Anonymität auf Anwendungsebene

Sebastian Clauß, Stefan Schiffner
TU Dresden
{sc2,Stefan.Schiffner}@inf.tu-dresden.de

Abstract: Zur Nutzung von Dienstleistungen im Internet werden oft personenbezogene Daten von Nutzern angefordert. Um Nutzern trotzdem zu ermöglichen, ihre Privatsphäre bestmöglich zu schützen, sollten Nutzer persönliche Identitätsmanagementsysteme einsetzen können. Um zu entscheiden, ob bzw. welche personenbezogenen Daten herausgegeben werden sollen, ist eine Anonymitätsbewertung der Datenherausgabe wünschenswert.

In dieser Arbeit wird ein Modell für Nutzerdaten vorgestellt und darauf aufbauend ein statistisches Angreifermodell sowie eine konkrete Umsetzung des Modells entwickelt. Es wird gezeigt, wie anhand dieses Modells die Anonymität eines Nutzers im Bezug auf die Herausgabe personenbezogener Daten gemessen werden kann. In diesem Zusammenhang wird die Shannon-Entropie als Anonymitätsmaß für einzelne Nutzer kritisiert, und ein lokales Anonymitätsmaß definiert.

Es wird erläutert, wie das vorgeschlagene Modell in einem Identitätsmanagementsystem eingesetzt werden kann. Als Teil dessen wird außerdem das Problem der Übermittlung von Falschinformation diskutiert, und es wird gezeigt, daß auch deren Auswirkung auf die Anonymität des Nutzers im vorgeschlagenen Modell abgebildet werden kann.

1 Einleitung

Im Internet offerieren eine Vielzahl von Dienstleistern eine unüberschaubare Zahl von Dienstleistungen. Abhängig vom Dienst können personenbezogene Daten des Nutzers benötigt werden, damit der Dienst erbracht werden kann. Einige dieser Informationen wird der Dienstleister zu seiner eigenen Sicherheit in einer für ihn auf Korrektheit überprüfbaren Form fordern. Der Nutzer hingegen hat ein gewisses Bedürfnis auf informationelle Selbstbestimmung. Er möchte nicht jedem Dienst seine volle Identität offenbaren.

Prinzipiell ist eine selektive Herausgabe von personenbezogenen Daten möglich, wenn Dienstleister attributbasierte Zugriffssysteme [WWJ04] einsetzen. Wenn Attribute hierbei in Form von anonymen Credentials ausgetauscht werden, kann eine minimale Verkettbarkeit der herausgegebenen Daten erreicht werden.

Es ist allerdings oft nicht klar, wie stark die Herausgabe einer bestimmten Information die Privatsphäre eines Nutzers kompromittieren kann. Der Nutzer benötigt hierzu ein Werkzeug, das eine möglichst objektive Einschätzung darüber ermöglicht, inwieweit ein Dienstleister den Nutzer anhand einer erhaltenen Information identifizieren kann.

Wie kann ein solches Werkzeug aussehen? In dieser Arbeit wird ein Modell vorgestellt, das das Wissen eines Dienstleisters (im Folgenden *Angreifer* genannt) statistisch beschreibt. Es wird gezeigt, wie dieses Wissensmodell genutzt werden kann, um die Einschränkung der Anonymität des Nutzers durch die Herausgabe seiner personenbezogenen Daten zu messen.

Die Arbeit ist in vier weitere Sektionen unterteilt. In der direkt folgenden werden themenverwandte Arbeiten vorgestellt und abgegrenzt. In Abschnitt 3 wird das zugrunde liegende System sowie ein allgemeines Angreifermodell beschrieben. Im Abschnitt 4 wird ein konkretes Angreifermodell angegeben, und dessen Konsistenz bewiesen. Es wird u.a. gezeigt, wie anhand dieses konkreten Modells Anonymität gemessen werden kann, und es wird die Shannon-Entropie als Anonymitätsmaß kritisiert. Fernerhin wird in Abschnitt 5 eine mögliche Anwendung des Modells in einem Identitätsmanagementsystem zum Schutz der

Privatsphäre von Nutzern benannt. Als Teil dessen wird außerdem das Problem der Übermittlung von Falschinformation diskutiert.

2 Rückblick — Anonymität messen und erhalten

Viele Wissenschaftler haben sich in der Vergangenheit mit Anonymität und Verkettbarkeit sowie deren Messbarkeit beschäftigt. Pfitzmann und Hansen definieren in [PH05] die *Anonymitätsmenge* als eine Menge von Nutzern, die sich ähnlich verhalten. Die *Mächtigkeit der Anonymitätsmenge* kann als Anonymitätsmaß dienen.

Sweeney [Swe02] bezieht sich bei der Definition einer k-*Anonymität* auf ein Szenario, in dem es darum geht, Datenbanken mit persönlichen Daten so zu anonymisieren, dass die Datensätze nicht reidentifiziert werden können. k-anonym heißt dabei, dass es mindestens k Einträge in der anonymisierten Datenbank gibt, die sich in den Teilen gleichen, die auch in einer öffentlichen Datenbank verfügbar sind. Das Maß der k-Anonmität gilt nur unter der Annahme, dass jede Identität, die in der öffentlichen Datenbank eingetragen ist, gleichwahrscheinlich zur anonymisierten Datenbank beigetragen hat. In diesem Punkt ähnelt sich das Konzept der Anonymitätsmenge und der k-Anonymität. In [GSW04] beschreiben Gilburd et al. ein verteiltes Verfahren zur Erzeugung k-anonymer Datenbanken.

Shannon führte in [Sha48] *Entropie* als Maß für den statistischen Informationsgehalt von Informationsquellen ein. Dieses im Folgenden als *Shannon-Entropie* bezeichnete Maß wird oft als Anonymitätsmaß verwendet, indem die Anonymitätsmenge als Quellenalphabet angenommen wird. Je größer der Informationsgehalt dieser Quelle ist, desto anonymer ist der einzelne Nutzer.

Unter der Annahme eines beobachtenden Angreifers wurden in der Literatur verschiedene Ansätze zum Messen der Anonymität in Netzen vorgestellt. Allgemein kann zwischen durch formale Sprachen und Logiken definierten Ansätzen, z.B. [SS99] und informationstheoretischen Ansätzen unterschieden werden. Da diese Arbeit auf letzterem aufbaut, soll hier nur auf solche Ansätze näher eingegangen werden. Bei den im Folgenden betrachteten Ansätzen handelt es sich jedoch um Anonymitätsmaße auf Netzebene, es wird also eine Betrachtung der Nachrichteninhalte, wie sie im vorliegenden Papier erfolgt, ausgeklammert. Angriffe werden nur aufgrund von Beobachtungen innerhalb der Netzebene durchgeführt.

In [SD02] stellen Serjantov und Danezis ein informationstheoretisches Maß für Anonymität vor. Sie belegen, dass Anonymitätsmengen nicht zur Quantisierung von Anonymität ausreichen, da die Nutzer mit unterschiedlich hoher Wahrscheinlichkeit zur Anonmitätsmenge beitragen können. Die Autoren wenden den von Shannon eingeführten Entropiebegriff an, um Anonymität zu messen. Letztlich zeigen sie, wie Anonymitätsmengen und Shannon-Entropien zusammenhängen. Der in Definition 12 verwendete Entropiebegriff ist äquivalent zu dem von Serjantov und Danezis benutzten.

Auch Díaz et al. verwenden in [DSCP02] die Shannon-Entropie für ein Messverfahren zur Qualitätsbewertung von Anonymitätsdiensten auf Netzebene. Es wird analysiert, wie viel ein Angreifer aus der Beobachtung eines Anonymitätsdienstes lernen kann. Ziel dieses Ansatzes ist es, Anonymitätsdienste an möglichst starken Angreifern zu messen. Um verschiedene Anonymitätsdienste besser vergleichen zu können, normieren Díaz et al. im Unterschied zu [SD02] das Anonymitätsmaß anhand der Anzahl der beteiligten Nutzer.

Steinbrecher und Köpsell verallgemeinern in [SK03] den Anonymitätsbegriff zur *Verkettbarkeit*. Bezüglich Anonymität wird analysiert, wie Nachrichten mit Sendern beziehungsweise Empfängern zusammenhängen, d.h. verkettbar sind. Dieses Konzept wird verallgemeinert, indem Nachrichten, Sender, Empfänger und Kommunikationsbeziehungen auf eine Stufe gestellt werden, d.h. als Objekte betrachtet werden, über deren Zusammenhänge Informationen existieren. Man kann somit untersuchen, wie stark einzelne Objekte verkett-

bar sind. Zwei Nachrichten sind zum Beispiel verkettbar, wenn man sagen kann, dass sie zur gleichen Kommunikationsbeziehung gehören. Anonymität stellt sich dabei als Spezialfall heraus. Ein Nutzer ist anonym, wenn er nicht mit seinen Aktionen verkettbar ist.

Bisher existieren nur wenige Ansätze für Anonymitätsbewertung unter Einbeziehung der Anwendungsebene, d.h. der Nachrichteninhalte. Einer davon wird von Irwin und Yu [IY04] beschrieben. Sie schlagen eine Funktion vor, die Policies generiert. In diese Funktion gehen Identifizierbarkeit und Sensibilität der Daten ein. Wie stark Daten zur Identifizierbarkeit beitragen, hängt dabei davon ab, wie stark die Anonymitätsmenge eingeschränkt wird. Die Sensibilität der Daten hängt von der Einschätzung des Nutzers ab. Da das Modell nicht auf statistischen Schlußverfahren basiert, ist es aber möglich, dass eine Auswahl von Daten, die der Nutzer als ungefährlich einstuft, ihn deanonymisiert. [1]

Eine Anonymitätsbewertung muss demzufolge abhängig von einer konkreten Kommunikationssituation, also des zum Zeitpunkt der Kommunikation vorhandenen Wissenstandes des Angreifers, vorgenommen werden. Der Nutzer benötigt demnach eine von Einzelattributen abstrahierte Metapher, die ihn darin unterstützt, eine konkrete Vorstellung davon zu bekommen, wie gefährlich einzelne Daten in einer gegebenen Situation für seine Privatsphäre sind.

3 Wissensmodell des Angreifers

In diesem Abschnitt wird zunächst ein abstraktes System zur Modellierung der Daten von Nutzern (von Dienstleistungen im Netz) definiert. Auf dieses aufbauend wird ein statistisches Wissensmodell des Angreifers spezifiziert.

3.1 Modellierung der Nutzerdaten

Nutzer differenzieren sich durch ihren Besitz, ihr Wissen oder ihr Verhalten. Diese Unterscheidungsmerkmale sollen im Folgenden als *Attributwerte* subsummiert werden.

Definition 1 (Attributwert) *Ein* Attributwert *ist eine Eigenschaft eines Nutzers, die zur Unterscheidung von anderen Nutzern dienen kann.*

Attributwerte werden entprechend der folgenden Definition durch *Attribute*[2] strukturiert:

Definition 2 (Attribute) *Ein* Attribut $\mathcal{A}$ *ist eine Menge von Attributwerten* $a \in \mathcal{A}$ *und einem speziellen Attributwert „nicht gesetzt“* $\ominus$*. Ein Nutzer hat nur genau einen Wert aus dieser Menge.*

Zum Beispiel enthält das Attribut „Familienstand“ die Werte „ledig“, „verheiratet“, „geschieden“ und „verwitwet“. Es kann aber auch Attribute geben, die nicht bei jedem Nutzer sinnvoll sind. So hat eine ledige Person (noch) keinen Hochzeitstag.

Mit Hilfe von Attributen werden enprechend [PH05] *digitale Identitäten* wie folgt definiert:

Definition 3 (Digitale Identität) *Eine* digitale Identität *ist ein vollständiges Bündel von Attributen. Das heißt, eine* digitale Identität *besteht aus genau einem Wert jedes Attributs.*

Beispiel: In Abbildung 1 wird der Zusammenhang zwischen Attributen, ihren Werten einer digitalen Identität an einem konkreten Beispiel gezeigt. ◇

[1] wie in [Swe02] gezeigt.

[2] der hier verwendete Begriff „Attribut“ wird oft auch als *Merkmal* bezeichnet, wobei ein Attributwert dann eine *Merkmalsausprägung* darstellt.

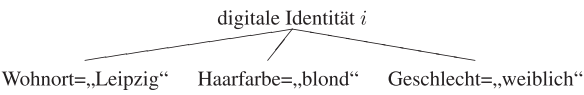

Abbildung 1: Digitale Identität mit zugeordneten Attributen und deren Wert

3.2 Modellierung der Wissensbasis des Angreifers

Durch Beobachtungen erhält der Angreifer einen begrenzten Einblick in die Daten der Nutzer und deren Zusammenhänge. Solche Informationen kann er sammeln und beliebig statistisch auswerten. Mit einer zunehmenden Anzahl von Beobachtungen erhält der Angreifer immer genauere Informationen über die Wahrscheinlichkeitsverteilung der digitalen Identitäten.

Definition 4 (Angreiferzustand) *Der* Zustand $Z^{\mathcal{X}}$ *eines Angreifers* $\mathcal{X}$ *ist ein Tripel* $(\mathcal{I}, h, t)$*, wobei:*

- $\mathcal{I}$ *die Menge aller möglichen digitalen Identitäten ist.*

  $$\mathcal{I} = \mathcal{A}_1 \times \mathcal{A}_2 \times \ldots \times \mathcal{A}_n$$

- $h : \mathcal{I} \mapsto \mathbb{R}$ *eine Funktion, die jeder digitalen Identität eine Wahrscheinlichkeit zuordnet.*
- t *ist die Zahl der Beobachtungen, die zu diesem* Zustand *führten.*
- *Die Summe aller Wahrscheinlichkeiten ist eins.*

  $$\sum_{(\mathcal{I})} h(i) = 1$$

Ein Angreiferzustand repräsentiert die aktuelle Wissensbasis des Angreifers. Durch *Beobachtungen* wird diese verändert.

Definition 5 (Beobachtung) *Eine* Beobachtung *besteht aus einem möglicherweise unvollständigen Bündel von Attributwerten. Dieses Bündel besteht wiederum aus höchstens einem Wert pro Attribut. Die Menge aller möglichen* Beobachtungen *ist das Kreuzprodukt aller Attribute mit einem zusätzlichen Element „nicht beobachtet“* $\perp$*.*

$$\mathcal{B} = (\mathcal{A}_1 \cup \perp) \times (\mathcal{A}_2 \cup \perp) \times \ldots \times (\mathcal{A}_n \cup \perp)$$

Intuitiv drückt Definition 5 aus, dass ein im Netz agierender Nutzer Attributwerte offenbart. Diese werden vom Angreifer beobachtet, wodurch er ein immer differenzierteres Bild der tatsächlich vorhandenen digitalen Identitäten (und damit der Nutzer) konstruieren kann.

Anhand von Beobachtungen kann ein Angreifer zwischen verdächtigen und nicht verdächtigen digitalen Identitäten unterscheiden. Die Menge der Verdächtigen, die zu einer Beobachtung gehören, lässt sich wie folgt definieren:

Definition 6 (Verdächtige) *Die Menge der* Verdächtigen $\mathcal{V}_b$ *bezüglich einer Beobachtung* $b = (x_1, .., x_n)$ *beinhaltet alle Identitäten* $i = (x'_1, .., x'_n)$*, deren Attributwerte entweder mit denen in* b *übereinstimmen oder nicht beobachtet wurden.*[3]

$$\mathcal{V}_b = \{i | \forall k : \mathbb{N}.(x'_k = x_k \vee x'_k = \perp)\} \quad (1)$$

[3] In dieser Arbeit wird davon ausgegangen, dass es nur die Zustände „verdächtig“ und „nicht verdächtig“ gibt. Als Verallgemeinerung ist auch ein kontinuierliches Maß vorstellbar, das heißt, unterschiedliche Grade des Verdachts.

Ein wichtiges Element des Modells ist, dass der Angreifer durch zusätzliche Beobachtungen immer mehr über seine Opfer lernt. Dieser Lernprozess wird in der nächsten Definition formalisiert.

Definition 7 (Aktualisierung des Angreiferzustands) *Sei b eine Beobachtung, $\mathcal{V}_b$ die Menge der Verdächtigen bezüglich b und $\mathcal{Z}$ eine Menge von Zuständen. Eine* Aktualisierungsfunktion *$\delta : \mathcal{Z} \times \mathcal{B} \rightarrow \mathcal{Z}$ erzeugt aus einem Zustand und einer Beobachtung einen neuen Zustand.*

Diese Definitionen stellen einen Rahmen dar, in dem verschiedene konkrete Beobachtungen und statistische Auswertungen basierend auf digitalen Identitäten formalisiert werden können. Um z.B. auch aktive Angriffe einbeziehen zu können wird bewusst offen gelassen, wie es zu einer Beobachtung kommt, etwa ob der Angreifer hierzu lediglich Nachrichten abhört, oder selbst Nachrichten einbringt und Reaktionen der Nutzer im System beobachtet.

Die nachfolgende Definition fasst die obigen Definitionen zu einem statistischen Angreifermodell zusammen.

Definition 8 (Statistisches Angreifermodell) *Ein* statistisches Angreifermodell *eines Angreifers A besteht aus einer Menge $\mathcal{I}$ von möglichen digitalen Identitäten, einer Menge möglicher Beobachtungen $\mathcal{B}$, einer Menge $\mathcal{Z}^A$ von Zuständen sowie einer Funktion δ, die Zustände und Beobachtungen in neue Zustände überführt.*

4 Ein konkretes Angreifermodell

In diesem Abschnitt wird ein Angreifermodell vorgestellt. Das hier vorgestellte Modell wurde mit folgenden Zielen entworfen:

1. Beobachtungen sollen dazu führen, dass sich die Häufigkeitsverhältnisse im Modellzustand den tatsächlich im System vorhandenen annähern.
2. Beobachtete Zusammenhänge zwischen Attributwerten verschiedener Attribute sollen im Modell abgebildet werden.
3. Es soll möglich sein, verteilte Angreifer zu modellieren.
4. Anhand eines Modellzustandes soll für eine zusätzliche Beobachtung b_x eine Entropie berechnet werden können, die aus Angreifersicht angibt, wieviel Information der Angreifer bei Kenntnis der Beobachtung b_x zusätzlich benötigen würde, um den in b_x beobachteten Nutzer (d.h. dessen digitale Identität) eindeutig zu identifizieren.

4.1 Konkrete Wissensbasis des Angreifers

Gegeben seien eine Menge digitaler Identitäten $\mathcal{I}$, eine Menge $\mathcal{B}$ aller möglichen Beobachtungen. Die Menge der zulässigen Zustände $\mathcal{Z}$ wird induktiv definiert. In einem Anfangszustand $Z_0 = (\mathcal{I}, h, t)$ soll gelten:

$$\forall j \in \mathcal{I}. \left(\left(\sum_{(i \in \mathcal{I})} h(i) = 1 \right) \wedge \left(h(j) \geqq 0 \right) \right) \tag{2}$$

$$t = 0 \tag{3}$$

Eine Funktion $\delta : \mathcal{Z} \times \mathcal{B} \rightarrow \mathcal{Z}$ berechnet aus einem alten Zustand $Z_k = (\mathcal{I}, h_k, t_k)$ und einer Beobachtung $b \in \mathcal{B}$ einen neuen Zustand $Z_{k+1} = (\mathcal{I}, h_{k+1}, t_{k+1})$ wie folgt:

$$h_{k+1} \quad : \quad i \mapsto \frac{h_k(i) * t_k + x}{t_k + 1} \tag{4}$$

$$x \quad = \quad \begin{cases} \frac{1}{|\mathcal{V}_b|} & \text{wenn } i \in \mathcal{V}_b \\ 0 & \text{sonst} \end{cases}$$

$$t_{k+1} \quad = \quad t_k + 1 \tag{5}$$

Beispiel: Es wird ein Modell angenommen, das aus drei Attributen mit je zwei Werten besteht. Es können also höchstens acht Identitäten unterschieden werden. Wenn man das Modell, wie in Abbildung 2, als Würfel darstellt, repräsentiert jede Ecke eine digitale Identität. In der Abbildung ist dargestellt, wie der Angreiferzustand sich bei der Beobachtung $b = (0, 0, \perp)$ verändert.

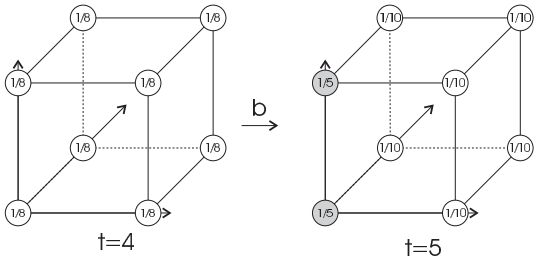

Abbildung 2: Verdächtige und nicht Verdächtige bezüglich einer Beobachtung b

◇

Dieses einfache Modell zählt relative Häufigkeiten. Mit steigender Beobachtungszahl kann angenommen werden, dass sich die relativen Häufigkeiten an die Wahrscheinlichkeiten annähern. Das Modell ist konsistent, wenn die Funktion h immer die Eigenschaften einer Verteilungsfunktion hat. Dies wird im Folgenden gezeigt.

Lemma 9 *Die Summe aller Wahrscheinlichkeiten $h(i)$ eines beliebigen Angreiferzustandes $Z = (\mathcal{I}, h, t)$ ist 1.*

$$\sum_{(i \in \mathcal{I})} h(i) = 1$$

Beweis: durch Induktion über t.

□

Lemma 10 *Die Funktion $h(i)$ ist für kein i negativ.*

Beweis: durch Induktion über t. □

Satz 11 *Das vorgeschlagene Modell ist konsistent.*

Beweis: Zu zeigen ist, dass $\forall i \in \mathcal{I}.(0 \leq h(i) \leq 1)$ für alle konstruierbaren Zustände gilt. Diese Aussage folgt direkt aus den Lemmata 9 und 10. □

4.2 Nutzung der Angreiferwissensbasis zum Messen von Verkettbarkeit

Wie bereits oben beschrieben, wird vielfach die Shannon-Entropie [Sha48] als Anonymitätsmaß verwendet. Ist ein Angreiferzustand Z bekannt, kann die Shannon-Entropie $H_\varnothing$ einer Information b bestimmt werden.

Verkettbarkeit der Informationen. In diesem Zusammenhang definieren wir Verkettbarkeit als Wahrscheinlichkeit, daß Objekte aufgrund von Beobachtungen miteinander in Relation gesetzt werden können. Objekte sind im Kontext dieser Arbeit digitale Identitäten (von Sender oder Empfänger) und Nachrichten. Identifizierbarkeit ist ein Spezialfall der Verkettbarkeit. Ist ein Sender mit Nachrichten verkettbar, ist er bezüglich dieser Nachrichten identifizierbar. Entsprechendes gilt für Nachrichtenempfänger.

Definition 12 (Shannon-Entropie) *Sei b eine Beobachtung und $\mathcal{V}_b$ die Menge der Verdächtigen bezüglich der Beobachtung. Die* Shannon-Entropie *von b bezüglich eines gegebenen Zustandes ist die* Shannon-Entropie *der Verdächtigen $\mathcal{V}_b$.*

$$
\begin{aligned}
H_\varnothing &= -\sum_{(v\in\mathcal{V})} p(v|b)\log_2 p(v|b) \\
p(v|b) &= \frac{p(v \wedge (\bigvee_{(w\in\mathcal{V})} w))}{p(\bigvee_{(w\in\mathcal{V})} w)} \\
&= \frac{p(v)}{\sum_{(i\in\mathcal{V})} p(i)}
\end{aligned}
$$

Die Shannon-Entropie $H_\varnothing$ bekommt die Pseudoeinheit Bit.

Aus der Shannon-Entropie kann die dazu äquivalente Mächtigkeit einer gleichverteilten Anonymitätsmenge $\mathcal{A}$ berechnet werden:

$$|\mathcal{A}| = 2^{H_\varnothing}$$

Beispiel: Die Shannon-Entropie der Beobachtung $b = (0, 0, \perp)$ in Abbildung 2 ist gleich ein Bit. Daraus folgt, dass die Verdächtigen so anonym sind wie in einer zweielementigen Anonymitätsmenge, die gleichverteilt ist. ◇

Die Shannon-Entropie $H_\varnothing$ gibt demzufolge an, wieviel Information zusätzlich zu b im Mittel benötigt wird, um eine digitale Identität eindeutig zu identifizieren. In Abschnitt 5 wird erläutert, wie dieses Maß in einem Identitätsmanagementsystem zum Schutz der Privatsphäre des Nutzers Anwendung finden kann.

Die Shannon-Entropie stellt allerdings einen Mittelwert der Anonymitätsmenge dar[4]. Der Einzelne, der seine Privatsphäre schützen will, ist aber mehr daran interessiert, wieviel Information nötig ist, um ihn selbst zu identifizieren. Folgendes Beispiel soll die Diskrepanz zwischen diesen beiden Zielen verdeutlichen:

Beispiel: Einem Beobachter ist bekannt, dass ein Quellzeichen A mit einer Wahrscheinlichkeit von 50% auftritt. Wenn der Beobachter genau daran interessiert ist, zu erkennen, ob das Zeichen A auftritt (d.h. wie anonym A ist), ist dies unabhängig von der Entropie der Quelle, sondern hängt nur von der Wahrscheinlichkeit des Auftretens von A ab. Die Entropie hängt demgegenüber auch von der Anzahl der Quellzeichen ab, d.h. auch wenn A eine Wahrscheinlichkeit von 50% hat, kann die Entropie bei entsprechender Anzahl und

[4] Das heißt, sie gibt an, wieviel Information durchschnittlich benötigt wird um irgend ein Element aus der Anonymitätsmenge zu identifizieren.

Verteilung der anderen Quellzeichen beliebig hoch sein. Die Informationsmenge zur Identifikation von A bleibt demgegenüber — unabhängig von der Entropie der Quelle — gleich. ◇

Dieses Problem kann umgangen werden, indem die *lokale Anonymität* als Anonymitätsmaß verwendet wird.

Definition 13 (lokale Anonymität) *Die Informationsmenge H_i einer Identität i, die noch nötig ist, um i zu* identifizieren*, falls die Beobachtung b von i veröffentlicht wurde, ergibt sich aus der bedingten Wahrscheinlichkeit $p(i|b)$.*

$$H(i) = \log_2 p(i|b) \tag{6}$$

Aus Sicht des einzelnen Nutzers ist die lokale Anonymität das bestmögliche Anonymitätsmaß.

4.3 Zusammenführung der Wissenstände mehrerer Angreifer

Dienstleister im Internet können Informationen austauschen bzw. zusammenführen, um genauere Nutzerprofile zu erstellen. Intuitiv einsichtig ist hierfür das Vorgehen, dass mehrere Angreifer ihre individuellen Einzelbeobachtungen an einer zentralen Stelle zusammenführen und diese zentrale Stelle daraus eine gemeinsame Wissensbasis erstellt.

Mit dem vorgestellten konkreten Angreifermodell ist es darüber hinaus möglich, mehrere Angreiferzustände zusammenzuführen, ohne die entsprechenden Einzelbeobachtungen berücksichtigen zu müssen. In diesem Abschnitt wird ein solches Vorgehen beschrieben und dessen Korrektheit bewiesen.

Die Zusammenführung geschieht durch Bildung eines neuen Zustandes, der an jeder Stelle $h(i)$ dem nach der Beobachtungsanzahl gewichteten Mittel der Zustände der Informanten entspricht. Dabei ist der Angreifer selbst auch ein Informant.

Definition 14 (Zustandsvereinigung) *Zwei auf der gleichen Identitätsmenge basierende Zustände $Z^A = (\mathcal{I}, h^A, t^A)$ und $Z^B = (\mathcal{I}, h^B, t^B)$ können zu einem neuen Zustand $Z^A \cup Z^B = (\mathcal{I}, h^C, t^C)$ wie folgt* vereinigt *werden:*

$$t^C = t^A + t^B \tag{7}$$

$$h^C : i \mapsto \frac{t^A h^A(i) + t^B h^B(i)}{t^C} \tag{8}$$

Beispiel: In Abbildung 3 ist illustriert, wie das Vereinigen von Zuständen funktioniert. Die Idee, Zustände zu vereinigen, ist Informationsaustausch zu modellieren. Sinnvoll ist dieser nur, wenn Austauschen genau so gut wie selbst beobachten ist[5] und in der Tat kann das bewiesen werden.

◇

Zur Vorbereitung werden noch eine weitere Definition und ein Hilfssatz benötigt:

Definition 15 (Zustandsgleichheit) *Zwei Zustände $Z^A = (\mathcal{I}^A, h^A, t^A)$ und $Z^B = (\mathcal{I}^B, h^B, t^B)$ sind* gleich *genau dann wenn $\mathcal{I}^A = \mathcal{I}^B$, $h^A = h^B$ und $t^A = t^B$.*

[5] Das stimmt im Allgemeinen in natürlichen Umgebungen nicht. Hier ist es aber so, dass Angreifer deterministisch und gleichartig sind, das heisst, wenn sie gleiches beobachten, dann folgern sie auch gleiches.

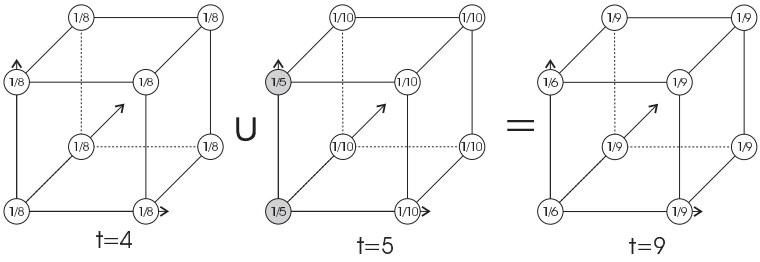

Abbildung 3: Vereinigung zweier Zustände

Lemma 16 (Beobachtungsreihenfolge) *Seien Z ein beliebiger, aber fester Zustand und b_1, b_2 Beobachtungen. Die Reihenfolge des Beobachtens hat keinen Einfluss auf den resultierenden Zustand.*

Beweis: Es sind drei Teilziele zu zeigen:

1. Die Zustände basieren auf der gleichen Identitätsmenge. Dies folgt direkt aus der Definition der Aktualisierungsfunktion.
2. Die Zustände enthalten die gleiche Anzahl von Beobachtungen.
3. Letztlich ist zu zeigen, dass die Funktion h sich im Ergebnis nicht ändert. Laut Definition des konkreten Modells gilt folgendes:

$$\frac{\frac{h(i)t+x_1}{t+1}(t+1)+x_2}{t+2} = \frac{\frac{h(i)t+x_2}{t+1}(t+1)+x_1}{t+2}$$

Wobei x_1 ausschließlich von b_1 abhängt und x_2 von b_2.

□

Satz 17 *Seien A und B Listen von Beobachtungen aus der gleichen Menge von Beobachtungen und Z^A,Z^B Zustände, die aus diesen Beobachtungen konstruiert sind. Sei nun Z^C der Zustand, der entsteht, wenn erst A und dann B beobachtet wird, dann gilt:*

$$Z^C = Z^A \cup Z^B$$

Beweis: Der Beweis wird per Induktion über t mit Hilfe des Beobachtungsreihenfolge-Lemmas (16) geführt. □

In Satz 17 wird keine Aussage über die konkreten Beobachtungen der beiden Angreifer gemacht. Im Extremfall könnten die Angreifer exakt dasselbe beobachtet haben. In der Tat ist Informationsaustausch zwischen Angreifern, die möglicherweise dieselben Nachrichten beobachtet haben, nicht sinnvoll. Es ist für einen gewinnbringenden Informationsaustausch wichtig zu wissen, auf welche Art und Weise der Informant seine Informationen gesammelt hat.

Diese Schwäche entsteht nicht dadurch, dass man nicht die gesamten Beobachtungen verarbeitet. Zwei Beobachtungen aus verschiedenen Quellen können gleich sein. In diesem Fall ist der Informationsaustausch sinnvoll. Wichtig ist nur, dass die Beobachtungen nicht derselben Nachricht entspringen, die zum Beispiel nur an verschiedenen Stellen des Datennetzes beobachtet wurde.

5 Nutzen des Modells für Identitätsverwaltungswerkzeuge

5.1 Bewertung der Anonymität des Nutzers

Werkzeuge zur Verwaltung digitaler Identitäten, wie sie zum Beispiel im PRIME Projekt [PRI] entwickelt werden, sollen dazu beitragen, die Privatsphäre eines Nutzers bei der Kommunikation im Netz zu schützen. Sie helfen vor allem dabei, nur die Information an eine bestimmte Instanz weiterzugeben, die der Nutzer auch weiter geben will.

Der mündige Nutzer muss dafür einschätzen, wie gefährlich die Herausgabe einer bestimmten Information ist, d.h. wie stark ein herauszugebendes Bündel von Attributwerten die Anonymität des Nutzers kompromittiert. Zu dieser Einschätzung kann die in Abschnitt 4.2 beschriebene Berechnung der Shannon-Entropie bzw. der minimalen Anonymität als Maß für Verkettbarkeit beitragen.

Unter der Annahme, dass der Nutzer genügend Informationen über den Wissenstand des Angreifers hat, kann der Nutzer ein zur Herausgabe vorgesehenes Bündel von Attributwerten als zusätzliche Beobachtung b_x modellieren. Für diese Beobachtung b_x kann er die in Abschnitt 4.2 angegebenen Anonymitätsmaße berechnen. Falls für eine Dienstleistung Alternativen bei der Herausgabe von Attributen bestehen, können diese mit Hilfe der Anonymitätsmessung verglichen und die beste ausgewählt werden.

Die Schwierigkeit für den Nutzer liegt darin, das Wissen des Angreifers abzuschätzen. Dies kann in zwei Teilprobleme untergliedert werden:

1. **Informationsbeschaffung:** Prinzipiell können Statistiken Dritter mittels Zustandsvereinigung eingebracht werden. So kann auf öffentliche Statistiken, wie die der Statistischen Bundes- und Landesämter oder kommerzieller Anbieter, zurückgegriffen werden. Zukünftig denkbar wären auch speziell für diese Zwecke zugeschnittene Services von Datenschutzbehörden. Problematisch hingegen ist die Informationsbeschaffung bei den Angreifern, also beispielsweise bei Dienstleistern im Internet. Hierzu wird eine Kooperationsbereitschaft des Angreifers benötigt, die etwa im Rahmen von vertrauensbildenden Maßnahmen bezüglich Datenschutz durchaus denkbar ist, aber nicht der Regelfall sein wird. Die Abschätzung wird im Allgemeinen immer unvollständig und zumindest geringfügig fehlerbehaftet sein, so daß eine Anonymitätsbewertung mit derartigen Methoden Hinweise für den Nutzer geben kann, aber nicht als hartes Entscheidungskriterium benutzt werden darf.

2. **Rechen- und Speicheraufwand:** Die Verwaltung einer Wissensbasis mit Hilfe des vorgestellten konkreten Modells erfordert die Speicherung eines relativen Häufigkeitswertes für jede digitale Identität. Weiterhin muß beim Hinzufügen von Beobachtungen mit jedem dieser Häufigkeitswerte eine Berechnung durchgeführt werden. Bei einer für ein Identitätsmanagementsystem realistischen Anzahl von Attributen und Attributwerten ist dies nicht bzw. nicht in einem vertretbaren Zeitrahmen möglich. Es bleibt zu untersuchen, wie entsprechende Effizienzverbesserungen mit minimalem Genauigkeitsverlust durchgeführt werden können.

5.2 Bewertung von Fehlinformationen

Fehlinformationen können einen gewissen Schutz der Privatsphäre bieten. Zusammen mit Kostenfunktionen für Lügen können wirtschaftliche Aspekte von Datenschutz untersucht werden. In diesem Abschnitt soll der Begriff *Lüge* formal definiert werden. Es soll außerdem unter Verwendung des oben beschriebenen konkreten Angreifermodells gezeigt werden, dass unter bestimmten Umständen Lügen tatsächlich ein Mittel ist, um die Privatsphäre des Nutzers zu schützen.

Definition 18 (Lügner, Lüge) *Ein Nutzer, der gefälschte, also nicht zu seiner eigentlichen Identität gehörende Attribute veröffentlicht ist ein* Lügner. *Eine Beobachtung, die ein gefälschtes Attribut enthält, wird* Lüge *genannt.*

In Bezug auf digitale Identitäten gilt für Lügen der folgende Satz:

Satz 19 *Die digitale Identität eines Lügners $i \in \mathcal{I}$ ist nicht in der Menge der Verdächtigen $\mathcal{V}_l$ bezüglich seiner Lüge l.*

Beweis: Sei $i \in \mathcal{I}$ die digitale Identität eines Lügners und $l \in \mathcal{B}$ eine Lüge, dann beinhaltet l einen Attributwert, welcher nicht Attributwert von i ist. Das heißt, dass es ein Attribut gibt, das in i und l verschieden belegt ist, also gilt die Bedingung 1 $\forall n : \mathbb{N}.(x_n = y_n \vee y_n = \perp)$ für i nicht. Somit ist der Lügner i nicht in der Menge der Verdächtigen bezüglich seiner Lüge l. □

Aus dem Satz 19 über Lügen folgt, das ständiges Lügen der beste Weg ist, um die Privatsphäre des Nutzers zu schützen. Für sich allein gesehen ist diese Feststellung zwar korrekt, aber das Ziel eines Nutzers bei der Nutzung von Dienstleistungen im Netz ist es außerdem (und üblicherweise vorrangig), eine bestimmte Dienstleistung in einer entsprechenden Qualität erbracht zu bekommen. Herausgeben falscher Daten führt aber oft dazu, daß eine Dienstleistung nicht erbracht werden kann, oder zumindest kostenintensiver wird, wie das folgende Beispiel zeigt:

Beispiel: Ein Kunde eines Internetbuchhandels möchte seine Interessensgebiete gegenüber seinem Buchhandel verbergen. Um dieses Ziel zu erreichen bestellt er auch Bücher, die er nicht lesen will, dadurch entstehen ihm zusätzliche Kosten. ◇

Es ist deshalb notwendig, auch durch Lügen entstehende Kosten zu modellieren. Dies kann geschehen, indem jedem Nutzer eine individuelle Kostenfunktion c zugeordnet wird, die Nachrichten auf Kosten abbildet. Intuitiv soll diese Funktion ausdrücken welche Kosten entstehen, falls ein Nutzer eine beliebige Nachricht emittiert. Die Kosten für eine Nachricht, die keine Lüge ist, sollten dabei null sein.

Kosten für den Schutz der Privatsphäre lassen sich mit Hilfe der Kostenfunktion als Optimierungsaufgabe mit Nebenbedingung formulieren:

$$\begin{aligned} c(b) &\rightarrow min \\ H(b) &> k \end{aligned}$$

Dabei ist $H(b)$ ein Maß für die Anonymität des Nutzers, die noch vorhanden ist, wenn er b preisgibt. In den Nebenbedienungen für obiges Optimierungsproblem kann der Nutzer angeben, welche Attribute er unbedingt preisgeben muss, damit der Dienst für ihn überhaupt brauchbar bleibt.

6 Zusammenfassung und Ausblick

Zur abstrakten Modellierung von Nutzern von Dienstleistungen im Netz wurde zunächst ein System definiert, in dem Nutzern digitale Identitäten zugeordnet sind. Mit Hilfe von Beobachtungen erhält ein Angreifer (Beispielsweise ein Dienstleister im Netz) Teilinformationen über die im System vorhandenen Nutzer (d.h. deren digitale Identitäten). Es wurde ein statistisches Modell für das Wissen des Angreifers vorgestellt, das es ermöglicht, Angreiferwissen aufzubauen und zu verwalten.

Ein konkretes Angreifermodell wurde beschrieben und dessen Konsistenz bewiesen. Weiterhin wurde gezeigt, dass das konkrete Modell als Grundlage für die Messung der Anonymität von Nutzern eingesetzt werden kann. Die Shannon-Entropie als Anonmitätsmaß

wurde kritisiert und mit der *minimalen Anonymität* ein neues Maß vorgeschlagen. Es wurde weiterhin gezeigt, daß das konkrete Modell eine effiziente Zusammenführung von Wissenständen mehrerer Angreifer im System ermöglicht.

Es wurde die Anwendbarkeit des Angreifermodells bzw. dessen Konkretisierung und der darauf basierenden Anonymitätsmaße im Kontext von Identitätsmanagementsystemen zum Schutz der Privatsphäre des Nutzers diskutiert. In diesem Zusammenhang wurde auch auf die Möglichkeit der Herausgabe von Falschinformationen zur Wahrung der Anonymität und Implikationen dieses Vorgehens eingegangen.

Abschließend kann festgestellt werden, dass das vorgeschlagene Angreifermodell und dessen Konkretisierung einen Mehrwert für nutzerorientierte Identitätsmanagementsysteme bieten können, wenn dem Nutzer ausreichende Informationen über das Angreiferwissen zur Verfügung stehen.

Um dem Ziel eines Einsatzes in Identitätsmanagementsystemen näherzukommen soll in weiterführenden Arbeiten u.a. untersucht werden, wie möglichst effiziente Abschätzungen des Angreiferwissens bei großen Mengen von Attributen und Attributwerten möglich sind. In dem vorgestellten Modell sind Attributwerte für jede Entität statisch. Hierzu muss untersucht werden, wie zeitlich veränderliche Attribute modelliert werden können. Weiterhin kann untersucht werden ob und unter welchen Bedienungen ein Angreiferzustand konvergiert, insbesondere welchen Einfluss Lügen auf die Konvergenz von Zuständen hat.

Literatur

[DSCP02] Claudia Díaz, Stefaan Seys, Joris Claessens und Bart Preneel. Towards measuring anonymity. In Roger Dingledine und Paul Syverson, Hrsg., *Proceedings of Privacy Enhancing Technologies Workshop (PET 2002)*. Springer-Verlag, LNCS 2482, April 2002.

[GSW04] Bobi Gilburd, Assaf Schuster und Ran Wolff. k-TTP: A New Privacy Model for Large-Scale Distributed Environments. In *ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, Seiten 563–568, Seattle, 2004.

[IY04] Keith Irwin und Ting Yu. An identifiability-based access control model for privacy protection in open systems. In *WPES '04: Proceedings of the 2004 ACM workshop on Privacy in the electronic society*, Seiten 43–43. ACM Press, 2004.

[PH05] Andreas Pfitzmann und Marit Hansen. Anonymity, Unlinkability, Unobservability, Pseudonymity, and Identity Management - A Consolidated Proposal for Terminology. Version 0.23 at http://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.23.pdf, 2005. Version 0.8: Anonymity, Unobservability, and Pseudonymity - A Proposal for Terminology,in: Hannes Federrath (Hg.): Designing Privacy Enhancing Technologies; Proc. Workshop on Design Issues in Anonymity and Unobservability; LNCS 2009; 2001; 1-9;.

[PRI] Privacy and Identity Management for Europe. http://www.prime-project.eu.org/.

[SD02] Andrei Serjantov und George Danezis. Towards an Information Theoretic Metric for Anonymity. In Roger Dingledine und Paul Syverson, Hrsg., *Proceedings of Privacy Enhancing Technologies Workshop (PET 2002)*. Springer-Verlag, LNCS 2482, April 2002.

[Sha48] C.E. Shannon. A Mathematical Theory of Communication. *The Bell System Technical Journal*, 1948.

[SK03] Sandra Steinbrecher und Stefan Köpsell. Modelling Unlinkability. In Roger Dingledine, Hrsg., *Proceedings of Privacy Enhancing Technologies workshop (PET 2003)*. Springer-Verlag, LNCS 2760, March 2003.

[SS99] Paul F. Syverson und Stuart G. Stubblebine. Group Principals and the Formalization of Anonymity. In *Proceedings of the World Congress on Formal Methods (1)*, Seiten 814–833, 1999.

[Swe02] Latanya Sweeney. *k*-Anonymity: A Model For Protecting Privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2002.

[WWJ04] Lingyu Wang, Duminda Wijesekera und Sushil Jajodia. A logic-based framework for attribute based access control. In *FMSE '04: Proceedings of the 2004 ACM workshop on Formal methods in security engineering*, Seiten 45–55. ACM Press, 2004.