Cryptographic Treatment of Private User Profiles

CAST-Workshop am 24.11.2011 (CAST-Förderpreis 2011)

TECHNISCHE UNIVERSITÄT DARMSTADT

Felix Günther

TU Darmstadt

Das Internet und seine „Benutzer“

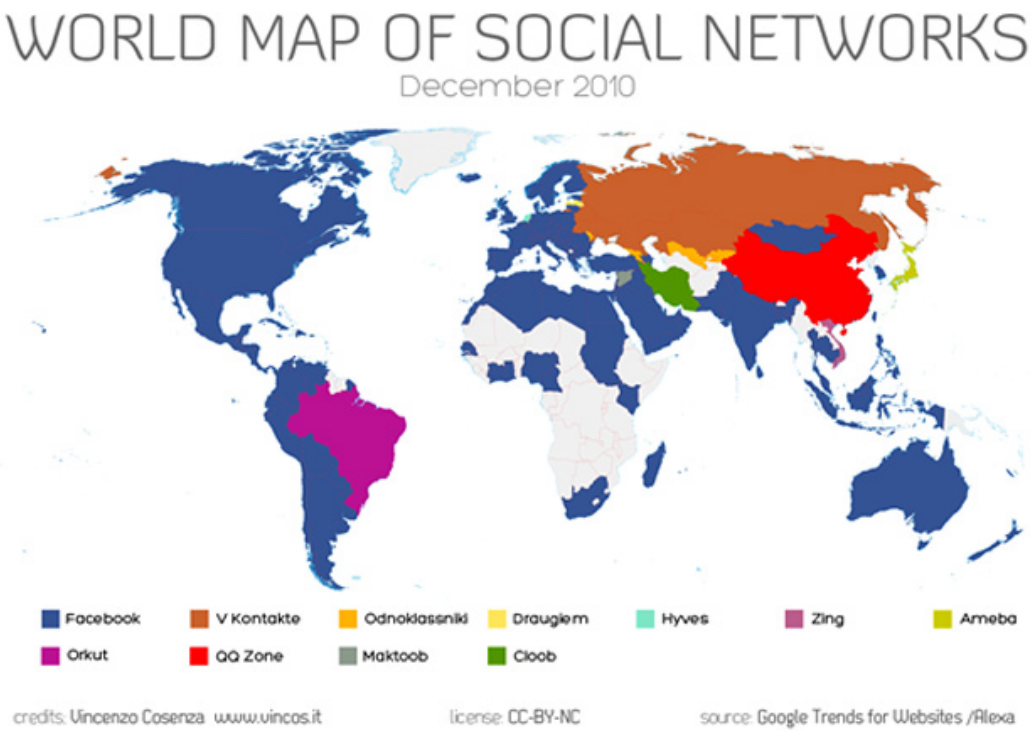
TECHNISCHE UNIVERSITÄT DARMSTADT

Soziale Netze, Interaktion und Online Communities dominieren Internet von heute.
Services von Benutzern für Benutzer – die Benutzer im Zentrum des Internets.

Soziale Netzwerke = Global Players

TECHNISCHE UNIVERSITÄT DARMSTADT

Soziale Netze sind überall auf der Welt populär wie nie.



Facebook allein hat über 800 Mio. aktive Nutzer, 50% loggen sich täglich ein.

Inhalte von Benutzern

TECHNISCHE UNIVERSITÄT DARMSTADT

Interaktion im Netz basiert auf von Benutzern eingestellten Inhalten.

Inhalte

- Personenbezogene Informationen: Name, Adresse, Tätigkeiten, ...
- Digitale Daten: Fotos, Videos, Kommentare, ...

Soziale Interaktion

- Daten publizieren (und verändern), von anderen publizierte Daten abrufen
- neue Kontakte knüpfen, (synchron/asynchron) kommunizieren

Inhalte + soziale Interaktion können Informationen über Benutzer verraten.

Inhalte von Benutzern

TECHNISCHE UNIVERSITÄT DARMSTADT

Interaktion i… …Inhalten.

Inhalte…

- Pe…
- Di…

Sozial…

- Date… …en
- neue… …ifizieren

Inhalte… …er Benutzer verraten.

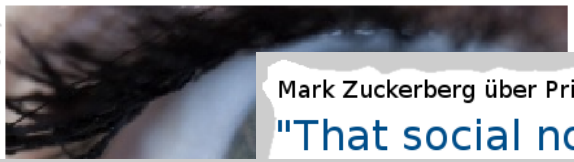
POST-PRIVACY

Cicero ONLINE MAGAZIN FÜR POLITISCHE KULTUR

Privatsphäre war gestern

VON **PAUL SOLBACH** 11. NOVEMBER 2011

picture alliance



Mark Zuckerberg über Privacy:

"That social norm is just something

Deine Facebook-Kommentare werden in Googles Suche angezeigt

t3n Open Web Business

Google hat damit begon… integrieren. Dies war bis… auf iFrames, bzw. AJAX u… diese Hürde überwunde… Kommentare ausgelesen…

Trotz Zensur gibt es im Iran 17 Millionen Facebook-Nutzer

vorlesen / MP3-Download

heise online

Trotz Internetzensur nutzen mindestens 17 Millionen Iraner das Social Netzwork …litärischen Organisation …gen der Regierung, den …rollieren, seien gescheitert, …ionsgarden angegliedert.

http://heise.de/-1356073

Karman widmet Nobelpreis arabischem Frühling

ZDF heute.de

Friedensnobelpreis an drei Frauenrechtlerinnen

Ein Preis für Frauen: Den Friedensnobelpreis erhalten Liberias Präsidentin Johnson-Sirleaf, die liberianische Aktivistin Gbowee und die jemenitische Journalistin Karman für den Kampf für Frauenrechte. Karman widmete den Preis dem arabischen Frühling.

http://www.heute.de/ZDFheute/inhalt/8/0,3672,8356296,00.html

Inhalte von Benutzern

TECHNISCHE UNIVERSITÄT DARMSTADT

Interaktion i

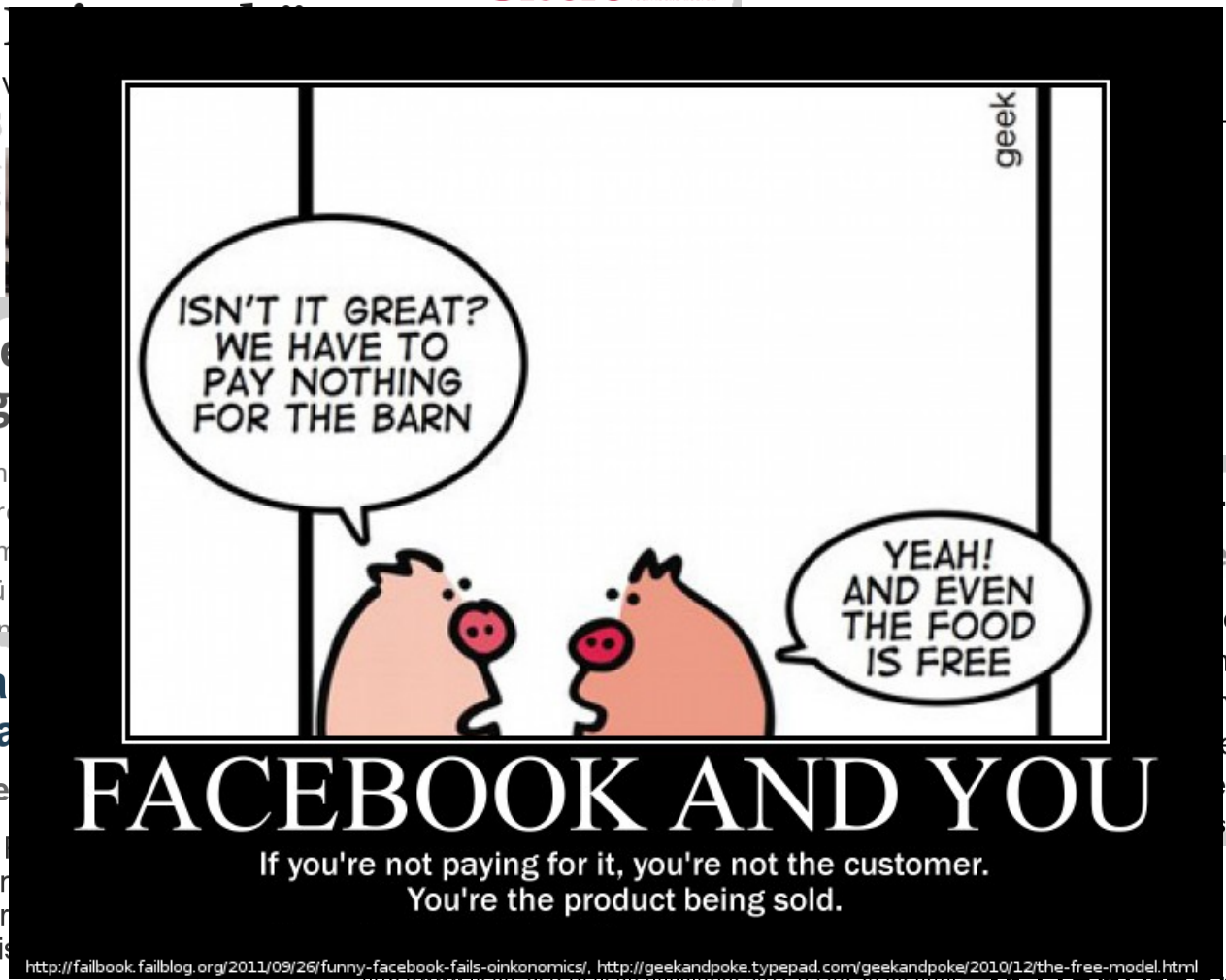
Inhalte

- Pe
- Di

Sozial

- Date
- neue

Inhalte



er Benutzer verraten.

Zentrale Aufgabe: Schutz von Privatsphäre und Benutzerdaten

Bisherige Ansätze

TECHNISCHE UNIVERSITÄT DARMSTADT

Nicht-kryptographische Ansätze

- [Carminati, Ferrari, Perego 2009] Zugriffskontrolle mit semantischen Regeln und Beweisen, semi-zentralisierte Infrastruktur, synchrone Kommunikation

Kryptographische Ansätze

- [Lucas, Borisov 2009] zentralisierter Ansatz, erfordert Vertrauen in Provider
- [Graffi et al. 2009, Baden et al. 2009 (OSN Persona)] Attribute-Based Encryption, ohne formales Modell, nur Vertraulichkeit der Daten
- [Jahid, Mittal, Borisov 2011 (EASiER)] Attribute-Based Encryption, semi-trusted Server, nur Vertraulichkeit der Daten

Fokus dieser Arbeit: Kryptographisches Modell und Benutzerprofile

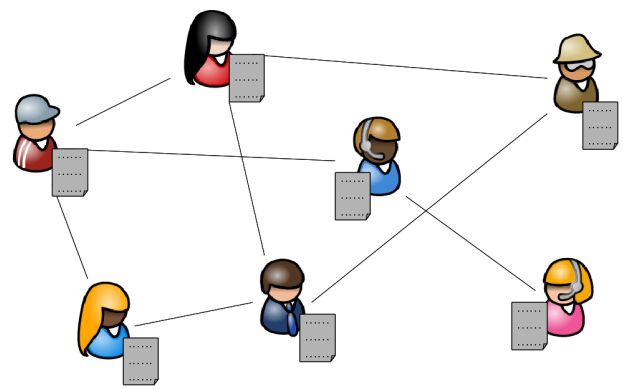
TECHNISCHE UNIVERSITÄT DARMSTADT

Vorteile eines kryptographischen Ansatzes

- Verfahren als Baustein für eigentliche Anwendung, vielfach einsetzbar
- Formales Modell: Was sind Profile? Was genau heißt Privatsphäre?
- Formale Beweise: Erreichen eines Schutzziels beweisbar.

Benutzerprofil

- zentrales Element jeder Online-Community
- „Benutzer = Profil“ in sozialen Netzen



Funktionalität

- Daten publizieren und abrufen
- Zugriffsrechte erteilen und entziehen

Formales Modell für Benutzerprofile

TECHNISCHE UNIVERSITÄT DARMSTADT

- Ein Profil P wird modelliert als Menge von Paaren

$$P \stackrel{def}{=} \left\{ (a, \bar{d}) \mid a \in \mathcal{I}, \bar{d} \in \{0,1\}^* \right\}$$

  - $\mathcal{I}$ ist eine Menge von eindeutigen Attribut-Indizes a
  - $\bar{d}$ ist der zugehörige, in P gespeicherte Wert

- P ist öffentlich und authentisiert durch seinen Besitzer U_P
- U_P besitzt einen Profil-Management-Schlüssel pmk
- U_P kennt Attribut d und autorisierte Gruppe G zu $(a, \bar{d}) \in P$

Besitzer U_P

pmk

Öffentliches Profil P

$(a_1, \bar{d}_1)$
$(a_2, \bar{d}_2)$
$(a_3, \bar{d}_3)$

Autorisierte Gruppen G_i

Indizes können auch Pseudonyme sein

Beispiele

a = Geburtstag	a = 2123
$\bar{d}$ = %$%§§“!	$\bar{d}$ = §#~&“;
d = 01.04.1982	d = Bart Simpson

Profil-Management-Schema

TECHNISCHE UNIVERSITÄT DARMSTADT

Ein Profil-Management-Schema besteht aus

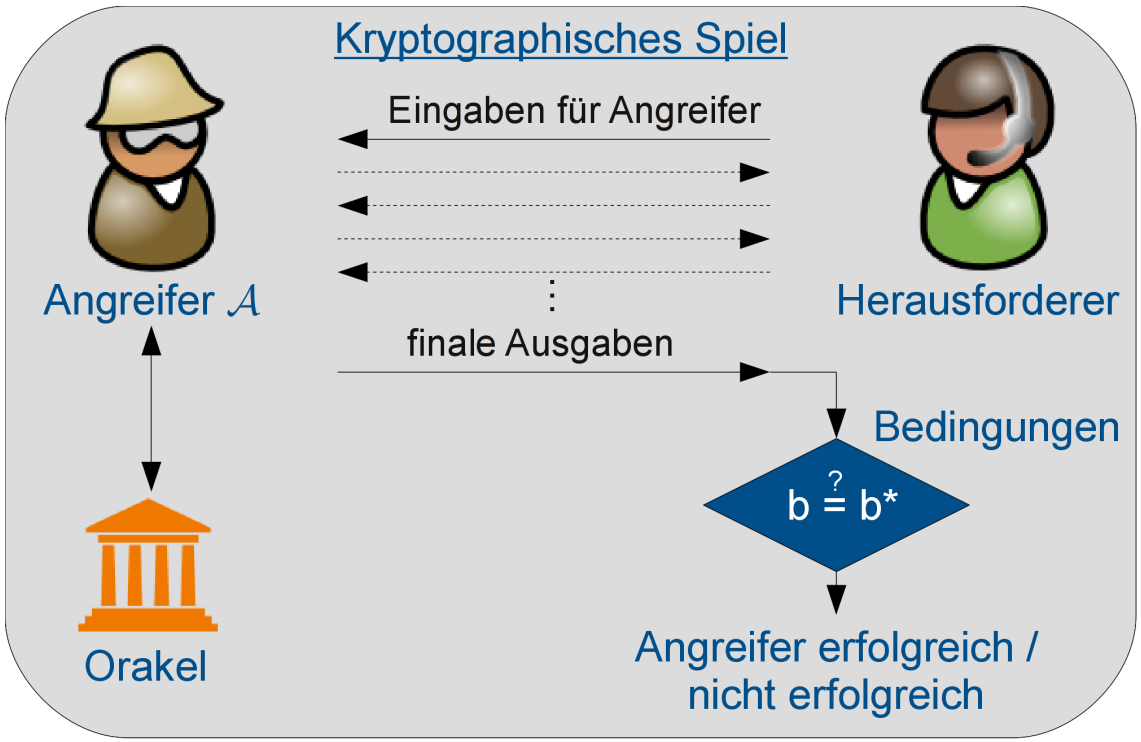
$\texttt{Init}(\kappa)$	Initialisiert P und erstellt pmk.
$\texttt{Publish}(pmk, P, (a, d), G)$	Fügt $(a, \bar{d})$ zu P hinzu, gibt Zugriffsschlüssel rk_U für jeden $U \in G$ aus.
$\texttt{Retrieve}(rk_U, P, a)$	Gibt entweder das Attribut d oder $\perp$ zurück.
$\texttt{Delete}(pmk, P, a)$	Löscht $(a, \bar{d})$ aus P.
$\texttt{ModifyAccess}(pmk, P, a, U)$	Gewährt oder widerruft Zugriffsrecht für U auf $(a, \bar{d}) \in P$.

Wenn U_P $(a, \bar{d})$ in P publiziert und nicht gelöscht hat
und U (nicht-widerrufene) Zugriffsrechte für a (als Teil seines rk_U) hat,
dann kann U das Attribut d abrufen.

Beweisbare Sicherheit

TECHNISCHE UNIVERSITÄT DARMSTADT

Wie Angriffe auf ein (formales) Krypto-Schema definieren?



Sicherheitsdefinition: Schema sicher $\Leftrightarrow \Pr[\mathcal{A}$ erfolgreich] vernachlässigbar

Angreifermodell für Profile

TECHNISCHE UNIVERSITÄT DARMSTADT

PPT-Angreifer $\mathcal{A}$ interagiert mit Nutzern und Profilen mittels Orakel-Anfragen:

Corrupt(U)	Korrumpiert U komplett, $\mathcal{A}$ erhält pmk und alle rk_U.
Publish(P, (a, d), G)	U_P veröffentlicht (a, d) und erteilt allen in G Zugriff.
Retrieve(P, a, U)	Ruft a aus P im Namen von U ab.
Delete(P, a)	U_P löscht (a, $\bar{d}$) aus P (falls existent).
ModifyAccess(P, a, U)	U_P gewährt oder widerruft Zugriffsrecht für U auf (a, $\bar{d}$) $\in P$.

$\mathcal{A}$ hat jederzeit (Lese-) Zugriff auf alle Profile im System.

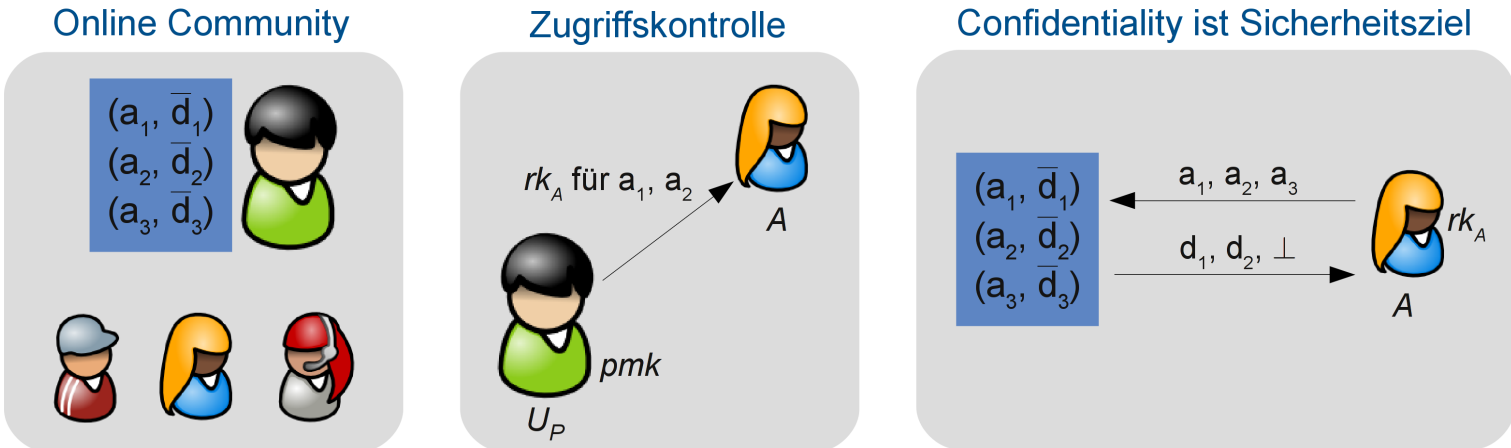
$\mathcal{A}$ ist adaptiv, erhält mittels Anfragen Kontrolle über (alle) Profile und Attribute.

Sicherheitsziel: Confidentiality

TECHNISCHE UNIVERSITÄT DARMSTADT

U_P veröffentlicht Paare $(a, \bar{d})$ in P und verteilt Zugriffsschlüssel rk_U an Benutzer U.

Confidentiality: Attribute d sind vor nicht-autorisierten Nutzern geschützt.



Indistinguishability-Ansatz:

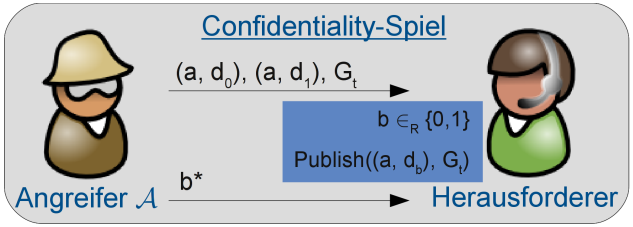
$\mathcal{A}$ ohne Zugriffsrechte für $(a, \bar{d})$ kann nicht unterscheiden, welches Attribut d in $\bar{d}$ verschlüsselt ist.

. . . sogar wenn $\mathcal{A}$ andere Attribute im gleichen Profil lesen kann.

Sicherheitsziel: Confidentiality

Formale Definition

TECHNISCHE UNIVERSITÄT DARMSTADT



Confidentiality-Spiel (high-level):

1. Führe $\texttt{Init}(\kappa)$ für jeden Benutzer U aus.
2. $\mathcal{A}$ interagiert mit Nutzern durch Anfragen und gibt aus:
   - zwei Index-Attribut-Paare (a, d_0), (a, d_1)
   - Gruppe von Benutzern G_t
   - Profilbesitzer U_P, der nicht in G_t ist
3. Wähle Bit $b \in_R \{0, 1\}$. Führe $\texttt{Publish}(pmk, P, (a, d_b), G_t)$ aus.
4. $\mathcal{A}$ interagiert mit Nutzern durch Anfragen und gibt ein Bit b^* aus.

$\mathcal{A}$ ist erfolgreich wenn:

- $b = b^*$
- $\mathcal{A}$ hat weder U_P noch einen jemals für a autorisierten Benutzer korrumpiert
- $\mathcal{A}$ hat d_b nicht trivial mittels Retrieve-Anfrage abgerufen

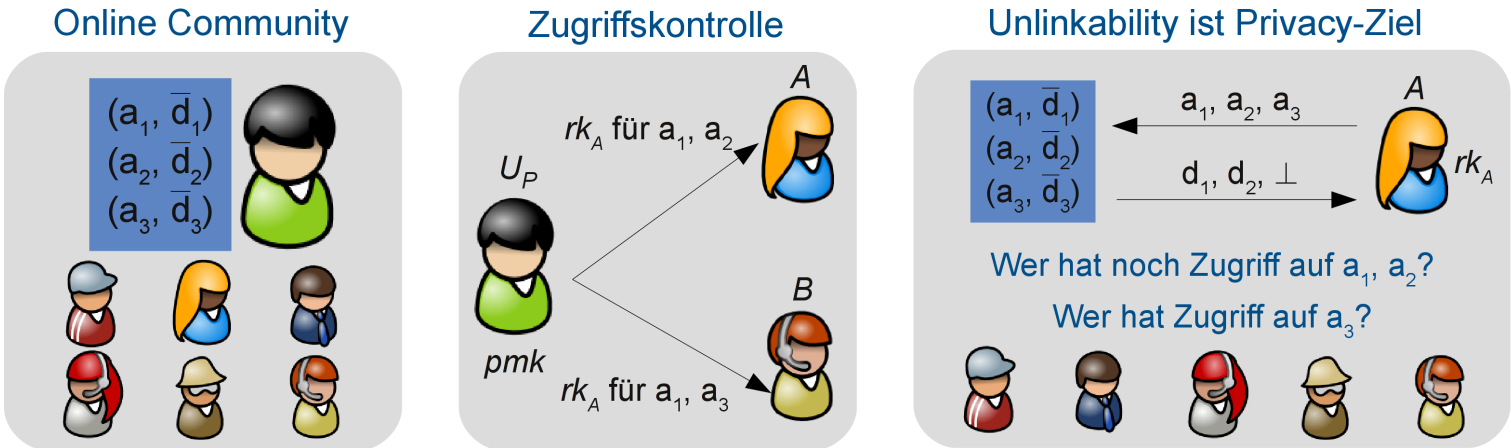
Ein Profil-Management-Schema ist confidential wenn für alle $\mathcal{A}$ gilt:
$|\Pr[\text{erfolgreicher Angriff}] - 1/2|$ ist vernachlässigbar in κ.

Privacy-Ziel: Unlinkability

TECHNISCHE UNIVERSITÄT DARMSTADT

Besitzer U_P weiß, welche Benutzer Zugriff auf welche Paare $(a, \bar{d})$ in P haben.

Unlinkability: Profile verraten nicht, wer auf welche Attribute zugreifen darf.



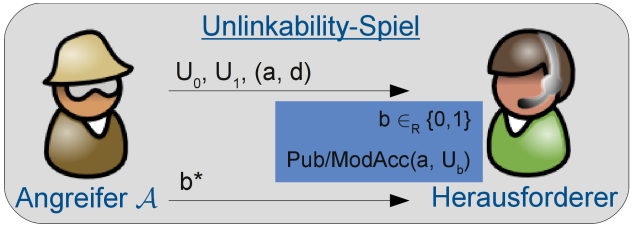
Indistinguishability-Ansatz:

$\mathcal{A}$ mit Zugriffsrechten für $(a, \bar{d})$ kann nicht unterscheiden, ob Nutzer A oder Nutzer B Zugriff auf a gewährt wurde.

Privacy-Ziel: Unlinkability

Formale Definition

TECHNISCHE UNIVERSITÄT DARMSTADT



Unlinkability-Spiel (high-level):

1. Führe $\texttt{Init}(\kappa)$ für jeden Benutzer U aus.
2. $\mathcal{A}$ interagiert mit Nutzern durch Anfragen und gibt aus:
   - zwei Benutzer U_0, U_1, Index-Attribut-Paar (a, d), Profilbesitzer U_P
3. Wähle Bit $b \in_R \{0, 1\}$.
   - Wenn $(a, \cdot) \notin P$, führe $\texttt{Publish}(pmk, P, (a, d), \{U_b\})$ aus.
   - Wenn $(a, \cdot) \in P$, führe $\texttt{ModifyAccess}(pmk, P, a, U_b)$ aus.
4. $\mathcal{A}$ interagiert mit Nutzern durch Anfragen und gibt ein Bit b^* aus.

$\mathcal{A}$ ist erfolgreich wenn:

- $b = b^*$
- U_P, U_0 und U_1 wurden nicht korrumpiert
- $\mathcal{A}$ hat weder Retrieve(P, a, U_0) noch Retrieve(P, a, U_1) angefragt

Ein Profil-Management-Schema ist unlinkable wenn für alle $\mathcal{A}$ gilt:
$|\Pr[\text{erfolgreicher Angriff}] - 1/2|$ ist vernachlässigbar in κ.

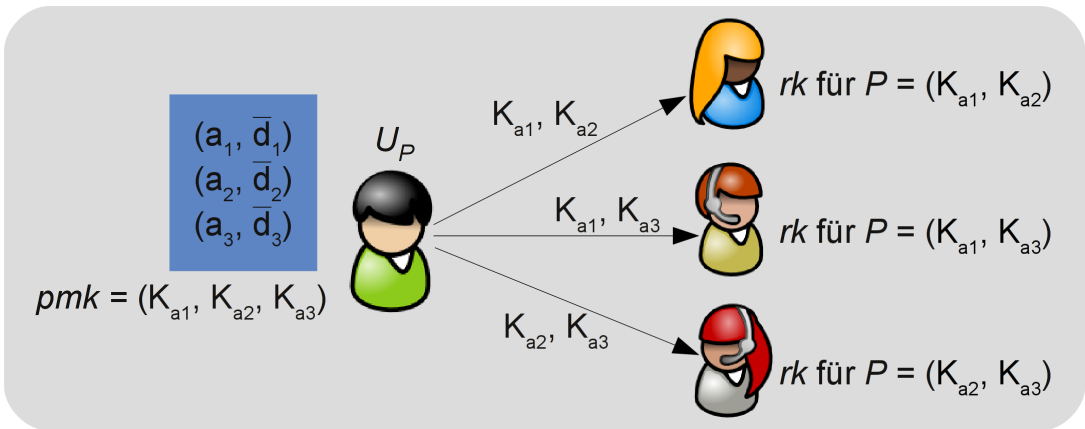
Shared Key (SK) - Verfahren

TECHNISCHE UNIVERSITÄT DARMSTADT

- Intuitiver Ansatz: Teile einen geheimen Schlüssel pro Attribut.
- Separate Schlüssel $K_a \leftarrow SE.KGen(\kappa)$ für jedes $(a, \bar{d})$: $\bar{d} = SE.Enc(K_a, d)$
- Revokation: erneute Verschlüsselung mit neuem K_a

(SE.KGen, SE.Enc, SE.Dec)
CCA-sicheres Sym. Enc. Verfahren

KGen(κ): gibt Schlüssel K aus
Enc(K, M): gibt Chiffrat C aus
Dec(K, C): gibt Nachricht M oder ⊥ aus



- Gewährleistet Confidentiality und perfekte Unlinkability
- Aber: Jeder Benutzer muss einen Schlüssel pro Attribut pro Profil speichern

Broadcast Encryption (BE) - Verfahren

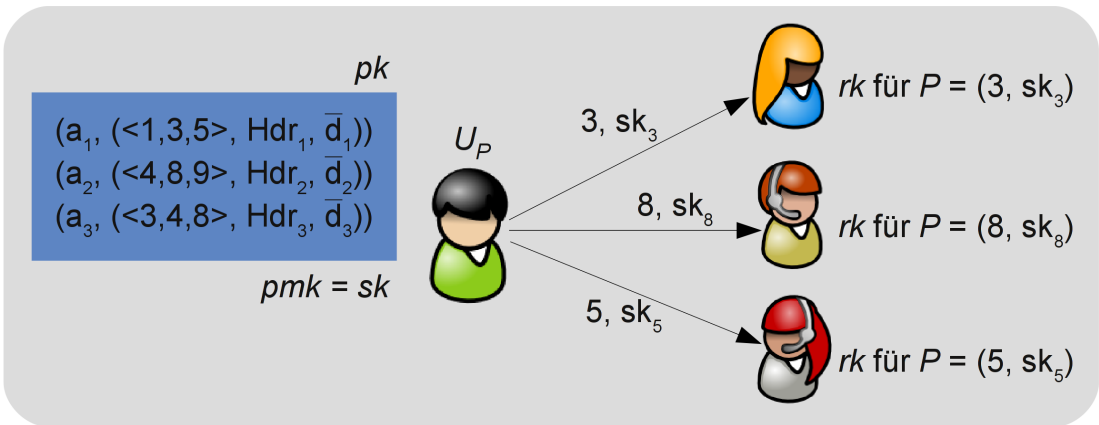
TECHNISCHE UNIVERSITÄT DARMSTADT

- Nutzer verwalten je eigene Broadcast-Gruppe mit $(pk, sk) \leftarrow BE.Setup(\kappa, n)$
- Jeder autorisierte Benutzer i erhält einen Schlüssel sk_i pro Profil
- Für jedes (a, d) und autorisierte Nutzer S: $(Hdr, K_a) \leftarrow BE.Enc(S, pk)$, $\hat{d} = SE.Enc(K_a, d)$ und dann $\bar{d} = (Hdr, S, \hat{d})$.
- Revokation: erneute Verschlüsselung mit neuem (Hdr, K_a) für modifiziertes S

(BE.Setup, BE.KGen, BE.Enc, BE.Dec)
adaptiv CCA-sicheres Br. Enc. Verfahren

Setup(κ, n): gibt Paar (sk, pk) aus
KGen(i, sk): gibt Schlüssel (i, sk_i) aus
Enc(S, pk): gibt (Hdr, K) aus
Dec(S, i, sk_i, Hdr): gibt K oder ⊥ aus

Schlüssel K verwendet mit SE-Verfahren
[Gentry, Waters 2009]



- Erreicht Confidentiality und perfekte Anonymity (schwächer als Unlinkability)
- Dafür: Jeder Benutzer muss nur einen Schlüssel pro Profil speichern

Overhead für die Speicherung der Schlüssel

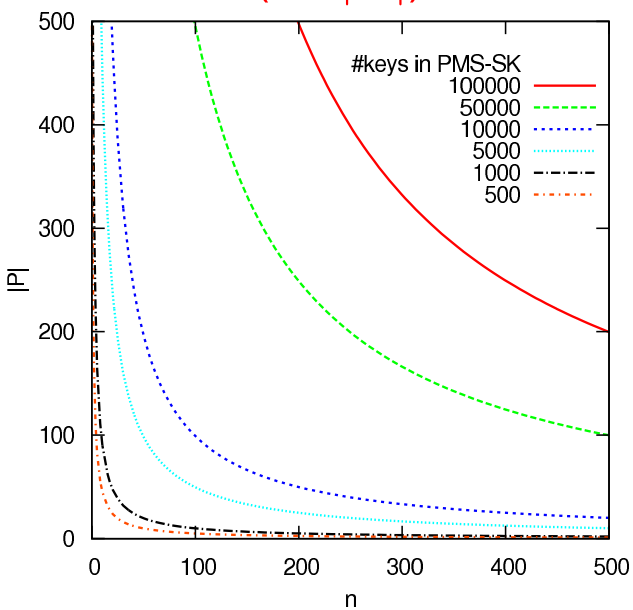
TECHNISCHE UNIVERSITÄT DARMSTADT

- Jeder Benutzer hat im Durchschnitt n Kontakte.
- Jeder Benutzer teilt im Schnitt $|P|$ Attribute mit jedem seiner Kontakte.

SK-Verfahren

$(n+1) \cdot |P|$ Schlüssel pro Nutzer

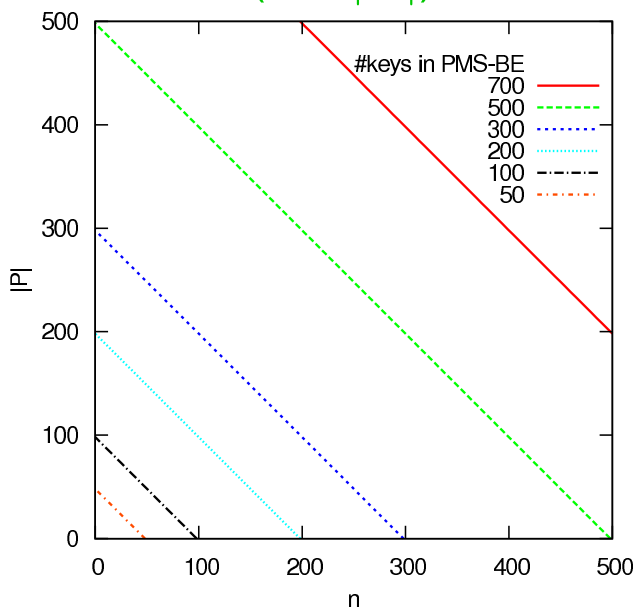
$O(n \cdot |P|)$



BE-Verfahren

$n + 2 + |P|$ Schlüssel pro Nutzer

$O(n + |P|)$



Overhead für die Speicherung der Schlüssel

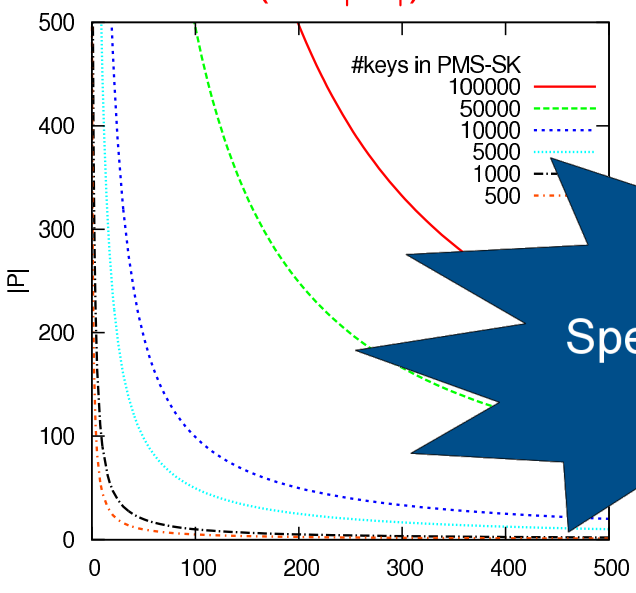
TECHNISCHE UNIVERSITÄT DARMSTADT

- Jeder Benutzer hat im Durchschnitt n Kontakte.
- Jeder Benutzer teilt im Schnitt $|P|$ Attribute mit jedem seiner Kontakte.

SK-Verfahren

$(n+1) \cdot |P|$ Schlüssel pro Nutzer

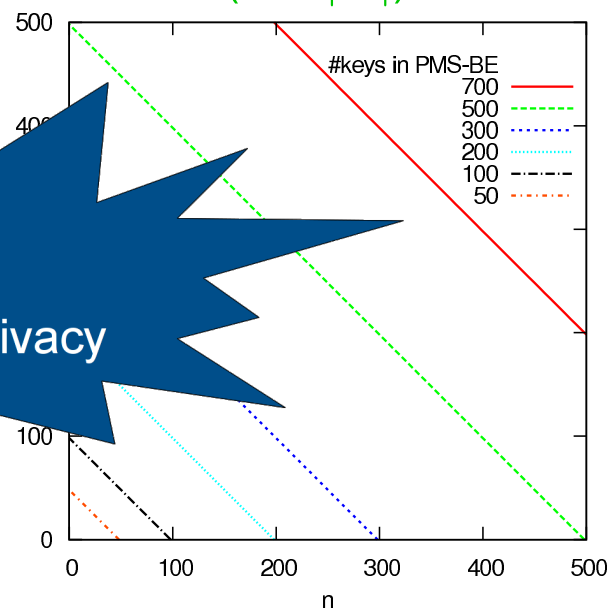
$O(n \cdot |P|)$



BE-Verfahren

$n + 2 + |P|$ Schlüssel pro Nutzer

$O(n + |P|)$



Trade-Off
Speicher vs. Privacy

Schlüssel-Updates mittels Group Key Management (LKH, OFT) optimierbar

Auswirkungen in realen Communities

TECHNISCHE UNIVERSITÄT DARMSTADT

Analyse für Facebook, Twitter, XING, Flickr (basierend auf deren Statistiken)

Community	# Kontakte	# Attribute	# Schlüssel		Speicherbedarf (KB)*	
			SK	BE	SK	BE
facebook	150	180	~27000	332	650	8
twitter	50	180	~9000	232	220	6
XING	168	~36	~8350	220	200	5
flickr	12	200	2000	214	62	5

* 192-bit-Schlüssel (SE und BE)

- Kosten für SK und BE differieren um Faktor 10 bis 80
- In SK und BE liegt Overhead unter 1 MB, was zumeist akzeptabel sein dürfte

Zusammenfassung

TECHNISCHE UNIVERSITÄT DARMSTADT

Erstes kryptographisches Modell für private Benutzerprofile

- Sicherheitsziel: Confidentiality von Profildaten (einzelnen Attributen)
- Privacy-Ziel: Unlinkability / Anonymity autorisierter Benutzer
- Sicherheit / Privacy in zentralisierten und dezentralisierten Netzen

Zwei Verfahren als Bausteine

- SK: Symmetrische Verschlüsselung, ein Schlüssel pro Attribut, Conf. + Unlink.
- BE: Broadcast-Verschlüsselung, ein Schlüssel pro Profil, Conf. + Anon.

Theoretische und Praktische Evaluation

- Trade-Off: Speicherbedarf vs. Privacy, beide Verfahren praktikabel

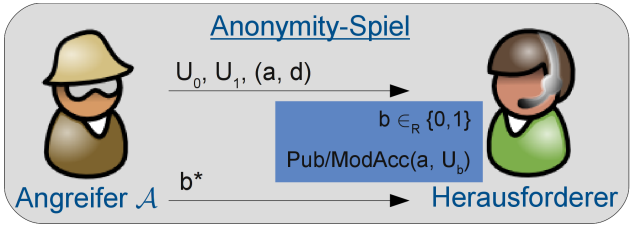
Ausblick

- Umsetzung in Browser-Plugin für existierendes Netzwerk
- Instantiierung mit Anonymous Broadcast Encryption [Libert et al. 2011]

Privacy-Ziel: Anonymity

Formale Definition

TECHNISCHE UNIVERSITÄT DARMSTADT



Anonymity-Spiel (high-level):

1. Führe `Init`(κ) für jeden Benutzer U aus.
2. $\mathcal{A}$ interagiert mit Nutzern durch Anfragen und gibt aus:
   - zwei Benutzer U_0, U_1, Index-Attribut-Paar (a, d), Profilbesitzer U_P
3. Wähle Bit $b \in_R \{0, 1\}$.
   - Wenn $(a, \cdot) \notin P$, führe `Publish`($pmk, P, (a, d_b), \{U_b\}$) aus.
   - Wenn $(a, \cdot) \in P$, führe `ModifyAccess`(pmk, P, a, U_b) aus.
4. $\mathcal{A}$ interagiert mit Nutzern durch Anfragen und gibt ein Bit b^*.

$\mathcal{A}$ ist erfolgreich wenn:

- $b = b^*$
- U_P, U_0 und U_1 wurden nicht korrumpiert
- $\mathcal{A}$ hat weder Retrieve(P, a, U_0) noch Retrieve(P, a, U_1) angefragt
- *U_0 zugriffsberechtigt für ein Attribut $\iff$ U_1 ist ebenfalls zugriffsberechtigt*

Ein Profil-Management-Schema ist anonym wenn für alle $\mathcal{A}$ gilt:
$|\Pr[\text{erfolgreicher Angriff}] - 1/2|$ ist vernachlässigbar in κ.