Unterstützung der Privatsphäre in mobiler IP-basierter Kommunikation

Christian Hauser
Institut für Kommunikationsnetze und Rechnersysteme
Universität Stuttgart
hauser@ikr.uni-stuttgart.de

08.06.2005
Outline

Motivation
Threat Analysis
A New Approach
Conclusions and Future Work
Collection of Context Data

- Ubiquitous use of platform → many different applications

- Detailed traces of context data combined by context management
  - Real World (e.g., location)
  - Applications

→ Privacy Risk!
Context-Use Without Protection

- Everybody could access all context
- Two dimensional problem
  - Rich trace
  - Long trace
Privacy Protection Approach (1)

Split of knowledge by restricting access to information

- Friend finder
- Diabetes service
- Webmiles
- Navigation service

Friends of X
Blood Sugar Level of X
Websites of X
Location of X
Privacy Protection Approach (2)
Privacy Protection Approach (3)
Privacy Protection Approach (4)

Split of knowledge by restricting access to information and use of different identities (to prevent collaboration)

- Friend finder
- Diabetes service
- Webmiles
- Navigation service

Friends of X
Blood Sugar Level of Y
Websites of Z
Location of A

www.uni-stuttgart.de
www.drug-assist.org
www.amazon.com
www.cancer-help.org
www.walmart.de
www.map-brussels.com
www.new-jobs.com
Privacy Protection Approach (5)
Privacy Protection Approach (6)

Split of knowledge by restricting access to information and use of different identities (to prevent collaboration)

- Friend finder
- Diabetes service
- Webmiles
- Navigation service

Split of knowledge by changing identities over time (and providers)
Example and Focusing

- **Privacy approach**
  - use of multiple (virtual) identities, VIDs
  - tune amount of disclosed data in context of each identity separately
Example and Focusing

- **Privacy approach**
  - use of multiple (virtual) identities, VIDs
  - tune amount of disclosed data in context of each identity separately

- **Pitfall: Augmentation of a VID**
  Two possibilities: **Linking** of several VIDs
Example and Focussing

- **Privacy approach**
  - use of multiple (virtual) identities, VIDs
  - tune amount of disclosed data in context of each identity separately

- **Pitfall: Augmentation of a VID**
  Two possibilities: Linking of several VIDs and inference of data
Example and Focusing

- **Privacy approach**
  - use of multiple (virtual) identities, VIDs
  - tune amount of disclosed data in context of each identity separately

- **Pitfall: Augmentation of a VID**
  Two possibilities: Linking of several VIDs and inference of data
  - application data
  - data of communication system

- **Focus on IP based communication system**
Protection Goals

- **Unlinkability of VIDs**
  
  trace cannot be enriched by information of several VIDs

- **Limitation of trace**
  
  short trace alleviate inference danger

Violation of both: More knowledge at the attacker than user wants

--> against right on informational self-determination

Potential Attackers

- **Communication partners**
  
  other (private) users or service providers

- **Providers of the communication systems**
  - can be forced to disclose information (legal interception)
  - can be hacked
  - may be not trustworthy (according to "Internet Model" everybody can be provider, i.e., provide a Home Agent)
Threat Analysis
Packet based communication: Two basic pieces of information

- **identifier**: indicates which device is addressed
  - may be chosen *arbitrarily* (thus without containing any sensitive information)
  - known to communication system and communication partner
- **locator**: indicates where packet must be delivered to
  - inherently contains location in terms of network topology which can be mapped to (sensitive) geographical location in IP
  - must be known to communication system
  - does not have to be known to communication partners

**Comparison: Classical IP**
both pieces of information collapse into the IP address

**Comparison: Mobile IP**
- home address is a kind of identifier
- care-of address is a kind of locator
- (but: home address is locator to user’s home and care-of address is known to communication partners in case of route optimization)
Abstraction of the linking problem

• (Many) VID contexts of the user are inherently merged
  - behind all VIDs is only one user
  ⇨ everything that leads to the (real) user is dangerous wrt. link of VIDs
    (and often regarding privacy in general)

• Real-world attributes, reflected in the system
  - location, location changes (movement), network connection, ...
  - global use patterns
    • sleeping times, working times, ...
  ⇨ attributes, which are identical for all VIDs of same user
  ⇨ danger rises with decrease of number of users having the attribute

• Contrast
  - communication sessions not dangerous wrt. to link
    ⇨ can be different for each VID
    ⇨ rather similar for VIDs of different users (e.g., when using same service)
Concretion of linking problem to communication

• **Real-world user behaviour reflected in locator, reflecting**
  - location, movement, network connection
  - (vertical handover models, ...)

• **Remarks**
  - there exist more unique attributes (e.g., one identifier/locator/interface per user)
  - could be solved by technical systems – the real-world things can’t

**Inference**

**Question: Where is sensitive information contained?**

1. In identifier: Home of user (usually)
2. In locator: Location, network connection
3. In locator changes: Movement behaviour
## Threat Analysis

<table>
<thead>
<tr>
<th>Linking of VIDs</th>
<th>Threats in fixed scenario</th>
<th>Additional threats in mobile scenario</th>
</tr>
</thead>
</table>
| **LinkF**: Identical data in context of VIDs  
*Example*: Identical identifier, identical locator | **LinkM(1)**: Identical behavior of VIDs observed by *identical* patterns of data or events  
*Example*: Change from identical old locator to identical new locator | |
| **LinkM(2)**: Identical behavior of VIDs observed by *similar* patterns of data or events  
*Example*: Simultaneous locator changes with unknown locators | |

<table>
<thead>
<tr>
<th>Inference of personal information</th>
<th></th>
<th></th>
</tr>
</thead>
</table>
| **InfFI**: Inference from the identifier  
*Example*: home of VID | No additional inference from the identifier | |
| **InfFL**: Inference from a *single* locator  
*Example*: Location of the user at communication time | **InfML(1)**: Inference from *several* locators  
*Example*: Location trace of a user over a period of time | |
| **InfML(2)**: Inference from user behavior by locator changes  
*Example*: Inference of activity by rate of locator changes | |

*Institute of Communication Networks and Computer Engineering  
University of Stuttgart*
A New Approach
• Different networks supposed to be operated by different parties
• Separate contexts for VIDs throughout packet’s path
• Two agents in a row: no entity knows both, identifier and locator
• Locator invisibly stored when not needed
• Home Agents HA2-x are changed frequently
• Identifiers not from home netw. but from different, arbitrary networks
  Each of those networks operates a "Home Agent"
• User can configure trade-off between performance and privacy
Conclusions and Future Work

- **Future context-aware systems need suitable privacy protection**
  - approach of multiple VIDs very promising
  - support by communication system necessary
  - new threat implied: Linking of VIDs

- **Threat analysis regarding communication system**
  - mobility adds significantly to threat
  - solution must be especially designed for multiple identities and mobility

- **Existing proposals not well prepared**

- **New approach**
  - solves or at least alleviates all identified problems
  - user in control of trade-off: costs vs. privacy

- **Future work**
  - realization of proof-of-concept
  - quantification of protection vs. costs
    - evaluation of sensible configurations