

Location Privacy

Dimitar Yordanov

Betreuer: Mirco Stern

Sicherheit und technischer Datenschutz in
Informationssystemen

Seminar im Sommersemester 2006

Motivation

- Genaues Standortbewusstsein ist heute sehr wichtig
 - Beispiel: (Navigation. Man will bestimmte Dienste in seiner Nähe in Anspruch nehmen)
- Eine Vielzahl von Systemen ermöglichen die Lokalisierung
 - Beispiel: GPS, Mobilfunknetz, Sensornetze
- Problem: Privatheit der Nutzer, die die Dienste Benutzen

Motivation (2)

- Aus der Positionsinformation sind andere wichtige Informationen ableitbar
 - Wo sich eine Person befindet steht im engen Zusammenhang damit, was sie gerade macht
- Location Privacy und Positionsbewusstsein stehen im Konflikt zu einander
 - Damit bestimmte Dienste benutzt werden können, muss man etwas über seine Position verraten

Überblick

- ✓ Motivation
- Definition
- Positionierung vs. Tracking
- Anonyme Herausgabe privater Daten
- Durch den Nutzer bestimmte Herausgabe
- Ein Gesamtsystem
- Zusammenfassung

Was ist „Location Privacy“?

- **Location Privacy** ist eine spezielle Art des Datenschutzes, die definiert wird als die Fähigkeit, andere nicht autorisierte Parteien an der Bestimmung der gegenwärtigen oder früheren Position einer Person zu hindern (R.Beresford und F.Stajano)
- **Location Based Services (LBS)** sind solche Dienste, bei denen das Wissen über die Position einer Person oder eines Objekts, benutzt wird um den Dienst zu personalisieren (E.Snekkenes)

Positionierung

- „Positionierung“ - der Nutzer kann seine Position selber bestimmen (GPS-Navigationssysteme)
 - Der Benutzer bestimmt selber, was mit den Daten passiert
 - Gleiche Vorgehensweise wie bei anderen sensitiven Daten möglich

Tracking

- „Tracking“ - die Position wird durch die Infrastruktur berechnet
 - Keine Kontrolle über die Sammlung von Positionsdaten
 - Zurückhalten privater Daten nicht möglich
 - Voraussetzung für die Erstellung von Privatheit:
 - Das Vertrauen von dem „Location System“
- Wir bleiben beim Tracking

Überblick

- ✓ Motivation
- ✓ Definition
- ✓ Positionierung vs. Tracking
- Anonyme Herausgabe privater Daten
- Durch den Nutzer bestimmte Herausgabe
- Ein Gesamtsystem
- Zusammenfassung

Anonymisieren

- Anonyme Herausgabe privater Daten
 - Bei vielen Szenarien, kann durch Anonymisieren Privatheit erreicht werden
 - Beispiel: Wenn ich an einem Geschäft vorbeilaufe sag mir den Preis von ...
 - **Anonymität** ermöglicht einem nicht identifiziert werden zu können in einer Menge von Objekten
 - **Pseudonyme Identität** ist eine, die die wirkliche Identität ersetzt

Anonymisieren (2)

- Problem: Oft nicht möglich, Anonymität zu garantieren
 - Aus Positionsinformation lässt sich die Identität herleiten
 - Beispiel: Die Person, die sich in einem Büro befindet, ist mit großer Wahrscheinlichkeit die, die dort arbeitet

Anonymisieren von Positionsinformation

- Marco Gruteser und Dirk Grunwald:
 - K-Anonymität
 - Anonymity Set - ist die Menge von Objekten, die die gleiche Identität haben können
 - Ein Objekt ist K-Anonym, wenn seine Positionsinformation von mindestens k-1 anderen Objekten nicht zu unterscheiden ist
 - Zwei Möglichkeiten Anonymität herzustellen:
 - Räumliche Lokalität
 - Zeitliche Lokalität

Adaptive-Interval Cloacking Algorithmus

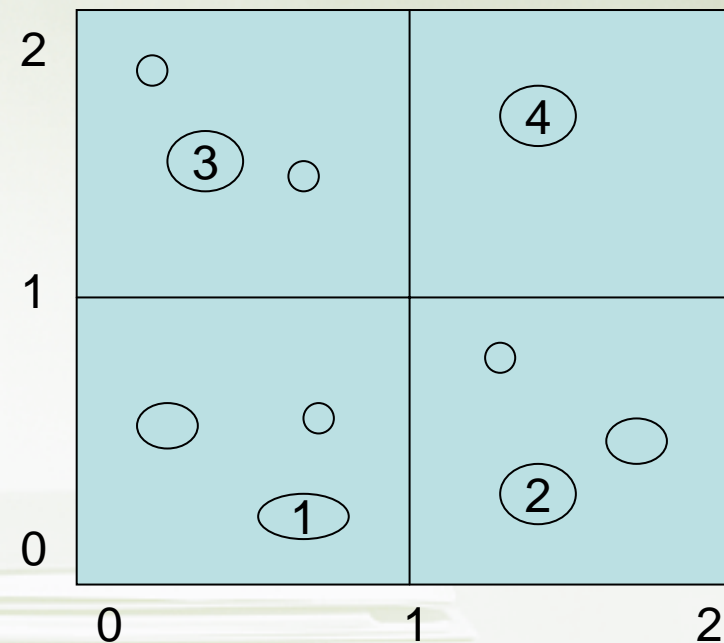
- Veränderung der räumlichen Genauigkeit
 - Eingaben:
 - K-min - spezifiziert das nötige Anonymitätsniveau
 - Die Position des Nutzers
 - Die Koordinaten des Bereiches, der durch den Anonymitätsserver bedient wird
 - Die Position von allen anderen Objekten in dem Gebiet
 - Der Algorithmus unterteilt den Bereich um die Position des Objekts bis die Anzahl der Objekte unter k-min sinkt, und gibt das zuvor berechnete Gebiet als Ergebnis zurück

Adaptive-Interval Cloacking Algorithmus(2)

- Erweiterung: Veränderung der zeitlichen Lokalität
 - Genauere Angaben über die räumliche Position, indem die Genauigkeit der zeitlichen verringert wird
 - Die Idee: Verzögere die Anfrage bis k-min Objekte das Gebiet besucht haben
 - Zusätzlicher Parameter - eine minimale Größe des Gebiets
 - Wenn die Größe erreicht wird, wird versucht, abzuwarten, bis sich genügend Objekte in dem Bereich aufgehalten haben

Mögliche Angriffe beim Modell

- Tupel, die sich in Zeit und Raum überschneiden ($k\text{-min} = 3$)
 1. $([0,1], [0,1], [t_1, t_2])$
 2. $([1,2], [0,1], [t_1, t_2])$
 3. $([0,1], [1,2], [t_1, t_2])$
 4. $([0,2], [0,2], [t_1, t_2])$
- Ungenaue Information könnte für manche Dienste unzureichend sein
 - Beispiel: Navigieren, Taxi rufen



Identitätsaufdeckung bei dauerhaften Pseudonymen

- Betrachtung von Pseudonymen
- Das gleiche Pseudonym zu benutzen ist problematisch
- Dienstanbieter könnten ihre gemeinsame Kenntnisse ausnutzen
 - Beispiel: Man will den Preis des Kaffees von jedem Geschäft an dem man vorbeiläuft wissen
- Lösung: Zu jedem Anbieter ein anderes Pseudonym geben

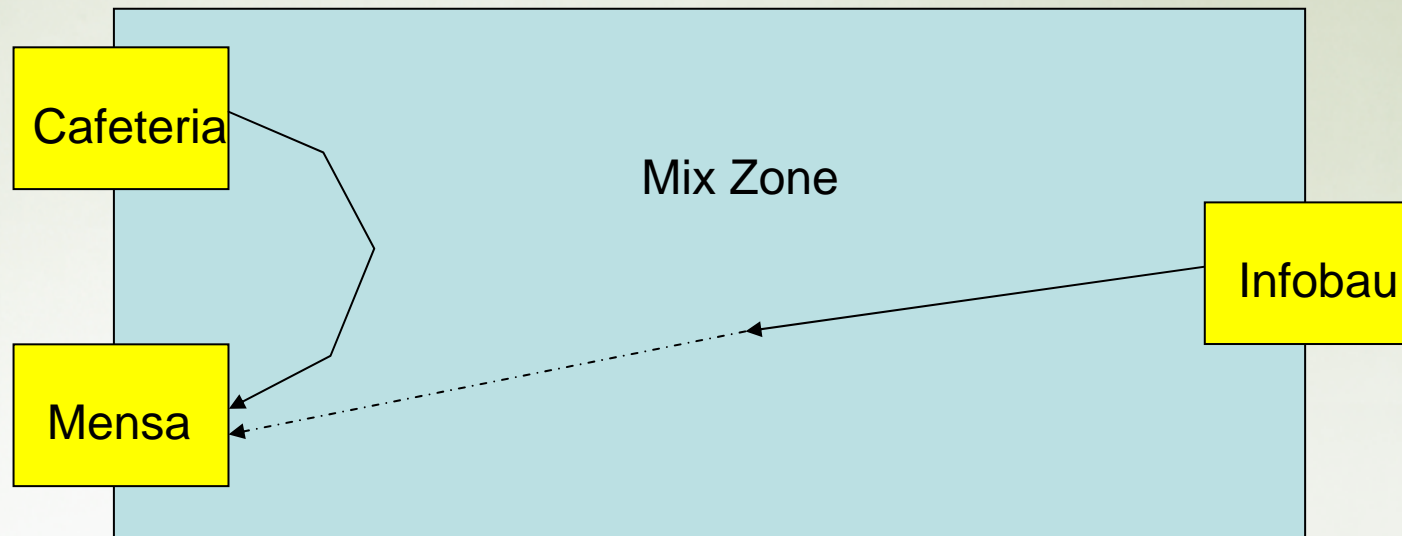
Dynamische Pseudonyme

- Problem: Die Identität kann aufgedeckt werden
 - Wenn der gleiche Dienst über längere Zeit benutzt wird
- Die Beantwortung einer einfachen Frage genügt
 - „Wo verbringt ein Pseudonym die meiste Zeit?“
 - Oder „Wer verbringt an einer bestimmten Position am meisten Zeit?“
- Die Lösung: Ständige Änderung des Pseudonyms
 - Herausforderung: Angreifer darf den Wechsel nicht mitbekommen

Mixed Zones und Application Zones (Beresford)

- „Mix Zone“ wird für eine Gruppe von Personen definiert als ein räumliches Gebiet mit maximaler Größe, in dem kein Nutzer sich für einen Dienst angemeldet hat
- „Application Zone“ ist ein Bereich, in dem ein Nutzer sich für einen Dienst registriert hat.
- Die Dienstgeber bekommen keine Positionsinformation von Objekten, die sich in einer „Mix Zone“ befinden
 - In diesem Bereich wird die Identität gewechselt

Mixed Zones und Application Zones (2)



- Problem: Benutzer werden nicht richtig gemischt
 - Wenn es zu wenige Benutzer gibt
 - Wenn es selten benutzte Dienste gibt
 - Angreifer könnten zusätzliches Wissen ausnutzen
- Dynamische Pseudonyme könnten existierende Anwendungen am Funktionieren hindern

Wie kann man die Anonymität messen?

- Den Grad der Anonymität messen können
 - Um spezifizieren zu können, wie anonym man sein möchte
 - Um sicherzustellen, dass eine Person anonym bleibt
- „Anonymity Set“ könnte als Messwert dienen
- Problem: „Anonymity Set“ reicht an sich nicht aus
 - Die Wahrscheinlichkeit eine Identität zu haben ist nicht gleich verteilt \Rightarrow Die Identität einer Person kann bestimmt werden
- Einsetzen von Entropie um Anonymität zu messen
 - Die Messungen (von Beresford) zeigen pessimistische Resultate

Wann soll anonymisiert werden?

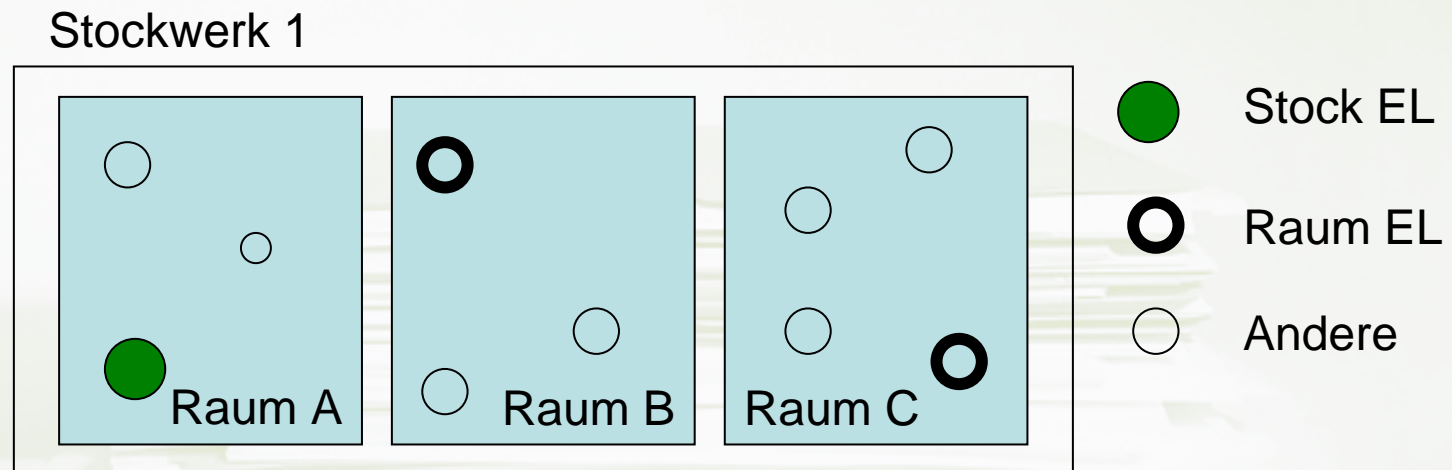
- Ich vertraue meinen Location Server
Probleme:
 - Einmal gespeicherte Daten sind für weitere Angriffe anfällig
 - Die Sicherheit des Systems ist sehr aufwendig
- Je weniger Instanzen die Position kennen desto besser
 - Im folgenden Ansatz entscheidet schon die Infrastruktur, ob Positionsdaten überhaupt erhalten werden

Privatheit sichern noch während der Beschaffung der Daten (Gruteser, Schelle)

- Das Sensornetz „stört“ die Positionsdaten, sodass sie die k-Anonymität erfüllen
- Hierarchische Aggregation
 - Netzwerkknoten organisieren die Beschaffung der Positionsdaten in einer baumartige Struktur
 - Mehrere Knoten (Ebenenleiter) im Baum verhüllen die Daten
 - Die Hierarchie reflektiert die räumlichen Eigenschaften des Ortes

Baumaufbau

- Datenfluss :
 - Knoten \Rightarrow EL (Ebenenleiter) in dem Raum \Rightarrow zum EL auf dem Stockwerk ... \Rightarrow und am Ende in den Location Server



Informationsverhüllung, Sicherheit

- Zwei Möglichkeiten
 - Liefere genaue Anzahl der Objekte, aber ungenaue Position
 - Gib die genaue Position zurück, liefere ungenaue Anzahl der Objekte
- Sicherheitsprobleme
 - Netzwerkverkehr Analysieren
 - Falsche Informationen Senden
 - Abhören

Überblick

- ✓ Motivation
- ✓ Definition
- ✓ Positionierung vs. Tracking
- ✓ Anonyme Herausgabe privater Daten
 - Durch den Nutzer bestimmte Herausgabe
 - Ein Gesamtsystem
 - Zusammenfassung

Durch den Nutzer bestimmte Herausgabe

- Manchmal muss die genaue Position bekannt gegeben werden
 - Beispiel: Man will seine Kollegen wissen lassen, wenn man im Büro ist
- Kontrolle über die Herausgabe ist nötig
- Man braucht also Methoden, um seine Anforderungen zu spezifizieren

Leonhardt's Ansatz

- Zwei mögliche Anfragen im Bezug auf LBS
 - Wo befindet sich eine Person?
 - Wer befindet sich an einer bestimmten Position?
- Untersucht Zugriffskontrollmethoden auf Einsetzbarkeit
 - Matrixbasiert (Lampson)

Matrixbasierte Zugriffskontrolle

- $\langle \text{Subjekt} \rangle \{ \langle \text{Aktion} \rangle \} \langle \text{Ziel} \rangle$
 - Zugriff zu einem Ziel, durch ein Subjekt, mit einer Aktion ist gewährt, wenn die entsprechende Kombination in der Matrix vorhanden ist

Beispiel:

Joe darf Fred sehen

Joe {testForCollocation} Fred

Joe darf sehen, dass Fred im Infobau ist

Joe {testForCollocation(PERSON)} Infobau WHEN
PERSON=Fred

Matrixbasierte Zugriffskontrolle(2)

- $\langle \text{Subjekt} \rangle \{ \text{Aktion} \} \langle \text{Ziel 1} \rangle \dots \langle \text{Ziel n} \rangle$
 - Das Subjekt ist autorisiert Aktionen über die Kombination von Zielen durchzuführen
 - Beispiel von zuvor:
Joe { testFürKollokation } Fred, Infobau
- Der Vollständigkeit halber können mehrere Subjekte definiert werden
 $\langle \text{Subjekt 1} \rangle \dots \langle \text{Subjekt m} \rangle \{ \text{Aktion} \} \langle \text{Ziel 1} \rangle \dots \langle \text{Ziel n} \rangle$
 - Mehrsubjektregelwerke fordern mehrere Subjekte die die Aktionen zusammen durchführen
 - Beispiel: Joe, Fred {sehen} Geld, Karlsruhe

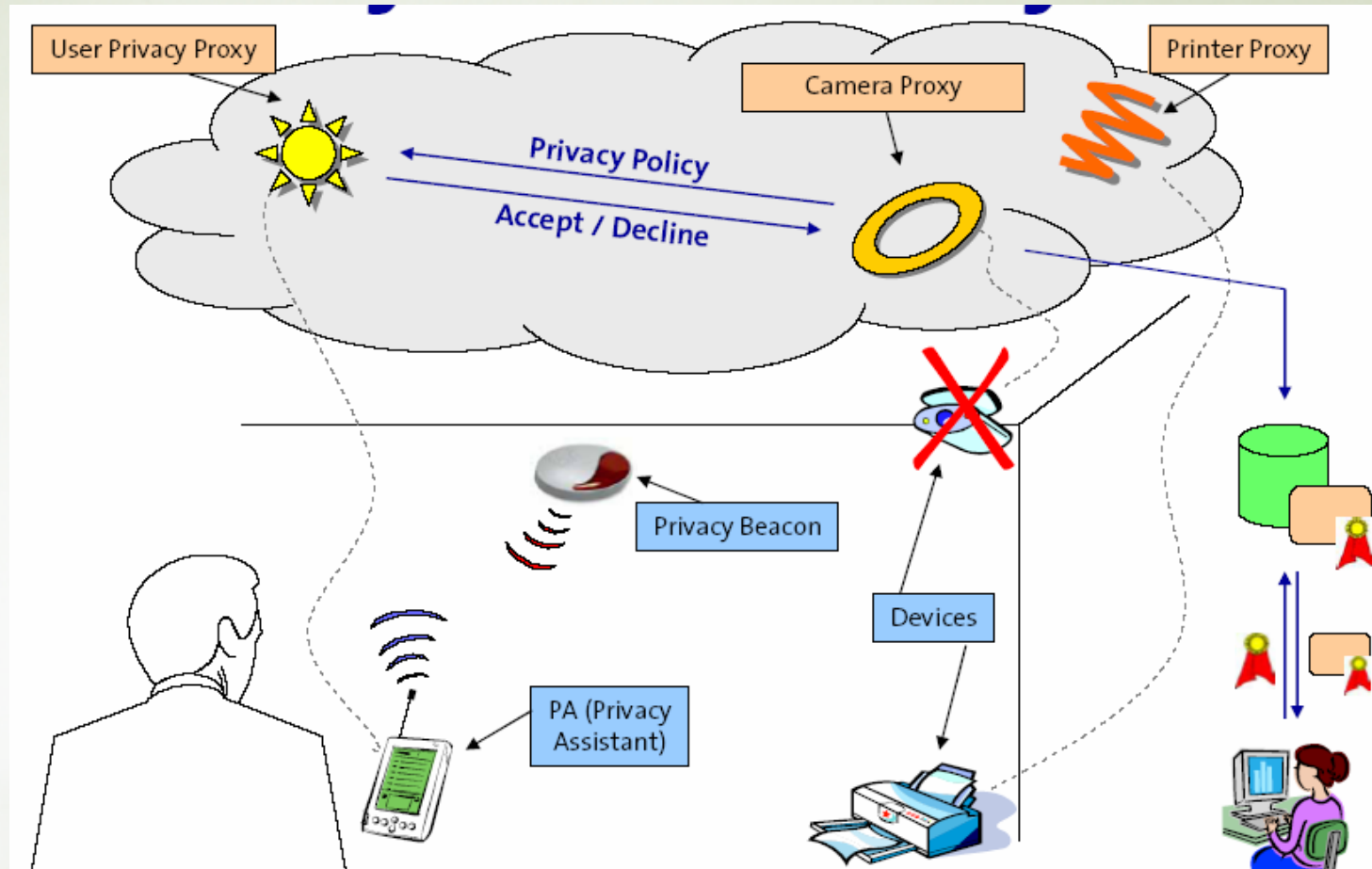
Überblick

- ✓ Motivation
- ✓ Definition
- ✓ Positionierung vs. Tracking
- ✓ Anonyme Herausgabe privater Daten
- ✓ Durch den Nutzer bestimmte Herausgabe
- Ein Gesamtsystem
- Zusammenfassung

Ein Gesamtsystem

- Bisher wurden einzelne Methoden vorgestellt, die Positionsinformation schützen können
- Wie passt das alles zusammen?
- \Rightarrow Rahmenwerk für den Einsatz der Methoden

Langheinrich's Modell



Quelle: M.Langheinrich

Privacy Awareness System (pawS)

- Privacy Proxy -ermöglichen Zugriff auf die Daten des Subjekts
 - Service Proxy - Proxy für einen Dienst
 - Personal Proxy - ermöglicht das Zusammenspiel zwischen Nutzer und dem Service Proxy
- Privacy Assistent - Mobiles Gerät (PDA ...)
- Privacy Policies
- Policies Bekanntgabe
 - Implizit
 - Aktive Bekanntgabe
- Privacy Beacon - Kündigt die Daten an, die für die einzelnen Dienste gesammelt werden würden
- Proxy Abkommen - Welche Vereinbarung wurde getroffen
- Datenzugriff, User Logs

Zusammenfassung

- Schutz der Positionsinformation ist für die Bewahrung unserer Privatheit sehr bedeutsam
- Anonyme oder pseudonyme Identität können in hohem Maße zu unserem Datenschutz bei LBS beitragen
 - Schwer herzustellen
 - Nicht immer einsetzbar
- Wichtig sind Möglichkeiten für den Nutzer, Sammlung, Verwendung, Speicherung und Weitergabe von Positionsinformation zu beeinflussen

**Vielen Dank
für die Aufmerksamkeit!**



Fragen ???



Literatur

- „Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking“ - Marco Gruteser, Dirk Grunwald
- „Location Privacy in Pervasive Computing“ - Alastair R. Beresford, Frank Stajano
- „Privacy-Aware Location Sensor Networks“ - Marco Gruteser et al.
- „A Privacy Awareness System for Ubiquitous Computing Environments“ - Marc Langheinrich
- „Supporting Location-Awareness in Open Distributed Systems“ - Ulf Leonhardt