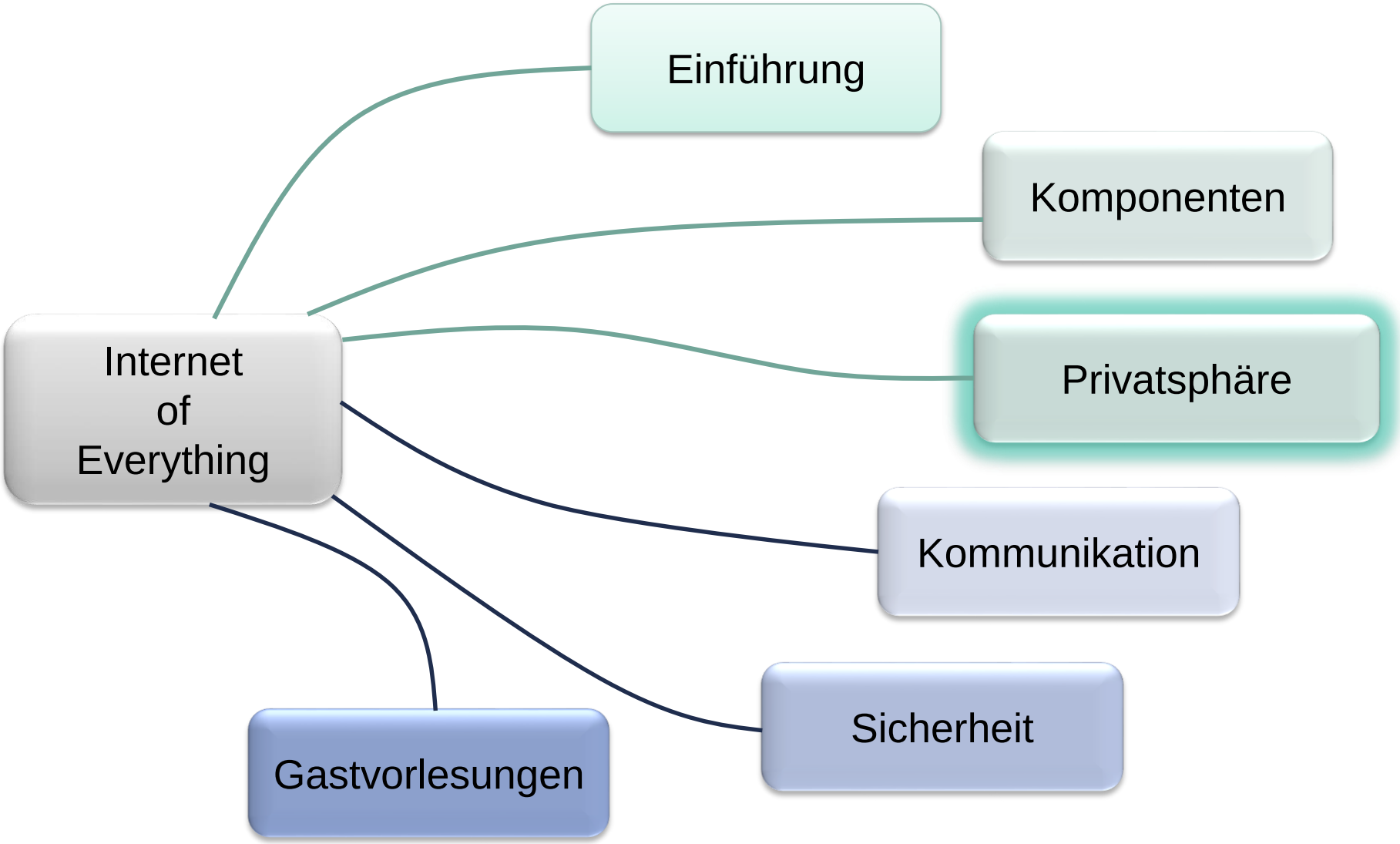
Karlsruher Institut für Technologie

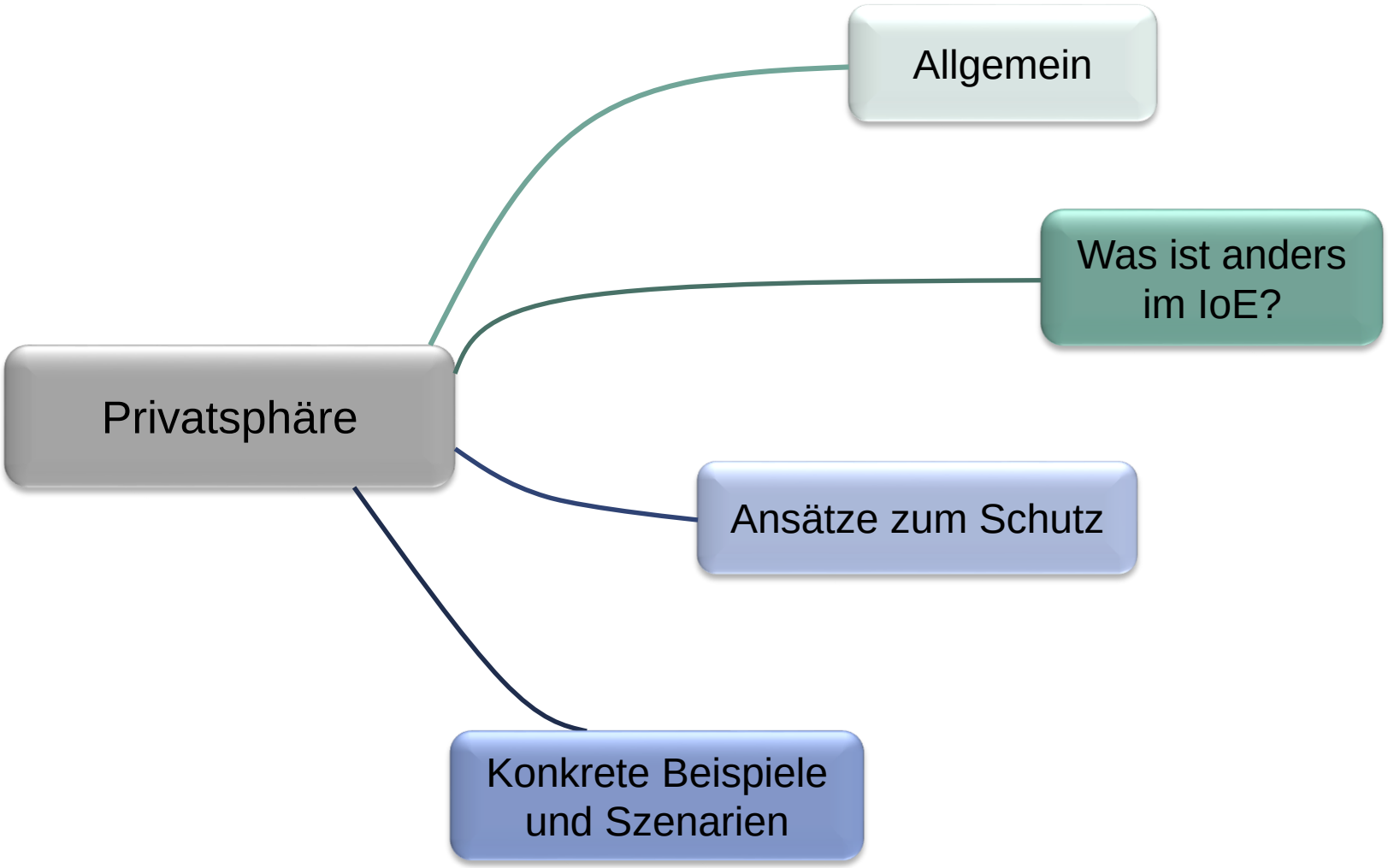
Vorlesung Internet of Everything
Kapitel 3 – Privatsphäre

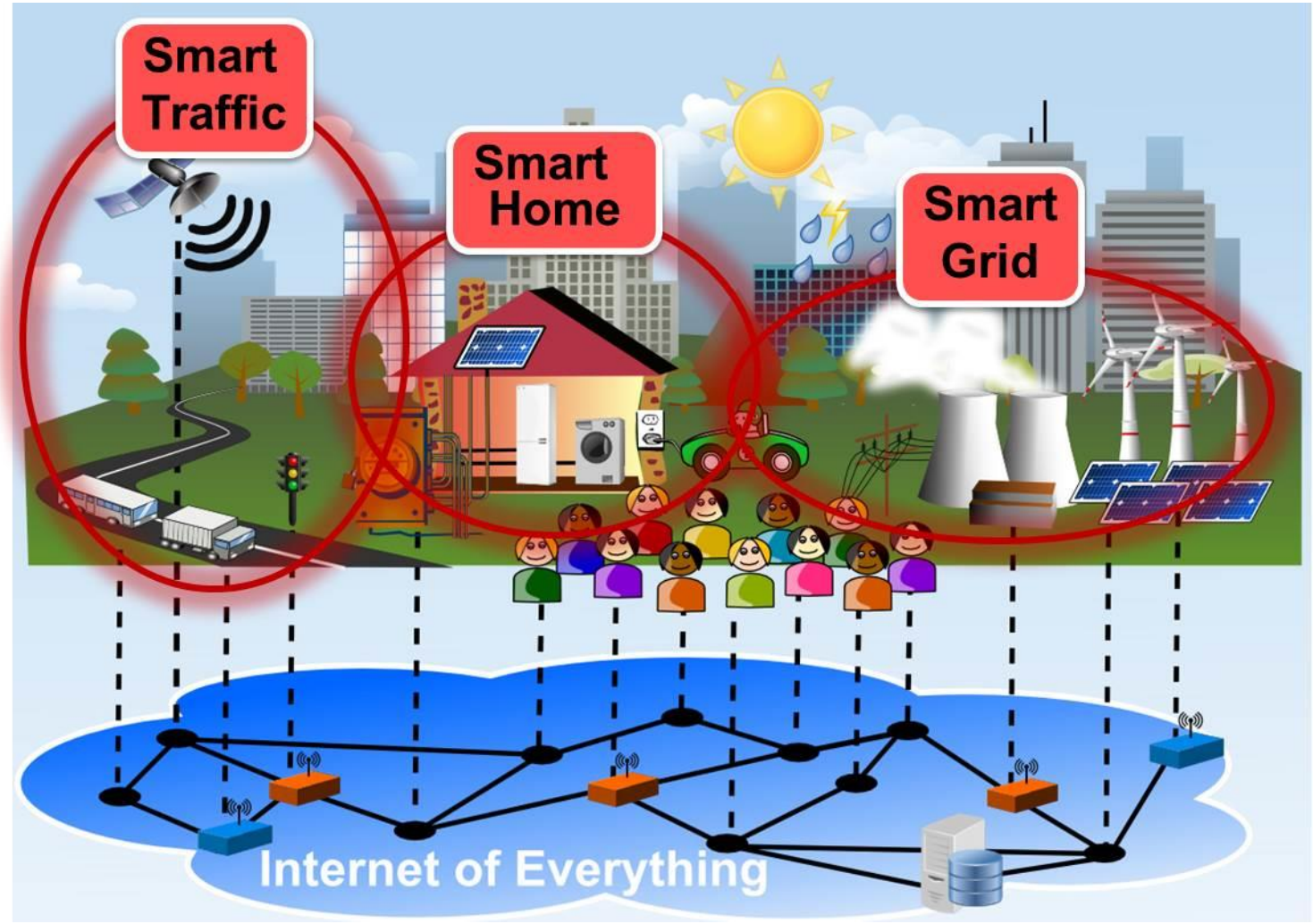
Prof. Dr. Martina Zitterbart, Martin Florian, Markus Jung
[zitterbart, florian, m.jung]@kit.edu

Institut für Telematik, Prof. Zitterbart

Inhalte der Vorlesung



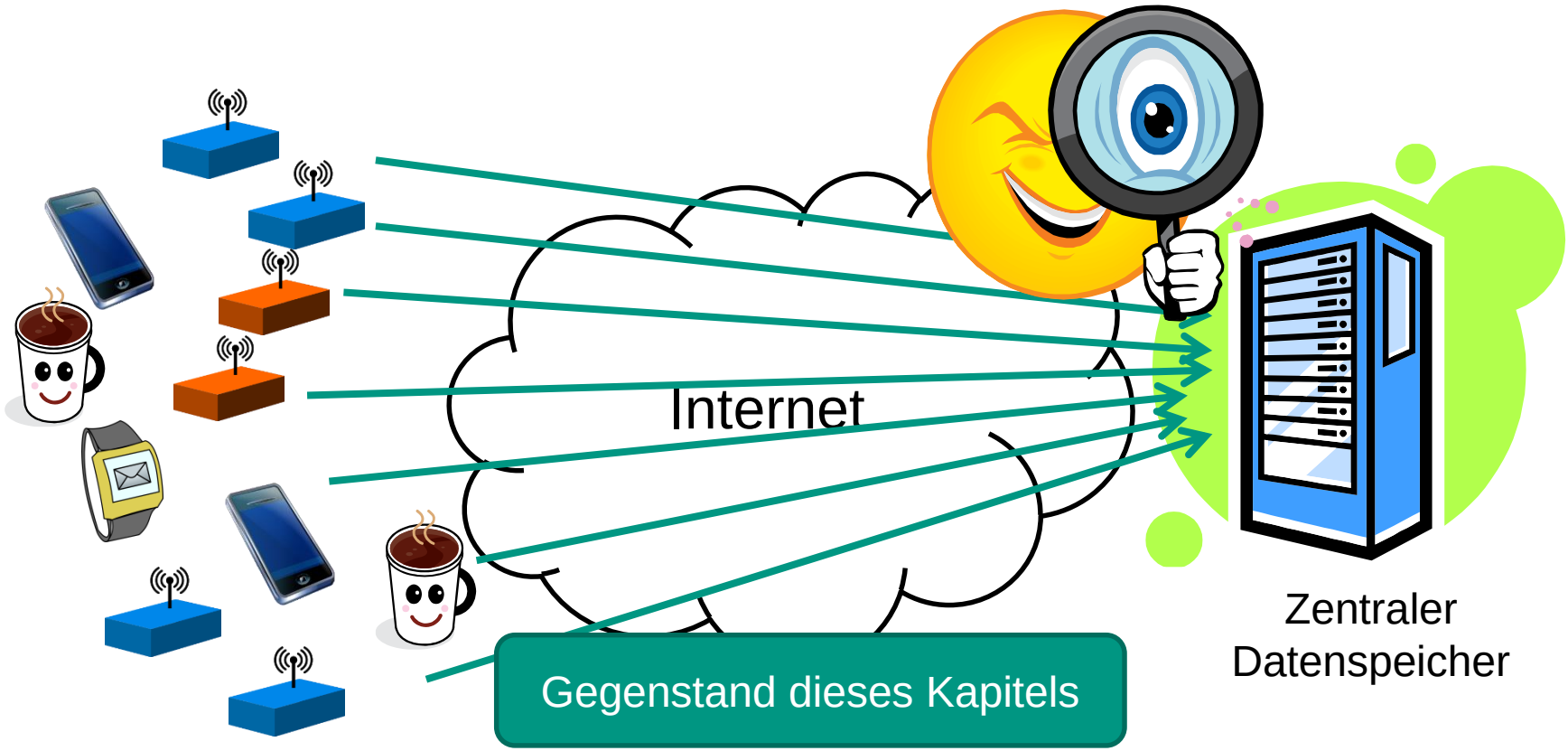

Allgemein
Was ist anders im IoE?
Privatsphäre
Ansätze zum Schutz
Konkrete Beispiele und Szenarien



→ Viele vernetzte Dienste und Anwendungen

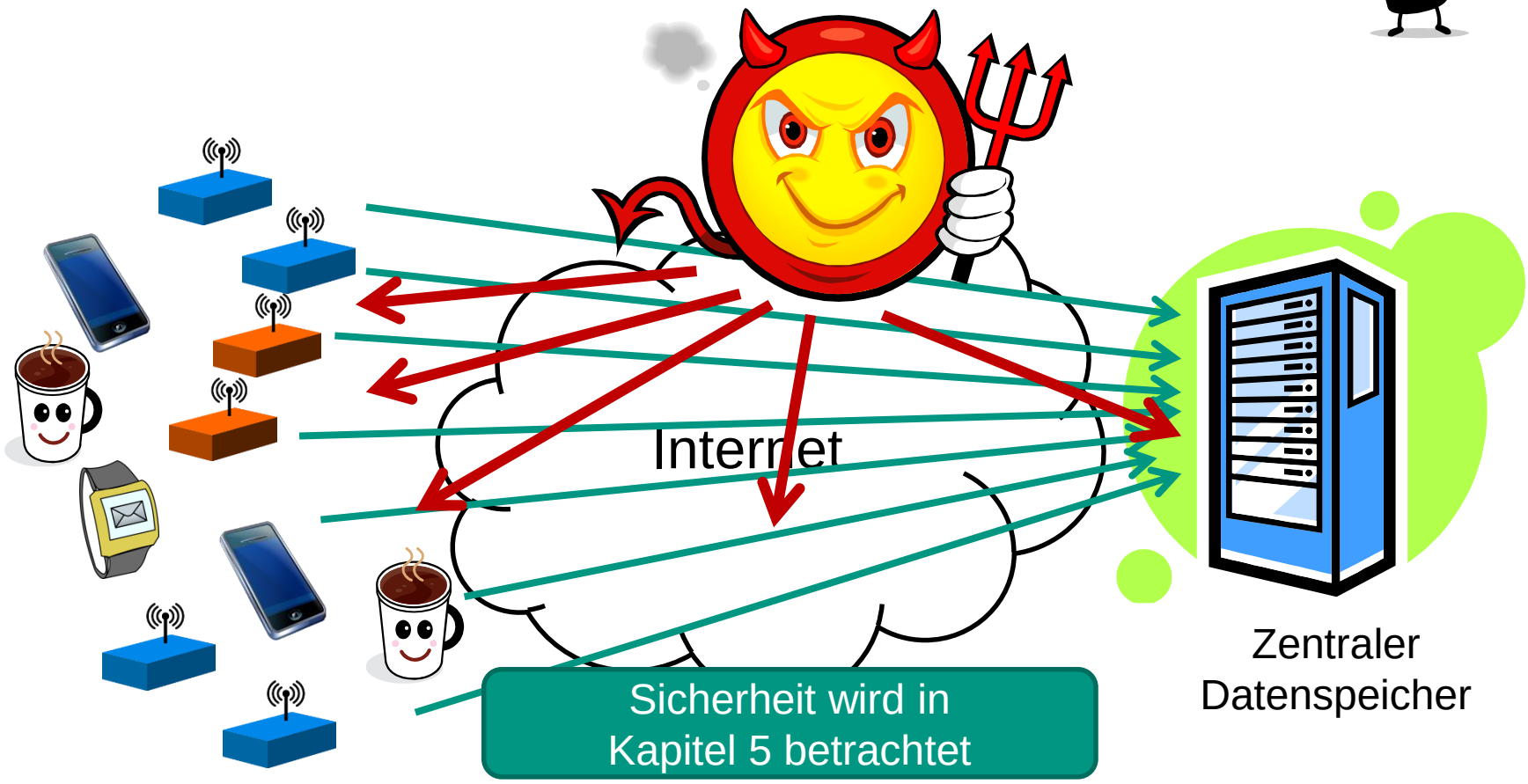
Aber …

- „Das Internet“ weiß mehr über mich als ich selbst!
  - Wo bleibt meine Privatsphäre?



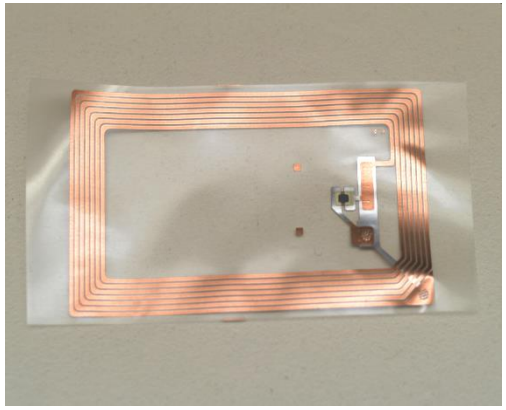
Aber …

- Sensoren können (z.B. über das Internet) angegriffen werden!
  - Sensoren sind auch häufig einfach physikalisch zugänglich

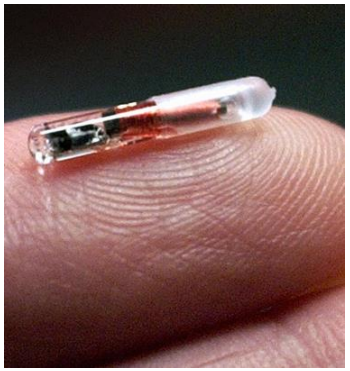


Beispiel: RFID

- RFID = Radio Frequency Identification
  - Drahtlose Übertragung von Informationen zwischen einem (kleinen) Transponder und einer Basisstation
  - „Nachfolger“ der Barcodes



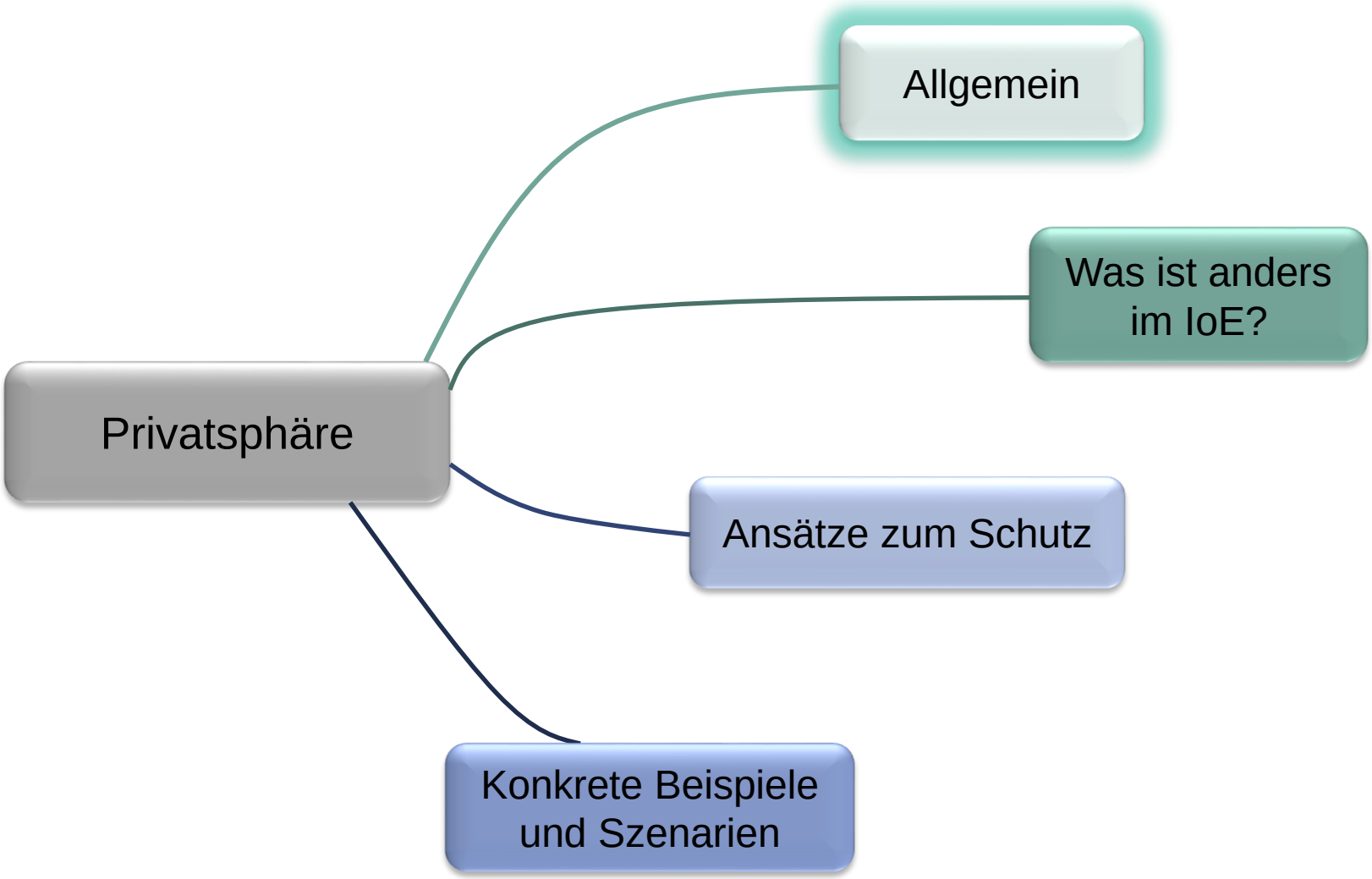
- Vielseitige Einsatzgebiete
  - Neue Personalausweise
  - Neue Reisepässe
  - Logistik
  - Einzelhandel
  - KITCard
  - Bücher aus der Bibliothek
  - …



RFID: Herausforderung für den Datenschutz

- Probleme
  - Eindeutige Identifikation jedes Objekts/Produkts möglich
    - Nicht mehr nur Identifikation der Produktgruppe
  - Identifikation erfordert keinen Sichtkontakt
  - Szenario: im Vorbeigehen Auslesen des Inhalts von Einkaufstaschen, der Historie von Kleidungsstücken, Fahrkarten,…

- Lösungsansätze
  - „Kill“-Befehl, Blocker-Tags
  - Authentifizierung für Zugriff
    - Beispiel Reisepass
      - Schlüssel wird aus aufgedrucktem maschinenlesbaren Teil des Passes berechnet
    - RFID Tags können einfache kryptographische Operationen
      - Hashing, XOR, teilw. symmetrische Kryptographie
      - Keine komplexen Public-Key-Operationen möglich
  - Physische Zerstörung
    - z.B. Antenne abreißen, Mikrowelle…


Allgemein
Was ist anders im IoE?
Privatsphäre
Ansätze zum Schutz
Konkrete Beispiele und Szenarien

Schutz der Privatsphäre (Datenschutz)

- Schutz des Persönlichkeitsrechts bei der Verarbeitung personenbezogener Daten
  - Recht auf informationelle Selbstbestimmung

- Welche Daten sind schützenswert?
  - Laut BDSG: Personenbezogene Daten
  - Weitreichender: Privacy-relevant data

- Im Kontext vom Internet of Everything: auch Metadaten!
  - Wer kommuniziert wann mit wem?
  - Wo hält sich ein Nutzer auf?
  - Wer nutzt einen bestimmten Dienst?

→ Schutz der Metadaten aus technischer Sicht besonders herausfordernd

Säulen zum Schutz der Privatsphäre

- Regulierung (z.B. Datenschutzgesetze)
  - Bundesdatenschutzgesetz
  - Einzelne Regelungen in Telekommunikationsgesetz und Telemediengesetz
  - Landesdatenschutzgesetze
- Selbstregulierung
  - TRUSTe
  - TÜV (Teil von S@ver Shopping)
  - Trusted Shop
- Selbstschutz

→ Privacy Enhancing Technologies (PETs)

Wie Systeme sicher gestalten?

- Bevor ein System abgesichert werden kann muss bekannt sein
  - Welcher Dienst soll erbracht werden, wie sieht die Architektur aus?
  - Was soll geschützt werden?
  - Gegen welche Angriffe bzw. was kann der Angreifer?
  - → Sicherheit ohne konkrete Beschreibung eines Systems nicht möglich

- Daher im Folgenden
  - Schutzziele
  - Angreifermodelle
  - …

Allgemeine Schutzziele (IT-Sicherheit)

- Schutzziel
  - Anforderungen an eine Komponente oder ein System, die erfüllt werden müssen, um schützenswerte Güter vor Bedrohungen zu schützen
- Häufige Kategorisierung in

  - Confidentiality (Vertraulichkeit)
    - Ein System bewahrt Vertraulichkeit, wenn es keine unautorisierte Informationsgewinnung ermöglicht
  - Integrity (Integrität)
    - Ein System bewahrt *starke* Integrität, wenn es nicht möglich ist, Daten unautorisiert zu manipulieren
    - Ein System bewahrt *schwache* Integrität, wenn unautorisierte Manipulationen an Daten nicht unbemerkt möglich sind
  - Availability (Verfügbarkeit)
    - Ein System bewahrt Verfügbarkeit, wenn es keine unautorisierte Einschränkung der Funktionalität des Systems ermöglicht
- Weitere Schutzziele
  - Authentizität
  - …

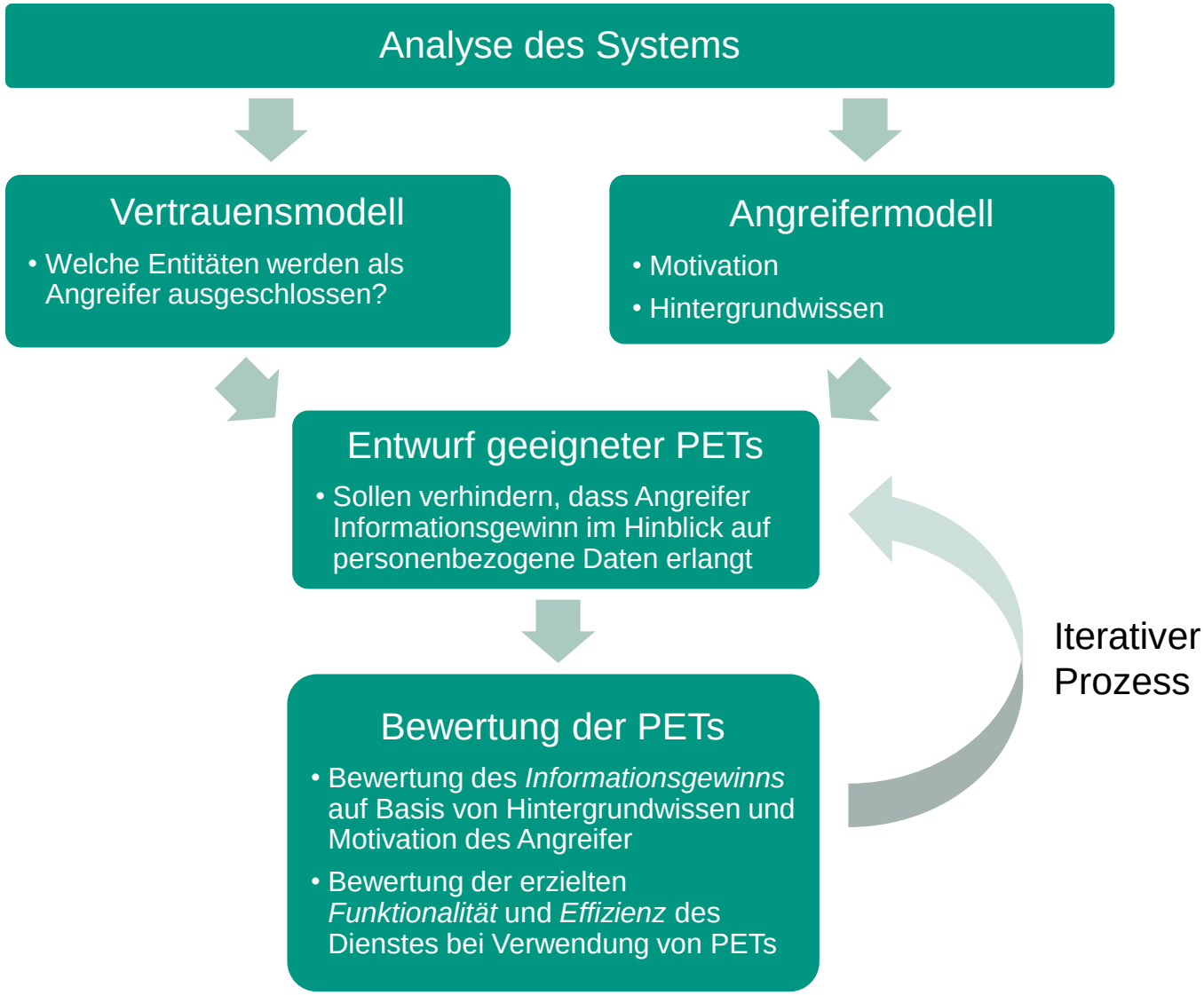
Spezifische Schutzziele für Privatsphäre

- Unverkettbarkeit (linkability)
  - Ein System bewahrt Unverkettbarkeit, wenn personenbezogene Daten aus zwei unterschiedlichen Kontexten für einen Angreifer nicht miteinander in Bezug gesetzt werden können
  - Technische Verfahren
    - Datenvermeidung, Anonymisierung, …
- Transparenz
  - Ein System bewahrt Transparenz, wenn die Verarbeitung von personenbezogenen Daten nachvollziehbar und überprüfbar ist
  - Technische Verfahren
    - Erstellung von Logdateien, …
- Intervenierbarkeit
  - Ein System bietet Intervenierbarkeit, wenn betroffene Personen über die Art der Erfassung und Verarbeitung ihrer personenbezogenen Daten selbst bestimmen können
  - Technische Verfahren
    - Schaffung von Wahlmöglichkeiten hinsichtlich der Verarbeitung von Daten, …

Weitere relevante Schutzziele

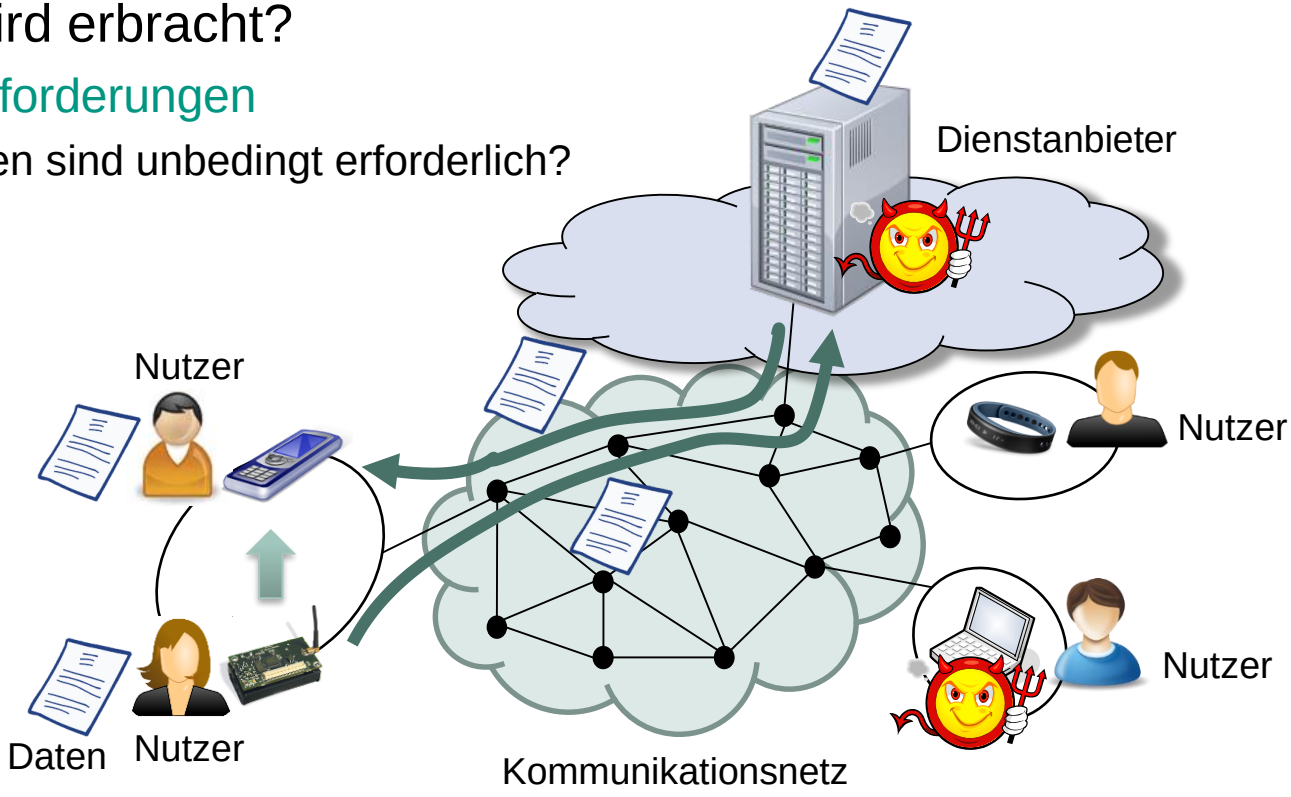
- Nicht-Identifizierbarkeit
  - Angreifer kann Daten keiner natürlichen Person zuordnen
  - Konflikt mit Schutzziel Authentizität…
  - Technische Verfahren
    - Pseudonymisierung – Daten mit fiktiver Identität assoziiert
    - Anonymisierung – Daten mit keinerlei Identität assoziiert
- Unentdeckbarkeit
  - Angreifer kann nicht eindeutig bestimmen, ob ein Datum existiert oder nicht
  - Oft Konflikt mit Anforderungen an Funktionalität…
  - Technische Verfahren
    - Datenvermeidung, geschicktes Routing
- Abstreitbarkeit (plausible deniability)
  - Angreifer kann dritter Partei („Richter“) nicht beweisen, dass ein Datum existiert oder nicht
  - Konflikt mit Schutzziel Nicht-Abstreitbarkeit…
  - Technische Verfahren
    - Datenvermeidung, kryptografische Bausteine wie Axolotl und OTR

Prozess für Entwurf und Bewertung von PETs



Analyse des Systems

- Welche Entitäten sind beteiligt?
  - Dienstanbieter
  - (Dienst-)Nutzer
- Welcher Dienst wird erbracht?
  - Funktionale Anforderungen
    - Welche Daten sind unbedingt erforderlich?



Vertrauensmodell

- Was ist Vertrauen?
  - Subjektive Bewertung in wie fern sich eine andere Entität in Bezug auf einen konkreten Sachverhalt erwartungsgemäß verhalten wird
  - Im Folgenden: Annahme, dass Entität kein Angreifer ist

Beispiele:

- Vollständiges Vertrauen
  - Nutzer vertraut allen Entitäten des Systems uneingeschränkt
  - „Anfänge des Internet“, bzw. „Post-Privacy“
- Vertrauen in zentrale Instanz
  - Nutzer vertraut zentralem Dienstanbieter (Google, Amazon…) oder
  - vertrauenswürdiger dritten Partei (trusted third party, TTP)
- Verteiltes Vertrauen
  - Nutzer vertraut, dass eine Teilmenge der beteiligten Entitäten nicht böswillig kooperiert
- Keinerlei Vertrauen
  - Nutzer vertraut keiner Entität des Systems

Angreifermodell

- Motivation?
  - Weshalb würde ein Angreifer einen Angriff vornehmen?
  - Wieviel darf ein Angriff den Angreifer kosten?
  - Was kann der Angreifer, was nicht?

→ Kategorisierung von Angreifern sinnvoll

  - Hilft der Vergleichbarkeit von Protokollen

Ziele des Angreifers (Bedrohungsanalyse)

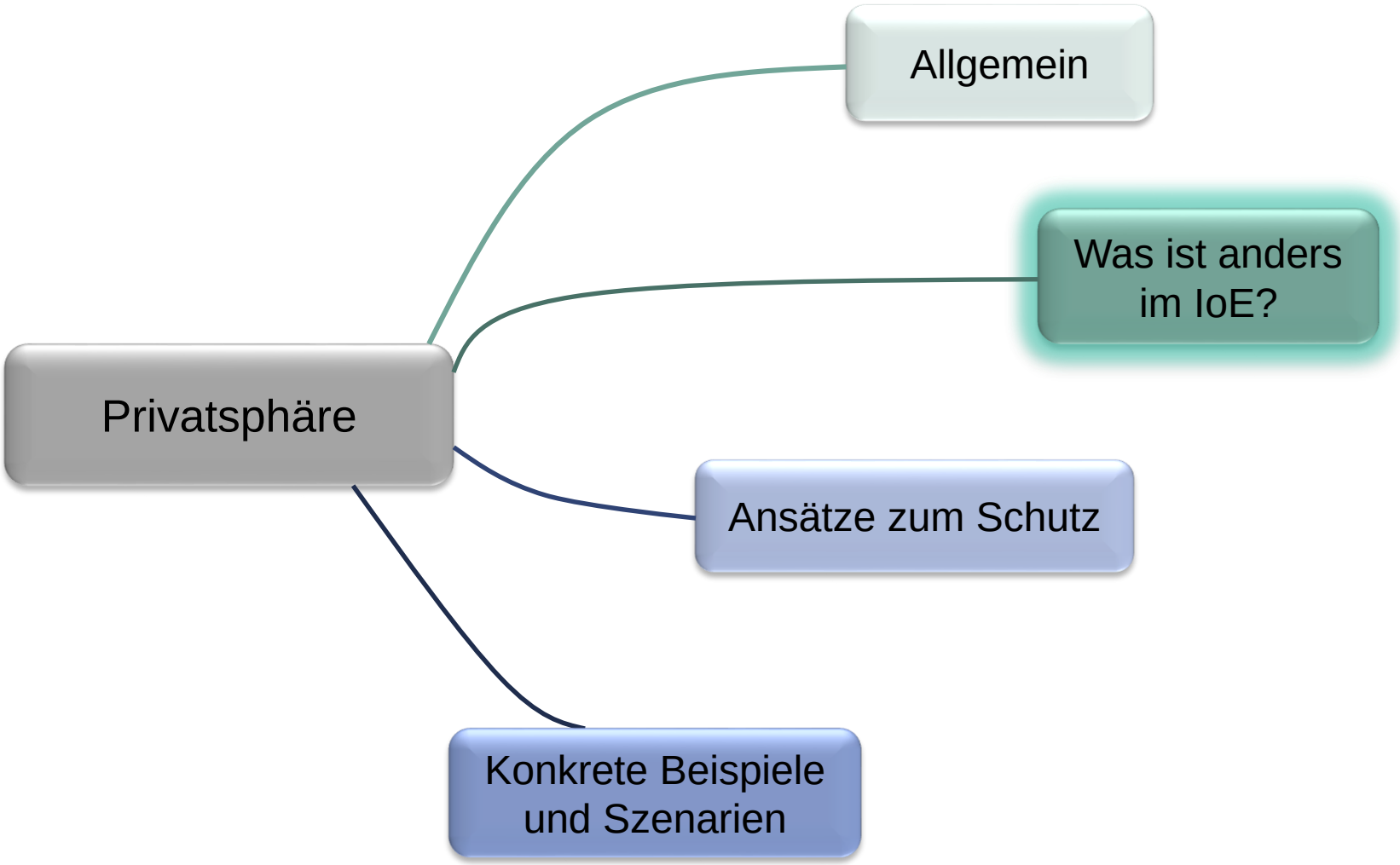
- Was will ein Angreifer mit seinem Angriff überhaupt erreichen?
- Typische „Bedrohungen“ für Netze sind
  - Abhören von Daten
    - Unbefugt in den Besitz geheimer Informationen gelangen – Angreifer erspäht vertrauliche Patientendaten über Konkurrenten
    - Schutzziel: Vertraulichkeit
  - Modifizieren von Daten
    - Daten so modifizieren, dass Angreifer einen Vorteil daraus erhält – Patientendaten so verändern, dass Konkurrent falsch behandelt wird.
    - Schutzziel: Integrität
  - Maskerade und Erzeugen von Daten
    - Daten erzeugen und im Sensornetz unter einer gefälschten Identität versenden
    - Schutzziel: Authentizität

„Klassisches“ Angreifermodell

- In klassischen Protokollen häufig nicht explizit genannt, z.B. TLS (RFC 4346) oder IPSec (RFC 4301, 4303).
  - Implizit geht man häufig von einem Angreifer aus, den man „Dolev-Yao“ Angreifer nennt
  - Angreifer ist omnipräsent im Netz, kann sämtliche Kommunikation abhören
  - Kann eigene Dateneinheiten erzeugen und versenden
  - Kann fremde Dateneinheiten modifizieren
  - Kann allerdings nicht Entschlüsseln oder Verschlüsseln, ohne den Schlüssel zu kennen

Angreifer = „Outsider“

[DY83]


Allgemein
Was ist anders im IoE?
Privatsphäre
Ansätze zum Schutz
Konkrete Beispiele und Szenarien

Was ist anders im IoE?

- Technologie greift viel stärker in das private Leben ein
  - Datenerfassung im privaten Umfeld
    - Unter Umständen sogar am Körper

- Beispiele aus der Einführung
  - Netatmo: Die Wetterstation für das Smartphone
  - Lifelogging Armbänder
  - ….

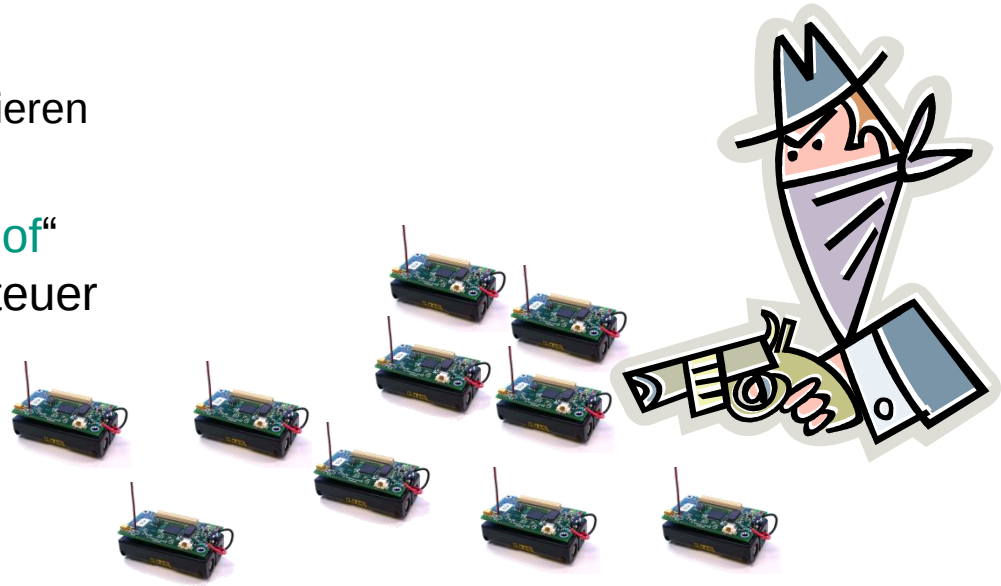
→ Privatsphäre stärker gefährdet als im klassischen Internet

IoE-spezifische Herausforderungen (I)

- ...an den technischen Schutz der Privatsphäre
  - Heterogene Geräte
    - Energie, Rechenkapazität, Speicher für klassische (kryptografische) Algorithmen nicht ausreichend
  - Häufig keine zentrale Infrastruktur nutzbar
    - Keine Public Key Infrastruktur, keine zentralen Vertrauensanker
  - Vertrauensmodell oft unklar
    - Gegen wen muss ich mich schützen?
    - Wer sind meine Vertrauensanker?
  - Angreifermodell anders
    - Geräte u.U. von Angreifer korrumpiert

Angreifermodell im IoE

- Geräte befinden sich häufig an öffentlichen oder leicht zugänglichen Orten
  - Angreifer kann physisch auf Geräte zugreifen
    - Kann Geräte evtl. einfach „klauen“ oder physisch zerstören
  - Manches wird auch einfacher
    - Speicher auslesen
    - Komplett re-programmieren
    - Korrumpieren
  - Sogenannte „tamper-proof“ Hardware für Sensoren teuer



Angreifermodell im IoE (II)

- Andere Möglichkeit: Viren und Würmer
  - Bei Geräten mit Internetzugang…
  - Angreifer findet Implementierungs- oder Konfigurationsfehler
  - Kann am Ende Geräte fernsteuern
  - Bsp.: Mirai: Botnetz aus Kameras und Videorekordern  [Krebs2016]

- Angreifer korrumpiert eine Menge von Geräten, die danach zusammenarbeiten (führt zu „byzantinischen“ Fehlern)

→ Herausforderung für Kommunikationsprotokolle

  - Erbringe einen Dienst sicher in Gegenwart von korrumpierten („bösen“) Geräten, z.B. sicher in Gegenwart von β=10% korrumpierten Geräten

Angreifer = „Insider“

Wichtig: Anzahl korrumpierter Geräte

- Wie viele Geräte korrumpiert Angreifer?
- Annahme: Netz mit n Geräten
  - Klar: Korrumpiert der Angreifer $n-1$ oder n Geräte, gibt es keine sinnvolle Definition von Sicherheit mehr (Benutzer sollte andere Probleme zuerst lösen…)
  - Auch klar: Je mehr korrumpierte Geräte im Netz, desto schwieriger wird Sicherheit
  - Angreifer wird in der Realität nicht einfach alle Geräte korrumpieren können
    - Korrumpieren von Geräten „kostet“ den Angreifer etwas
    - Z.B. Zeit, Geld in Form der notwendigen Hardware, usw.
    - Angreifer verfügt nur über begrenzte Ressourcen oder will nur begrenzte Ressourcen für seinen Angriff ausgeben

→ Häufig geht man davon aus, dass der Angreifer $n' < n$, d.h. einen Prozentsatz $\beta\%$ korrumpiert

Angreifermodell im IoE (III)

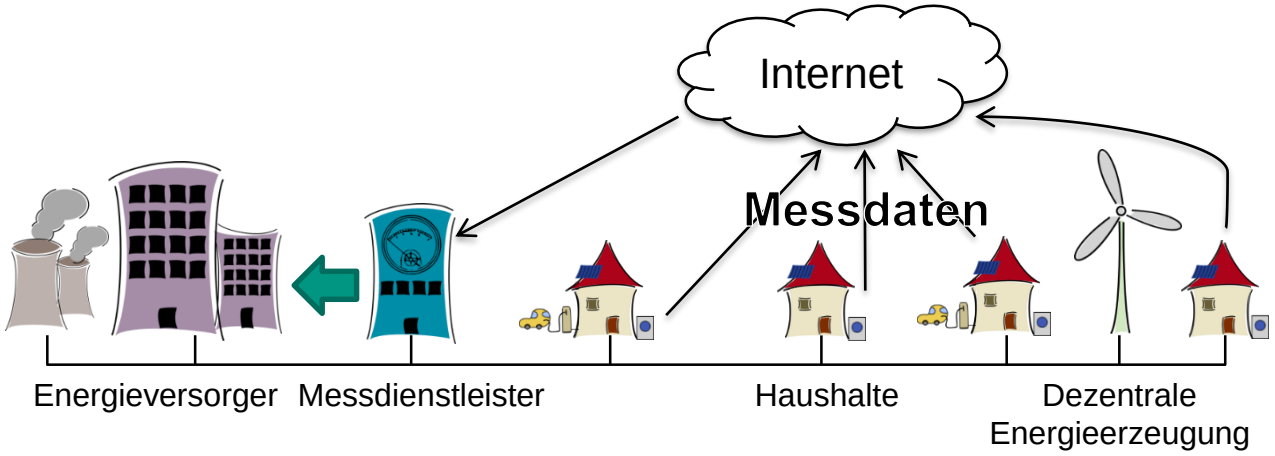
- Häufig Übertragung von personenbezogenen Daten an (zentralen) Dienstanbieter
  - Wie werden die Daten dort gespeichert?
    - … und welche?
  - Was passiert wenn der Anbieter bösartig mit den Daten agiert?
    - Im besten Fall „nur“ personalisierte Werbung
  - Angriffe auf Vertraulichkeit der Daten
    - Datensätze beim Anbieter wirklich sicher?
  - Angriffe auf Verfügbarkeit der Daten (einfacher) möglich
  - I.d.R. auch: keine Transparenz

IoE-spezifische Herausforderungen (II)

- ...an den technischen Schutz der Privatsphäre
  - Mehr Daten
    - Internet of Everything
      - Viele Dinge sind beteiligt
      - Datenerfassung ist allgegenwärtig (überall)
    - Datenzusammenführung bei Dienstanbietern
      - Bspw. Positionsdaten und Tagesplanung für Fahrplanauskunft
  - Kontinuierliche Datenerfassung
    - Sensoren messen 24/7 und überall
      - Fitnessarmbänder messen Schlafqualität
      - Intelligente Toilette erfasst „gesundheitsrelevante Daten“
  - Sensiblere Daten
    - Erfassung im eigenen Wohnraum (Privatsphäre!)
      - Stromverbrauch, Raumtemperatur, Anwesenheit, …
    - Beobachtung der eigenen Gesundheit
      - Puls, Bewegung
    - Positionsbezogene Daten
  - Vielfalt von gemessenen Phänomenen
    - Verkettung von Daten erlaubt detaillierte Profilbildung

Beispiel Smart-Metering

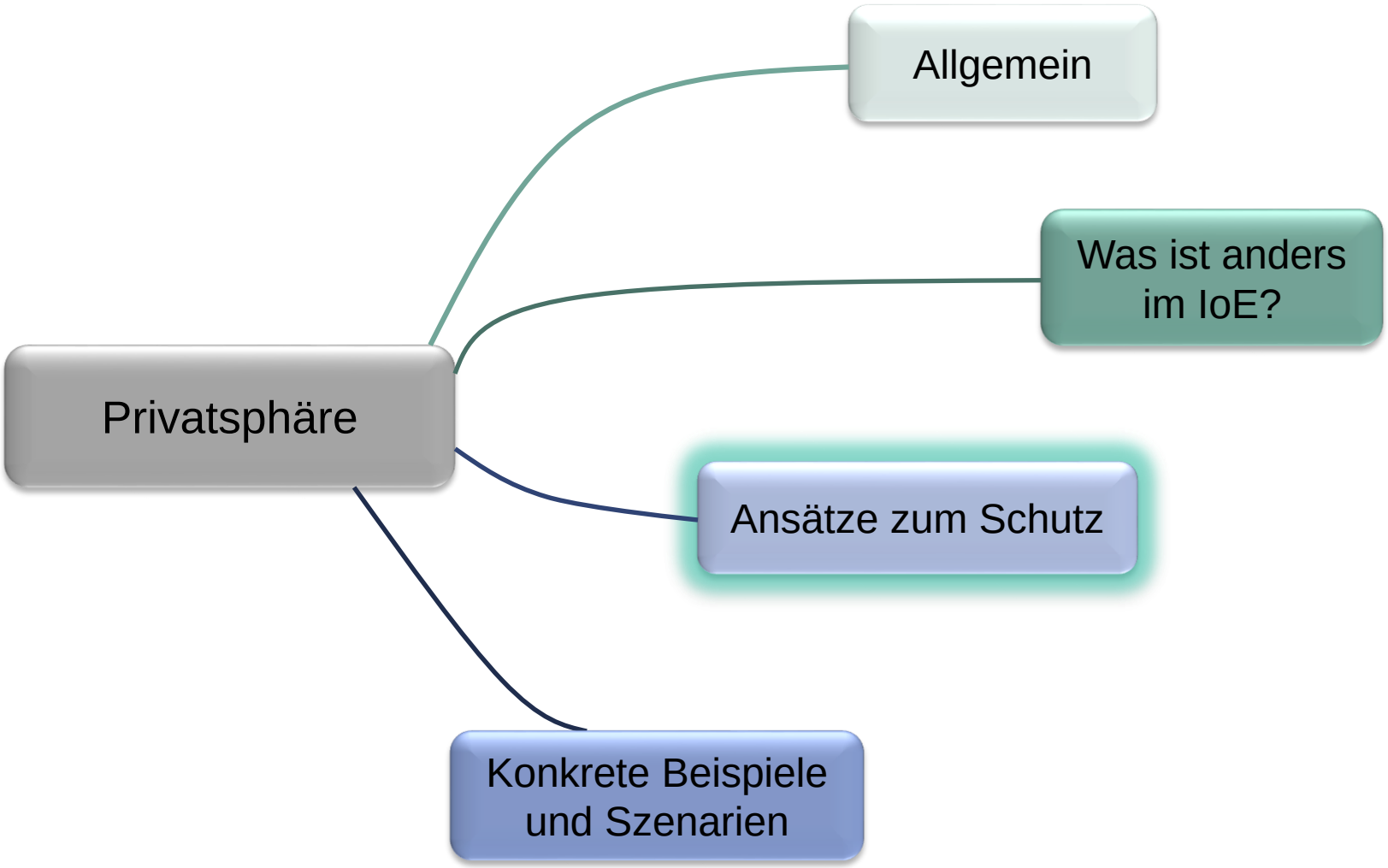
- Informationsbedarf für Stromnetz der Zukunft ist enorm
  - Wo, Wie, Wann, Wieviel Energieverbrauch
- Lösung: Smart Metering – intelligente Stromzähler
  - „Echtzeit“-Auslesen aus der Ferne
- Problem: detailliertes Verbrauchsprofil ermöglicht Rückschlüsse über Privatleben → Einschnitt in Privatsphäre
- Klassische Verschlüsselung keine Lösung
  - Schutzbedarf nicht nur vor Outsider sondern auch vor „Datensenke“



Beispiel Smart-Traffic

- Hier werden für die Diensterbringung u.a. Positionsdaten benötigt
- Bsp: Google Live Traffic
  - Positionsupdates von Android-Smartphones [Jeske2013]
    - Bestimmung des Verkehrsaufkommens / Stauerkennung
    - Nutzung für verbesserte Navigation (Google Maps)
    - Daten werden andauernd gesammelt, auch wenn keine App offen ist
  - Smartphone enthält eindeutige, gleichbleibende ID

- Problem 1: Google kann* immer herausfinden, wo man sich befindet
  - *aus technischer Sicht steht dem nichts im Wege
- Problem 2: Selbst ohne Zuordnungsmöglichkeit von ID zu Smartphone/Nutzer…
  - Zuordnung über Zusatzinformation, z.B. Wohnort oder Arbeitsplatz
  - Gerade bei gleichbleibenden IDs sehr problematisch [Golle2009] [Tock2014]


Allgemein
Was ist anders im IoE?
Privatsphäre
Ansätze zum Schutz
Konkrete Beispiele und Szenarien

Allgemeines Szenario im Folgenden

- Was wird erfasst? Samples
  - Beobachtung einer Messgröße zu einem bestimmten Zeitpunkt
  - Erforderlich für Diensterbringung
  - Enthalten meist Zeitstempel zusätzlich zur Messgröße

- Was soll (nach Möglichkeit) erreicht werden?
  - Samples geben nur wenig Privates preis
    → Datensparsamkeit
  - Samples werden nicht an potentielle Angreifer kommuniziert
    → Unentdeckbarkeit
  - Samples können nicht mit Nutzern in Verbindung gebracht werden
    → Nicht-Identifizierbarkeit
  - Samples können nicht untereinander in Verbindung gebracht werden
    → Unverkettbarkeit
    - Würde Profilbildung und ggf. Identifizierung ermöglichen

Beispiele für Ansätze zum Schutz (I)

- Verschleierung von Sampling-Werten
  - Präzision auf ein nötiges Minimum herabsetzen
  - Störwerte hinzufügen
  - Beispiele
    - Stromverbrauchserfassung in Kategorien (0-100, 101-500, 501-2000 Watt)
    - Positionsdaten zufällig um wenige Metern verschieben

- Zentralisierte Datensenken vermeiden
  - Peer-to-Peer-Netze zur Diensterbringung
  - Lokale Diensterbringung
  - Mehrere, nicht-kooperierende Dienstanbieter mit jeweils beschränkter Sicht
  - Beispiele
    - Peer-to-Peer Filesharing
    - Datenablage auf mehreren unabhängigen Cloud-Providern

Beispiele für Ansätze zum Schutz (II)

- Identität der Quelle verschleiern
  - Pseudonyme
  - Datenaggregation
  - Beispiele
    - „Spitznamen“ in sozialen Netzwerken
    - Durchschnitt / Summe der Messwerte über mehrere Individuen

- Unverkettbarkeit von Samples gewährleisten
  - Samplingrate und Samplingzeitpunkte geschickt wählen
  - Beispiele
    - Stromabrechnung nicht im Minutentakt sondern niedrigere Samplingrate, bspw. stündlich
    - Zeitliche Muster bei Übertragung vermeiden → stündlicher Sample um XX:12:13 Uhr ist immer derselbe Nutzer

„Guter Ansatz“ ist anwendungsspezifisch – keine universelle Lösung!

Entwurfsstrategien (I)

[Hoe2014]

Datenorientiert

- MINIMISE
  - Die Verarbeitung (und Erfassung) von personenbezogenen Daten soll auf das erforderliche Minimum reduziert werden
    - Datenvermeidung (select before collect), Anonymisierung, Pseudonymisierung, …
- HIDE
  - Personenbezogene Daten und deren Beziehung untereinander sollen verborgen werden
    - Verschlüsselung, Schutz der Metadaten, Pseudonymisierung, Mix-Netze, …
- SEPARATE
  - Unterschiedliche personenbezogenen Daten sollen möglichst separat und dezentral gespeichert und verarbeitet werden
    - Dezentralisierung, lokale Datenverarbeitung, …
- AGGREGATE
  - Personenbezogenen Daten sollen soweit wie möglich aggregiert verarbeitet werden
    - Aggregation über die Zeit, *k*-Anonymität, …

Entwurfsstrategien (II)

[Hoe2014]

Prozessorientiert

- INFORM
  - Betroffene Personen sollen über die Verarbeitung ihrer personenbezogenen Daten informiert werden (→ Schutzziel Transparenz)
- CONTROL
  - Betroffene Personen sollen die Kontrolle über die Verarbeitung ihrer personenbezogenen Daten behalten (→ Schutzziel Intervenierbarkeit)
- ENFORCE
  - Eine rechtskonforme Datenschutzrichtlinie soll erstellt und durchgesetzt werden
- DEMONSTRATE
  - Bei der Verarbeitung von personenbezogenen Daten muss die Einhaltung der Datenschutzrichtlinie demonstriert werden können

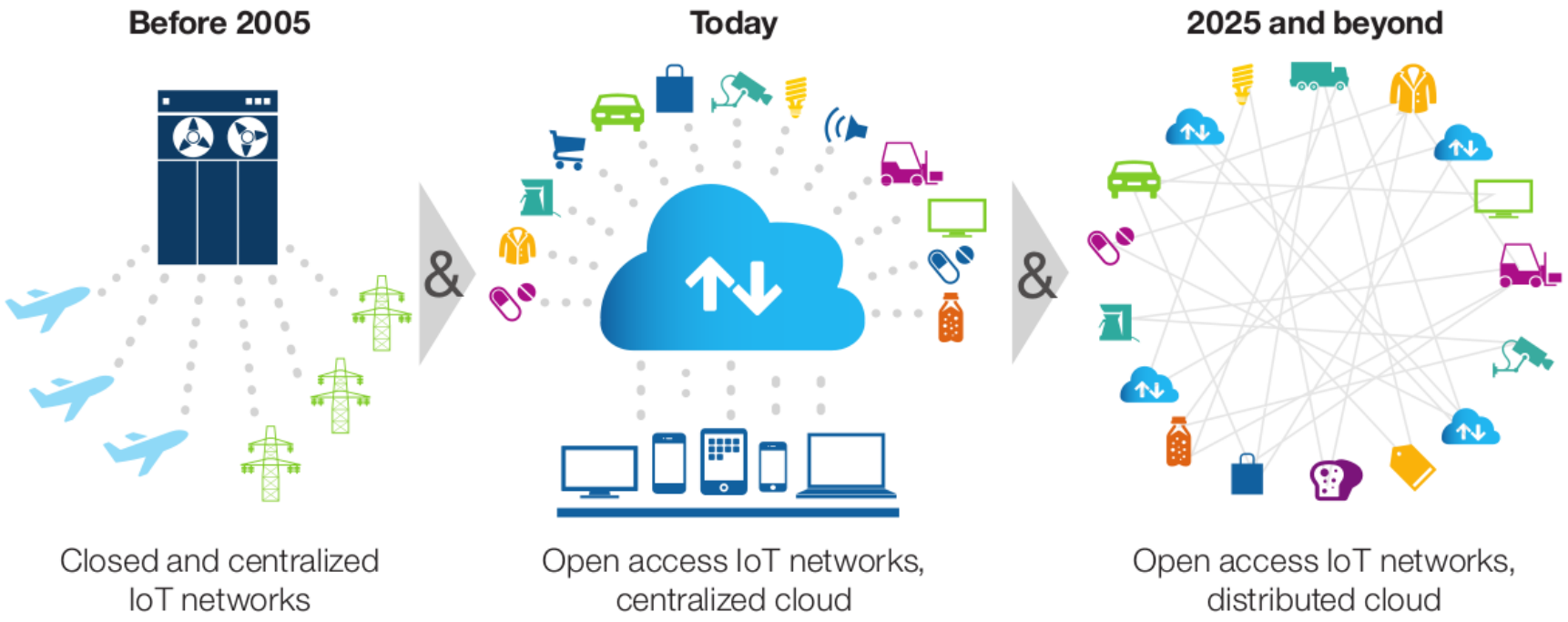
Device democracy (I)

- Studie von IBM [Pure2015]

- Identifiziert (u.a.) folgende Herausforderungen für das IoE:
  - „Cost of connectivity“
    - Das Betreiben einer zentralisierten Cloud verursacht hohe Kosten für Infrastruktur, Wartung und Sicherung
    - Geräte sollen billig sein aber lange Lebensdauer haben – wie den Langzeitsupport garantieren?
  - „The Internet after trust“
    - Wie Privatsphäre und Anonymität gewährleisten?
    - Thema auch: wie Abhängigkeit von zentralen Vertrauensankern lösen?

Device democracy (II)

- Lösungsansatz: Dezentralisierung → verteilte Cloud



[Pure2015]

Anforderungen an die verteilte IoE-Cloud

- “Trustless” Nachrichtenaustausch zwischen Geräten
  - Gewährleistung von: Privatheit, Integrität von Nachrichten…
  - Technische Garantien statt Vertrauen
- Verteilen von Daten
  - Verteilen von Anfragen
  - Sammeln von Sensorwerten und sonstigen Samples
  - Gewährleistung von: Privatheit, Integrität von Messwerten…
- Robuste und skalierbare Koordination von Geräten
  - Herstellen eines konsistenten Zustands
  - Zentralisierte Vertrauensanker nicht mehr geeignet

Technologien für die verteilte IoE-Cloud (I)

- Sicherer Nachrichtenaustausch zwischen Geräten
  - Lokal →  [s. Kapitel 5,6]
  - Global → über IP?
    - Problem Konnektivität: direkte Verbindungen in vielen Netzen (z.B. Mobilfunk) heute nur schwer möglich (wegen NAT, restriktiven Firewalls…)
    - Problem Privatsphärenschutz: Metadaten (wer redet mit wem?) für alle Entitäten auf dem Datenpfad einsehbar
  - Eigene Arbeiten: Peer-Tor-Peer (PTP)
    - Peer-to-Peer-Kommunikation über Tor NS
      - Netz-Restriktionen werden weitgehend umgangen
      - Metadaten werden geschützt
    - PTP kapselt Tor-Management, bietet einfaches Nachrichten-basiertes API
    - Work in Progress. Mehr Infos auf unseren Webseiten ;)

Technologien für die verteilte IoE-Cloud (II)

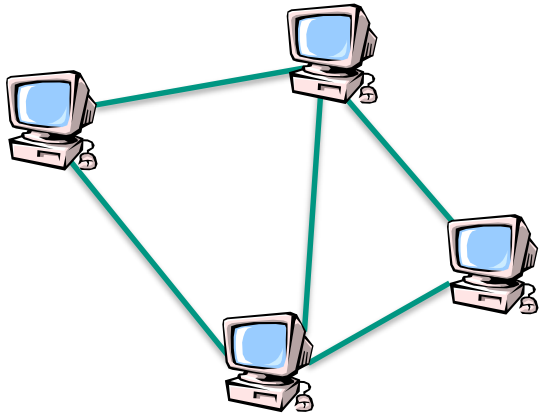
- Verteilen von Daten
  - Sehr anwendungsspezifisch!
  - → Wird im restlichen Kapitel genauer betrachtet

- Robuste und skalierbare Koordination von Geräten
  - Ohne Vertrauensanker → dezentrale Konsistenz
  - Vielversprechender Ansatz laut [Pure2015]: Blockchain-Netze
    - Erste und bekannteste Ausprägung: Bitcoin [Naka2008]
    - → dezentrales Zahlungssystem ohne Vertrauensanker
    - Im Folgenden: Blockchain-Crashkurs

Blockchain-Crashkurs: Überblick

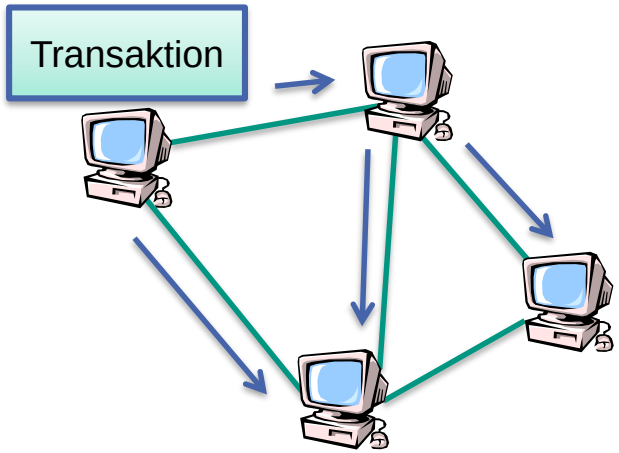
- High-level Vorstellung
  - Eine Blockchain ist ein schwarzes Brett
  - Man kann nur schreiben, nichts löschen
  - Jeder Eintrag hat einen Zeitstempel

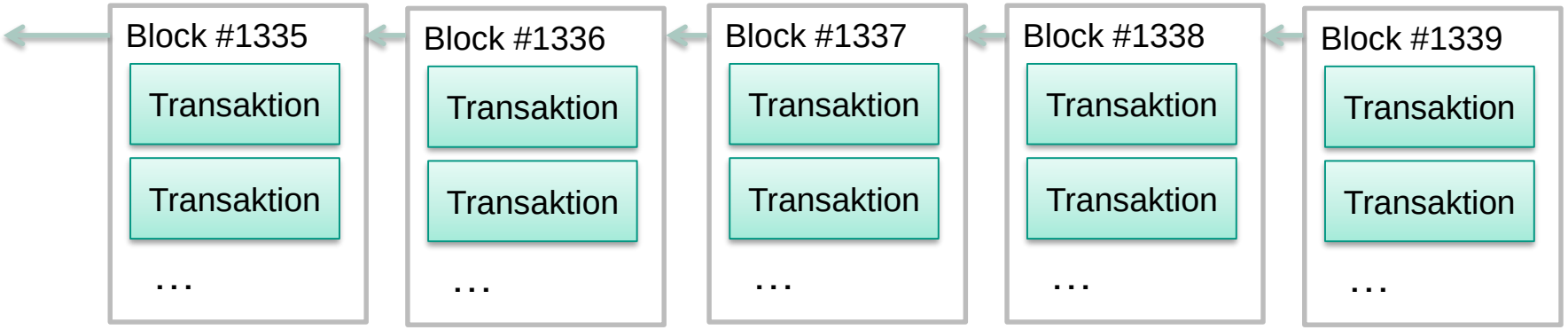
  - Realisiert wird die Blockchain durch ein Peer-to-Peer-Netz
  - Alle Peers kennen die komplette Blockchain
  - Kein Vertrauen in individuelle Peers nötig
    - Stattdessen Vertrauen, dass keine Gruppe von böswillig kooperierenden Nutzern existiert, die zusammen über die Mehrheit einer bestimmten Ressource im Peer-to-Peer-Netz verfügt



Blockchain-Crashkurs: Neue Einträge (I)

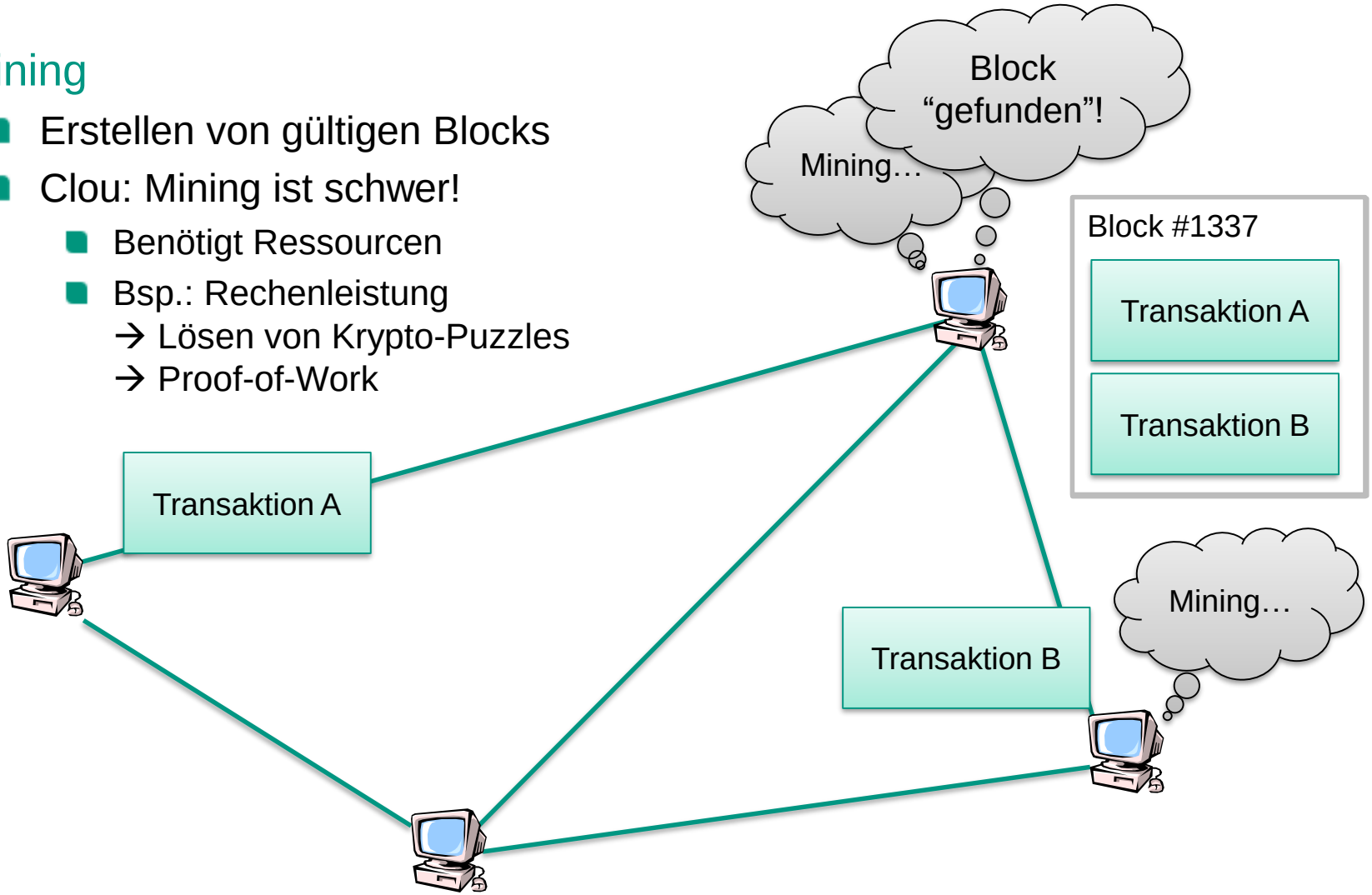
- Blockchain-Einträge werden typischerweise Transaktionen genannt
  - (aus dem Kontext: Zahlungssystem)
- Neue Transaktionen
  - Werden zunächst an alle Peers geflutet
  - Dann aber noch nicht Teil der Blockchain!
- Blocks
  - Eigentliche Bausteine der Blockchain
  - Enthalten mehrere Transaktionen
  - Referenz auf vorherigen Block → Blockchain





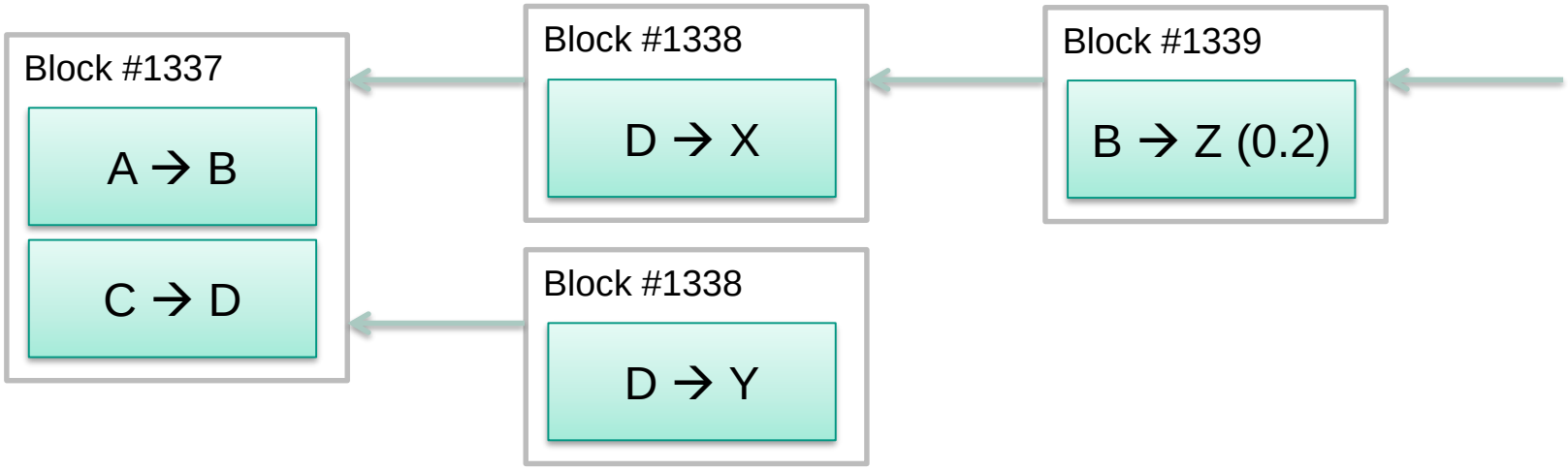
Blockchain-Crashkurs: Neue Einträge (II)

- Mining
  - Erstellen von gültigen Blocks
  - Clou: Mining ist schwer!
    - Benötigt Ressourcen
    - Bsp.: Rechenleistung
      → Lösen von Krypto-Puzzles
      → Proof-of-Work



Blockchain-Crashkurs: dezentrale Konsistenz

- Widersprüchliche Transaktionen
  - Werden beim Erstellen von Blocks ausgeschlossen
  - Problem: “parallel” erstellte Blocks → Blockchain-Forks



- Auflösung (vereinfacht): die Blockchain ist die längste bekannte Kette von validen Blocks
  - Kann durch jeden Peer verifiziert werden

Blockchain-Netze für das IoE

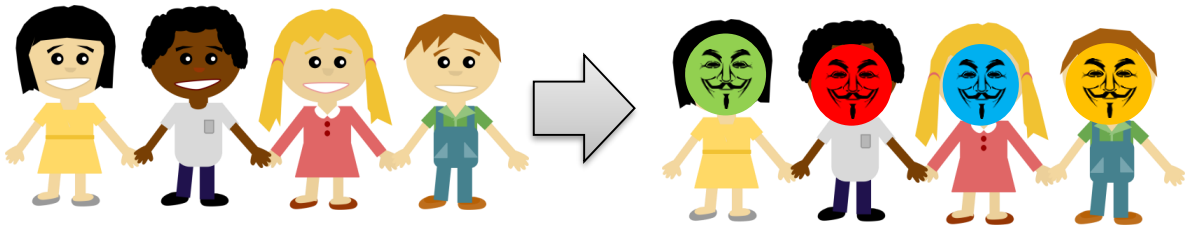
- Potential der Technologie noch nicht voll ausgeschöpft
  - Jetzt schon vielseitig einsetzbarer Baustein
    - Beispiel: Realisierung von Namensdiensten
    - Beispiel: Robuste Zeitstempel für Messwerte o.ä.
    - Beispiel: Finanzielle Gegenleistungen durch Micropayments
  - Zusätzlich: Transaktionen können beliebig komplexe Regeln enthalten, die von Minern überprüft werden
    - Ermöglicht z.B. “Smarte Verträge”, die von allein ausgeführt werden…
    - Aktives Experimentier- und Forschungsfeld!

Eigene Arbeiten: BitNym (I)

- Problem in vielen IoE-Anwendungen
  - Nutzer brauchen eindeutige Identifier, die fair verteilt werden

  → Das Erstellen von Phantom-Identitäten (Sybils) sollte nicht möglich sein
  - Zum Schutz der Privatsphäre sollten echte Identitäten verschleiert werden

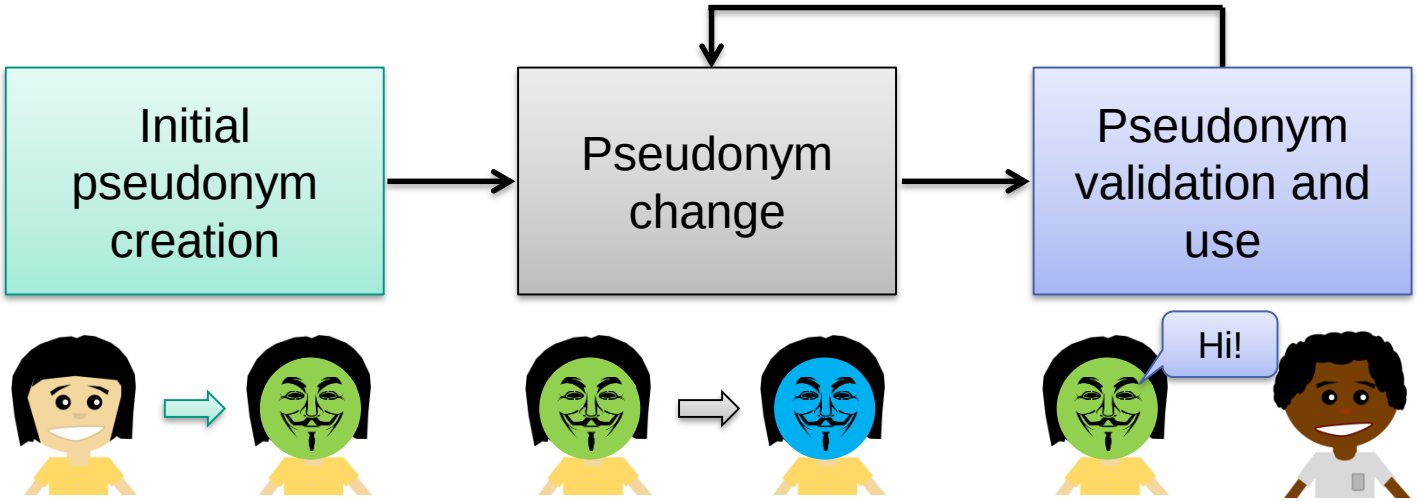
  → Verwendung von leicht zu wechselnden Pseudonymen



- Status quo: Pseudonyme werden durch Vertrauensanker verteilt
- Unser Ansatz: Vertrauensanker ersetzen mithilfe von Blockchain

Eigene Arbeiten: BitNym (II)

- Bausteine

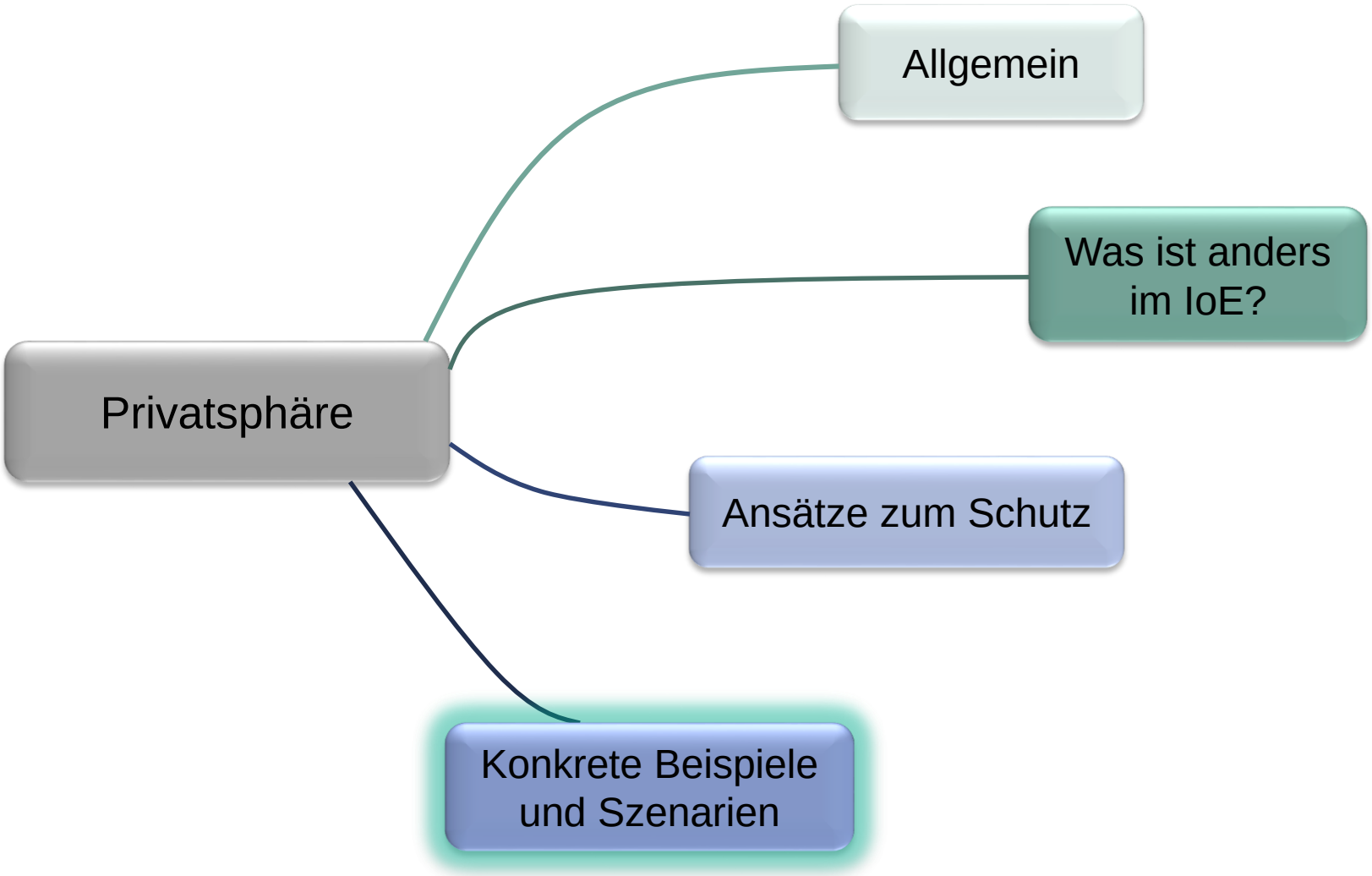


- Pseudonyme auf Blockchain abspeichern
- Sybil-resistente Zugangskontrolle

- Unverkettbare Pseudonym-wechsel

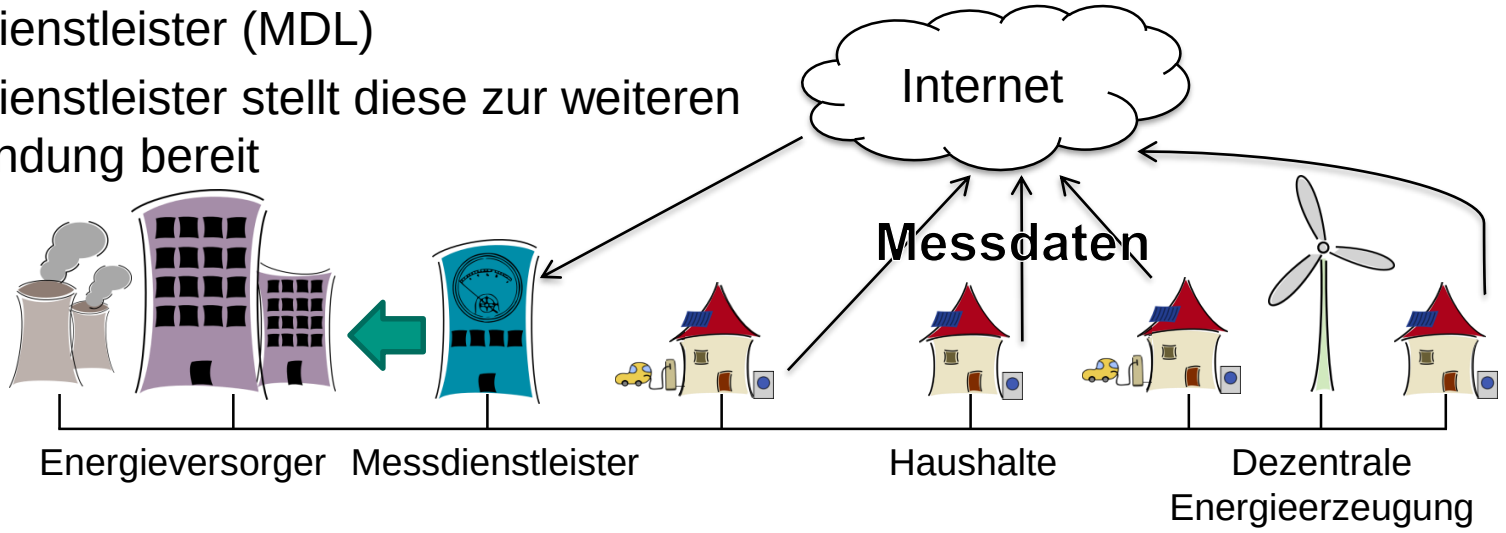
- Sybil-Resistenz beibehalten trotz Wechsel

- Blacklisting?
- Reputation?

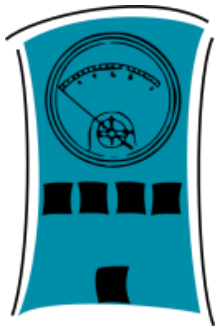

Allgemein
Was ist anders im IoE?
Privatsphäre
Ansätze zum Schutz
Konkrete Beispiele und Szenarien

Beispiel: Privatheit beim Smart-Metering

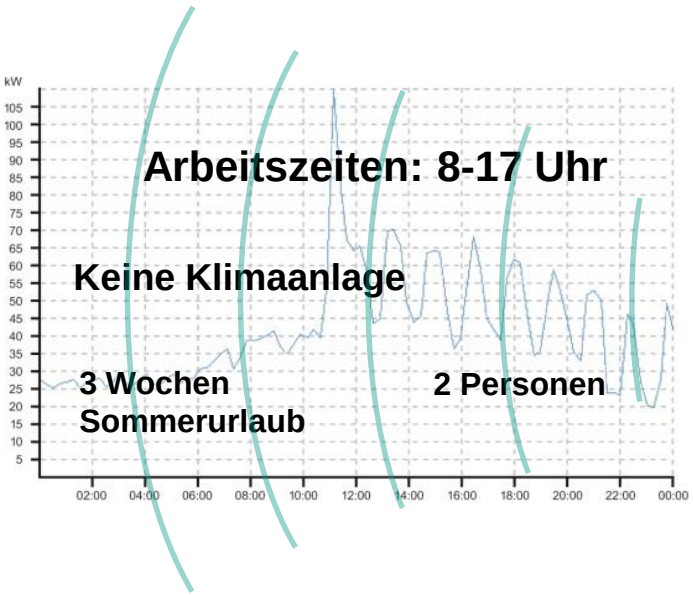
- Smart-Metering Szenario (in Deutschland)
- Ziel
  - Verbrauch von Energie in Energienetzen (topologisch) oder Kundenspezifisch (demographisch) in „Echtzeit“ nachvollziehbar machen
    - Beispiel topologisch: Last in Energienetzabschnitt „Karlsruhe-Mitte“
    - Beispiel demographisch: Energieverbrauch aller Kunden der EnBW
- Methodik
  - Stromzähler senden über Internetanbindung des Kunden Messdaten an Messdienstleister (MDL)
  - Messdienstleister stellt diese zur weiteren Verwendung bereit

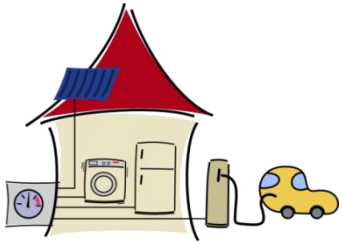


Gefahr für die Privatsphäre



Messdienstleister



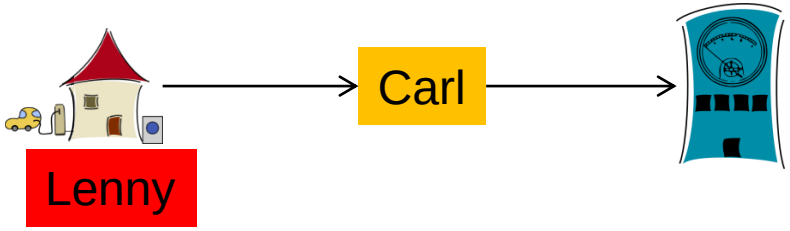


Haushalt

- Periodisches Senden von Messwerten
  - Beispielsweise alle 15 Minuten
  - Liefert detailliertes Verbrauchsprofil
  - Bietet Einblicke in Privatsphäre!

Generische Lösungsansätze (I)

- Pseudonymisierung
  - Messdaten mittels „falscher“ Identitäten, Pseudonymen, übertragen



- Probleme
  - Pseudonymverwaltung aufwändig
    - Benötigt Vertrauensanker o.ä.
  - Pseudonymisierte Daten sind verkettbar
    - Gleiches Pseudonym → gleicher Nutzer
    - Verkettete Daten (Profile) mittels externer Daten identifizierbar
      - Arbeitszeiten, Urlaube
    - Häufige Pseudonymwechsel nötig (und selbst dann…)
  - Übertragung selbst problematisch
    - IP-Adressen und Latenzen erhöhen Identifizierbarkeit

Generische Lösungsansätze (II)

- Modifikation des Energiebedarfs
  - Energiespeicher (bspw. Akkumulatoren) im Haushalt
  - Nach außen sichtbaren Energiebedarf „verharmlosen“ durch gezieltes Laden und Entladen der Akkumulatoren

- Problem
  - Potential der „Verharmlosung“ von Akkukapazität abhängig
    - Kostenintensive Anschaffung
  - Laden und Entladen verbraucht Energie und „verbraucht“ Akku
    - Kostenintensiver Betrieb / Wartung
  - Lade- / Entladestrategie komplex
    - Konstanten Stromverbrauch anvisieren oder randomisieren?
    - Was wenn Akkus leer / voll?
      - → Privatsphäre in Gefahr

Anwendungsspezifischer Ansatz nötig

- Analyse anwendungsspezifischer Anforderungen
- Ziel von Smart Metering sind aggregierte Daten
  - ∑ Energieverbräuche aller Haushalte in „Karlsruhe-Mitte“
  - ∑ Energieverbräuche der Kunden der EnBW
  - ∑ Energieverbrauch eines Haushalts im Monat

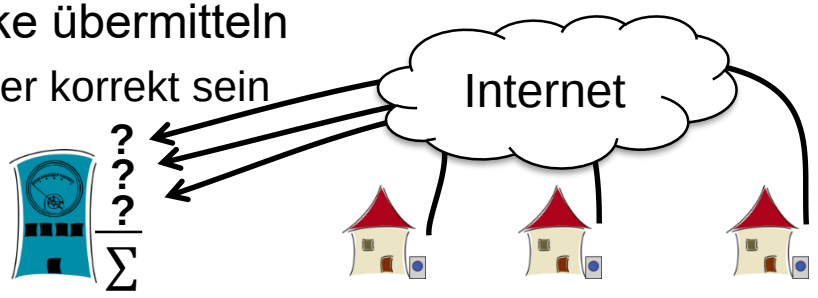
- Idee: Aggregation bevor Daten übermittelt werden
  - Aggregation über viele Haushalte → Privatsphäre geschützt
    - Wie viele? Offen.
  - Aggregation über langen Zeitraum → Privatsphäre geschützt
    - Wie lange? Offen.

- Heutiges langes Ableseintervall ist impliziter Schutz der Privatsphäre
  - Aggregation über Zeit
  - Jahresweise

Möglichkeiten zur Aggregation

- Aggregation über Zeit einfach zu realisieren
  - Stromzähler aggregiert auf Register
  - Vertrauenswürdige Zähler (Manipulationssicher / Verbot der Manipulation)
  - Komplexe Tarife (Nachtstrom) mittels mehrerer Register

- Aggregation über Haushalte schwierig zu realisieren
  - Problem: Daten, die aggregiert werden, darf Datensenke nicht erfahren
    - Möglichkeit: Vertrauenswürdige, dritte Partei aggregiert
      - Aber: Wodurch Vertrauen gerechtfertigt?
  - Lösung: „Falsche“ Daten an Datensenke übermitteln
    - Nach Aggregation muss Ergebnis wieder korrekt sein
  - Zwei grundlegende Vorgehen
    - Ohne Kooperation der Stromzähler untereinander
    - Mit Kooperation der Stromzähler untereinander



Beispiel ohne Kooperation: „Rauschen“

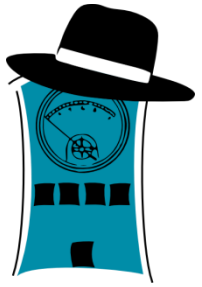
- Hinzufügen von planbarem „Rauschen“
  - Laplace-Rauschen auf Messwerte
  - Einzelner Rauschwert zufällig
  - Rauschwerte „eliminieren“ sich gegenseitig bei Summenbildung und vielen Teilnehmern

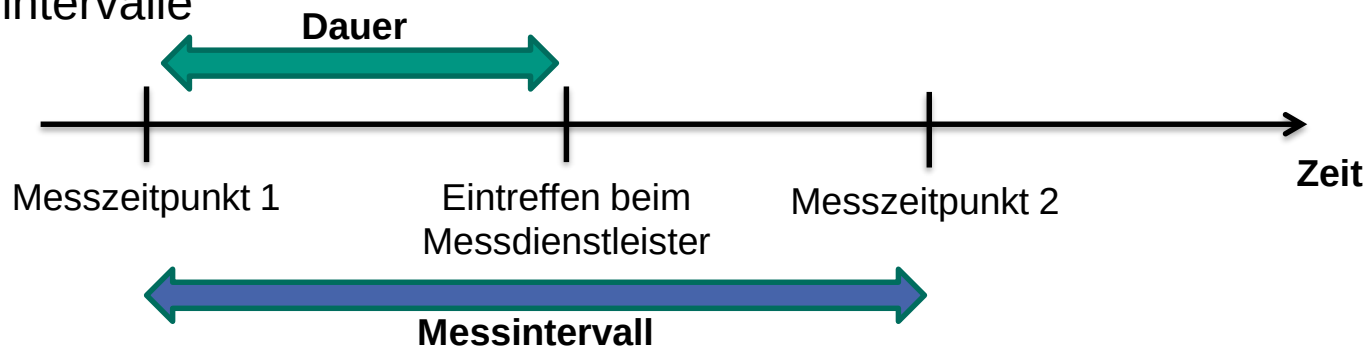

[Bohli2010]

- Problem: Sehr viele Teilnehmer notwendig
  - Keine feingranulare Messung (topologisch oder demografisch) möglich
  - Beispielrechnung
    - Ungenauigkeit von 400 Watt des Messwertes als Privatsphärenschutz
      - Genug um Alltagsverbrauch zu verschleiern
      - Zu wenig für Kochen/Waschen/Trocknen/Staubsaugen
    - 99,9% Genauigkeit des Aufsummierten Messwertes beim MDL
    - → Mindestens 3,8 Millionen Haushalte nötig
  - Nicht praktikabel

Mit Kooperation der Stromzähler untereinander

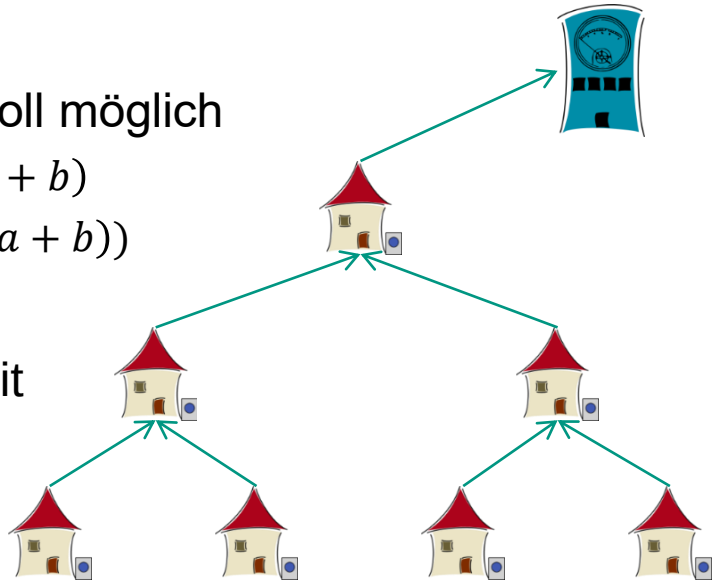
- Benötigt Protokoll zur Kommunikation zwischen Stromzählern
- Anforderungen
  - Einzelne Messwerte vor Messdienstleister geschützt
    - Nur Aggregat ermittelbar, auch bei Angriff durch bspw. MDL
  - Robuster Betrieb
    - Ausfall einzelner intelligenter Stromzähler
    - Ausfall der Kommunikationsanbindung
  - Realisierbarkeit auf ressourcenbeschränkter Hardware
    - Geringe Rechen- und Speicherkapazität
  - Aktuelle Ergebnisse = Kurze Dauer
  - Kurze Messintervalle





Beispiel: Homomorphe Verschlüsselung

- Homomorphe Verschlüsselung
  - „Rechnen“ mit verschlüsselten Daten sinnvoll möglich
    - Beispiel: $Enc_{PK}(a)\ (+)\ Enc_{PK}(b) = Enc_{PK}(a+b)$
    - Mit Secret Key entschlüsseln $Dec_{SK}(Enc_{PK}(a+b))$
- Beispiel Vorgehen: [Li2011]
  - Stromzähler verschlüsseln eigenen Wert mit Public Key des MDL
  - Verschlüsselter Wert wird entlang einer Baumstruktur weitergegeben
  - Jeder eintreffende (verschlüsselte) Wert wird aggregiert mit eigenem und weitergegeben
  - MDL erhält einen verschlüsselten Wert, den er entschlüsseln kann
  - Kein Stromzähler erfährt Wert eines anderen

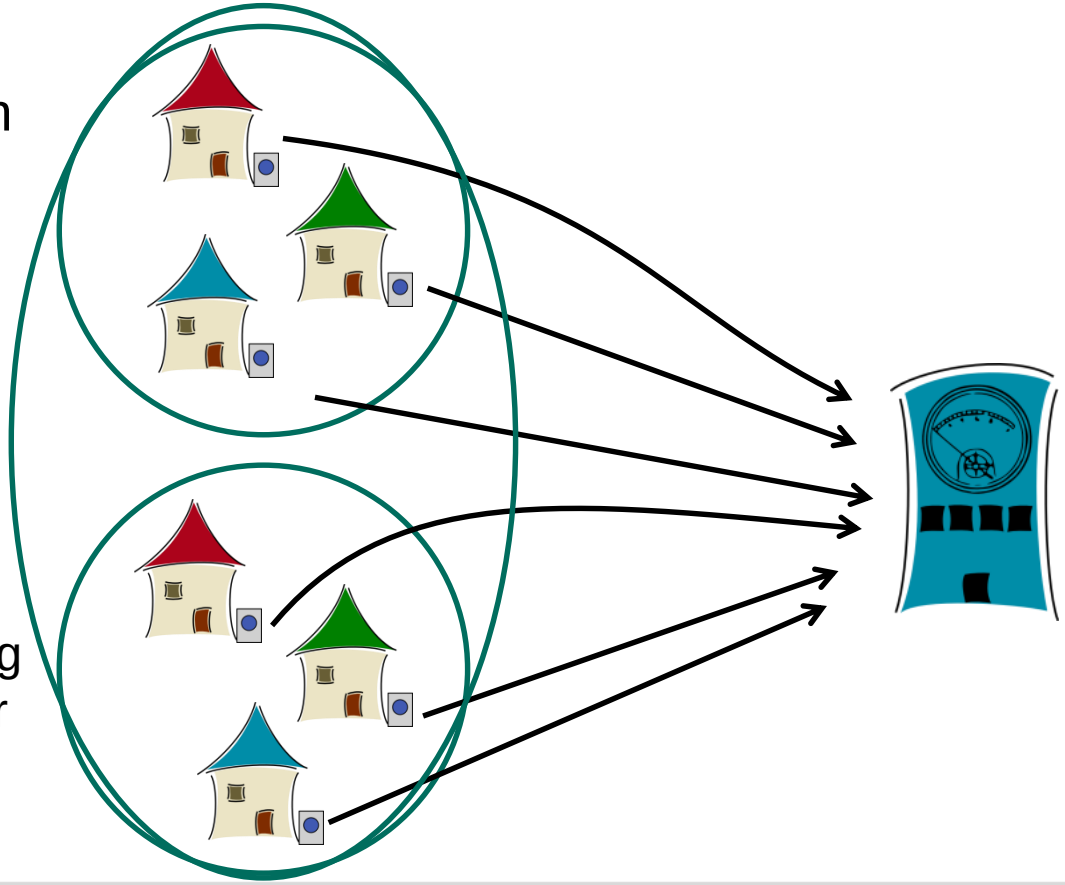


Probleme vieler Ansätze zur Kooperation von Stromzählern untereinander

- Teure Kryptografie
  - Ressourcenbeschränkte Hardware ungeeignet für benötigte Rechenoperationen
- Robustheit gegen Störungen
  - Störungen
    - Hardware / Software Probleme von Stromzählern
    - Kommunikationsinfrastruktur (DSL-Anschluss)
  - Mögliche Folgen
    - Verfälschte Messergebnisse
    - Totalausfall der gesamten Messung
    - Eingeschränkter Schutz der Privatsphäre
- Strukturvorgaben (bspw. Baum) ermöglichen Attacken durch MDL
  - Platzierung von korrumpierten Stromzählern an sensiblen Punkten
- Lange Dauer der Kooperation pro Messwert
  - Nur lange Messintervalle realisierbar

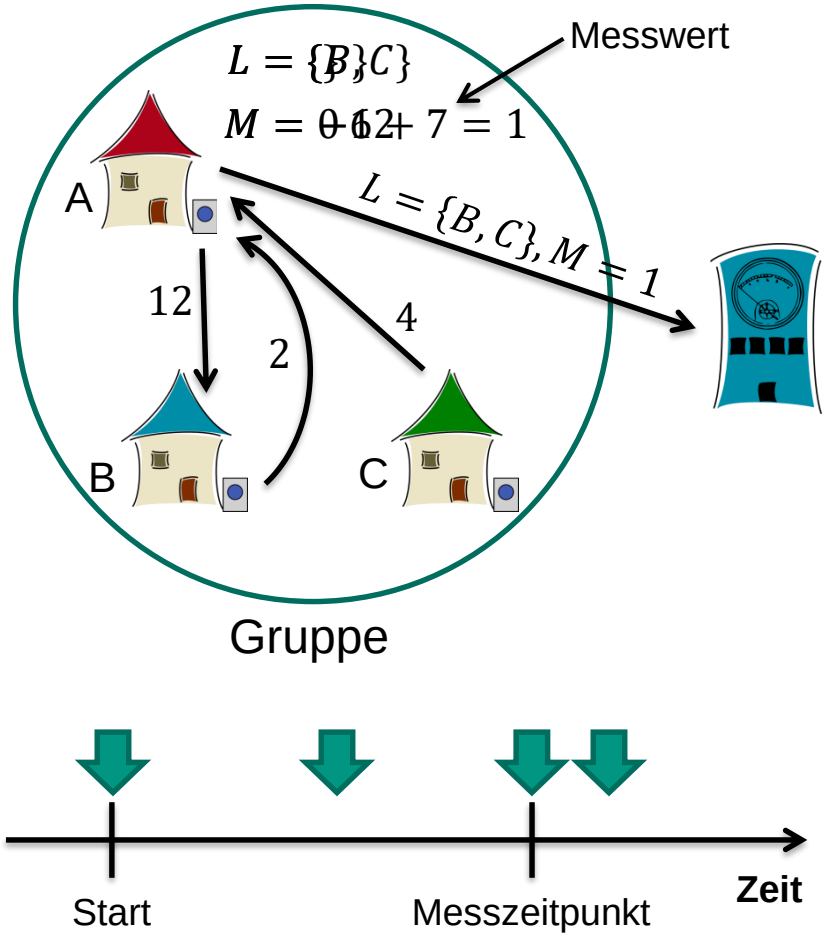
Eigene Arbeiten

- Privatsphärengerechtes Smart Metering Protokoll SMART-ER: SMART with Exactness and Robustness [Fin2014]
  - Basiert auf SMART, Verfahren für drahtlose Sensornetze
- Grundkonzept: Einteilung der Stromzähler in Gruppen
  - Vom Messdienstleister durchgeführt
  - Konfigurierbare Gruppengröße
  - Höhere Robustheit
  - Weniger Overhead
- Innerhalb von Gruppen
  - Kooperation zur Ermittlung maskierter, ungefährlicher Messwerte mit korrektem Aggregat



SMART-ER

- Pro Messintervall
  - Austausch von Zufallswerten innerhalb von Gruppen
  - Speichern der Kommunikations-partner (Abhängigkeiten)
  - Berechnen maskierter Messwerte
  - Senden maskierter Messwerte und Abhängigkeiten an Messdienstleister
  - Eventuelle Bereinigung empfangener Messwerte durch Messdienstleister
- Alle Zahlen aus Restklassenring $\mathbb{Z}/q\mathbb{Z}$
  - Bspw. $q = 2^{64}$



SMART-ER Animation in Einzelschritten

1. Ausgangssituation: L leer und $M = 0$
2. A sendet 12 an B
   1. → $M = -12$
   2. → $L = \{B\}$
3. A empfängt 2 und 4 von B, bzw. C
   1. → $M = -6$
   2. → $L = \{B, C\}$
4. Messwert wird ermittelt → 7
5. Datenübertragung an Messdienstleister
   1. $L = \{B, C\}$
   2. $M = -6 + 7 = 1$

→ Messdienstleister bildet Summe über alle (verschleierten) Messwerte und erhält dadurch das Aggregat

Evaluation von Privatsphärenschutz

- Smart Meter Privacy Break Game (SMPBG)
- Vorgehen
  - Angreifer stellt zwei Szenarien mit gleichem Aggregat

[Bohli2010]

  - Verteidiger wählt geheim ein Szenario und führt Smart Metering durch
  - Angreifer erhält Informationen entsprechend Angreifermodell
    - Bspw. Alle Kommunikation abhören, Messdienstleister korrumpieren

- Maß für Privatsphärenschutzes: Wahrscheinlichkeit für Angreifer richtiges Szenario zu wählen
  - Wahrscheinlichkeit $\frac{1}{2}$ = bester Schutz → Angreifer kann nur raten
  - Wahrscheinlichkeit 1 = kein Schutz → Angreifer kann immer richtig wählen

Beispiel: Korrumpierter Messdienstleister

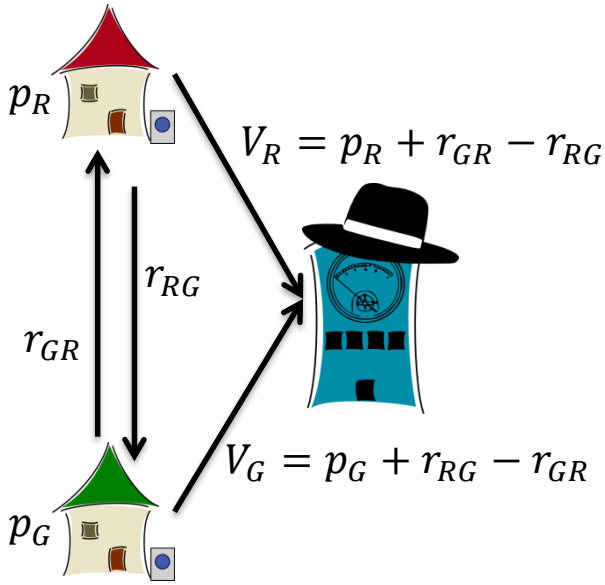
- Messdienstleister ist korrumpiert (Symbol Hut)
- Angreifer kann mit abgegebenen, maskierten Werten arbeiten
  - Roter Haushalt gibt V_R ab
  - Besteht aus Messwert p_R, empfangener Zufallszahl r_{GR} und versendeter Zufallszahl r_{RG}
  - $V_R = p_R + r_{GR} - r_{RG}$
  - Summe $V_R + V_G =$ gewünschtes Aggregat
  - Differenz d interessant

$$d = V_R - V_G = p_R + r_{GR} - r_{RG} - p_G - r_{RG} + r_{GR}$$
$$= p_R - p_G + 2r_{GR} - 2r_{RG} \quad \text{mit } r_{GR,}\, r_{RG} \in \mathbb{Z}/q\mathbb{Z}$$

Sei $d^{(1)} := d$ in Szenario 1, $d^{(2)} := d$ in Szenario 2

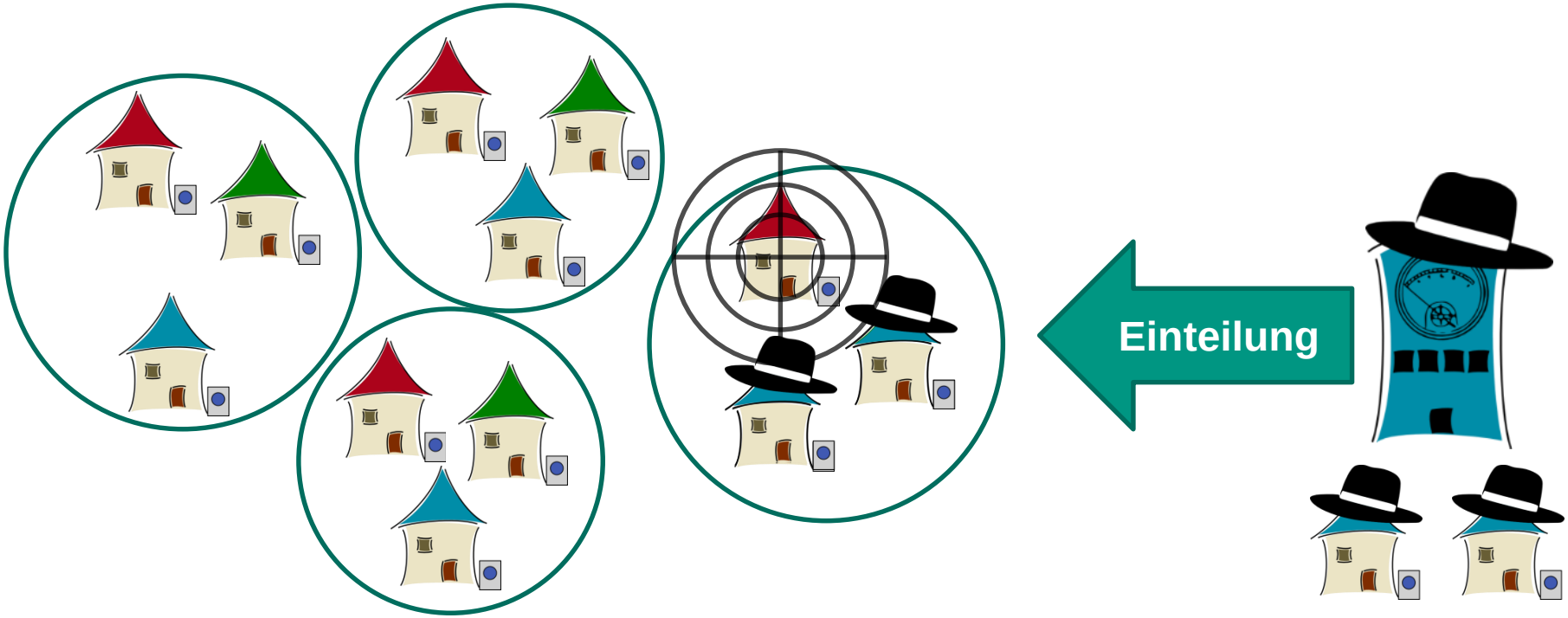
$$\Rightarrow P\left(d^{(1)} = y\right) = P\left(d^{(2)} = y\right) \forall\, y \in \mathbb{Z}/q\mathbb{Z}$$

$\Rightarrow$ Angreifer kann nur raten, $q.e.d.$



SMART-ER: Verbleibendes Problem

- Problem: Messdienstleister nimmt Gruppenbildung vor
  - Angriff mittels korrumpierter Stromzähler möglich
  - Benötigt Gruppengröße−1 korrumpierte intelligente Stromzähler



- Bei Gruppengröße 3 kann ein korrumpierter Messdienstleister mit zwei korrumpierten Stromzählern einen Angriff durchführen.

Gegenmaßnahme: Dezentrale Gruppenbildung

- Vermeiden von Einfluss auf Gruppenbildung durch Messdienstleister
- Smart Meter Speed Dating
  - Gruppen werden von Stromzählern selbst, dezentral ermittelt
  - Ohne Einfluss durch Messdienstleister / wenig Einfluss der Stromzähler
  - Hohe Robustheit durch kleine Gruppengröße
  - Wenig Anforderungen an Hardware (Implementiert in Sensornetz)
  - Skalierbarkeit eingeschränkt durch Speicher (linearer Aufwand)
- Elderberry
  - Baumbasierter Ansatz mit strukturiertem P2P-Overlay
  - Dezentrale Aggregation → Entlastung der Datensenke
  - Deutlich komplexer als Smart Meter Speed Dating
  - Sehr gute Skalierbarkeit

Beispiel: Privatheit im Smart-Traffic

- Das smarte Auto

Sensoren

- GPS
- Straßenbelag
- Wetter
- Video
- …

Kommunikation

- 3G/LTE
- Car2X Radio
- FM Radio

Intelligenz

- Navigation
- Automatische Notbremsung
- Bald smart genug, um selbst zu fahren?

Smart-Traffic-Anwendungen (I)

- Verkehrssicherheit
  - Kleiner geografischer Scope
  - → Lokale Car2Car / Car2X Kommunikation

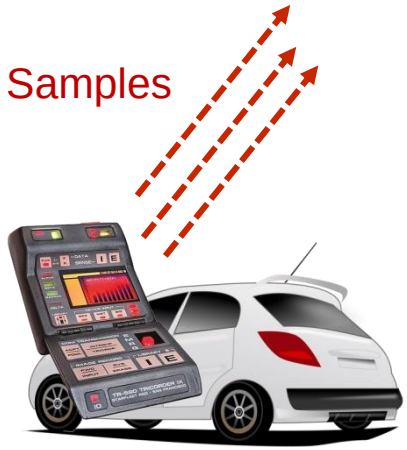
Die Straße hier ist rutschig!

Ich komme um die Ecke!

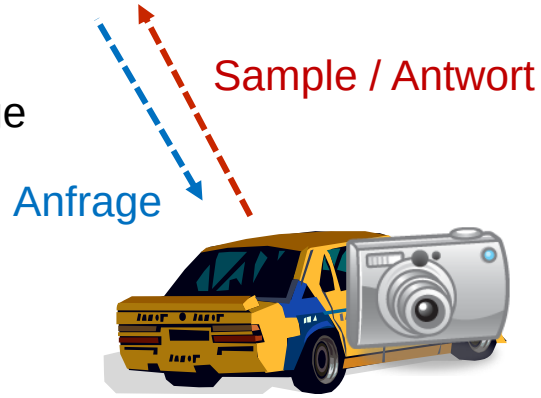
- Verkehrsoptimierung
  - Entwicklung der automatischen Fahrzeugnavigation
    - Statisch – Routenplanung mit fixem Kartenmaterial
    - Adaptiv – Einbeziehung von aktuellem Verkehrszustand
    - Koordiniert – Einbeziehung von geplanten Routen anderer Verkehrsteilnehmer, Veröffentlichung eigener Pläne

Smart-Traffic-Anwendungen (II)

- Mobiles Erfassen von Sensordaten
  - Wetter, Luftqualität…
  - Auch Verkehr
    → Datenquelle für adaptive Routenplanung
  - Herausforderung: Samples müssen Position erhalten



- Vehicular Clouds
  - Fahrzeuge als Diensterbringer
  - Beispiel: Pics-On-Wheels [Weng2013]
    - Dienstnutzer will aktuelles Foto von bestimmten Koordinaten (z.B. sein eigenes Haus)
    - Vorbeifahrende Fahrzeuge schießen Foto auf Anfrage
  - Herausforderung: wie Anfragen an geeignete Fahrzeuge weiterleiten?
    - Eignung hängt von Position der Fahrzeuge ab



Positionsbezogene Daten

- Benötigt bei vielen Smart-Traffic-Anwendungen
  - Bestimmung des Verkehrszustands → Floating Car Data (FCD)
  - Herkunft von Sensordaten
  - Bestimmung geeigneter Fahrzeuge für Anfragen

- Aber nicht nur
  - Beispiel Lifelogging
  - Beispiel Location Based Services
    - Suche nach Inhalten basierend auf geografischen Lokationen
      - Verrät oft eigenen Standort!
    - Check-In Anwendungen (z.B. Foursquare, Facebook)
    - Anzeige der Positionen von Freunden (z.B. Google Latitude)
  - Beispiel Crowdsensing

Positionssamples

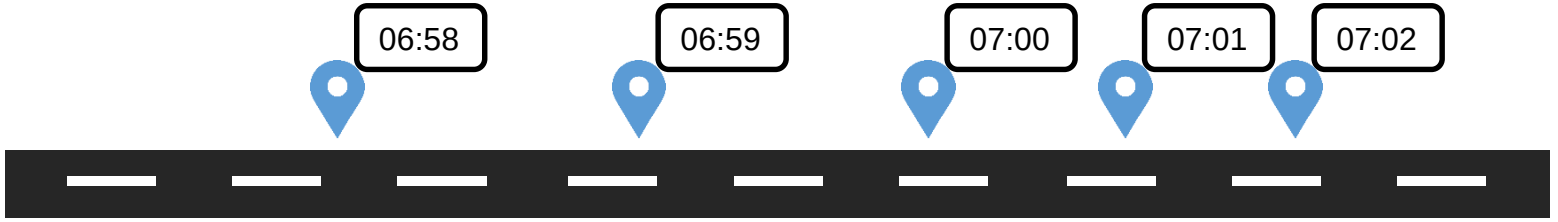
- Position
  - Geografische Repräsentationen
    - Breitengrad/Längengrad, z.B. (49.012421, 8.408077)
  - Kontext-spezifische Repräsentationen
    - „Im SCC-Gebäude“
    - Hierarchische Modelle, z.B. „Deutschland/Karlsruhe/KIT/Gebäude 20.20“
  - variable Präzision
- Zeit
  - …zu der Sample erstellt wurde (muss nicht geteilt werden)
  - …zu der Sample geteilt wurde (muss nicht gleich Erstellungszeit sein)

Positionssamples: Beispiel

- Alice fährt eine bestimmte Strecke ab

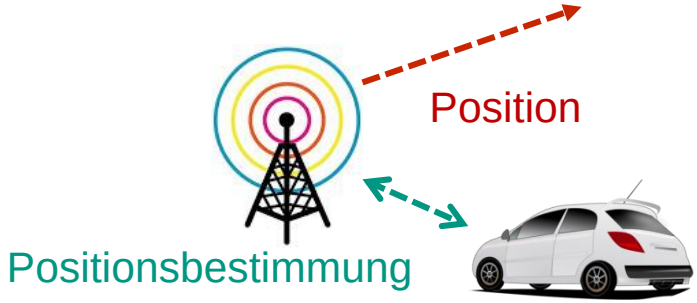
- Zu bestimmten Zeitpunkten teilt sie Ihre aktuelle Position

- Der Dienstanbieter erhält also z.B. folgende Informationen

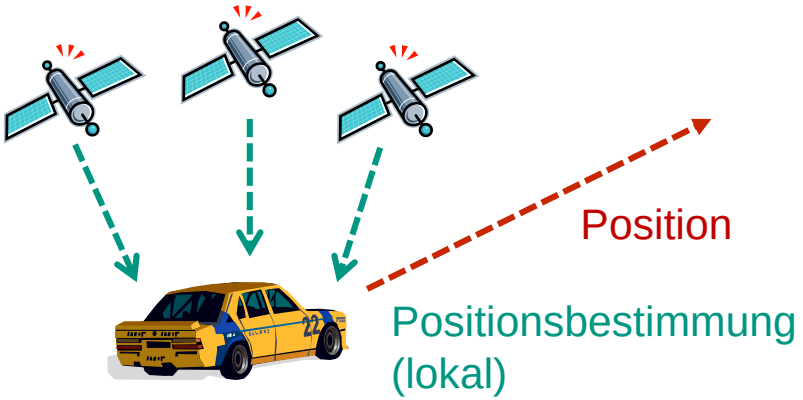


Positionsbestimmung

- Implizit
  - Bsp.: Nutzer kommuniziert mit Mobilfunk-Basisstation → ist in der Nähe
  - Bsp.: automatische Erfassung des KFZ-Kennzeichens
  - Nutzer hat keinen Einfluss auf Präzision
    - Potentiell jedoch auf Samplingrate!



- Explizit
  - Bsp.: GPS
  - Nutzer hat Einfluss auf Genauigkeit → Verschleierung o.ä. möglich



Herausforderungen für die Privatsphäre

- Präzise, zu Nutzern zuordenbare Positionssamples lassen potentiell weitläufige Rückschlüsse über Lebensgewohnheiten zu
  - Wohnort, Arbeitsplatz
  - Hobbies und soziale Kontakte
  - Krankenhausbesuche
  - Politische Orientierung

  …

→ Besonders schutzwürdig!

- Selbst wenn Positionssamples nicht explizit zu Nutzern zuordenbar…
  - Implizite Zuordnung mithilfe von Kontextwissen möglich
    - z.B. Wohnort und Arbeitsplatz  [Golle2009]
    - z.B. Fotos  [Tock2014]

Zum Ausprobieren

- Folgende Positionssamples wurden u.a. von einem unbekannten Kraftfahrzeug aufgezeichnet (an einem Arbeitstag)
  - t sei Zeitpunkt an dem Sample geteilt wurde
  - p sei Position als (Breitengrad, Längengrad) in dezimaler Schreibweise

t = 06:57;
p = (52.520266, 13.395238);

t = 07:05;
p = (52.525265, 13.368201);

t = 07:17;
p = (52.520652, 13.369433);

- Hausaufgabe
  - Was für Rückschlüsse erlauben diese Samples auf die Insassen des Fahrzeugs? (Antwort über E-Mail oder http://tmfeedozvp6v65nf.onion ☺)

Privatsphärenschutz in Smart Traffic (Überblick)

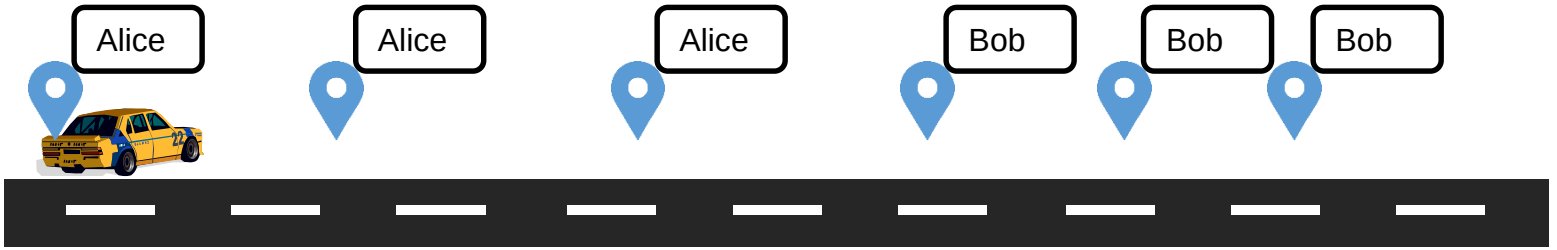
- Nicht-Identifizierbarkeit von Positionssamples
  - Verwendung von Pseudonymen
- Unverkettbarkeit von Positionssamples
  - Pseudonyme müssen oft gewechselt werden
  - Effektiver Pseudonymwechsel nicht trivial
    - Positionssamples vor und nach Wechsel können Zuordnung ermöglichen
- Verschleierung von Positionssamples
  - Idee: selbst bei einer existierenden Zuordnung zu Nutzern Bedenken für Privatsphäre minimieren

Trennung von Positionssamples u. Identitäten

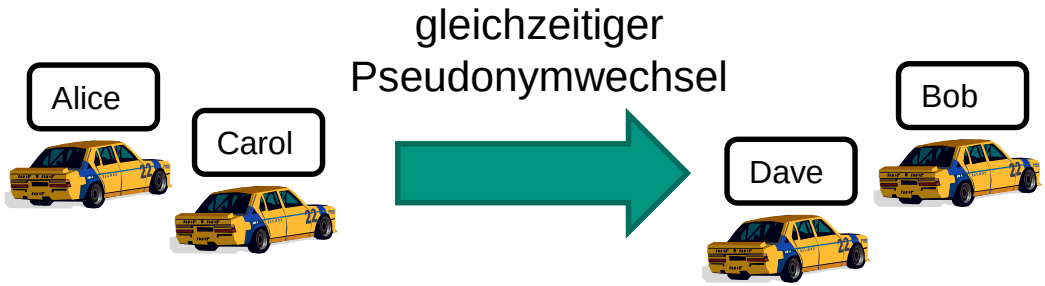
- Im Grunde ähnliche Ansätze wie bei Samples allgemein
- Vollständige Anonymität nicht immer möglich
  - Ermöglichen unbestrafbaren Missbrauch der Dienste
- Verwendung von Pseudonymen
  - Limitierte Anzahl pro Teilnehmer
  - Verfahren zum Ausschließen böswilliger Teilnehmer
    - z.B. indem Zuordnung durch Vertrauensanker möglich bleibt
      - Bsp.: Car2X-Kommunikationssysteme für die Verkehrssicherheit
  - Vorsicht beim Pseudonymwechsel!
    - Zuordnung über Kommunikationsadresse
    - Zuordnung über Positionssamples vor und nach Wechsel

Pseudonymwechsel (I)

- Pseudonymwechsel nötig, um Positionssamples desselben Nutzers nicht leicht zusammengruppieren zu können
- Problem: wenn Angreifer den Pseudonymwechsel beobachtet, ist dieser ggf. zwecklos
  - Bsp.: wer ist wohl Bob?



  - Besser: zufällig lange „Funkstille“ nach Pseudonymwechsel
  - Noch Besser: ein Mix



Pseudonymwechsel (II)

- Mix Zonen

[Beres2004]

  - Vordefinierte Gebiete mit
    - hohem Verkehrsaufkommen
    - niedriger relativer Geschwindigkeit von Verkehrsteilnehmern zueinander
  - Ideal geeignet sind z.B.
    - Kreuzungen
    - Parkplätze

- Mix-Zone-Ansatz
  - Verkehrsteilnehmer wechseln nur dann ihre Pseudonyme, wenn Sie eine Mix Zone passieren
    - Innerhalb der Mix Zone werden keine Positionssamples geteilt (Funkstille)
  - Dadurch ganz natürliches mixen der Pseudonyme
  - Je mehr Mix Zonen passiert werden, desto häufiger wird gemixt, desto geringer ist die Sicherheit, mit der ein Angreifer Positionssamples miteinander in Verbindung bringen kann

Verschleierung von Positionssamples

- Anwendungsabhängig
  - Bei Staumeldungen brauche ich genaue Position zur richtigen Zeit
  - Bei der Erfassung der durchschnittlichen Luftqualität pro Monat nicht
- Zeitliche Verschleierung
  - Positionsupdates zufällig verspäten
  - Update-Intervalle vergrößern
- Räumliche Verschleierung
  - Positionswerte um zufälligen Faktor „verschieben“
  - Präzision von Positionen verringern
    - „Deutschland“ statt „Karlsruhe“
  - Dummy-Werte
    - Mehrere falsche Positionen zusammen mit echter
    - Nicht immer sinnvoll einsetzbar

Zeitliche Verschleierung von Positionssamples

- Beispiel: Virtual Trip Lines [Hoh2008]
  - Vordefinierte „virtuelle Induktionsspulen“ (Trip Lines) im Straßengraph
    - an privatsphärentechnisch unbedenklichen Orten (Autobahn, Kreuzung…)
  - Positionssamples werden nur beim Passieren von Trip Lines geteilt
    - also nie an privatsphärentechnisch bedenklichen Orten

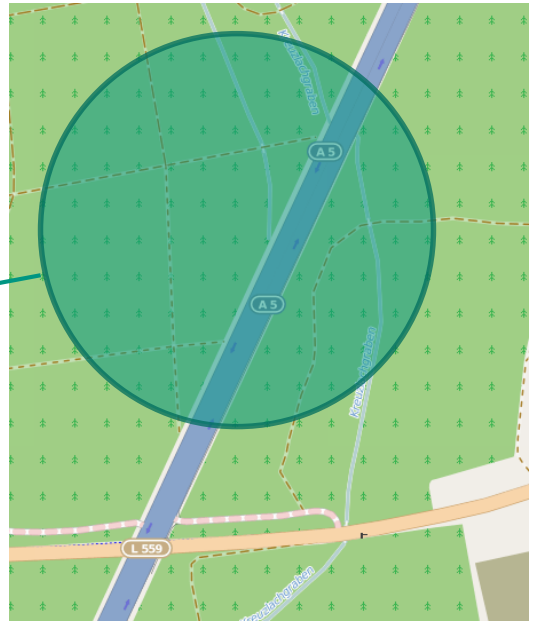
  - Dadurch auch ganz natürliches Mixen von Pseudonymen möglich
    - Alles außer den Trip Lines ist eine Mix Zone
    - Durch geringe Samplingrate wird Identifikation über Fahrverhalten erschwert

Räumliche Verschleierung v. Positionssamples

- Herausforderungen im Smart-Traffic-Kontext
  - Map-Matching
    - Fahrzeuge fahren i.d.R. auf Straßen → vermindert Verschleierung
  - Kontext-Wissen
    - Auf Autobahnen parkt niemand
    - Auf Landstraßen fahren nur wenige über 120 km/h

    → Kontext-Wissen vermindert Verschleierung

Das beobachtete Fahrzeug bewegt sich mit ~130 km/h in südlicher Richtung und befindet sich irgendwo in diesem Gebiet…



Quelle: openstreetmap.org

  - Gute räumliche Verschleierung…
    - …verringert am Ende Dienstqualität

→ Räumliche Verschleierung (vor allem im Smart-Traffic-Kontext) nicht ganz so einfach…

Eigene Arbeiten: Geocast

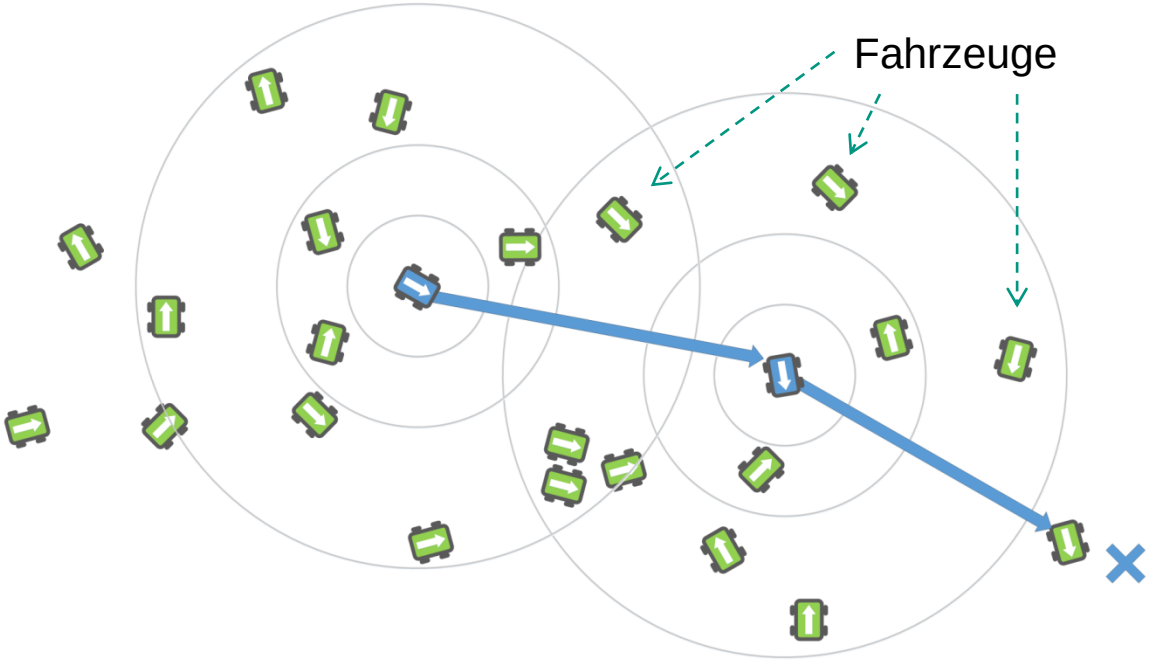
- Problem: ich möchte Anfragen an (mir unbekannte) Fahrzeuge innerhalb eines bestimmten geografischen Gebiets schicken
  - z.B. um lokales Verkehrsbild zu bekommen oder während Parkplatzsuche
  - z.B. im Kontext von Vehicular Clouds
  - Auch als Geocast bezeichnet

- Problem mit Geocast über lange Distanzen: Knoten (Fahrzeuge) müssen ihre genaue Positionen mit jemandem teilen, um korrekt erreichbar zu sein
  - Naiver Ansatz: zentraler Server
    - Empfängt Positionsupdates von allen Teilnehmern
    - Verteilt alle Nachrichten
    - → zentrale Datensenke, skaliert schlecht und ist attraktives Angriffsziel

Eigene Arbeiten: Geocast mit OverDrive

[Flor2016]

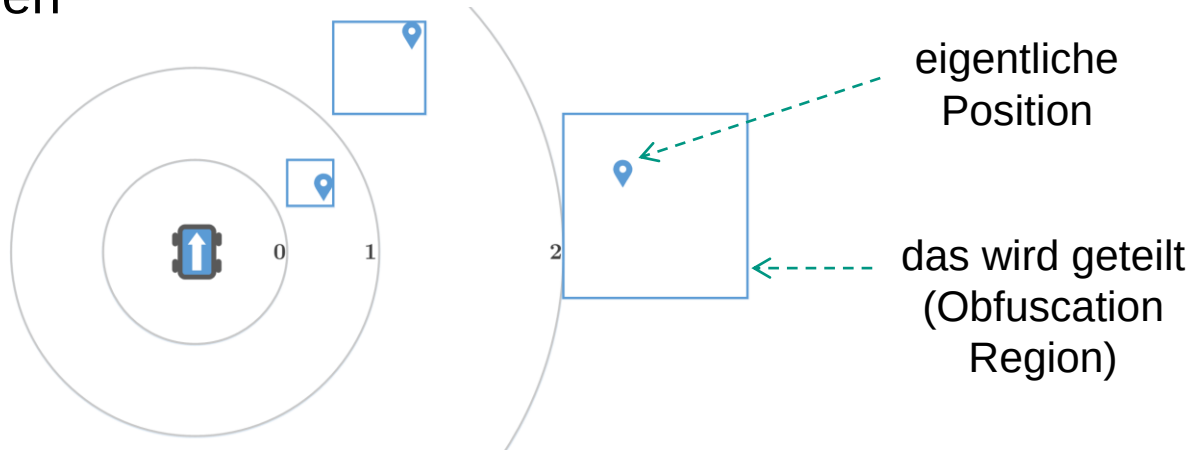
- Unser Ansatz: OverDrive
  - Peer-to-Peer Overlay-Netz direkt zwischen Fahrzeugen
  - Overlay-Nachbarn tauschen Positionsdaten aus
    → keine zentrale Datensenke, niemand hat „volle Sicht“
  - Nachrichten werden über mehrere „Hops“ weitergeleitet



Eigene Arbeiten: Privatheit in OverDrive

[Flor2016]

- Pseudonyme pro Fahrzeug
- Genaue Positionsdaten werden nur mit nahgelegenen Nachbarn geteilt
  → Datenlokalität
- Weit entfernte Nachbarn bekommen stark (um mehrere Kilometer) verfälschte Positionen



- Von Angreifern kontrollierte Knoten, die ihre Position fälschen, werden mithilfe von Proximity Tests erkannt
  - z.B., indem Zugehörigkeit zur selben GSM-Zelle getestet wird

Eigene Arbeiten: vorausschauende Navigation

- Vorausschauende, kooperative Routenplanung
  - Teilnehmer veröffentlichen geplante Routen und beziehen Pläne der restlichen Teilnehmer mit ein
    - Somit können z.B. Staus besser vorhergesehen und verhindert werden

- Problem
  - Wie eigene Pläne bezüglich einer knappen Ressource (z.B. Platz auf einer beliebten Straße oder einem Verkehrsnadelöhr) veröffentlichen…
  - …ohne dass meine Veröffentlichung auf mich zurückzuführen ist?
  - …ohne für Missbrauch durch bösartige Nutzer anfällig zu sein?

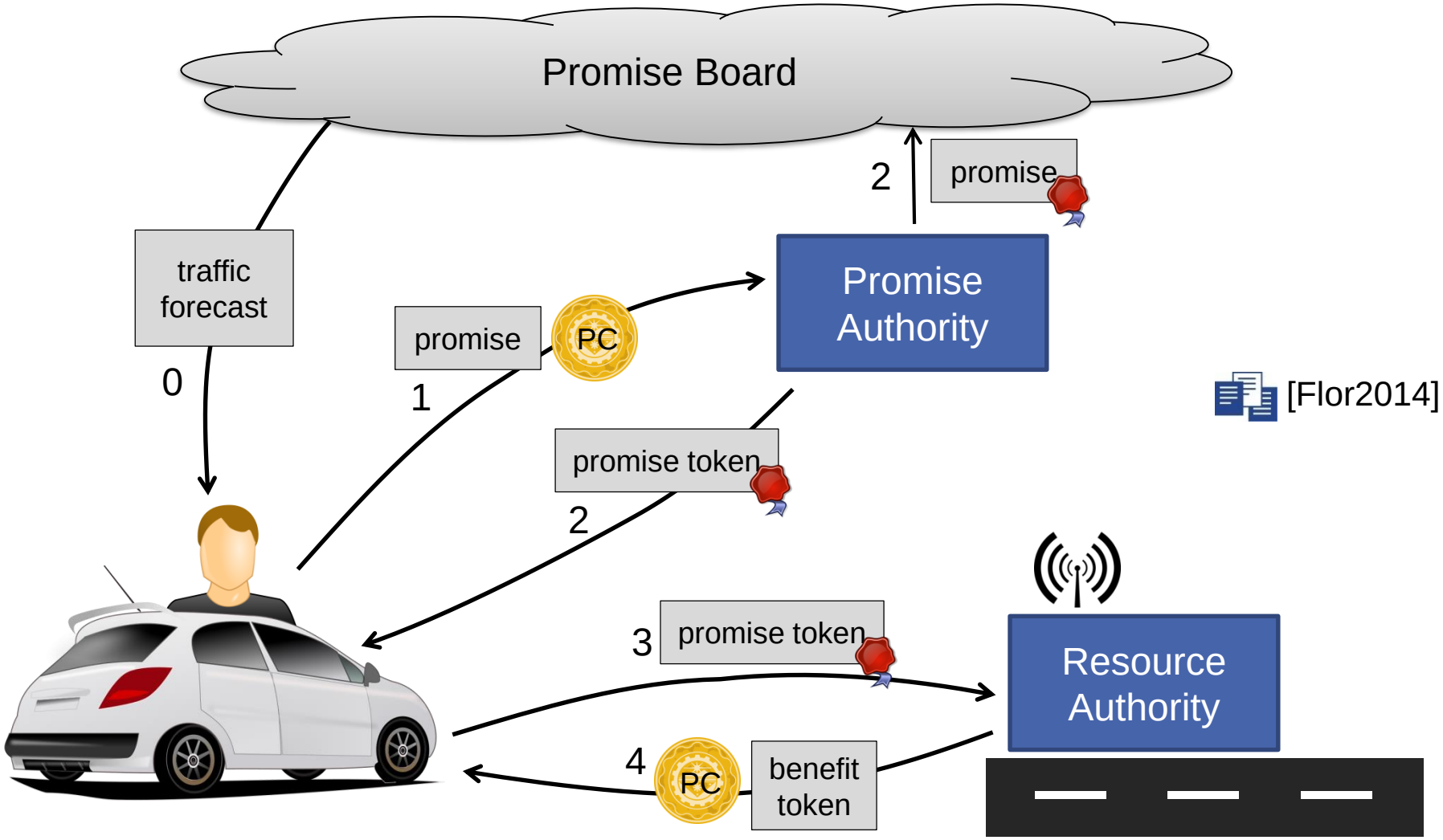
Eigene Arbeiten: vorausschauende Navigation

- Unser Ansatz: Privacy-Preserving Cooperative Route Planning
- Verwendung von virtuellen, anonymen Tokens zur Zugangskontrolle
  - Promise Coins (PCs)
  - Verwendung von blinden Signaturen, somit bleiben Nutzer anonym

- Ablauf
  - Bei der Anmeldung bekommt jeder legitimer Nutzer einen Pool an PCs
  - Das veröffentlichen von Plänen „kostet“ PCs
  - Wenn Pläne so erfüllt wurden, wie versprochen, werden verwendete PCs wieder erstattet
    - Evtl. zusammen mit Belohnung, z.B. einer Maut-Ermäßigung
  - Lügende Nutzer bekommen keine neuen PCs und werden somit schnell ausgeschlossen

[Flor2014]

Eigene Arbeiten: Privacy-Preserving Coop. Route Planning



[Flor2014]

Vergleich der betrachteten Szenarien

	Smart Traffic / Geocast	Smart Grid / Smart Metering
Zu schützende Daten	Lokationsdaten der Nutzer	Personenbezogene Messwerte
Motivation des Angreifers	Erstellung von Bewegungsprofilen	Rückschlüsse auf Nutzerverhalten
Hintergrundwissen des Angreifers	Zuordnung einzelner Lokationsdaten (z.B. Wohnort) zu Identitäten	Zuordnung von Messwerten zu Nutzerverhalten
Funktionale Anforderungen des Dienstes	Adressierung von Nutzer anhand deren Lokation	Zeitnahe Erfassung eines korrekten Aggregats
Geeignete PETs	Kooperative Verschleierung (OverDrive) + Pseudonymisierung ohne TTP (BitNym)	Kooperative Aggregation (SMART-ER, Elderberry)

Zusammenfassung

- Schutz der Privatsphäre stellt große Herausforderung im IoE dar
  - Ubiquitäre Erfassung personenbezogener Daten rund um die Uhr
  - Erfassung in besonders sensitiven Lebensbereichen (Smart Home, …)
  - Geräte mit beschränkten Ressourcen
  - Angreifer kann Geräte korrumpieren

- Schutz der Privatsphäre durch Privacy Enhancing Technologies (PETs)
  - Abhängig vom konkreten Anwendungsszenario
    - Funktionale Anforderungen des angebotenen Dienstes
    - Angreifermodell und Vertrauensmodell

- Bisher meist Vertrauen in zentralen Dienstanbieter oder vertrauenswürdige dritte Partei erforderlich
  - Aktuelle Forschungsarbeiten zu PETs, die lediglich verteiltes Vertrauen erfordern

Die von uns zur Erstellung der Folien genutzte

LITERATUR

Literatur

[Bohli2010] J.-M. Bohli, C. Sorge, and O. Ugus, “A Privacy Model for Smart Metering” in IEEE International Conference on Communications Workshops (ICC), 2010, pp. 1–5.

[Beres2004] A. R Beresford und F. Stajano. Mix zones: User privacy in location-aware services. In Proceedings of the 2nd IEEE Annual Conference on Pervasive Computing and Communications Workshops (PERCOMW04), pages 127–131. IEEE, März 2004.

[Danez2015] Danezis, George, et al.: Privacy and Data Protection by Design - from policy to engineering; Report, European Union Agency for Network and Information Security (ENISA), January 2015.

[DY83] D. Dolev, A. Yao, On the security of public key protocols, IEEE Transactions on Information Theory, Vol. 29(2), S. 198–208, 1983

[Fin2014] S. Finster and I. Baumgart, “SMART-ER: peer-based privacy for smart metering” in IEEE INFOCOM Workshop on Communications and Control for Smart Energy Systems, 2014, pp. 642–647.

[Flor2014] M. Florian, S. Finster, und I. Baumgart: Privacy-Preserving Cooperative Route Planning; IEEE Internet of Things Journal, 1(6):590–599, Okt. 2014

[Flor2016] M. Florian, F. Pieper, und I. Baumgart: Establishing location privacy in decentralized long-distance geocast services; Ad Hoc Networks, 37, Part 1:110–121, Feb. 2016

[Golle2009] Golle, Philippe und Kurt Partridge; On the Anonymity of Home/Work Location Pairs; Pervasive Computing, Springer, 2009

Literatur

[Hoe2014] J.-H. Hoepman, ”Privacy Design Strategies”; SEC 2014, IFIP AICT 428, pp. 446-459, 2014.

[Hoh2008] Hoh, Baik, et al.; Virtual trip lines for distributed privacy-preserving traffic monitoring; 6th Int. Conf. on Mobile systems, applications, and services (MobiCom), ACM, 2008

[Jeske2013] Jeske, Tobias; Floating Car Data from Smartphones: What Google And Waze Know About You and How Hackers Can Control Traffic; BlackHat Europe, 2013; https://media.blackhat.com/eu-13/briefings/Jeske/bh-eu-13-floating-car-data-jeske-slides.pdf

[Krebs2016] Brian Krebs: Hacked Cameras, DVRs Powered Today’s Massive Internet Outage; Krebs on Security, Okt. 2016; https://krebsonsecurity.com/2016/10/hacked-cameras-dvrs-powered-todays-massive-internet-outage/

[Li2011] F. Li, B. Luo, and P. Liu, “Secure and privacy-preserving information aggregation for smart grids,” Int. Journal on Security and Networks, vol. 6, no. 1, p. 28, 2011.

[Naka2008] Satoshi Nakamoto; Bitcoin: A Peer-to-Peer Electronic Cash System; 2008

[Pure2015] V. Pureswaran und P. Brody; Device democracy: Saving the future of the Internet of Things; IBM Global Business Services Executive Report , IBM, 2014

[Tock2014] Tockar, Anthony; Riding with the Stars: Passenger Privacy in the NYC Taxicab Dataset; Neustar Research, 2014; http://research.neustar.biz/2014/09/15/riding-with-the-stars-passenger-privacy-in-the-nyc-taxicab-dataset/

[Weng2013] Weng, Jui-Ting, Ian Ku und Mario Gerl; Surveillance service on the open mobile cloud; 10th Conf. on Wireless On-demand Network Systems and Services (WONS), IEEE, 2013