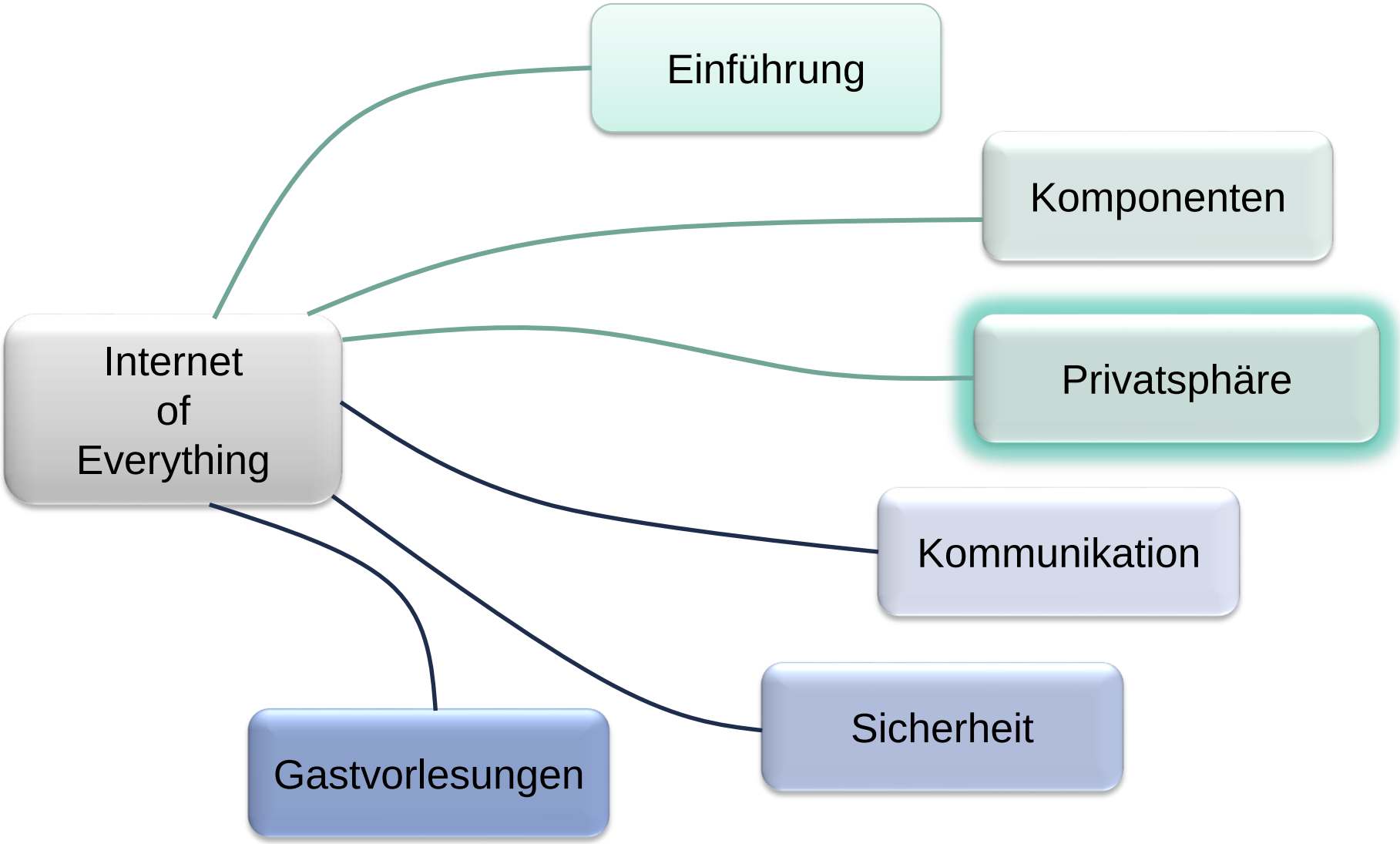
Karlsruher Institut für Technologie

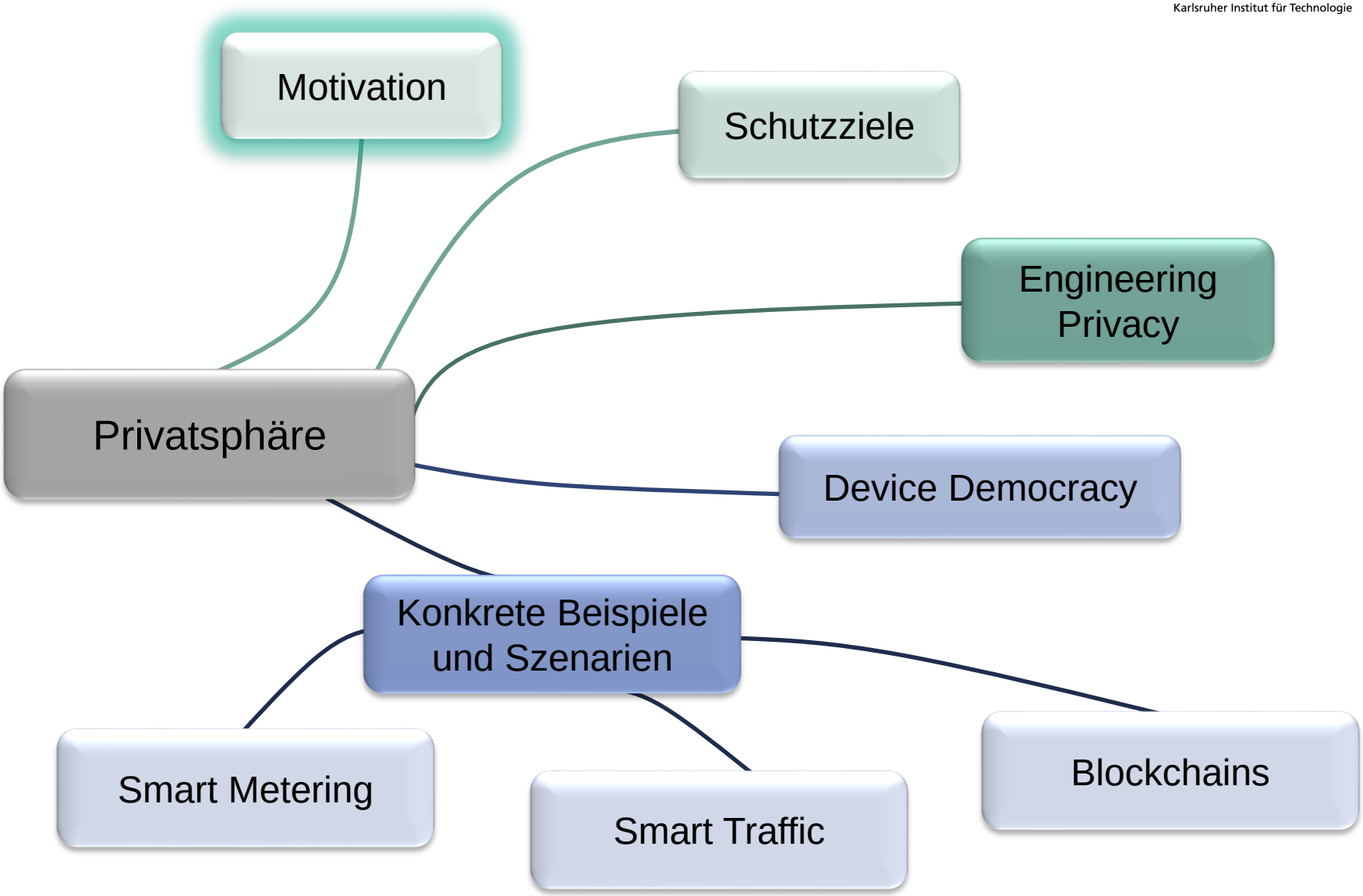
Vorlesung Internet of Everything
Wintersemester 2017/18

3. Privatsphäre

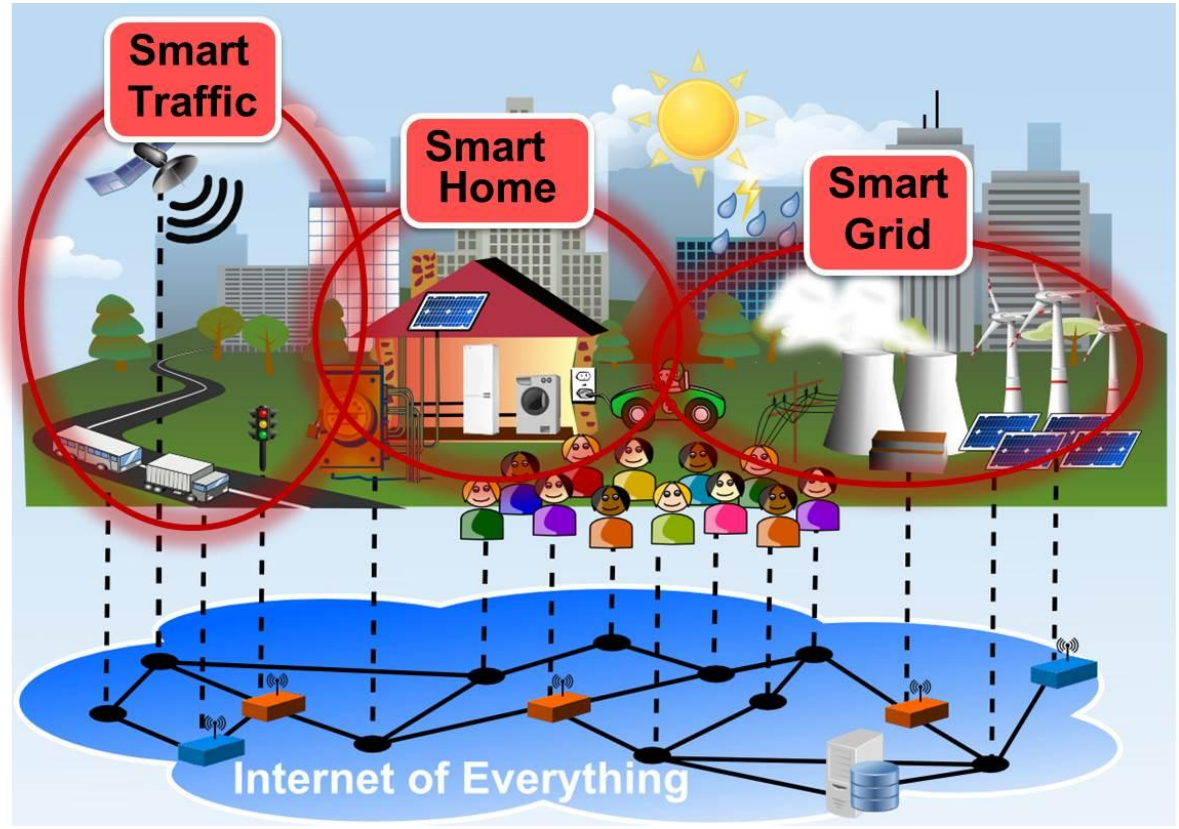
Institut für Telematik, Prof. Zitterbart

Inhalte der Vorlesung




Motivation
Schutzziele
Engineering Privacy
Privatsphäre
Device Democracy
Konkrete Beispiele und Szenarien
Smart Metering
Smart Traffic
Blockchains

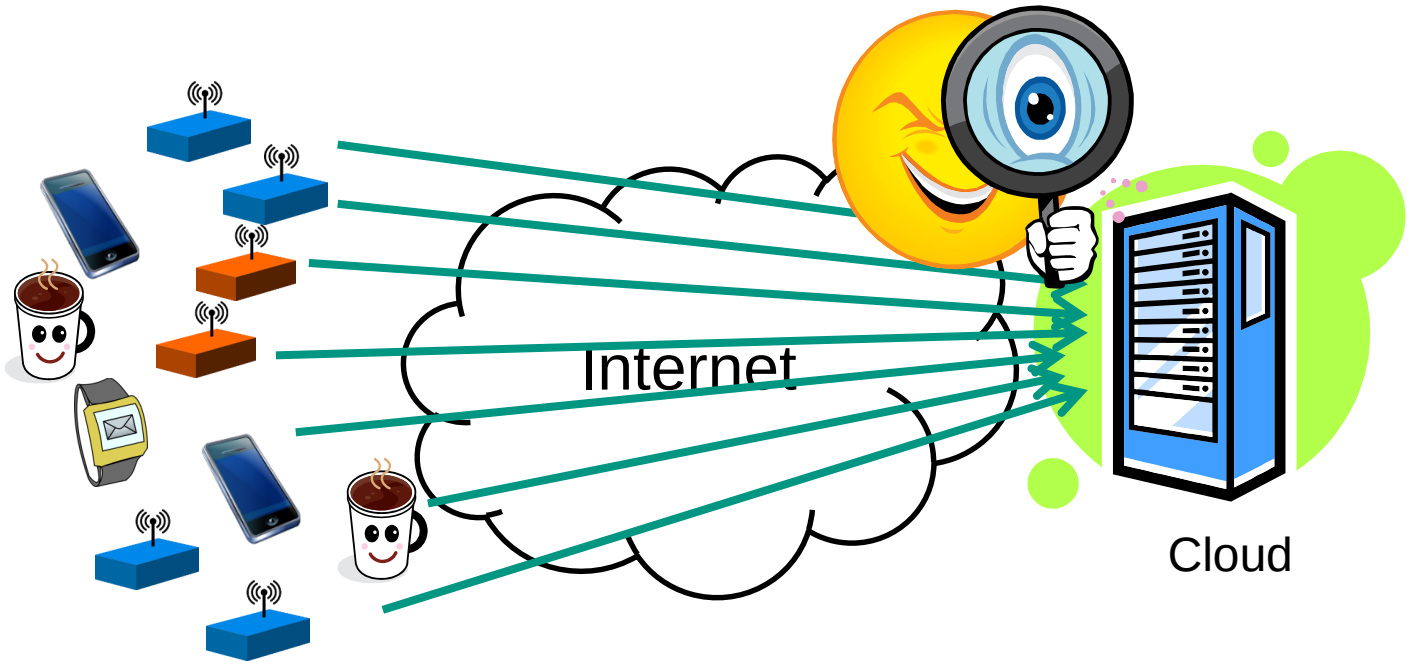
Dienste im Internet of Everything



→ Vielzahl von Geräten („Dingen“) erfassen Daten aus ihrer Umgebung, tauschen Daten miteinander aus, kombinieren und analysieren im System vorhandene Daten

„Internet“ weiß mehr über mich als ich selbst!

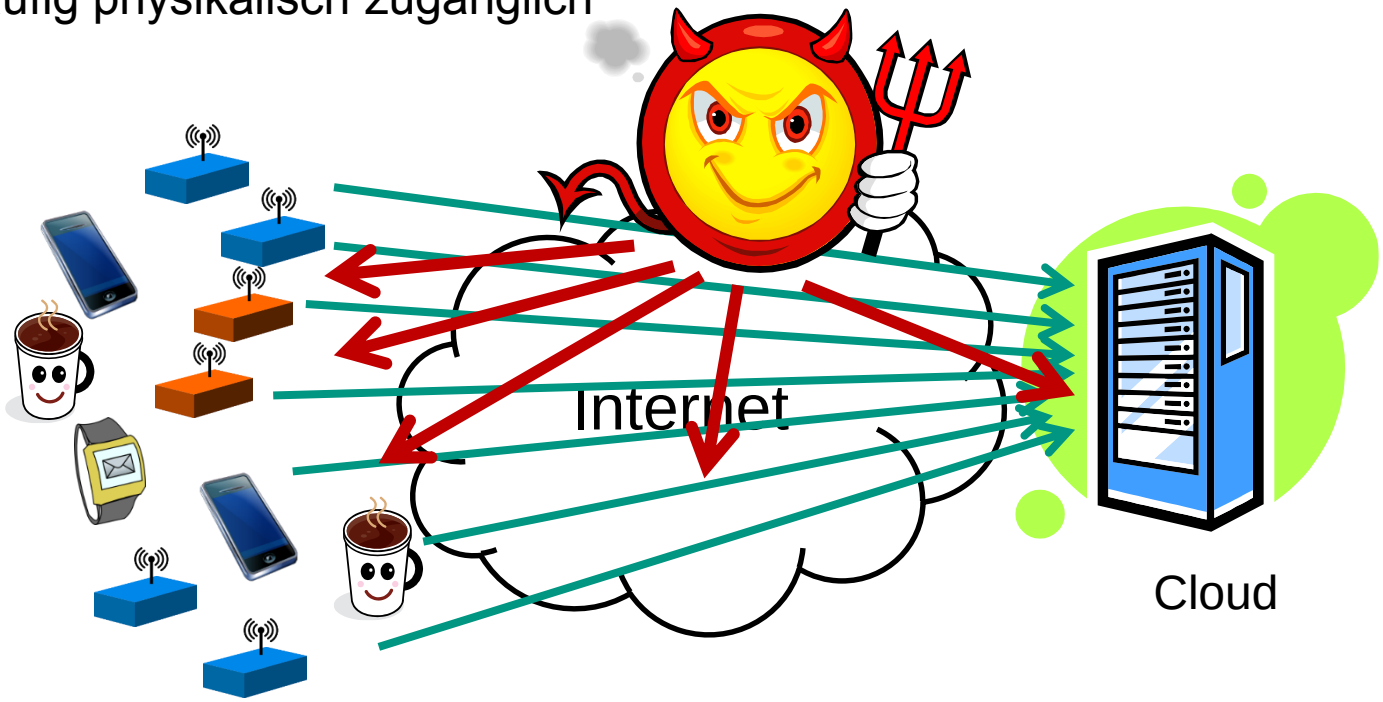
→ Wo bleibt meine Privatsphäre?



Gegenstand dieses Kapitels

„Dinge“ greifen an bzw. werden angegriffen

- Mirai-Botnet … bisher größter Angriff dieser Art
  - → Sehr viele „Dinge“ greifen an
  - „Dinge“ häufig physikalisch zugänglich



Sicherheit wird in Kapitel 5 behandelt

Datenerfassung im privaten Umfeld

- Technologie greift viel stärker in das private Leben ein
  - Datenerfassung sogar am Körper

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

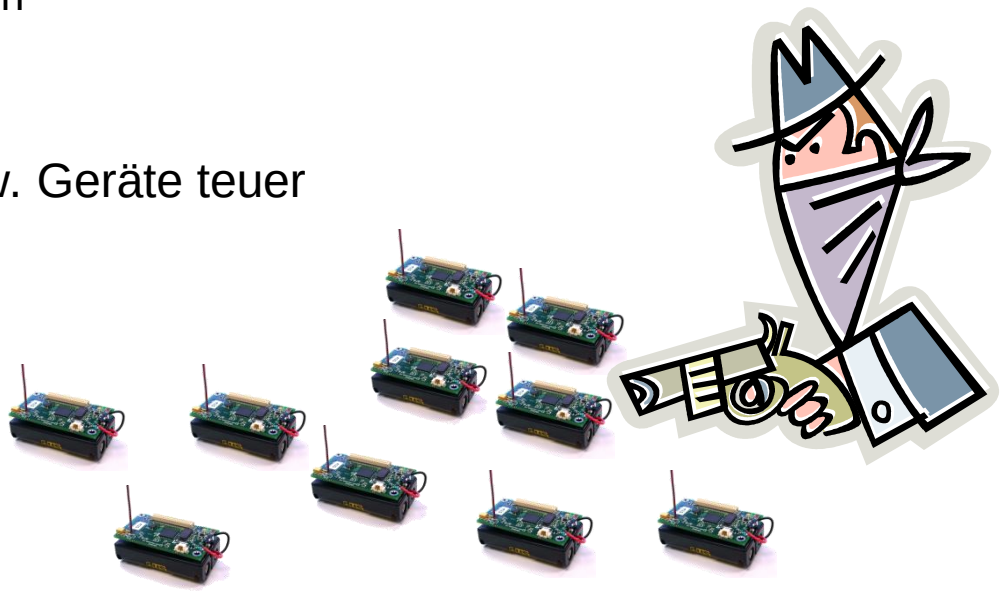
→ Im IoE ist die Privatsphäre noch stärker gefährdet als im klassischen Internet

Allgegenwärtige, ständige Datenerfassung

- Enorme Menge an Daten
  - Internet of Everything
    - Viele Dinge sind beteiligt
- Kontinuierliche Datenerfassung
  - Sensoren messen 24/7 und überall
    - Fitnessarmbänder messen Schlafqualität
    - Intelligente Toilette erfasst „gesundheitsrelevante Daten“
- Sensible Daten
  - Erfassung im eigenen Wohnraum (Privatsphäre!)
    - Stromverbrauch, Raumtemperatur, Anwesenheit …
  - Beobachtung der eigenen Gesundheit
    - Puls, Bewegung
  - Tracking der eigenen Bewegung
- Vielfalt von gemessenen Phänomenen
  - Verkettung von Daten (z.B. beim Dienstanbieter) erlaubt detaillierte Profilbildung

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 bzw. Geräte teuer



ReCap: „Klassisches“ Angreifermodell

- Als „Dolev-Yao“ Angreifer bezeichnet

- Eigenschaften
  - Angreifer ist omnipräsent im Netz, kann sämtliche Kommunikation abhören
  - Kann eigene Pakete erzeugen und versenden
  - Kann fremde Pakete modifizieren

  - Kann allerdings nicht Entschlüsseln oder Verschlüsseln, ohne den Schlüssel zu kennen

Angreifer = „Outsider“

[DY83]

Angreifermodell im IoE

- 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

- 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“


[Krebs2016]

Wichtig: Anzahl korrumpierter Geräte

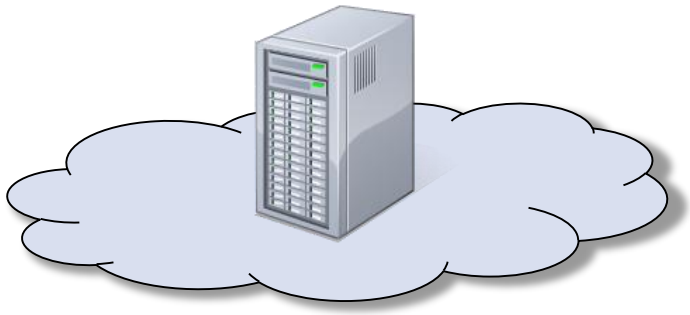
- Wie viele Geräte korrumpiert Angreifer?
- Annahme: Netz mit n Geräten
  - Klar: Korrumpiert der Angreifer $n-1$ oder n Geräte, gibt es keine sinnvolle Definition von Sicherheit mehr (Benutzer sollte andere Probleme zuerst lösen…)
  - Auch klar: Je mehr korrumpierte Geräte im Netz, desto schwieriger 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

Häufig: Cloud-basierte Architektur im IoE

- Übertragung personenbezogener 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.: kaum Transparenz über den Umgang mit den Daten



Zwei Beispieldienste im IoE

- Smart Traffic

Intelligenz

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

- Smart Grid

Smart-Traffic

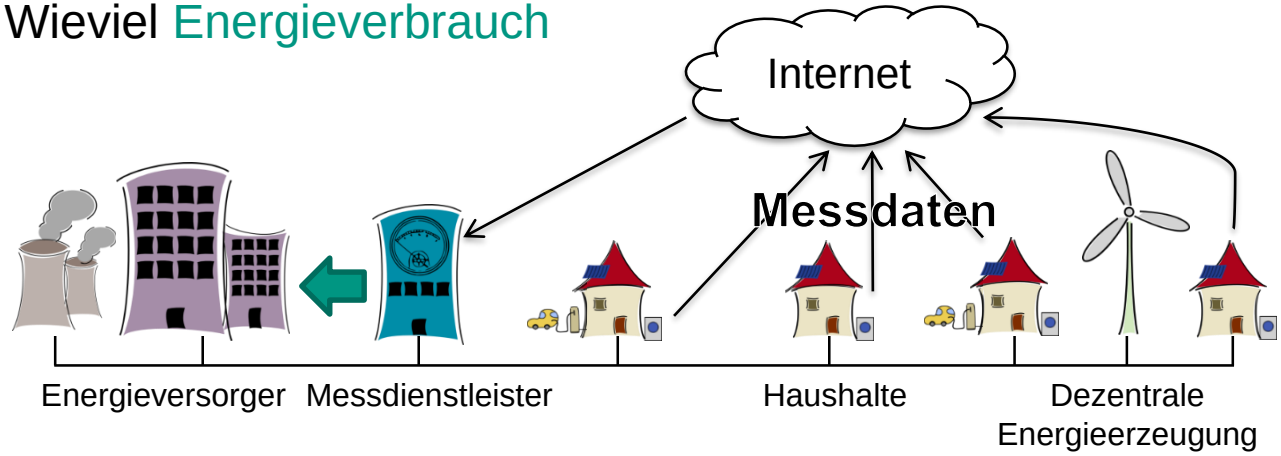
- Hier werden für die Diensterbringung u.a. Positionsdaten benötigt
- Bsp: Google Live Traffic
  - Positionsupdates von Android-Smartphones
    - Bestimmung des Verkehrsaufkommens / Stauerkennung
    - Nutzung für verbesserte Navigation (Google Maps)
    - Daten werden andauernd gesammelt, auch wenn 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

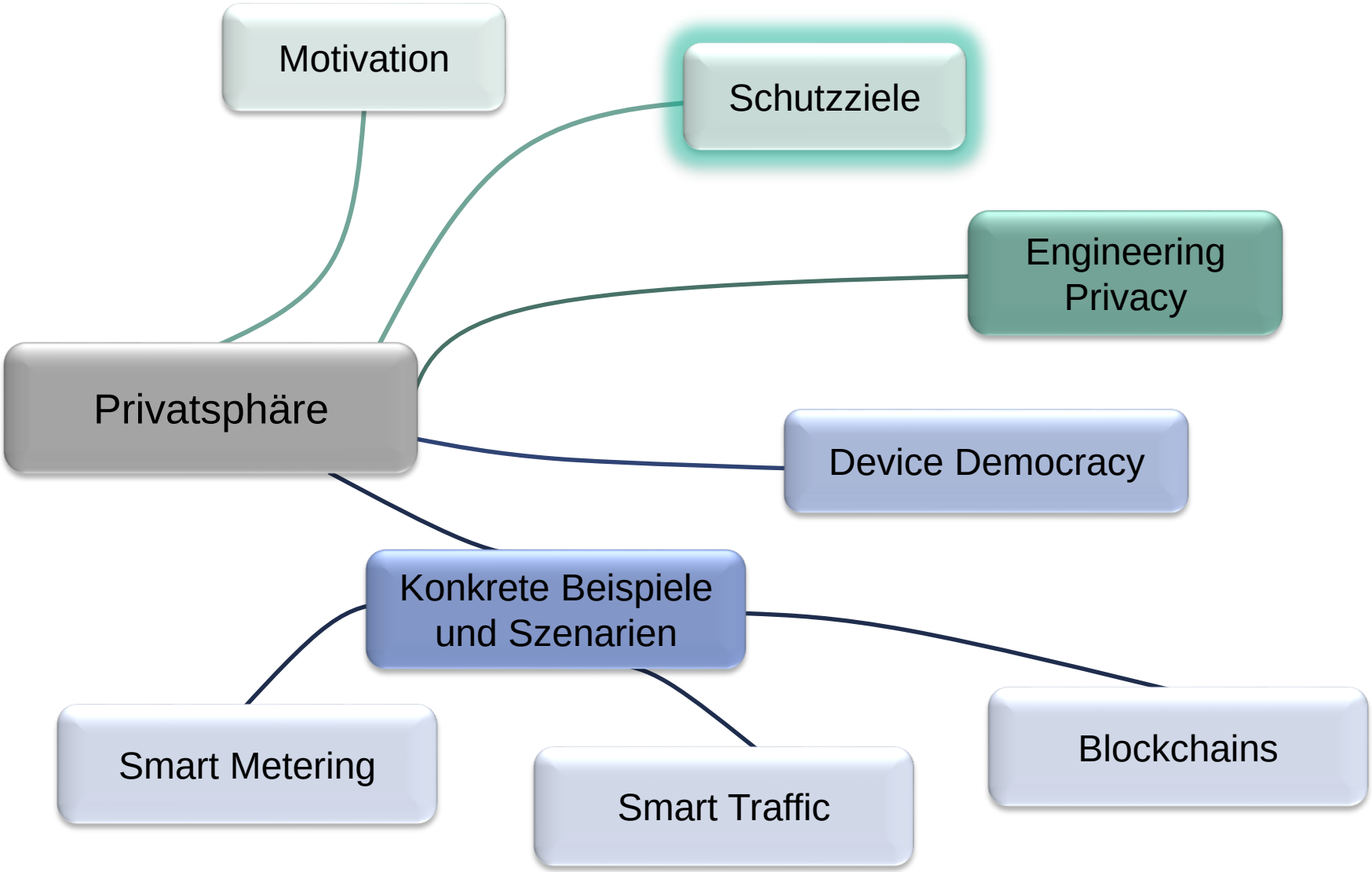
[Jeske2013] [Tock2014] [Golle2009]

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
- Schutzbedarf nicht nur vor Outsider sondern auch vor „Datensenke“


Motivation
Schutzziele
Engineering Privacy
Privatsphäre
Device Democracy
Konkrete Beispiele und Szenarien
Smart Metering
Smart Traffic
Blockchains

Schutzziele (IT-Sicherheit)

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

Schutzziele (IT-Sicherheit)

- Häufige Kategorisierung in

  - **C**onfidentiality (Vertraulichkeit)
    - Ein System bewahrt Vertraulichkeit, wenn es keine unautorisierte Informationsgewinnung ermöglicht
  - **I**ntegrity (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
  - **A**vailability (Verfügbarkeit)
    - Ein System bewahrt Verfügbarkeit, wenn es keine unautorisierte Einschränkung der Funktionalität des Systems ermöglicht
- Weitere Schutzziele, u.a.
  - Authentizität

IoE-spezifische Herausforderungen

- Teilweise ressourcenarme 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?

Privatsphäre (Privacy)

- Privacy is a fundamental human right
  - Which provides a right to respect for one´s „private and family life, his home and his correpsondence“
- Privacy protection is not only to be regarded as an individual value, but also as an essential element in the functioning of democratic societies
- In the digital world
  - Massive power imbalance between data processing entities and the individuals whose data is at stake

[Danez2014]

Schutz der Privatsphäre (Datenschutz)

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

- Welche Daten sind schützenswert?
  - Laut BDSG: Personenbezogene Daten
    - Personally identifiable information, personal data
    - Daten können zu einer Person in Bezug gesetzt werden
  - Weitreichender: Privacy-relevant data
    - Daten, die nicht zu einer einzelnen Person zugeordnet sind, können für Privatsphäre einer Gruppe relevant sein
    - Daten, die Verknüpfung mit anderen Daten ermöglichen können und so eine Beziehung zu einer Person herstellen, die sonst nicht erkennbar gewesen wäre
      - Beispiel: Verknüpfung von anonymen Positionssamples mit Informationen über Wohnort und Arbeitsplatz ermöglicht Identifizierung

BDSG: Bundesdatenschutzgesetz

Schutz der Privatsphäre (Datenschutz)

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

→ Metadaten können privacy-relevant sein

→ Schutz der Metadaten aus technischer Sicht besonders herausfordernd

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

Sind Forennutzer XYZ und ABC die gleiche Person?

Ist dies der Twitter-Account meines Angestellten?

  - Technische Verfahren, z.B.
    - Datenvermeidung
    - Separierung von Domänen (Verschlüsselung, Nutzung unterschiedlicher Identifier, Zugangskontrolle)
    - Anonymisierung und Pseudonymisierung

Spezifische Schutzziele für Privatsphäre

- Transparenz
  - Ein System bewahrt Transparenz, wenn die Verarbeitung von personenbezogenen Daten nachvollziehbar und überprüfbar ist
  - Transparenz muss vor, während und nach der Verarbeitung erfüllt sein

Welche Daten hat Firma YWK über mich gespeichert?

Wie wird meine Einkaufsliste vom Händler genutzt?

  - Technische Verfahren, z.B.
    - Erstellung von Logdateien
    - Nachvollziehbare Dokumentation
    - Source code
    - Privacy policies

Spezifische Schutzziele für Privatsphäre

- Intervenierbarkeit
  - Ein System bietet Intervenierbarkeit, wenn betroffene Personen über die Art der Erfassung und Verarbeitung ihrer personenbezogenen Daten selbst bestimmen können

Nein, bitte keine Cookies von dieser Webseite!

Ich möchte, dass sämtliche Daten über mich gelöscht werden!

  - Technische Verfahren, z.B.
    - Schaffung von Wahlmöglichkeiten hinsichtlich der Verarbeitung von Daten
    - Manuelles Überschreiben automatischer Entscheidungen

→ Verfahren erfordern die Unterstützung der Dienstprovider

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

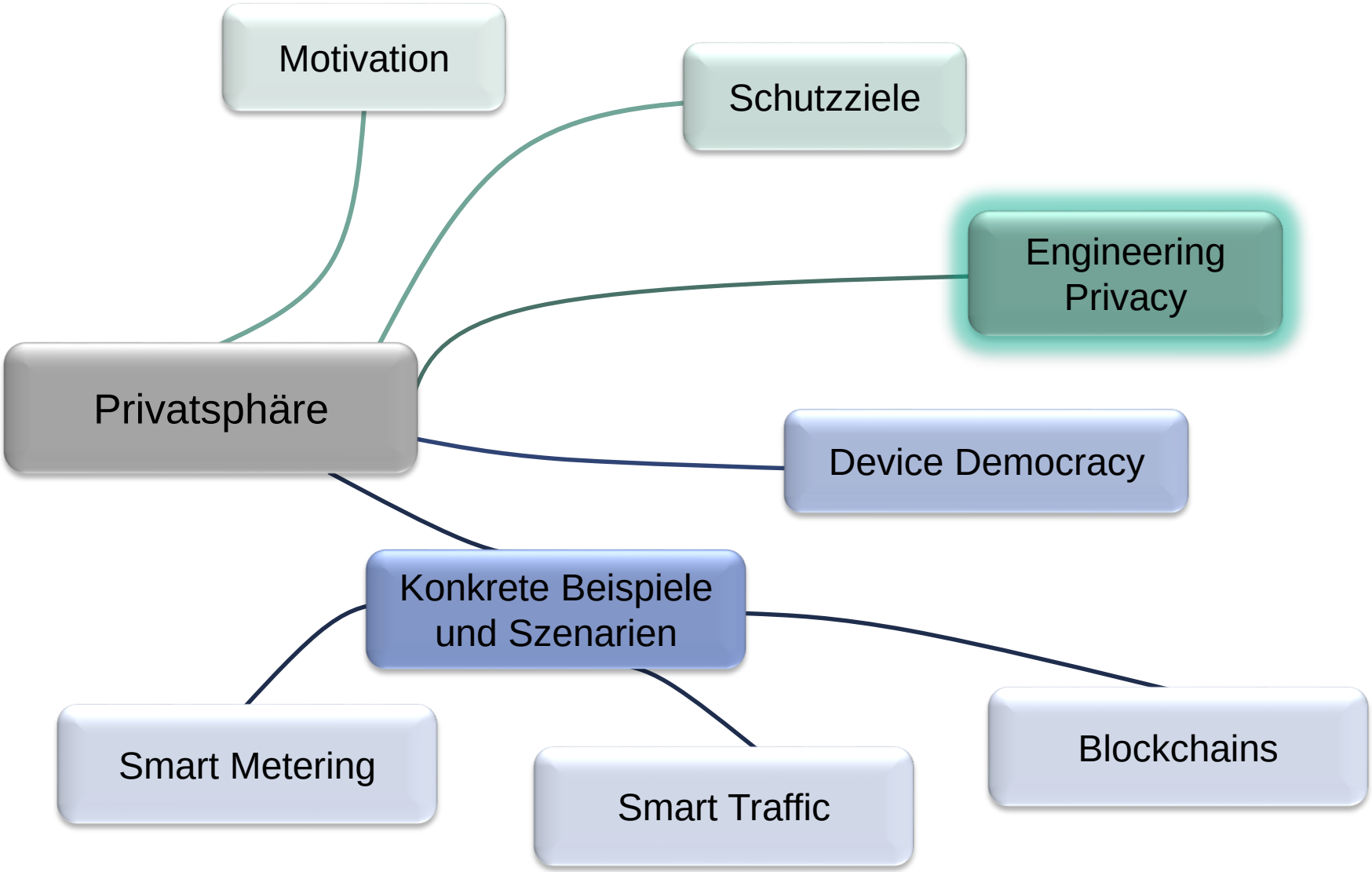
Wer hat die schlechte Produktbewertung geschrieben?

- Unentdeckbarkeit
  - Angreifer kann nicht eindeutig bestimmen, ob ein Datum existiert oder nicht
  - Oft Konflikt mit Anforderungen an Funktionalität
  - Technische Verfahren
    - Datenvermeidung, geschicktes Routing

Hat Herr Meier meine Email schon gelesen?

- 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 Off-the-Record messaging (OTR)

Ich bin überzeugt, dass Herr Müller Gedanken lesen kann!


Motivation
Schutzziele
Engineering Privacy
Privatsphäre
Device Democracy
Konkrete Beispiele und Szenarien
Smart Metering
Smart Traffic
Blockchains

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)

Privacy by Design

- … in legal documents
  - Described in broad terms as a general principle
- … by computer scientists and engineers
  - Equated with the use of specific PETs

- But … it is neither a collection of general principles nor can it be reduced to the implementation of PETs

→ It is a process
  - Involving various technological and organisational components

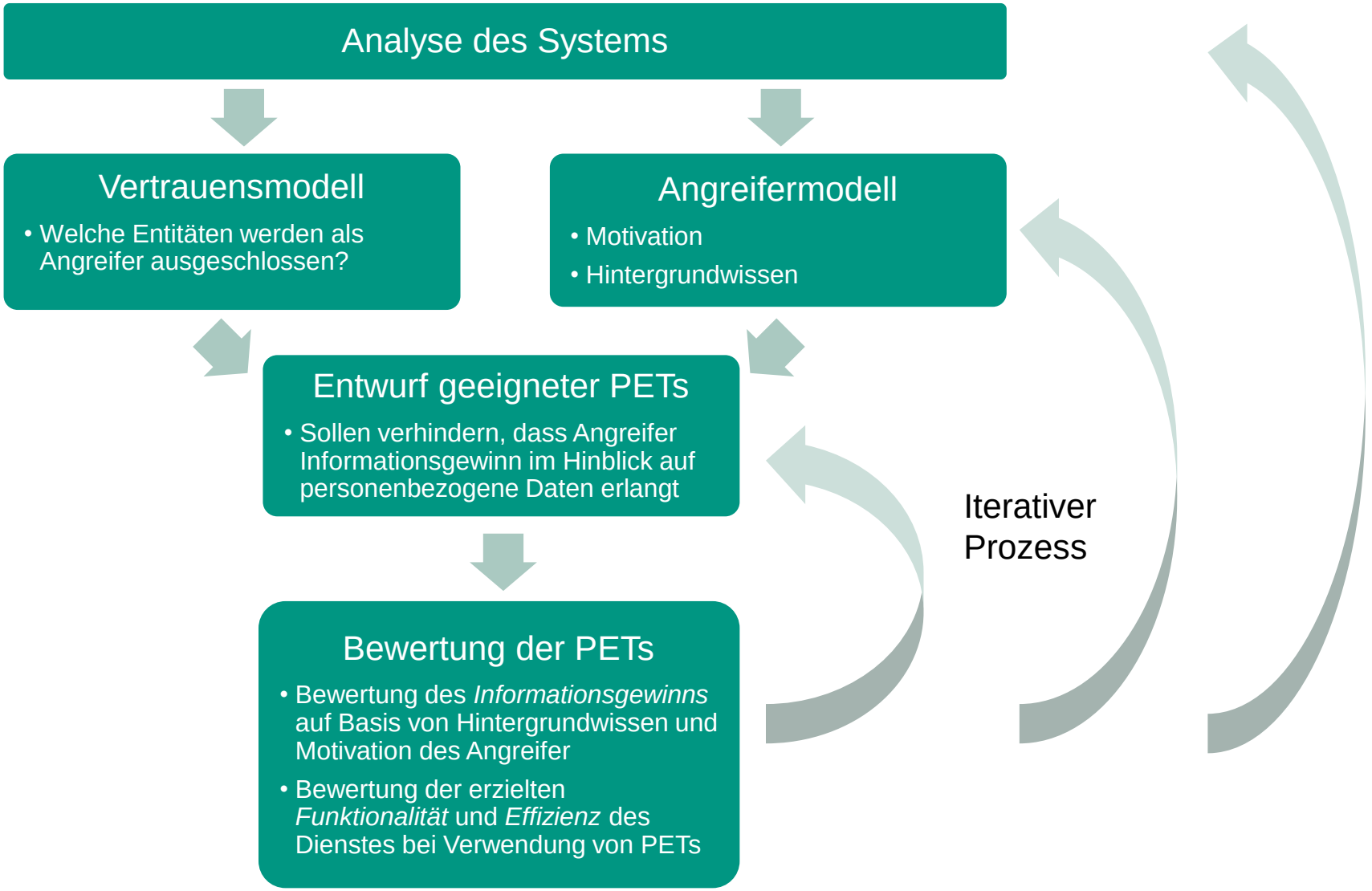
  - Needs well-defined objectives, methodologies, and evaluation means

[Danez2014]

Privacy by Design

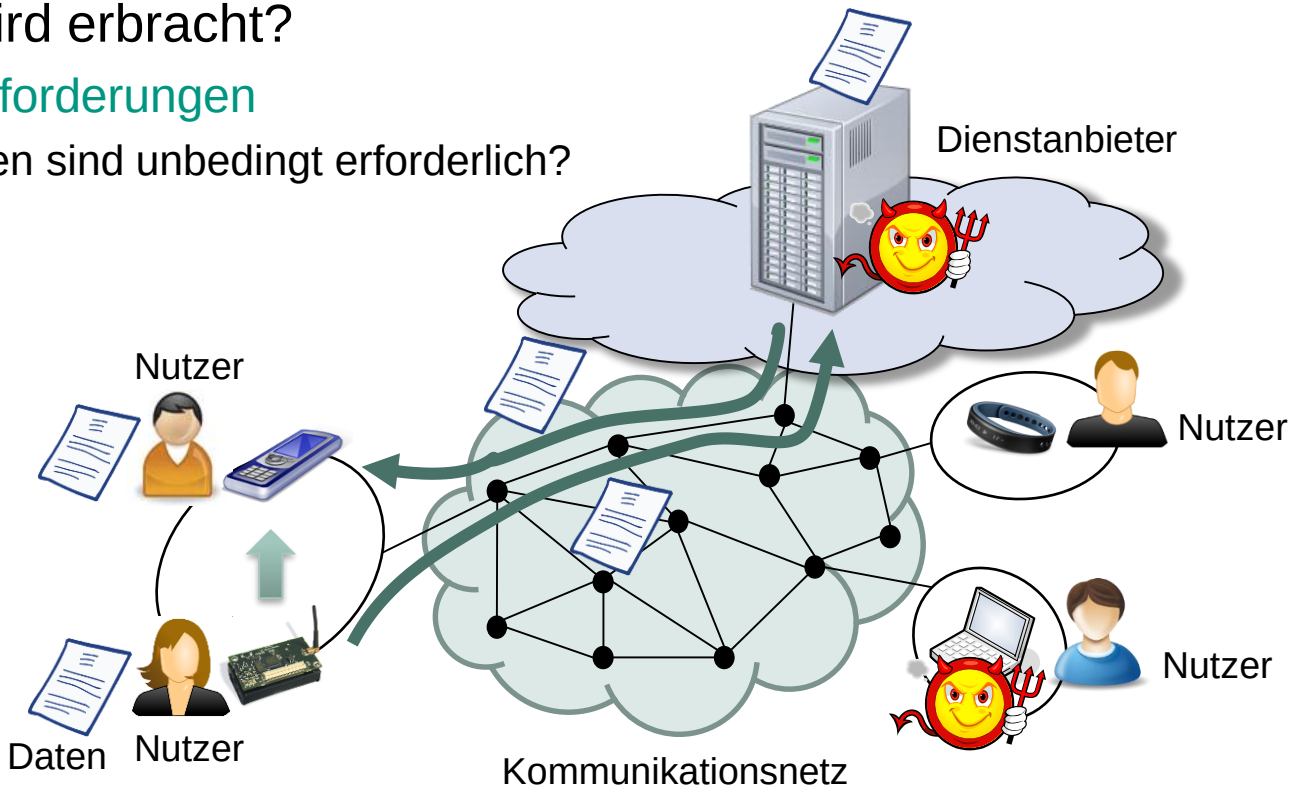
- Generally not the primary requirement of a system
- It may even come into conflict with other (functional or non-functional) requirements
- Proper identification of the objectives requires a privacy risk analysis

„Grober“ Prozess



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 PETs

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

Beispiel: Anonymes Feedback-Formular

- Beteiligte Entitäten: Studenten und Institutsmitarbeiter
- Erbrachter Dienst: Feedback von Studenten erhalten
- Schutzziel: Nicht-Identifizierbarkeit, Anonymität

- Ziel des Angreifers
  - Übungsleiter des Instituts will wissen, welcher Student die schlechte Bewertung verfasst hat
- Wissen des Angreifers
  - Zusätzlich zur Bewertung ist die IP-Adresse verfügbar
  - Kann mit IP-Adresse der Praktikumsanmeldung verknüpft werden

- Lösung: Zugriff über Tor
  - IP-Adresse wird verschleiert, kein Angriff mehr möglich
  - Weiterhin: Privacy-by-default! Nutzer wird zur Anonymität gezwungen → besser für ihn, besser für andere (größeres Anonymitätsset)
  - Zum Vergrößern des Anonymitätssets: http://tmfeedozvp6v65nf.onion

Privacy Design Strategies

- MINIMISE
  - The amount of personal data that is processed should be restricted to the minimal amount possible
  - Is the processing of the data proportional with respect to the purpose?
  - Do no less invasive means exist to achieve the purpose?
  - Design patterns
    - Select before you collect
    - Anonymisation
    - Pseudonymisation

[Danez2014]

Privacy Design Strategies

- HIDE
  - Any personal data, and their interrelationships, should be hidden from plain view
    - This way, they can not easily be abused
    - Hidden from whom? … depends on the context
  - Examples
    - Hide information that spontaneously emerges from the use of the system: e.g., communication pattern → hide information from anybody
    - Information is collected, stored or processed legitimately by one party → hide from any other party: apply encryption
  - Design patterns
    - Encryption
    - Mix networks
    - Pseudonymisation
    - …

[Danez2014]

Privacy Design Strategies

- SEPARATE
  - Personal data should be processed and stored in a distributed fashion, if possible
    - Data should be processed and stored locally
  - In practice often disregarded → centralized web-based services, cloud
  - Design patterns
    - No specific patterns known

[Danez2014]

Privacy Design Strategies

- AGGREGATE
  - Personal data should be processed at the highest level of aggregation
  - Design patterns
    - Aggregation over time
    - K-anonymity
    - …

[Danez2014]

Generisches Szenario im IoE

- 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

- 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

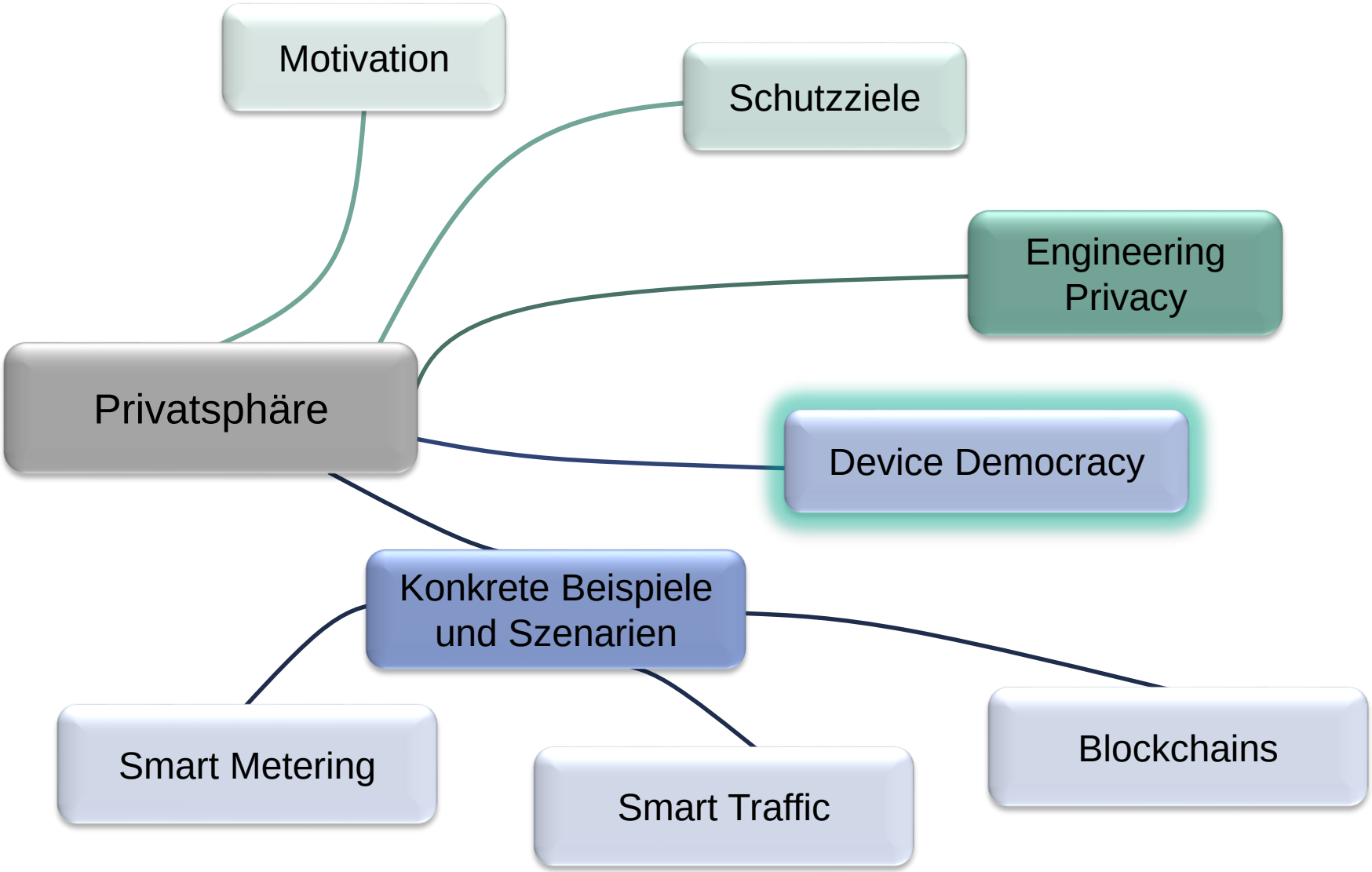
- Zentralisierte Datensenken vermeiden
  - Peer-to-Peer-Netze zur Diensterbringung
  - Lokale Diensterbringung
  - Mehrere, nicht-kooperierende Dienstanbieter mit jeweils beschränkter Sicht
  - Beispiele
    - Peer-to-Peer Filesharing
    - Datenablage auf mehreren unabhängigen Cloud-Providern

Beispiele für Ansätze zum Schutz

- Identität der Quelle verschleiern
  - Pseudonyme
    - Beispiel: „Spitznamen“ in sozialen Netzwerken
  - Datenaggregation
    - Beispiel: 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!


Motivation
Schutzziele
Engineering Privacy
Privatsphäre
Device Democracy
Konkrete Beispiele und Szenarien
Smart Metering
Smart Traffic
Blockchains

Device democracy

T.J Watson (IBM): „Ich glaube, dass es auf der Welt einen Bedarf von vielleicht fünf Computern geben wird“

… nicht belegt

- IBM study, 2014 
- Challenges or why the IoT already needs a reboot
  - Cost of connectivity
  - The Internet after trust
  - Not future-proof
    - Common things last for decades (e.g., door lock, light bulb) … most consumers replace smart phones or PCs every 18-36 months
  - A lack of functional value
    - Lack of meaningful value creation
  - Broken business models
    - … most are based on the use of analytics to sell user data or targeted advertising
    - … users may be open to sharing data – enterprises not

[Pure2015]

Cost of connectivity

- Many existing IoT solutions are expensive: high infrastructure and maintenance costs
  - Centralized cloud, large server farms
  - The cost of supporting and serving billions of smart devices will be substantial
    - E.g., the servers that distribute regulare software updates

[Pure2015]

The Internet after trust

Post-Snowden

→ The notion of IoT solutions built as centralized systems with trusted partners is now something of a fantasy

- In a network of the scale of the IoT, trust can be very hard to engineer and expensive, if not impossible, to guarantee
- Privacy and anonymity must be integrated into its design by giving users control of their own privacy

[Pure2015]

Transaction-based Systems

- Phone-calls, electricity metering, airline reservations …

Old fashioned?

- Message, tweet …

NY stock exchange:
5 million trades a day

5 billion social media transactions per day

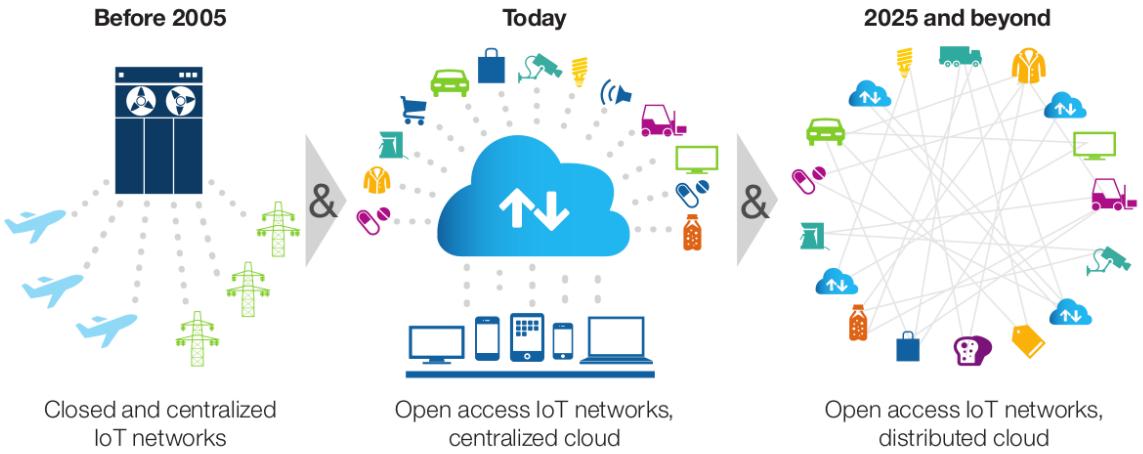
Scale and volume of transactions have exploded !

- Systems (new and old once) can be viewed as transaction-based

[Pure2015]

Apply Design Strategy SEPARATE

- … it´s time for the cloud to move from the data center to our doorknob
  - Centralized → decentralized



[Pure2015]

Anforderungen an die verteilte IoE-Cloud

- „Trustless” Transaktionen zwischen Geräten
  - Gewährleistung von: Privatheit, Integrität von Nachrichten…
  - Technische Garantien statt Vertrauen
- Sicheres, verteiltes Teilen 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

[Pure2015]

Challenge

- Building a decentralized IoT that can scale universally while maintaining private, secure and trustless transactions

IoT: billions of players, not all of which can be trusted.

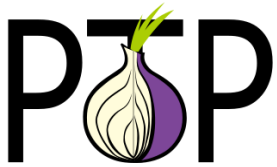
- … the blockchain may offer an elegant solution

[Pure2015]

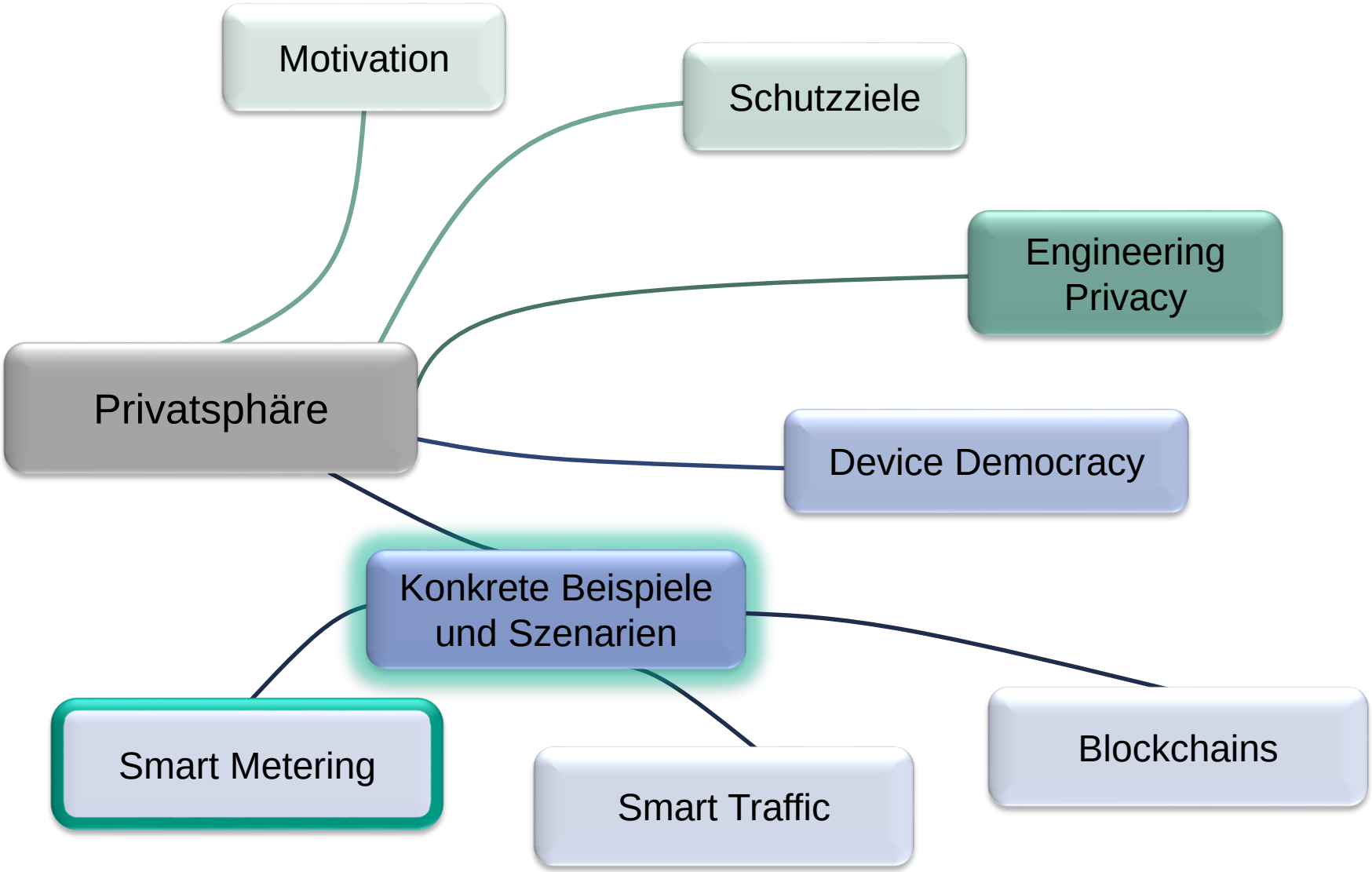
Note

- In the following we present examples of decentralized approaches that increase privacy
  - Without and with the utilization of the blockchain

Beispiel: Peer-Tor-Peer (PTP)

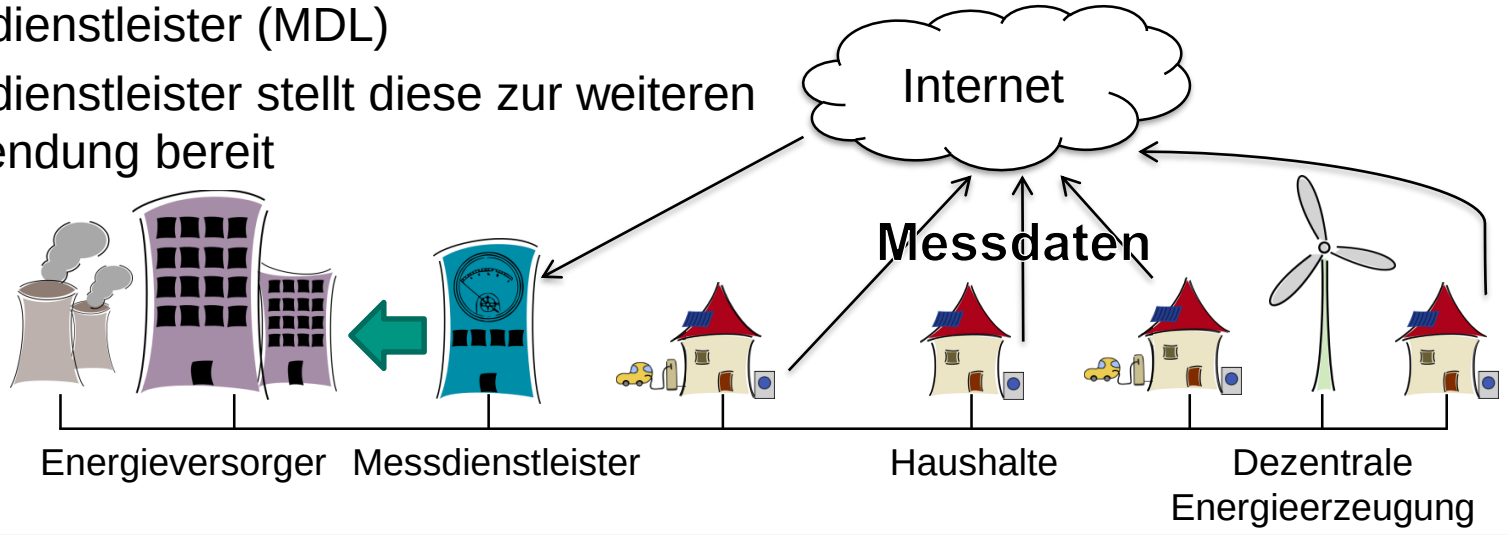


- Dezentralisierter Ansatz für pseudonyme Peer-to-Peer Kommunikation
  - Basierend auf Tor
- Netz-Restriktionen werden weitgehend umgangen
  - Kommunikation auch im Mobilfunknetz und hinter NATs möglich
- Metadaten werden geschützt
  - Nur die Verbindung zum Tor-Netzwerk ist erkennbar
  - Ob und mit wem kommuniziert wird ist unbekannt
- PTP kapselt Tor-Management
  - Teilnehmer verwenden Tor Hidden Services
- Bietet einfache Nachrichten-basierte API
- Open Source: https://github.com/kit-tm/PTP

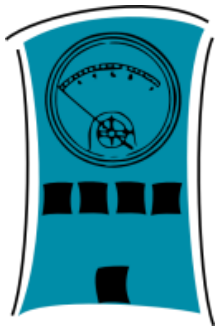

Motivation
Schutzziele
Engineering Privacy
Privatsphäre
Device Democracy
Konkrete Beispiele und Szenarien
Smart Metering
Smart Traffic
Blockchains

Beispiel: Privatheit beim Smart-Metering

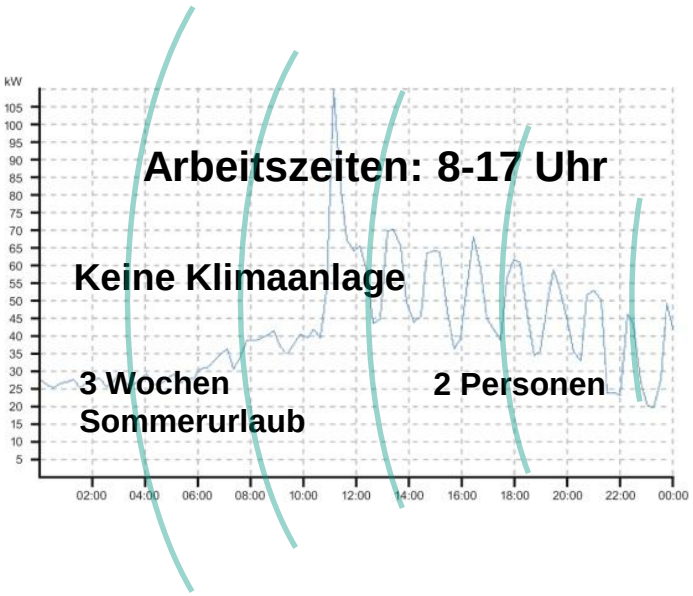
- Smart-Metering Szenario (in Deutschland)
- Ziel
  - Verbrauch von Energie in Energienetzen (topologisch) oder Kundenspezifisch (demographisch) in „Echtzeit“ nachvollziehbar machen
    - Beispiel topologisch: Last in Energienetzabschnitt „Karlsruhe-Mitte“
    - Beispiel demographisch: Energieverbrauch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

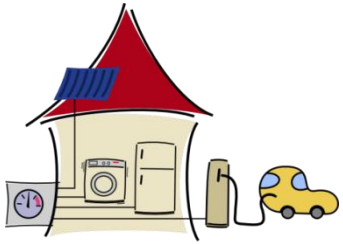


Gefahr für die Privatsphäre



Messdienstleister



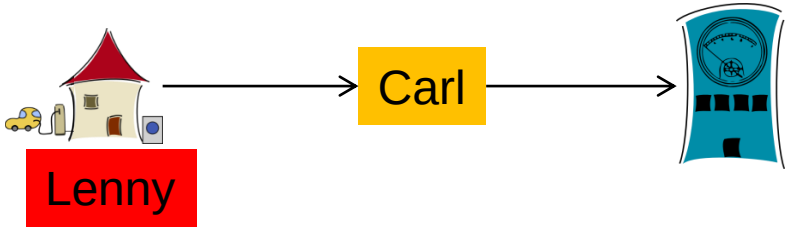


Haushalt

- 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

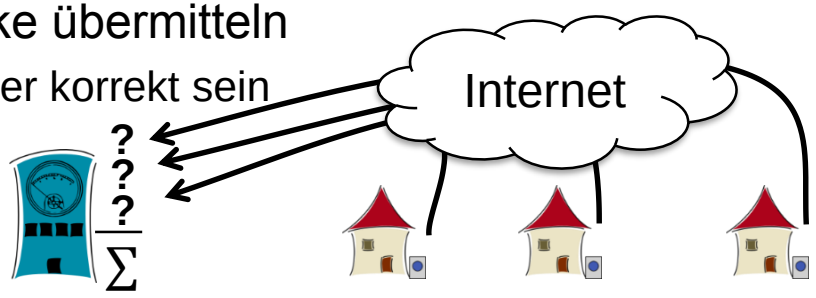
- Idee: Aggregation bevor Daten übermittelt werden
  - Aggregation über viele Haushalte → Privatsphäre geschützt
    - Wie viele? Offen.
  - Aggregation über langen Zeitraum → Privatsphäre geschützt
    - Wie lange? Offen.

- 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



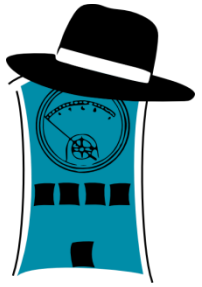
Beispiel ohne Kooperation: „Rauschen“

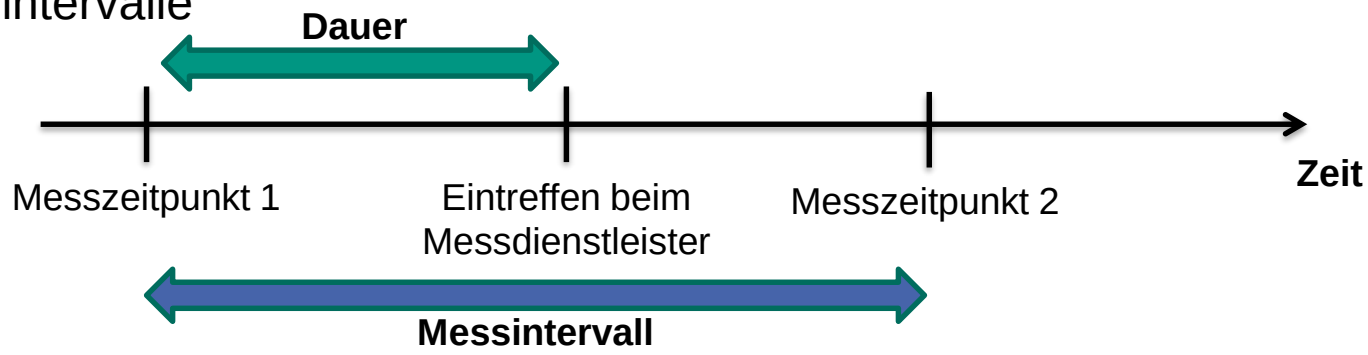
- Hinzufügen von planbarem „Rauschen“
  - Laplace-Rauschen auf Messwerte
  - Einzelner Rauschwert zufällig
  - Rauschwerte „eliminieren“ sich gegenseitig bei Summenbildung und vielen Teilnehmern
- Problem: Sehr viele Teilnehmer notwendig
  - Keine feingranulare Messung (topologisch oder demografisch) 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

[Bohli2010]

Mit Kooperation der Stromzähler untereinander

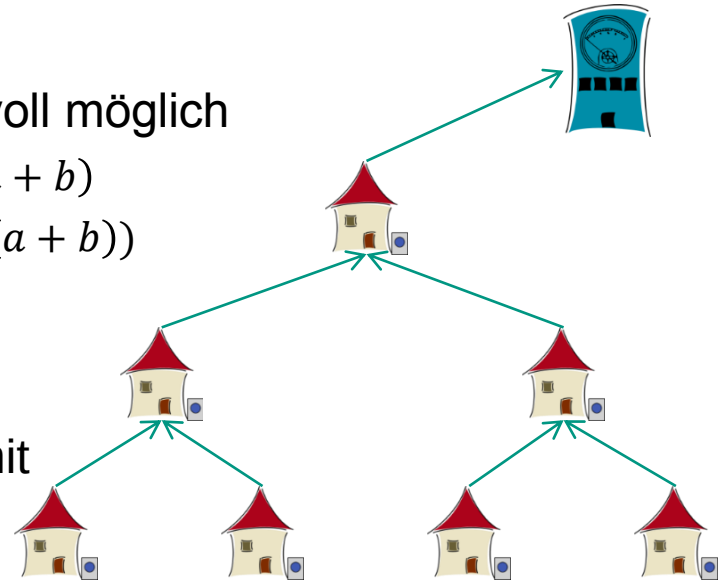
- Benötigt Protokoll zur Kommunikation zwischen Stromzählern
- Anforderungen
  - Einzelne Messwerte vor Messdienstleister geschützt
    - Nur Aggregat ermittelbar, auch bei Angriff durch bspw. MDL
  - Robuster Betrieb
    - Ausfall einzelner intelligenter Stromzähler
    - Ausfall der Kommunikationsanbindung
  - 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
  - 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

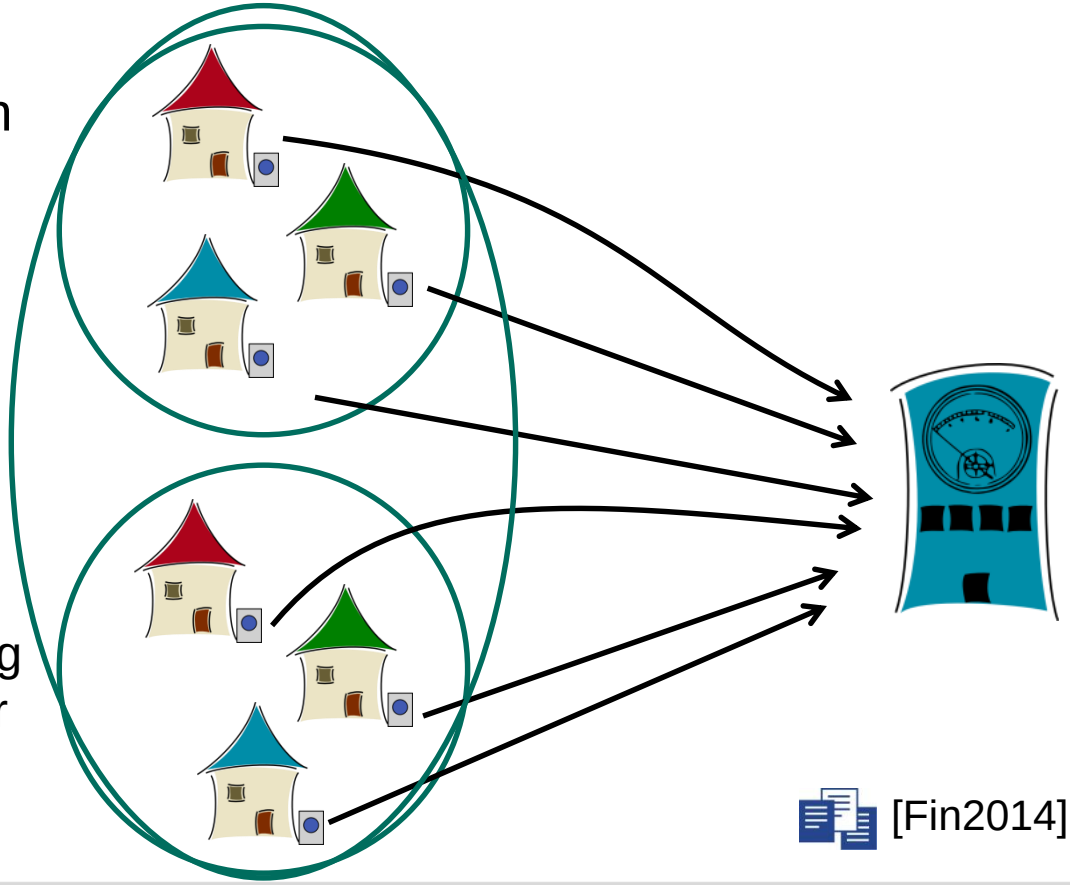
[Li2011]

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

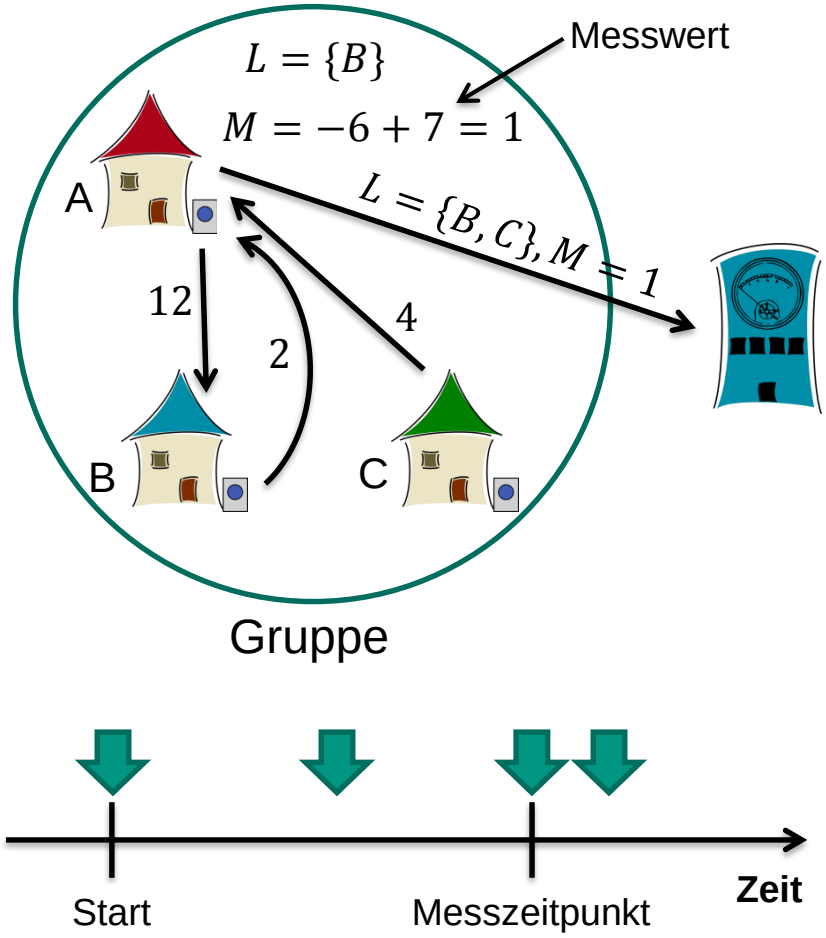
- Privatsphärengerechtes Smart Metering Protokoll SMART-ER: SMART with Exactness and Robustness
  - Basiert auf SMART, Verfahren für drahtlose Sensornetze
- Grundkonzept: Einteilung der Stromzähler in Gruppen
  - Vom Messdienstleister durchgeführt
  - Konfigurierbare Gruppengröße
  - Höhere Robustheit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



[Fin2014]

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

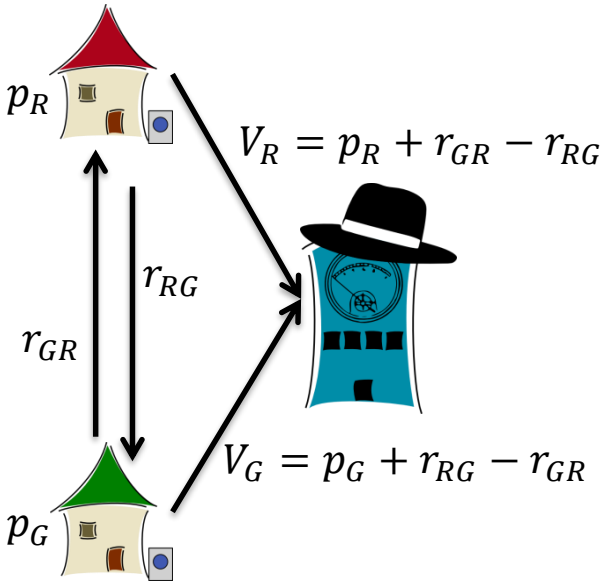
- Messdienstleister ist korrumpiert (Symbol Hut)
- Angreifer kann mit abgegebenen, maskierten Werten arbeiten
  - Roter Haushalt gibt V_R ab
  - Besteht aus Messwert p_R, empfangener Zufallszahl r_{GR} und versendeter Zufallszahl r_{RG}
  - $V_R = p_R + r_{GR} - r_{RG}$
- Summe $V_R + V_G =$ gewünschtes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

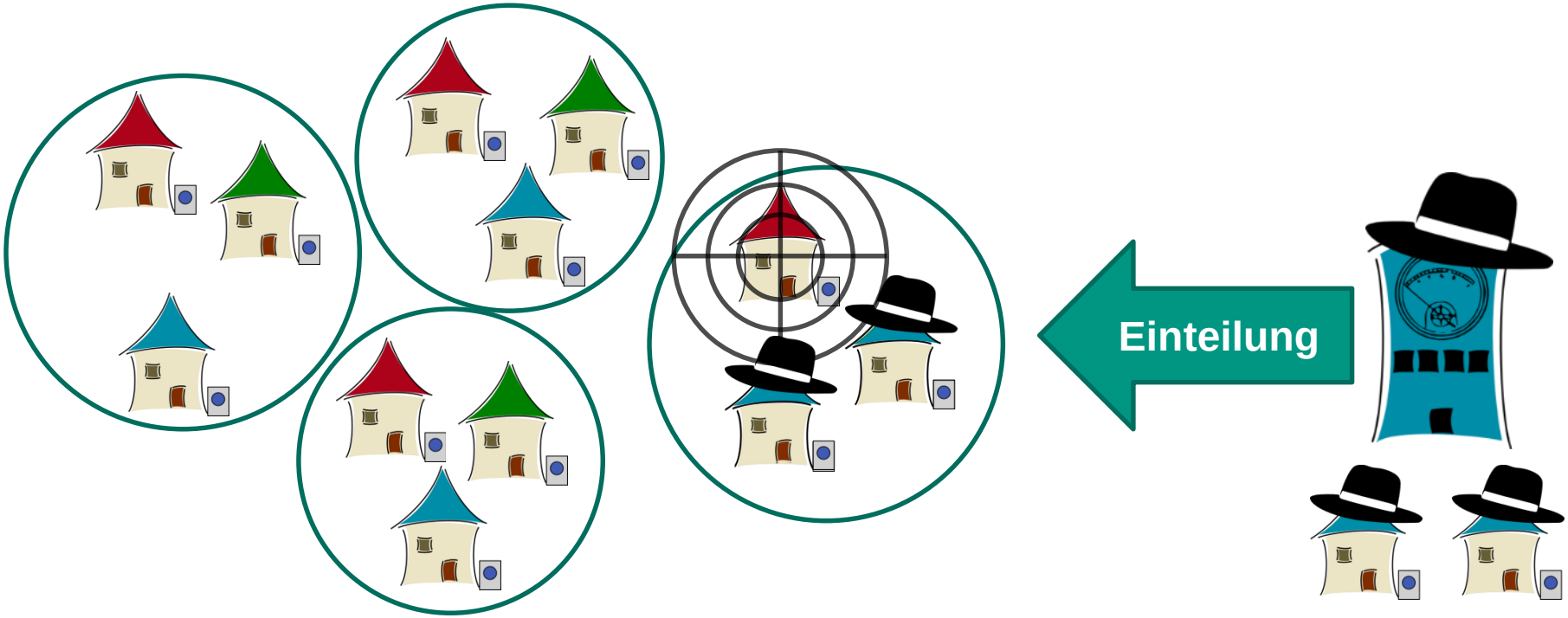
$$\Rightarrow P\left(d^{(1)} = y\right) = P\left(d^{(2)} = y\right) \forall\, y \in \mathbb{Z}/q\mathbb{Z}$$

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



SMART-ER: Verbleibendes Problem

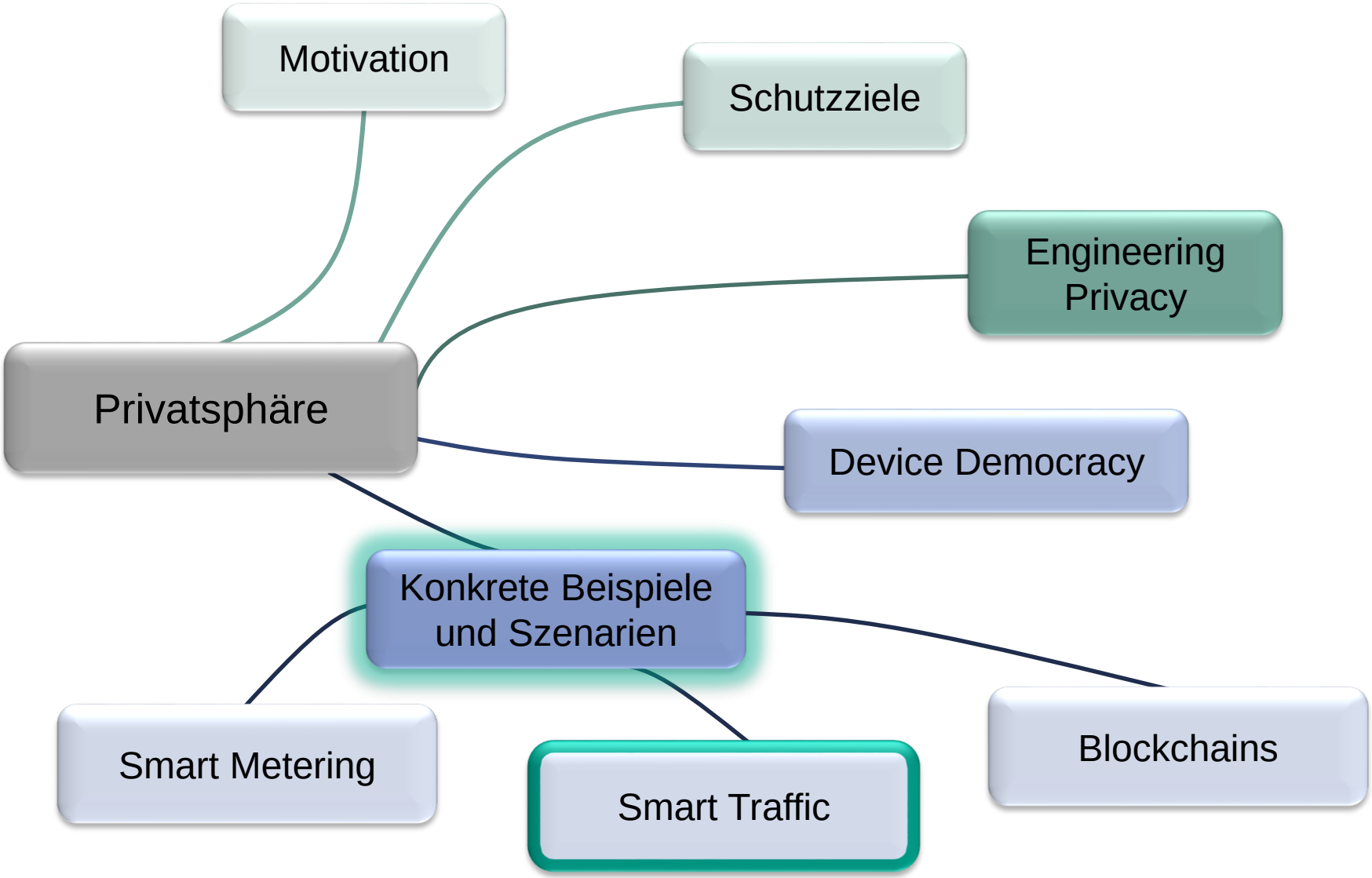
- Problem: Messdienstleister nimmt Gruppenbildung vor
  - Angriff mittels korrumpierter Stromzähler möglich
  - Benötigt Gruppengröße−1 korrumpierte intelligente Stromzähler



- 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


Motivation
Schutzziele
Engineering Privacy
Privatsphäre
Device Democracy
Konkrete Beispiele und Szenarien
Smart Metering
Smart Traffic
Blockchains

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

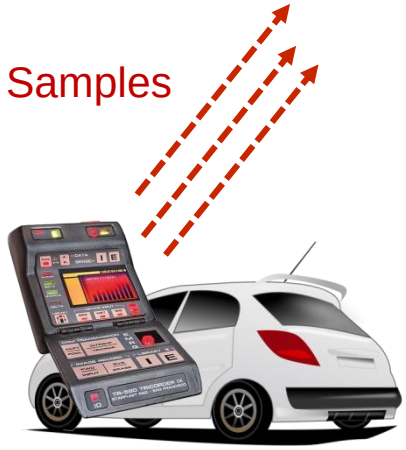
Die Straße hier ist rutschig!

Ich komme um die Ecke!

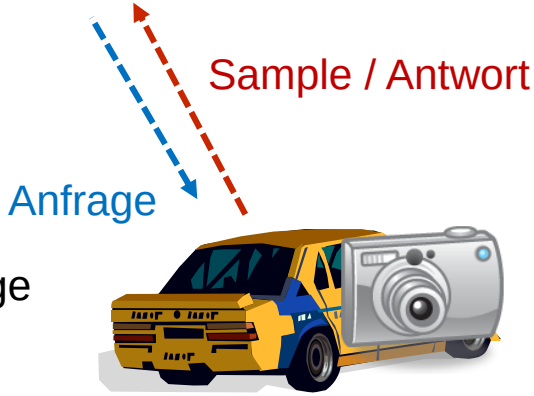
- Verkehrsoptimierung
  - Entwicklung der automatischen Fahrzeugnavigation
    - Statisch – Routenplanung mit fixem Kartenmaterial
    - Adaptiv – Einbeziehung von aktuellem Verkehrszustand
    - Koordiniert – Einbeziehung von geplanten Routen anderer Verkehrsteilnehmer, Veröffentlichung eigener Pläne

Smart-Traffic-Anwendungen (II)

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



- Vehicular Clouds
  - Fahrzeuge als Diensterbringer
  - Beispiel: Pics-On-Wheels
    - 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



[Weng2013]

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

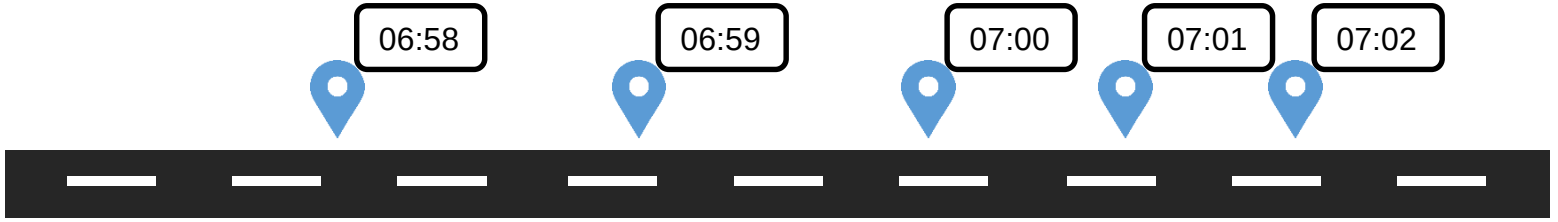
- Position
  - Geografische Repräsentationen
    - Breitengrad/Längengrad, z.B. (49.012421, 8.408077)
  - Kontext-spezifische Repräsentationen
    - „Im SCC-Gebäude“
    - Hierarchische Modelle, z.B. „Deutschland/Karlsruhe/KIT/Gebäude 20.20“
  - variable Präzision
- Zeit
  - …zu der Sample erstellt wurde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

Positionssamples: Beispiel

- Alice fährt eine bestimmte Strecke ab

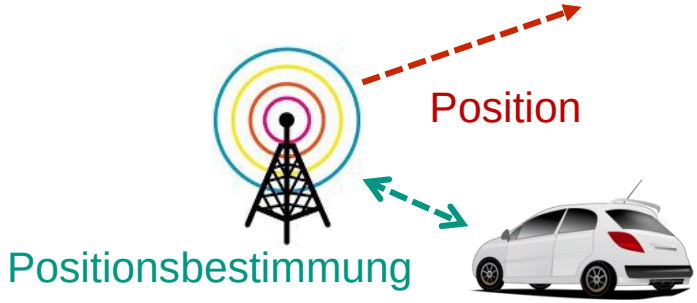
- Zu bestimmten Zeitpunkten teilt sie Ihre aktuelle Position

- Der Dienstanbieter erhält also z.B. folgende Informationen

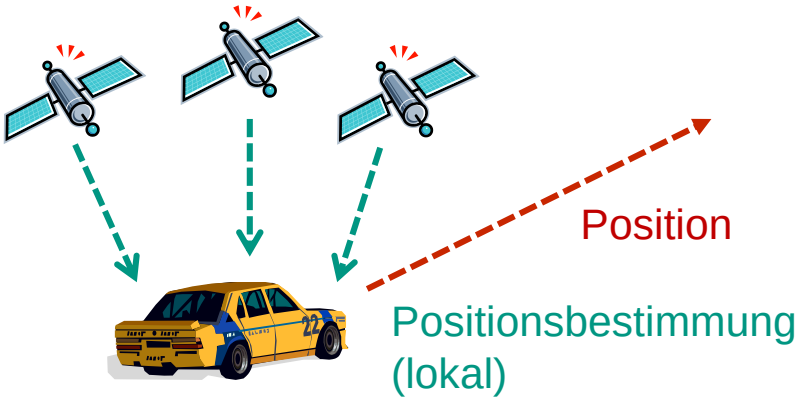


Positionsbestimmung

- 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
    - z.B. Fotos

[Tock2014] [Golle2009]

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);

- Was für Rückschlüsse erlauben diese Samples auf die Insassen des Fahrzeugs?

Privatsphärenschutz in Smart Traffic (Überblick)

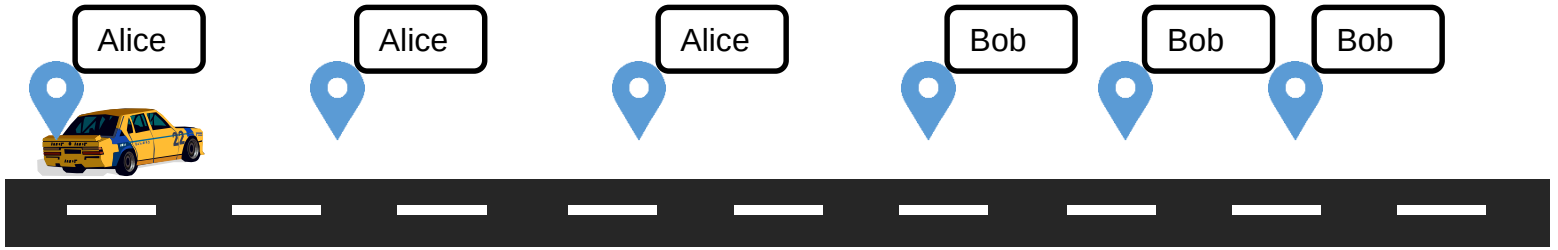
- Nicht-Identifizierbarkeit von Positionssamples
  - Verwendung von Pseudonymen
- Unverkettbarkeit von Positionssamples
  - Pseudonyme müssen oft gewechselt werden
  - Effektiver Pseudonymwechsel nicht trivial
    - Positionssamples vor und nach Wechsel können Zuordnung ermöglichen
- Verschleierung von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

Trennung von Positionssamples u. Identitäten

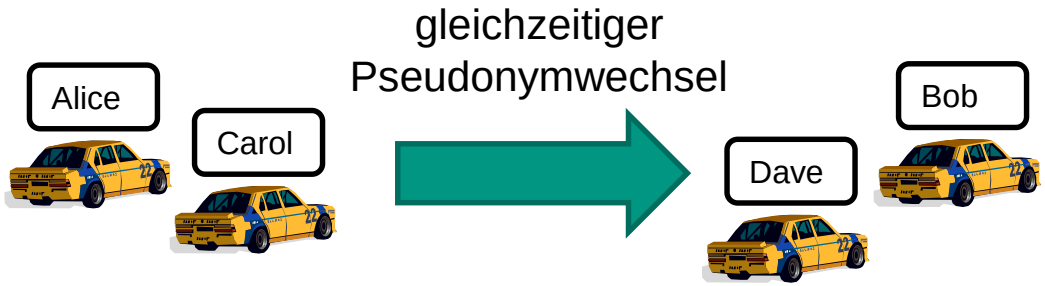
- Im Grunde ähnliche Ansätze wie bei Samples allgemein
- Vollständige Anonymität nicht immer möglich
  - Ermöglichen unbestrafbaren Missbrauch der Dienste
- Verwendung von Pseudonymen
  - Limitierte Anzahl pro Teilnehmer
  - Verfahren zum Ausschließen böswilliger Teilnehmer
    - z.B. indem Zuordnung 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
  - 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

[Beres2004]

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
  - 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

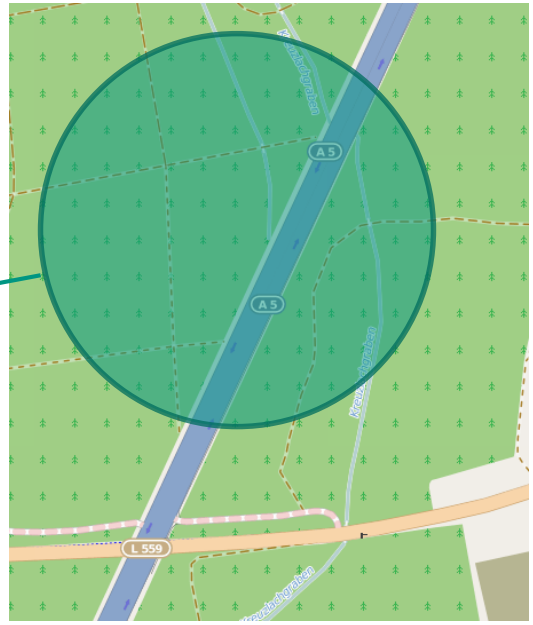
[Hoh2008]

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…

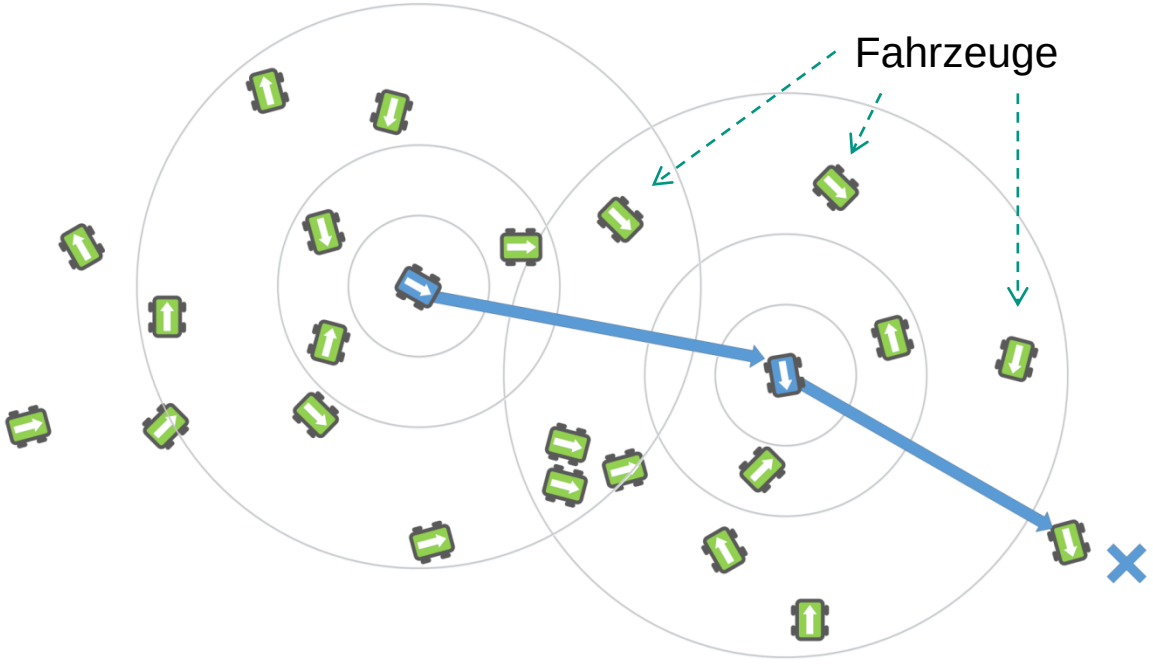
Eigene Arbeiten: Geocast

- 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

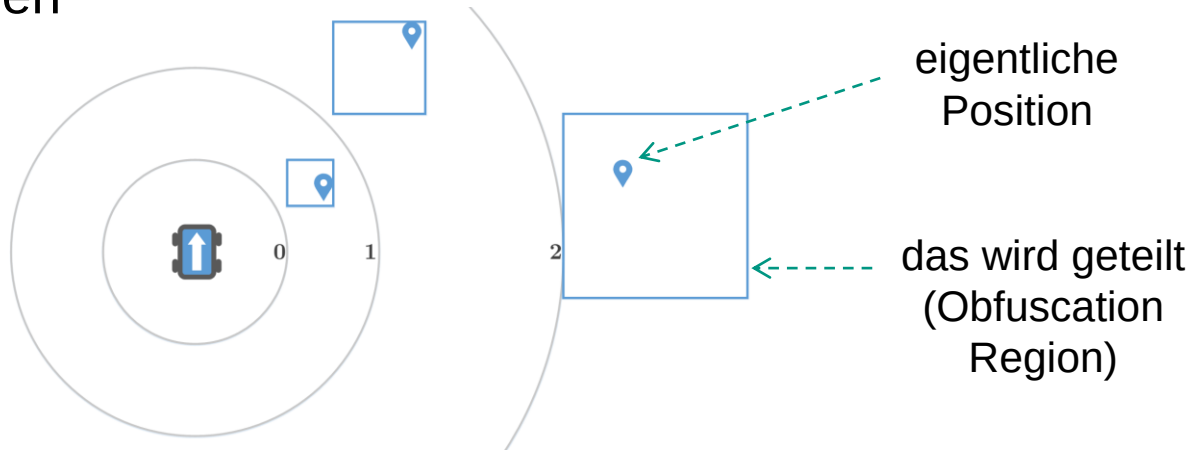
- Unser Ansatz: OverDrive
  - Peer-to-Peer Overlay-Netz direkt zwischen Fahrzeugen
  - Overlay-Nachbarn tauschen Positionsdaten aus
    → keine zentrale Datensenke, niemand hat „volle Sicht“
  - Nachrichten werden über mehrere „Hops“ weitergeleitet



[Flor2016]

Eigene Arbeiten: Privatheit in OverDrive

- 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?

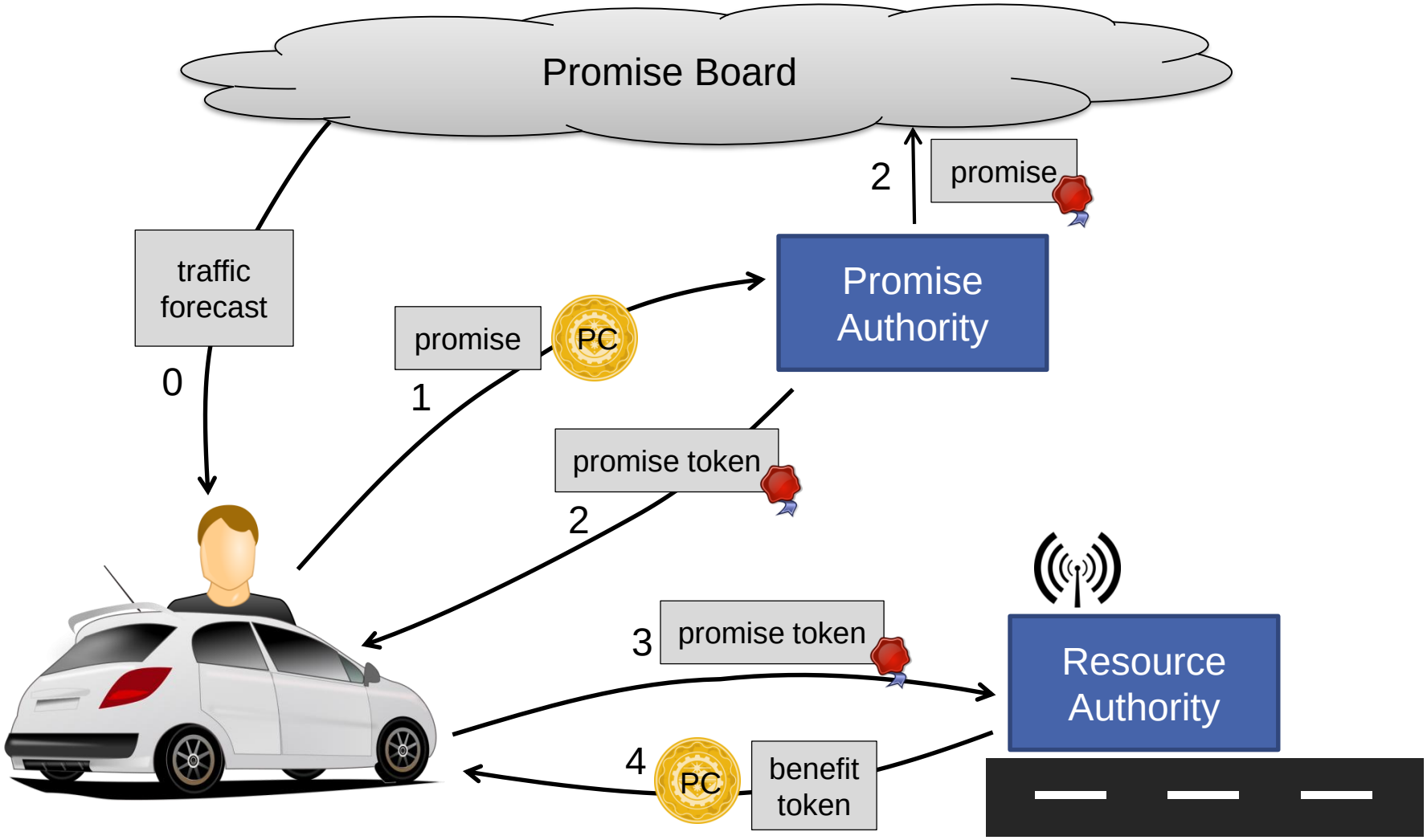
Eigene Arbeiten: vorausschauende Navigation

- 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

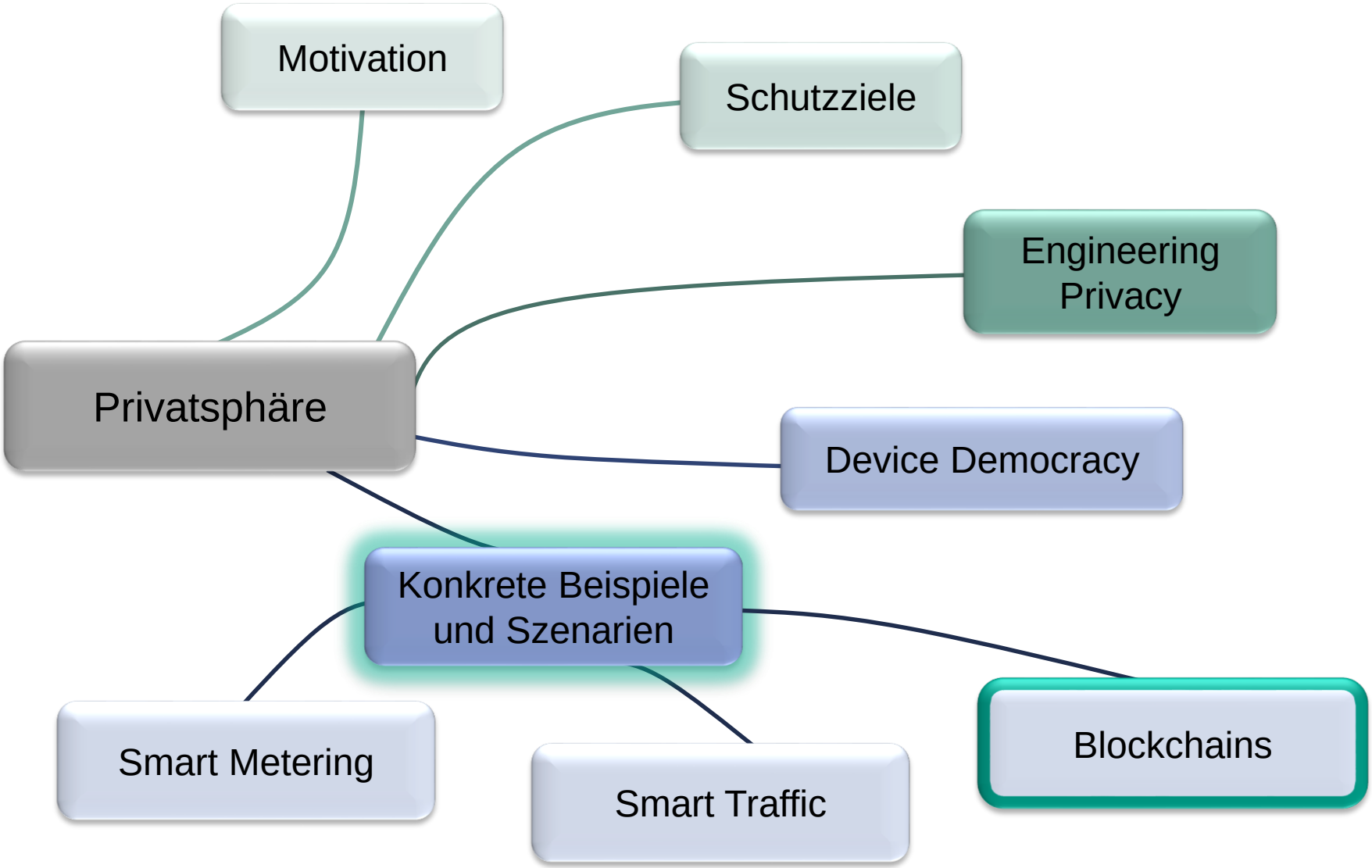
[Flor2014]

Eigene Arbeiten: Privacy-Preserving Coop. Route Planning



Vergleich der betrachteten Szenarien

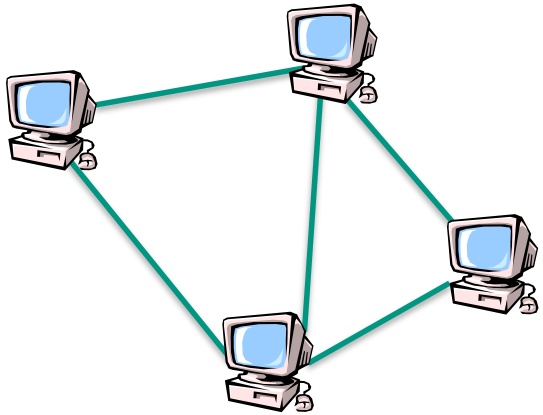
	Smart Grid / Smart Metering	Smart Traffic / Geocast
Zu schützende Daten	Personenbezogene Messwerte	Lokationsdaten der Nutzer
Motivation des Angreifers	Rückschlüsse auf Nutzerverhalten	Erstellung von Bewegungsprofilen
Hintergrundwissen des Angreifers	Zuordnung von Messwerten zu Nutzerverhalten	Zuordnung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


Motivation
Schutzziele
Engineering Privacy
Privatsphäre
Device Democracy
Konkrete Beispiele und Szenarien
Smart Metering
Smart Traffic
Blockchains

Blockchain-Crashkurs: Überblick

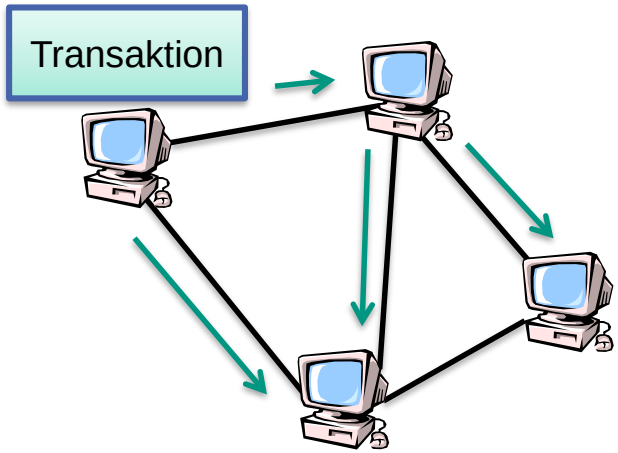
- High-level Vorstellung
  - Eine Blockchain ist ein schwarzes Brett
  - Man kann nur schreiben, nichts löschen
  - Jeder Eintrag hat einen Zeitstempel

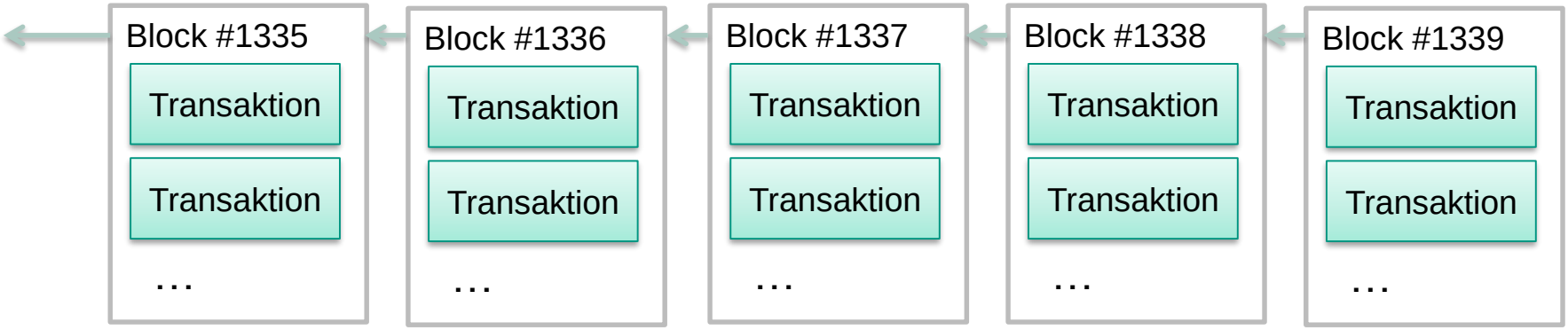
  - Realisiert wird die Blockchain durch ein Peer-to-Peer-Netz
  - Alle Peers kennen die komplette Blockchain
  - Kein Vertrauen in individuelle Peers nötig
    - Stattdessen Vertrauen, dass keine Gruppe von böswillig kooperierenden Nutzern existiert, die zusammen über die Mehrheit einer bestimmten Ressource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



Blockchain-Crashkurs: Neue Einträge (I)

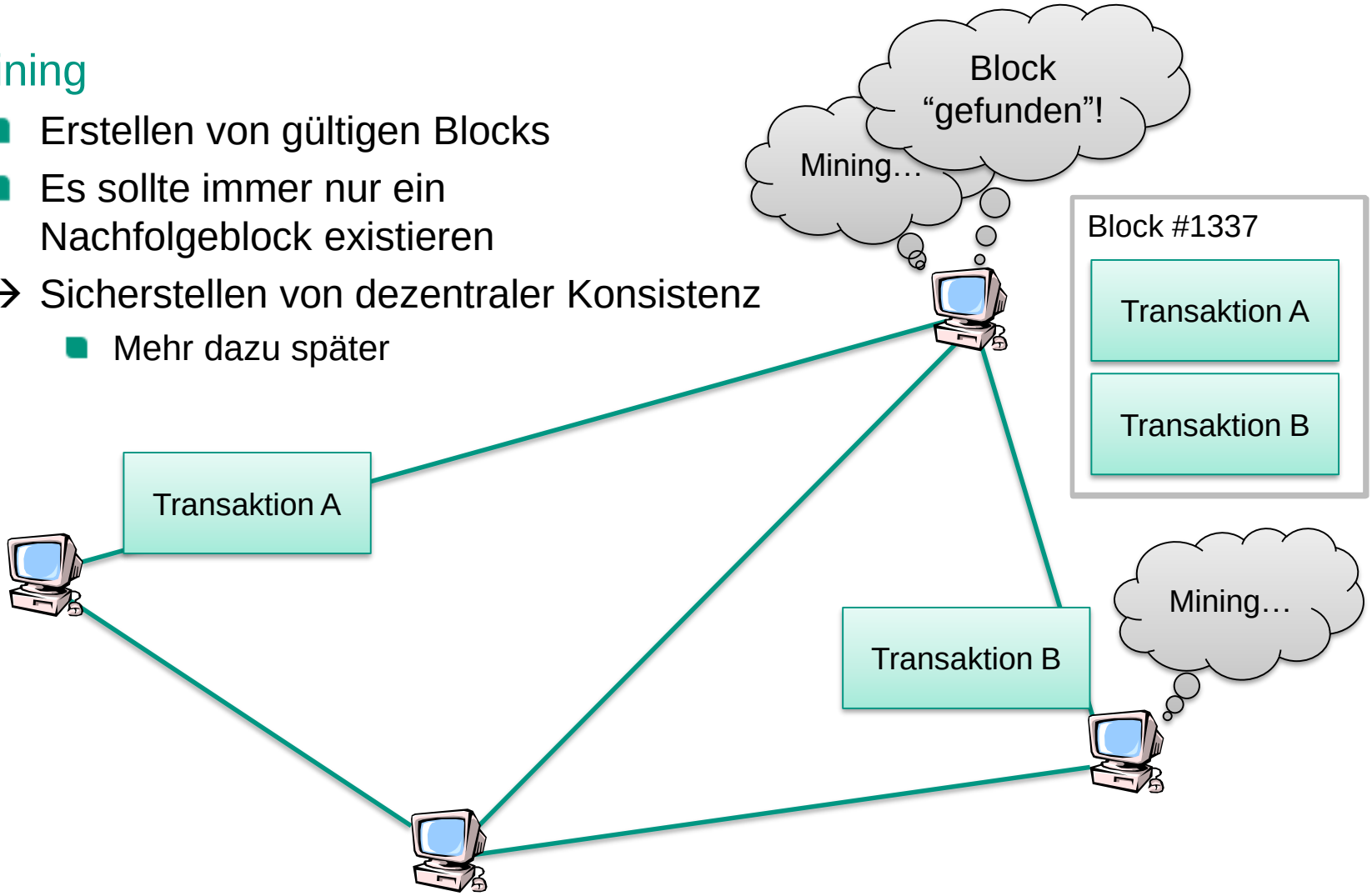
- Blockchain-Einträge werden typischerweise Transaktionen genannt
  - (aus dem Kontext: Zahlungssystem)
- Neue Transaktionen
  - Werden zunächst an alle Peers geflutet
  - Dann aber noch nicht Teil der Blockchain!
- Blocks
  - Eigentliche Bausteine der Blockchain
  - Enthalten mehrere Transaktionen
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





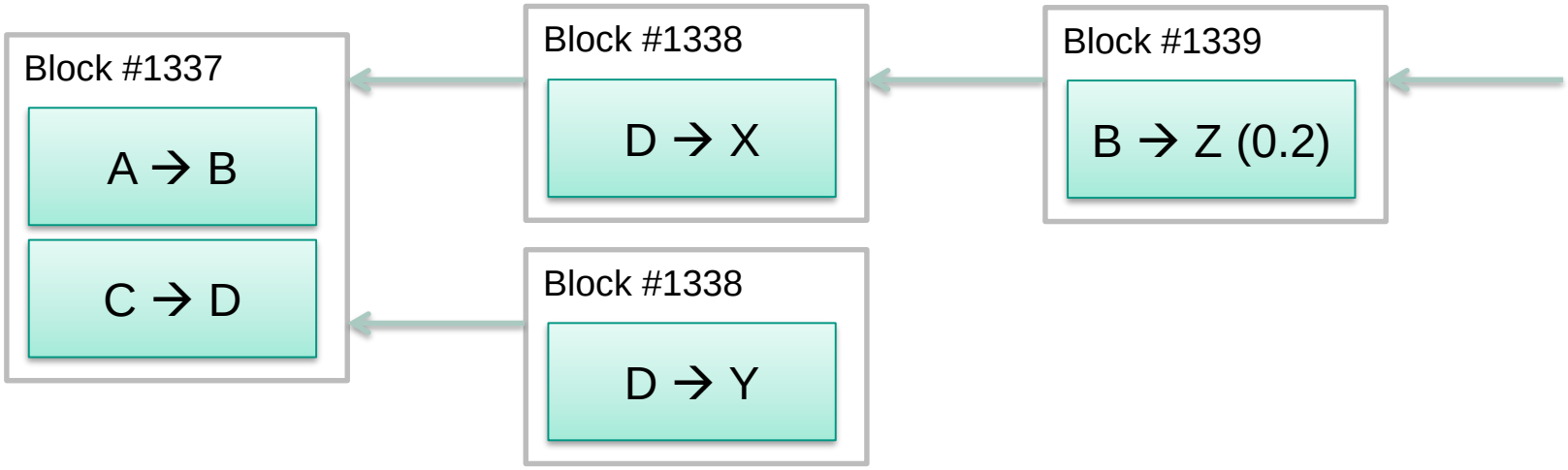
Blockchain-Crashkurs: Neue Einträge (II)

- Mining
  - Erstellen von gültigen Blocks
  - Es sollte immer nur ein Nachfolgeblock existieren
  - → Sicherstellen von dezentraler Konsistenz
    - Mehr dazu später



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

Bitcoin

- Dezentrales Zahlungssystem
- Basierend auf einer Blockchain
- „Konten“ entsprechen asymmetrischen Schlüsselpaaren

- Überweisungen erfolgen als Transaktionen
  - Können nur am Stück ausgegeben werden
  - Enthalten Verweise auf frühere, noch nicht verwendete, Transaktionen
    - Gültige kryptographische Signatur muss angegeben werden, um die eingehenden Transaktionen auszugeben
  - Enthalten Adresse(n), an die überwiesen werden soll

- Neue Bitcoins werden erzeugt, indem Blöcke erzeugt werden

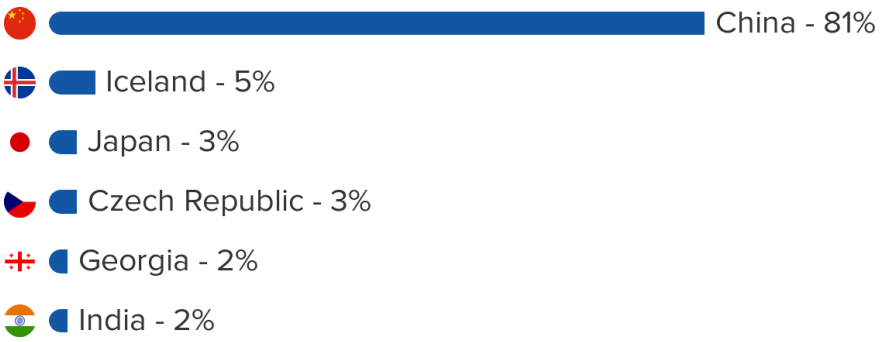
- Skalierung des Verfahrens problematisch
  - Es passen nur begrenzt viele Transaktionen in jeden Block

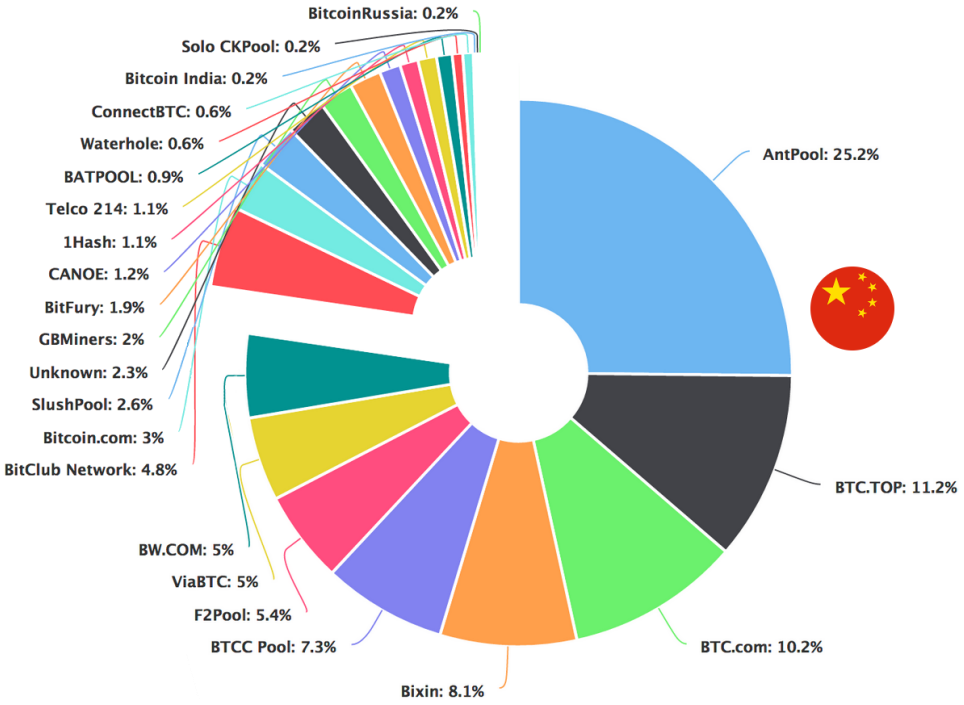
[Naka2008]

Bitcoin: Eine dezentrale Währung?

- Mining für Bitcoin soll dezentral erfolgen
  - Keine einzelne Instanz, die das ganze Netzwerk kontrollieren kann
- Individuelles Mining aber nicht mehr profitabel
  - → Zusammenschluss zu Mining Pools
  - Weiterhin Aufbau von Mining Farmen

- Theoretisch mal dezentralisiert, aber mittlerweile kaum mehr





Privatsphäre

- „Bitcoin ist eine anonyme Währung!“

- Naja…
  - Alle Transaktionen sind öffentlich sichtbar
  - Geldflüsse sind somit nachvollziehbar

- Anonymität ist somit nur begrenzt gegeben
  - Enttarnung möglich über Kombination mit anderen Informationsquellen, zum Beispiel regelmäßige Einkäufe

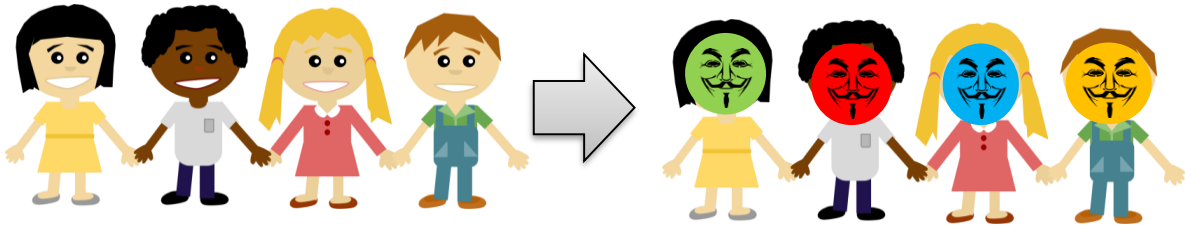
- Mögliche Lösung: Mixen von Transaktionen
  - Mehrere Leute erzeugen eine gemeinsame Transaktion
  - Danach ist nicht mehr unterscheidbar, welche Auszahlung mit welcher Einzahlung der Transaktion verknüpft ist

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

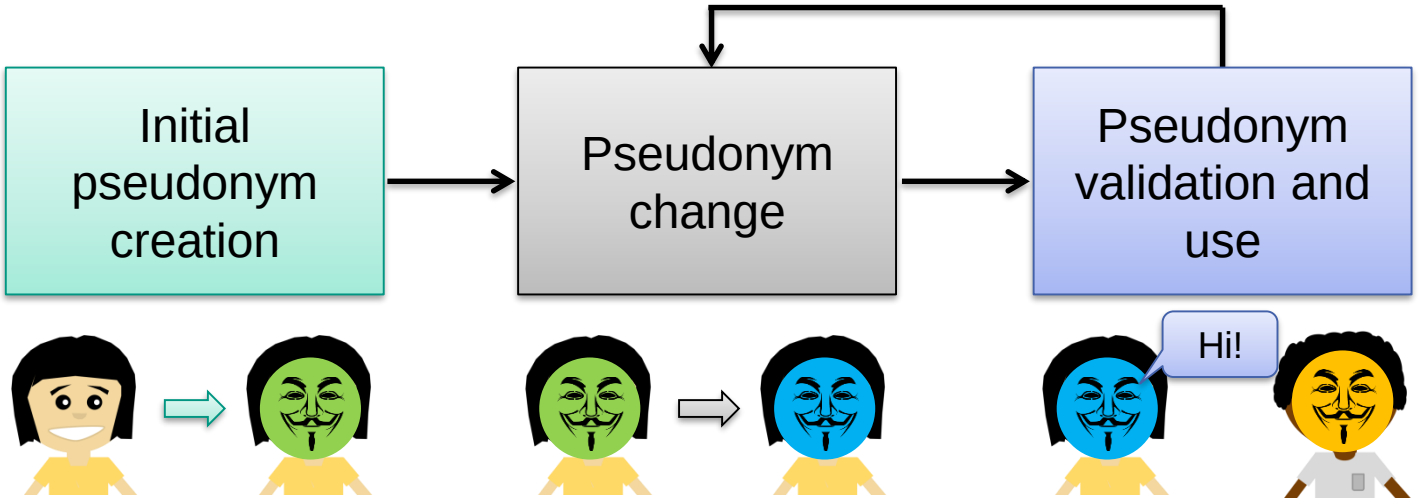


- Status quo: Pseudonyme werden durch Vertrauensanker verteilt
- Unser Ansatz: Vertrauensanker ersetzen mithilfe von Blockchain

[Flor2015]

Eigene Arbeiten: BitNym (II)

- Bausteine



- Pseudonyme auf Blockchain abspeichern
- Sybil-resistente Zugangskontrolle

- Unverkettbare Pseudonym-wechsel

- Sybil-Resistenz beibehalten trotz Wechsel

- Blacklisting?
- Reputation?

Smart Contracts

- Jeder Miner überprüft sämtliche Transaktionen und Blöcke dezentral
  - Damit auch die Skripte der Zahlungsbedingungen der Transaktionen
- Idee: Warum nicht die Skripte erheblich leistungsstärker gestalten?
- Smart Contracts
  - Komplexere Berechnungen dezentral durchführen
  - „Smarte Verträge“: Automatisierte Zahlungen basierend auf Zustand
  - Beispiel: Ein Gewinnspiel
    - Jeder Teilnehmer zahlt eine gewisse Summe ein, bis genügend Nutzer eingezahlt haben
    - Bei der letzten Einzahlung: Wähle einen zufälligen Nutzer aus
      - Zufall?? Aber das System muss doch deterministisch sein!
    - Der Gewinner bekommt das angesammelte Geld überwiesen

Ethereum

- Bitcoin ist primär für Zahlungen ausgelegt
- Im Gegensatz dazu: die Blockchain Ethereum

- Primäre Aufgabe ist die Verwendung von Smart Contracts
  - Zahlungsverkehr ist ebenfalls möglich
- Smart Contracts werden in der EVM (Ethereum Virtual Machine) ausgeführt
- Jede Operation der Smart Contracts kostet Gas
  - Maximale Menge an zu verwendendem Gas wird dem Aufruf mitgegeben
  - Verhindert Probleme mit endlosen Skripts
    - Wenn das Gas aufgebraucht ist, wird die Ausführung abgebrochen
  - Soll Spam auf der Blockchain vermeiden

- Verwendet aktuell Proof-of-Work, soll zukünftig Proof-of-Stake werden

Einschub: Konsistenzverfahren für Blockchains

- Proof-of-Work
  - Benötigt Rechenressourcen
  - Wer ein Krypto-Puzzle zuerst löst, darf den Block erstellen
  - Idee bei Bitcoin:
    - Blöcke enthalten eine frei wählbare Nonce
    - Global bekannte Difficulty: Mit wie vielen 0en muss der Hash beginnen
    - „Gewinnen“ tut der Miner, der seine Nonce so wählt, dass der Hash über den Block mit genügend vielen 0en beginnt

- Proof-of-Stake
  - Benötigt einen Einsatz (z.B. Bitcoins)
  - Grundlegende Idee:
    - Je mehr Einsatz der Miner bereit stellt, desto höher die Chance, dass er den nächsten Block erzeugen darf
  - Im Gegensatz zu Proof-of-Work: Keine verschwendete Rechenleistung

- Noch weitere Verfahren denkbar…

Beispiel: Smart Contract in Ethereum

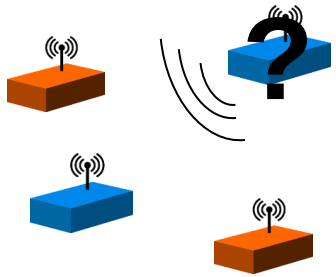
- Einfacher Smart Contract geschrieben in Solidity
- Speichert einen String
- Ist in der Lage, diesen später zurückzugeben
- Zusätzlich: Der Ersteller des Contracts ist in der Lage, diesen später wieder aus dem Netz zu löschen
  - Löschen entfernt den Contract aus der Zustands-Datenbank
  - Die Blockchain wird nicht verändert

  → Der Zustand kann wiederhergestellt werden!

```
contract mortal {
    address owner;

    /* Constructor */
    function mortal() { owner = msg.sender; }

    /* Recover the funds on the contract */
    function kill() {
        if (msg.sender == owner)
            selfdestruct(owner);
    }
}

contract greeter is mortal {
    string greeting;

    function greeter(string _greeting) public {
        greeting = _greeting;
    }

    function greet() constant returns (string) {
        return greeting;
    }
}
```

Eigene Arbeiten: Identitäten mit Blockchains (I)

- IoE-Bezug: Geräte brauchen eindeutige Identität
  - Z.B. für authentifizierte Kommunikation mit anderen Geräten
  - Zusätzlich zur Identität noch Attribute speichern, z.B. Gerätetyp
  - Außerhalb des IoE auch auf Personen anwendbar
- Zentralisiert kritisch wegen Privatsphäre und Einrichtungsaufwand
  - → Ein dezentraler Ansatz ist vorzuziehen



- Widersprüchliche Ziele:
  - Vertrauenswürdige Identitäten sollen dezentral erzeugbar sein
    - Möglicherweise auch mehrere Identitäten pro Gerät
  - Attribute einer Identität sollen mit anderen Geräten geteilt werden können
  - Unbekannte Geräte sollen die Attribute nicht lesen dürfen

Eigene Arbeiten: Identitäten mit Blockchains (II)

- Ansatz: Identitäten als Smart Contracts in der Blockchain abbilden
  - Die Adressen der Contracts bildet feste „Namen“
  - Durch die Logik der Contracts kann Zugangskontrolle implementiert werden

- Privatsphäre kann durch symmetrische Verschlüsselung der Attribute erreicht werden
  - Schlüssel können sicher in der Blockchain hinterlegt werden
  - Erspart erneuten Schlüsselaustausch, wenn der Schlüssel geändert wird

- Zugriffsrecht auf ein Attribut erfolgt durch Zugriff auf den Schlüssel
  - Durch die Smart Contracts kann zwischen Lese- und Schreibrecht unterschieden 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: Smart Contracts können beliebig komplexe Regeln enthalten, die von Minern überprüft und ausgeführt werden
    - Aktives Experimentier- und Forschungsfeld!

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.

[Danez2014] Danezis, George, et al.: Privacy and Data Protection by Design - from policy to engineering; Report, European Union Agency for Network and Information Security (ENISA), Dezember 2014.

[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

[Flor2015] M. Florian, J. Walter, und I. Baumgart. Sybil-Resistant Pseudonymization and Pseudonym Change without Trusted Third Parties; Proceedings of the 14th Workshop on Privacy in the Electronic Society (WPES), Denver, Colorado, USA, Oktober 2015

[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