personal data
- Not only access to it

Usage control vs. access control



- Examples of obligations:
  - Credit card details to be deleted after payment received
  - Data about product purchase to be kept until warranty period is over
  - Picture may be viewed maximum five times

Usage control mechanisms

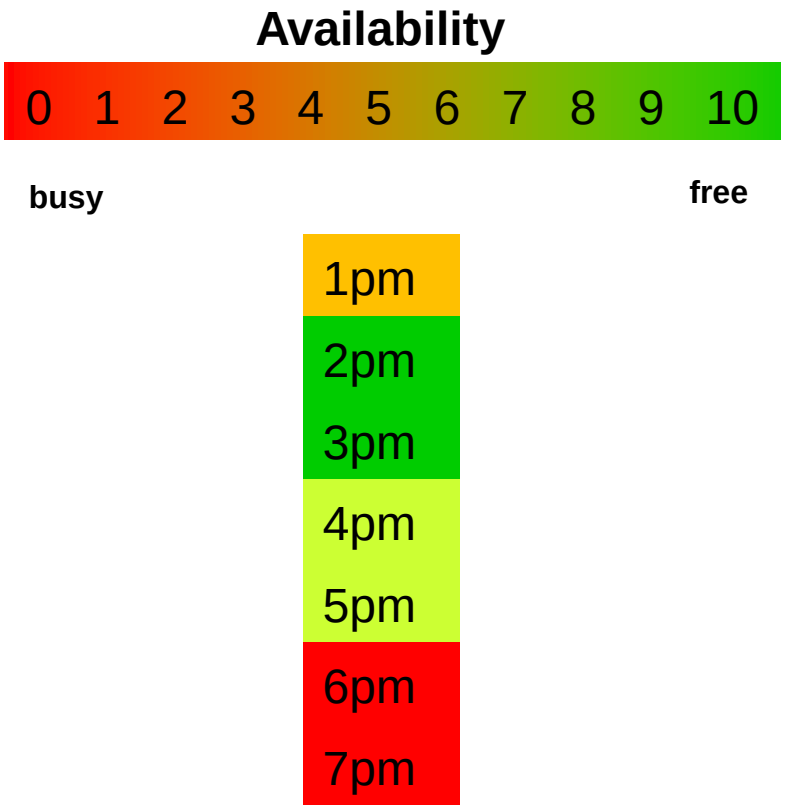
- Monitoring
  - Only reactive enforcement possible: e.g. penalties if violations detected
  - Relies on monitors working correctly (e.g. trusted computing)
- Run-time control
  - Prevents violations
  - Example: Digital Rights Management (DRM)
- Policy definition
  - Non-trivial task
  - Could be semi-automated with “template” approach

Experiences with NIC Trial

- Usage Control monitor runs (at least) faster than new data is created
- Policy set is not complicated nor large in such research data collection activities
  - Fairly easily constructed manually
- However, violations were found in both cases
  - Typical reasons: testing/debugging after new feature introduced, “teething problems” with new features, also some bugs were found,…
- Monitor development continues in a Nokia Open Source project (together with ETH Zürich)

Privacy concerns in meeting scheduling

- Scheduling algorithms using non-binary availability information can provide more optimal results
- But availability info. may lead to compromise of user privacy
- Including location info. makes this tradeoff even more extreme

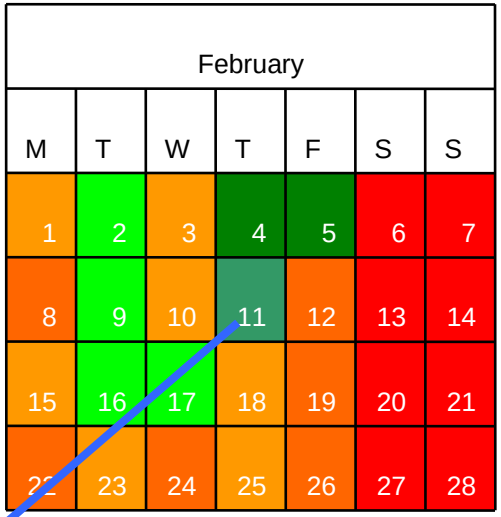


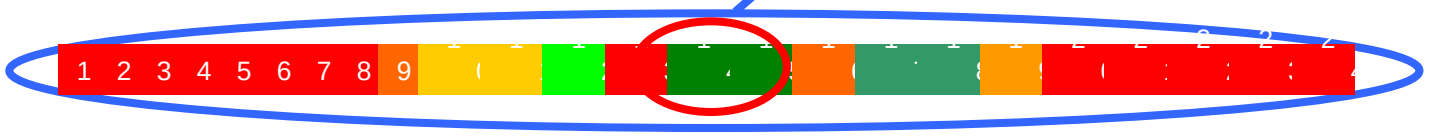
Privacy-preserving scheduling

[Bilogrevic, Jadliwala, Hubaux, Aad, Niemi, ’11]

Example usage scenario – Semi-automatic negotiation

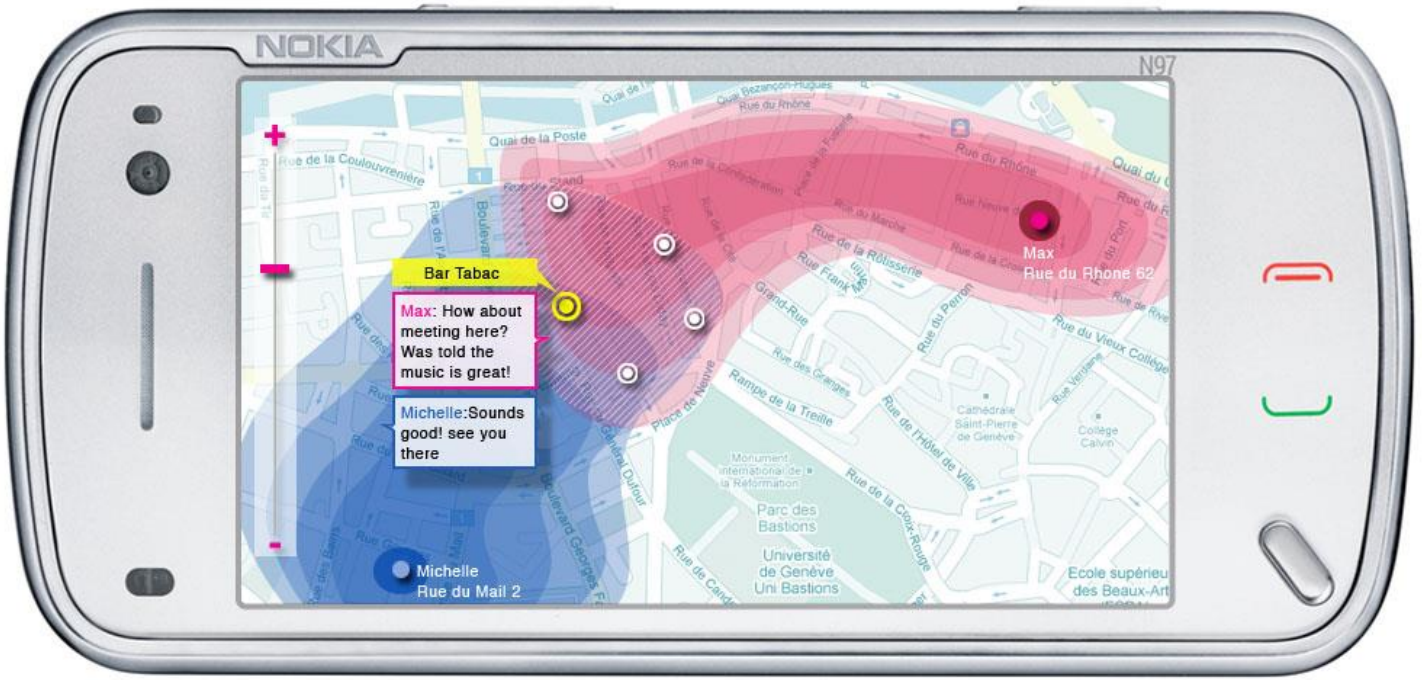
- For scheduled events:
  - The *semi-honest/semi-trusted* server suggests some suitable time slots
    - Cumulative cost function visualization
  - The user makes the final choice manually
- The system is adaptive and predictive -> user acceptance
  - The system adaptively learns about user's preferences, habits, and behaviors
  - The user can shift the coordination responsibility of selected events to the automatic negotiation system at anytime





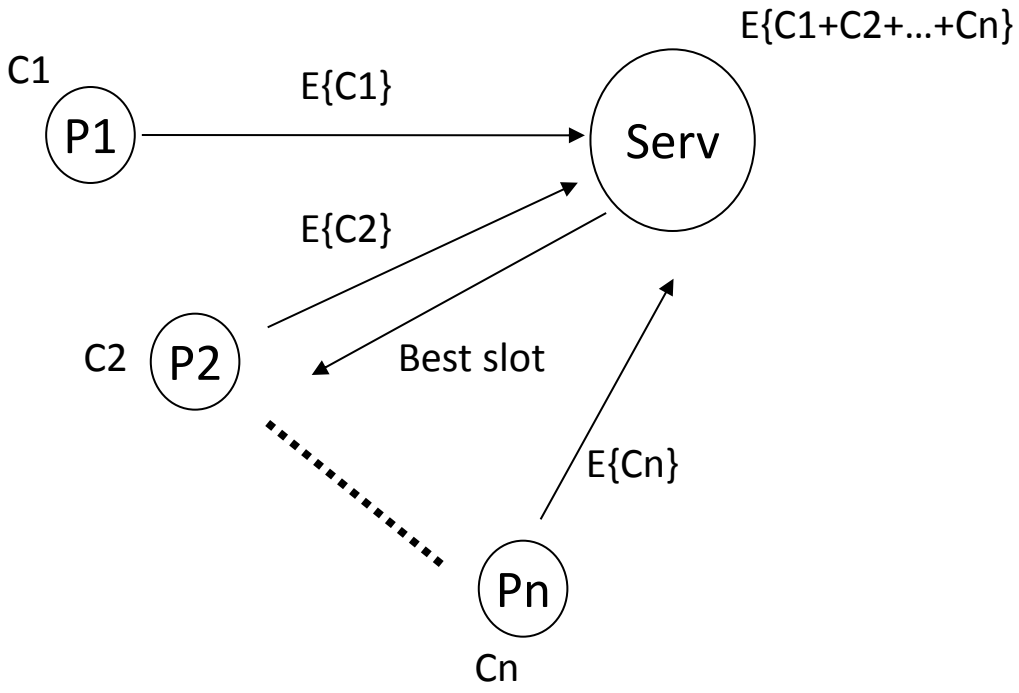
Time-Space optimizer

- The events will be optimized not only for time, but for location as well
- Avoiding meetings in different locations to be scheduled without adequate time in between



Multiparty computations

Make use of homomorphic function for privacy-preserving scheduling optimizations



Mobile Devices are PET-friendly

- PETs have significant computational needs
- PETs can benefit from platform security
- Examples of PETs
  - Privacy-preserving scheduling
  - Usage Control
  - Anonymous Credentials,
  - Pseudonym systems
  - Direct Anonymous Attestation

Mobility/Portability Helps Privacy

- Wireless links are inherently “anonymous”
- Mobile devices are personal and trusted
- Cues from context/history can help

Mobility/Portability Helps Privacy

- Wireless links are inherently “anonymous”
  - Caveat: vulnerable to radio fingerprinting
  - Link-layer address change policy depending on presence/behavior of neighbors
    - Similar strategy for granularity of location reporting
    - “Privacy-Triggered Communications in Pervasive Social Networks” AOC 2011

Privacy-Triggered Networking

[Jadliwala, Freudiger, Aad, Hubaux, Niemi, ’11]

Privacy-Triggered Networking

One wants to communicate (broadcast a message) without being

exposed -> “Hiding in the crowd”

- Micro-blogging
- Personal safety
- Local area social networking
- Dating
- Politically incorrect jokes...

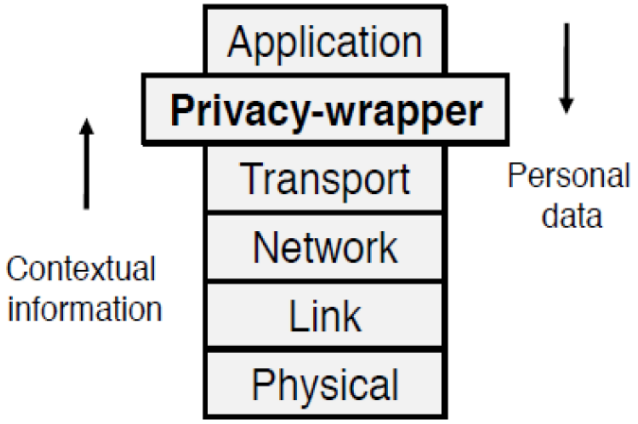
We developed tools and techniques to:

- Visualize privacy level
- Control communication based on it
- Check how the network performs if privacy-triggered

Privacy-Triggered Networking

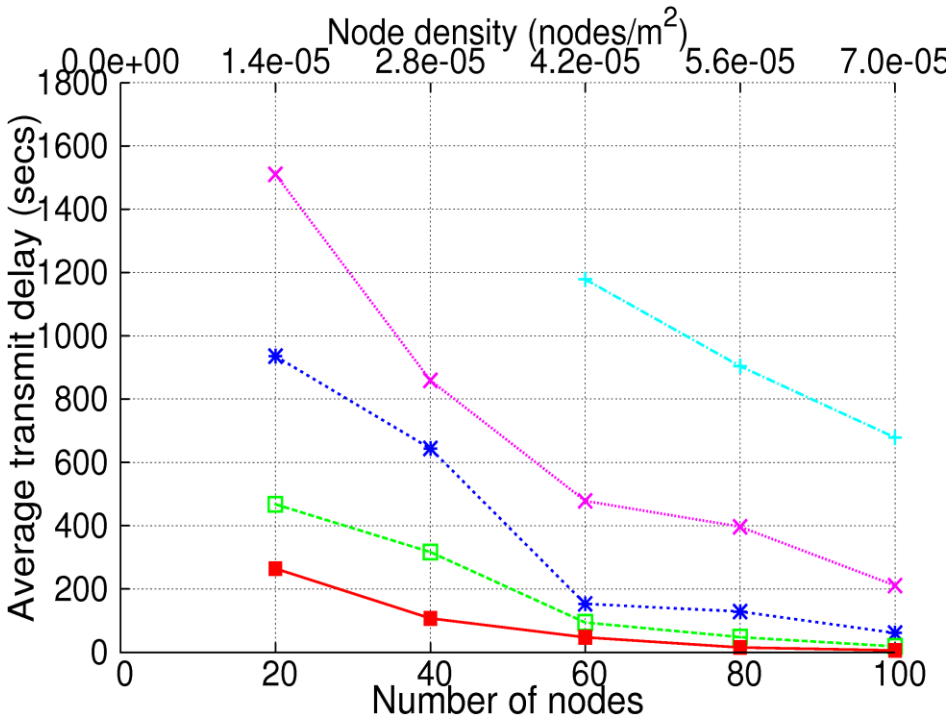
Anonymizing the transmitted information may not be enough!

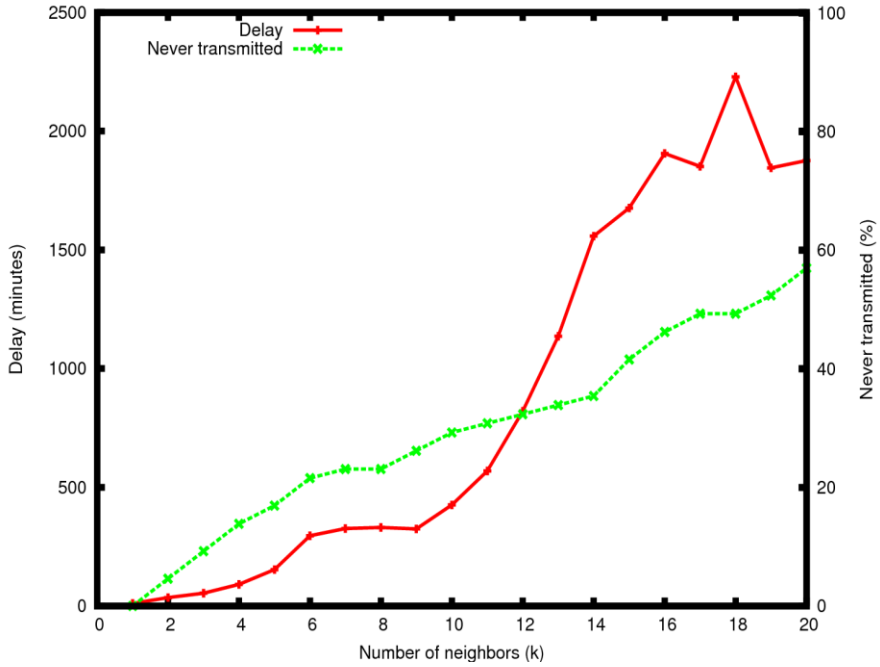
- Mobile device computes the “contextual” privacy level
- User sets a “triggering” threshold for privacy-sensitive messages
- Let's not limit ourselves to any specific privacy metric



Privacy-Triggered Networking

Any compromises made / price to pay for privacy?



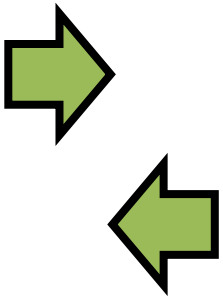


Privacy-Triggered Networking in NIC trial

- Usable in *ClassForum* application:
  - Questions, comments, votes could be sent during a lecture
- Participants could join an *anonymity* pool
  - Messages were forwarded via a random node in the anonymity pool
- Observation: approx. 50 % of the messages were sent via anonymity pool
- Privacy-triggered metric was the size of the anonymity pool
  - Observation: participants did not bother with limiting the size

Mobility/Portability Helps Privacy

- Wireless links are inherently "anonymous"
- Mobile devices are personal and trusted
  - Can host a "privacy broker" that measures privacy exposure over time, and across services



Mobility/Portability Helps Privacy

- Wireless links are more “anonymous”
- Mobile devices are personal and trusted
- Cues from context/history can help
  - in setting policies for sharing, access control…

Context-aware policy management

[Gupta, Miettinen, Asokan, ’11]

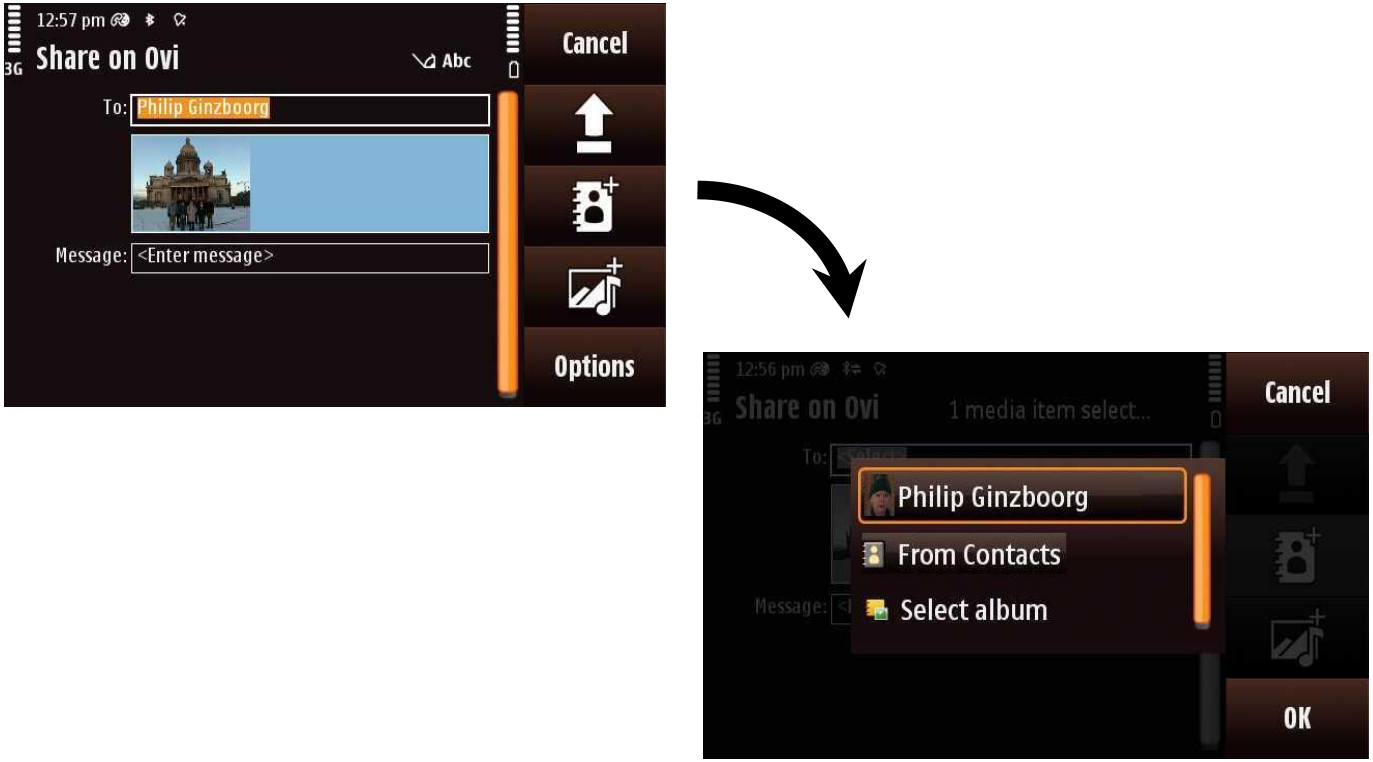
Clues in Data, Metadata, Context

Record nearby devices while taking a

Nearby Bluetooth Devices: 1
Philip N97m

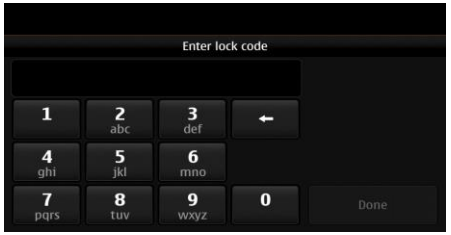
People
andreas
Elena
Paul
Pekka
Philip
Sampo
Zheng

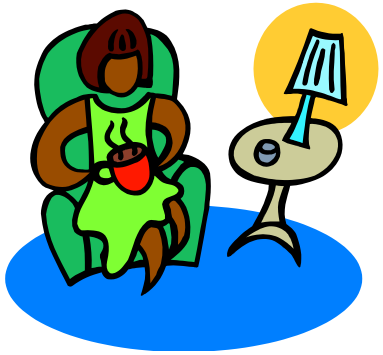
Infer Likely Sharing Targets



Choosing Depth of Device Locking

What local authentication method to use when?





Home

Work Cafeteria

Unknown

Other research topics

- Disruption-tolerant networking (together with Aalto university)
  - No guaranteed delivery → implications on security, esp. Integrity
  - Fragmentation of messages ; optimization
- 3GPP security
  - Machine-type communications
  - Device-to-device communications
- Applying (fully) homomorphic encryption
  - Cloud services
  - ”classical” services; e.g. voting

Summary

- A mobility vs. privacy data sets collected in NRC/Lausanne & EPFL
  - 80-person 3-month ad hoc network data set
- Several privacy protection mechanisms developed and tested
  - Nokia Instant Community pseudonym change algorithm
    - Not effective against attacker with large coverage
  - Privacy-triggered networking
    - Message is hold until hidden in an anonymity pool
  - Privacy-preserving scheduling
    - Efficient automated system with the help of semi-trusted server
  - Usage control
    - Provides means to protect privacy even after data transferred out of reach
  - Intuitive Context-based Policy Management
    - Reduces the burden of the user while protecting her better