Pseudonymes Einkaufen physischer Güter

Matthias Enzmann · Claudia Eckert

Fraunhofer Institut für Sichere Telekooperation (SIT)
{enzmann, eckert}@sit.fraunhofer.de

Zusammenfassung

Einkaufen im Internet stellt für viele potenzielle Kunden immer noch ein Wagnis dar. Oft geht es dabei nicht nur um die Sorge, was passiert, wenn Waren bezahlt aber nicht geliefert wurden, sondern auch darum, was mit den eigenen persönlichen Angaben passiert, die der Händler durch einen Internet-Einkauf erhält. Der in dieser Arbeit vorgestellte Lösungsansatz ermöglicht es, dass Waren im Internet eingekauft werden können, ohne dass einem Händler hierzu personenbezogene Angaben übermittelt werden müssen. Der gewählte Pseudonymansatz macht sich ausschließlich existierende Infrastrukturen zu Nutze, d.h. es müssen keine neuen Parteien dem Geschäftsprozess hinzugefügt werden, sondern lediglich die Verarbeitungsprozesse der beteiligten Parteien angepasst werden.

1 Einleitung

Datenschutz im Internet ist eine der großen Herausforderungen für den elektronischen Handel. Von einigen wird er lediglich als lästiges Hemmnis oder Wettbewerbsnachteil für den freien Handel im Internet gesehen, für andere stellt er eine der wesentlichen Akzeptanzvoraussetzungen für den Online-Einkauf dar. Tatsache ist, dass die einst in den kommerziellen Erfolg des Internets gesetzten Erwartungen nur in den seltensten Fällen erfüllt werden konnten. Dies liegt sicher nicht zuletzt an den Datenschutzbedenken von Nutzern, die befürchten, durch den Einkauf im Internet zu gläsernen Kunden zu werden, deren persönlicher Hintergrund, Vorlieben, Kaufgewohnheiten, etc. jedem Anbieter im Internet bekannt werden [7].

Auch wenn solche Ängste etwas übertrieben erscheinen, so spiegelt sich in ihnen die Verunsicherung von Nutzern hinsichtlich dessen wider, was mit ihren beim Online-Einkauf anfallenden Daten möglich ist. Beispiele aus der Vergangenheit zeigen auch, dass Anbieter bzw. Händler vielfach zu sorglos mit den personenbezogenen Daten ihrer Kunden umgegangen sind. Es bedurfte bspw. etlicher Missbrauchsfälle im Zusammenhang mit Kreditkartenzahlungen, bis die Übertragung von persönlichen sowie von Finanzdaten durch den Einsatz von Verschlüsselungsmechanismen abgesichert wurde. Damit ist dem Datenschutz aber noch nicht Genüge getan, denn hiermit wird lediglich die Vertraulichkeit der Datenübermittlung zwischen Kunde und Händler gegenüber Dritten sichergestellt. Für datenschutzbewusste Kunden bleibt immer noch der Händler selbst als potenzielles

P. Horster (Hrsg.): Sichere Geschäftsprozesse, 55 ff., 2002
Arbeitskonferenz Elektronische Geschäftsprozesse mit Schwerpunkt IT-Sicherheit

Datenschutzrisiko bestehen.

In diesem Beitrag stellen wir einen praktikablen Pseudonym-Ansatz vor, der sich ausschließlich vorhandener Infrastrukturen bedient, d.h. dem „normalen“ Ablauf eines Einkaufsprozesses keine neuen Parteien wie etwa Zwischenhändler hinzufügt. Der vorgestellte Ansatz erlaubt es dem Nutzer, während des gesamten Einkaufsprozesses, einschließlich der Auslieferungsphase, seinen Namen mit Hilfe eines Pseudonyms vor dem Händler zu verbergen. Der vorgestellte Lösungsansatz soll einen Ausgleich schaffen zwischen den Interessen der Kunden, ihre Privatsphäre vor Händlern zu schützen, und den Interessen von Händlern, Informationen über ihre Kunden für die Bewertung des eigenen Angebots zu erlangen. Der Lösungsansatz realisiert damit die zunehmend gestellte Forderung nach mehrseitiger Sicherheit.

Der Rest dieses Beitrags gliedert sich wie folgt. Im nächsten Abschnitt gehen wir auf gängige Methoden der Sammlung von Nutzerdaten sowie den damit einhergehenden Datenschutzproblemen ein, die sich bei typischen Einkäufen im Internet ergeben können. In Abschnitt 3 untersuchen wir, welche Datenströme bei heute üblichen Interneteinkäufen auftreten, bevor wir in Abschnitt 4 unseren Lösungsansatz vorstellen. Eine mögliche Architektur dieser Lösung wird in Abschnitt 5 skizziert. In Abschnitt 6 werden wir unsere Lösung sowie die zu Grunde gelegten Annahmen diskutieren. Schließlich gehen wir in Abschnitt 7 noch kurz auf verwandte Arbeiten ein und schließen mit einer kurzen Zusammenfassung.

2 Datenschutzproblematik und Händlerinteressen

Anders als beim Einkauf im realen Kaufhaus, muss der Kunde im Internet stets Name und Adresse offenlegen, um seine bestellten Waren geliefert zu bekommen. Der Einkauf im Internet wird deshalb oft mit der Bestellung bei klassischen, d.h. Papier-basierten, Katalogversendern verglichen. Dieser Vergleich ist jedoch aus Datenschutzsicht i.Allg. nicht zutreffend, da einem Internet-Anbieter ungleich mehr Möglichkeiten zur Verfügung stehen, seine Kunden bei ihrem Einkauf zu beobachten und die so erhaltenen Daten, mit Informationen aus anderen Quellen zu kombinieren und zu bewerten.

2.1 Nutzerverfolgung

Ein Internet-Anbieter erfährt im Gegensatz zu einem Katalogversender bspw. welcher Kunde sich welche Artikel wie lange angesehen hat, unabhängig davon, ob er diese später kauft. Mit Hilfe dieser Daten kann er somit auch feststellen, ob ein Kunde bestimmte Artikel einige Tage, Wochen oder Monate zuvor schon angesehen hat und damit Rückschlüsse auf das Interesse des Kunden für bestimmte Waren ziehen. Prinzipiell ist dies auch in einem kameraüberwachten Kaufhaus möglich, jedoch wäre dies mit viel größerem technischen Aufwand verbunden. Kameras müssten hierzu einzelnen Personen „folgen“, was in der Praxis noch schwierig ist, da die Kameras Personen wiedererkennen müssten, wenn sich diese bspw. um eine Ecke bewegen oder das nächste Mal einkaufen. Zudem wäre dies auch nur für wenige Personen möglich und nicht wie im Internet für alle Kunden.

Die Verfolgung von Nutzern ist durch verschiedenste Techniken möglich, meist unter Ausnutzung bestimmter Datenelemente der zu Grunde liegenden Transportprotokolle,

im *World Wide Web* i.d.R. des Internet-Protokolls (IP) und des Hypertext-Transfer-Protokolls (HTTP). Normalerweise können aufeinanderfolgende Zugriffe des selben Nutzers innerhalb einer Sitzung von einem Anbieter anhand der IP-Adresse des Nutzerrechners erkannt werden. Die Verfolgung von Nutzern kann auch durch Elemente von Protokollen „höherer“ Schichten wie des HTTP-Protokolls realisiert werden. Beispiele für solche Elemente im HTTP-Protokoll sind *Cookies*, das *Referer*-Datenelement oder durch die dynamische Zuteilung einer nutzerspezifischen URL (*session binding*) beim Aufruf der Anbieterseiten — Details hierzu finden sich bspw. in [8].

2.2 Profilbildung

Bewegungsdaten von Nutzern, im Folgenden auch *Clickstream*-Daten genannt, können zu einem Profil zusammengefasst und u.U. mit Daten aus anderen Quellen kombiniert und schließlich ausgewertet werden. Profildaten des Kunden können jedoch für den Anbieter wichtig sein. Sie können der Bewertung der Attraktivität des Warenangebots, der Prüfung der Erreichung beabsichtigter Zielgruppen und Ähnlichem dienen, um als Ergebnis daraus, ggf. Veränderungen des Portfolios vorzunehmen. Profildaten werden zum Teil heimlich gesammelt, aber mitunter auch durch elektronische Fragebögen direkt beim Kunden abgefragt. Über den Verwendungszweck der gesammelten Daten wird der Kunde jedoch häufig nicht oder nur sehr vage informiert, ebensowenig darüber, an welche Parteien außer dem Anbieter die Informationen noch übermittelt werden. Im Fall von Clickstream-Daten erfährt der Kunde noch nicht einmal, welche Daten über ihn erfasst werden. Dies hat zum einen zur Folge, dass viele Nutzer Internet-Anbietern generell misstrauen, da sie nicht wissen ob, und wenn ja, welche Daten für welchen Zweck gesammelt und ausgewertet werden. Zum anderen führt das dazu, dass Nutzer bei der Abfrage von Daten, die für den Online-Einkauf nicht relevant sind, konsequent lügen, um ihre Privatsphäre zu schützen. Diese Situation ist sowohl für Nutzer als auch für Anbieter unbefriedigend. Kunden erkennen keinen unmittelbaren Nutzen in der Preisgabe ihrer persönlichen Daten und sehen nur die Datensammelwut des Anbieters. Dies führt dazu, dass Nutzer noch weniger Vertrauen in die Anbieterseite haben. Anbieter müssen aber bei ihren Kunden um Vertrauen werben, um von ihnen nicht nur sinnlose Daten wie „Donald Duck, wohnhaft Schlossallee 1, 12345 Entenhausen, verheiratet, 4 Kinder, Einkommen >1.000.0000 EUR“ zu bekommen. Den Beteuerungen von Anbietern, Profilinformationen nur zu statistischen Zwecken auszuwerten, d.h. diese nicht mit dem Namen des Nutzers zusammenzuführen, dürfte i.d.R. wenig Glauben geschenkt werden, da dies für einen Kunden nur schwer überprüfbar ist.

2.3 Preisdiskriminierung

Unter dem Begriff der *Preisdiskriminierung* werden Aktivitäten des Anbieters verstanden, die dazu dienen, Produkte abhängig vom Käufer bzw. dessen Profil zu unterschiedlichen Preisen anzubieten. Dies ist jedoch nicht zu verwechseln mit Geschäftsmodellen, in denen Rabatte für „gute“ Kunden gewährt werden. In letzterem Fall sind allen Beteiligten die „Spielregeln des Marktes“ bekannt, d.h. jeder weiß, dass er ab einer gewissen Größenordung von getätigten Transaktionen verbesserte Konditionen bekommen kann. Mit Preisdiskriminierung ist vielmehr gemeint, dass Kunden bspw. abhängig von ihrem sozialen oder beruflichen Hintergrund unterschiedliche Preise für das gleiche Produkt zahlen. Bspw. könnte ein Anbieter einem laut Profil Computer-unerfahrenen Nutzer ein Aus-

laufmodell eines Personal Computers teurer verkaufen als einem „Profi-Nutzer“, der den PC zum tatsächlichen Marktwert angeboten bekäme. Somit kann Datenschutz bzw. die Kontrolle über das eigene Profil auch ein Ausdruck von Verbraucherschutz sein [12].

2.4 Datenschutz durch Pseudonyme

Wenn der Kunde dem Händler nicht namentlich bekannt wird, sondern lediglich über ein bestimmtes Merkmal vom Händler erkannt bzw. wiedererkannt wird, ergibt sich, nicht nur aus Sicht des Datenschutzes, eine andere Ausgangslage. Ein Pseudonym stellt ein solches Wiedererkennungsmerkmal dar. Durch die Verwendung eines Pseudonyms kann der Anbieter Profilinformationen dem Pseudonym, nicht aber dem Träger des Pseudonyms, dem Nutzer, zuordnen. Aus datenschutzrechtlicher Sicht ist dies ein Vorteil, da es sich bei pseudonymen Profilen nicht mehr um personenbezogene Daten handelt. Das Problem der Preisdiskriminierung kann durch Pseudonyme jedoch nicht gänzlich gelöst werden, da es zur Preisdiskriminierung nur eines Wiedererkennungsmerkmals mit dazugehörigen Daten bedarf. Durch einen regelmäßigen Wechsel von Pseudonymen kann dieses Problem aber abgemildert werden, da ein Händler die Grundlagen für die Preisdiskriminierung, d.h. seine Datenbasis bzgl. eines neuen Pseudonyms, auch wieder neu aufbauen muss. Weiterhin können Nutzer bei jedem Anbieter andere Pseudonyme verwenden, sodass das Kombinieren von Nutzungsprofilen verschiedener Händler deutlich erschwert wird. Dies setzt voraus, dass der Nutzer kein ihn eindeutig identifizierendes Datum, wie bspw. die gleiche E-Mail-Adresse, für verschiedene Pseudonyme nutzt. Da die Handhabung verschiedener (pseudonymer) Identitäten mitunter schwierig sein kann, können technische Hilfsmittel wie ein Identitätsmanagement [2] sinnvoll sein, um den Einsatz und die Verwaltung mehrerer Pseudonyme zu erleichtern.

Von einem pseudonymen Dienstangebot können auch Händler profitieren. Denn es ist zu erwarten, dass ein solches Angebot das Vertrauen der Kunden in den Händler stärkt, wenn dieser sich der Datenschutzproblematik aktiv annimmt, d.h. sich nicht auf das Veröffentlichen von Datenschutzerklärungen beschränkt. Ein Online-Anbieter kann ein pseudonymes Dienstangebot als ein Alleinstellungsmerkmal bspw. zur Neukundenwerbung vermarkten oder zur Bindung des vorhandenen Kundenstamms nutzen. Weiterhin entsteht bei der Sammlung und Verarbeitung pseudonymer Daten kein Datenschutzproblem beim Anbieter, da Datenschutzregelungen nur für personenbezogene Daten gelten. Sollte ein Händler seinen Kunden die Möglichkeit geben, demographische Daten wie Altersgruppe, Interessengebiete oder Geschlecht zur kundenspezifischen Web-Seitenanpassung (*Customization*) anzugeben, ist weiterhin zu erwarten, dass Nutzer mit ihren Angaben ehrlicher sind, da diese Daten nicht mehr mit ihrem wahren Namen, sondern lediglich mit einem Pseudonym in Verbindung gebracht werden.

3 Interneteinkauf

Zur Analyse der Datenströme, die normalerweise beim Einkauf im Internet anfallen, teilen wir den Einkaufsvorgang in vier Phasen auf, und untersuchen, welche Informationen von welcher der am Einkauf beteiligten Parteien erlangt werden. Die für diesen Beitrag zu Grunde gelegten Phasen sind: (1) Suchphase, (2) Bestellphase, (3) Bezahlphase, (4) Auslieferungsphase.

Eine Sequenz der Phasen bezeichnen wir als Sitzung. Werden nicht alle Phasen durchlaufen, sprechen wir auch von einer abgebrochenen Sitzung. Jede Phase wird in einer Sitzung genau einmal durchlaufen. Die Menge $\mathcal{A}$ bezeichne alle Akteure, die am Einkaufsprozess teilnehmen. Mit p_0 bezeichnen wir den Startzustand, der den Ausgangspunkt eines jeden Einkaufsvorgangs darstellt. Für den Beginn aller nachfolgenden Phasen p_i, $i > 0$, ist der Abschluss der vorhergehenden Phase p_{i-1} erforderlich. Der Übergang von p_i zu p_{i+1} lässt sich durch eine Transition $t_i(S, E, \delta_{i+1})$ beschreiben wobei $S \in \mathcal{A}$ einen Sender und $E \in \mathcal{A}$ einen Empfänger von Informationen δ_{i+1} bezeichne. Die Identität des Senders ist jedoch nicht *per se* dem Empfänger bekannt. Die Transition dient der Beschreibung eines, nicht immer technischen, Informationsflusses von einem Sender zu einem Empfänger. Innerhalb einer Phase können bspw. durch Anklicken eines Hyperlinks mehrfach Informationen zwischen einem Sender und einem Empfänger ausgetauscht werden. Dies verstehen wir als Zustandsänderungen innerhalb einer Phase und berücksichtigen die ausgetauschten Informationen erst beim Phasenübergang als Untermenge von δ. Weiterhin können für den Beginn einer Phase p_i Daten η_i notwendig sein, d.h., die Bedingung $\eta_i \subseteq \delta_i$ muss in diesem Fall gelten.

Im Folgenden sollen anhand dieser vier Phasen die Datenströme eines typischen Einkaufs im Internet analysiert werden. Jedem Phasenabschnitt sind die Daten bzw. Informationen vorangestellt, die in der jeweiligen Phase ausgewertet werden.

3.1 Start

Der Start p_0 stellt die Initialisierung des Einkaufsvorgangs und damit gleichzeitig auch der Suchphase p_1 dar. Hierunter ist das Aufrufen der Web-Seiten des Anbieters $A \in \mathcal{A}$ durch den Nutzer $N \in \mathcal{A}$ zu verstehen. Die Suchphase p_1 wird durch die Start-Transition $t_0(N, A, \delta)$ eingeleitet. Hierbei gilt $\delta = \{c_0\}$, falls beim Aufrufen der Web-Seiten des Anbieters A Daten c_0 einer früheren Sitzung wie bspw. Cookies oder andere nutzerspezifische Daten übermittelt werden, andernfalls $\delta = \emptyset$.

3.2 Suchphase

▶ (N, A, δ)

In der Suchphase p_1 durchsucht der Nutzer N zunächst das Warenangebot $\mathcal{W}$ des Anbieters A. Hierbei wird ihm vom Anbieter eine Sitzungsidentifikation sid zur Verwaltung des Warenkorbs zugewiesen. Entschließt er sich zum Kauf der Waren $\{w_1, \ldots, w_n\}$, $w_i \in \mathcal{W}$, leitet er durch die Transition $t_1(N, A, \{sid, id(N), w_1, \ldots, w_n, c_1\})$ die nächste Phase ein, wobei c_1 Clickstream- oder sonstige Inhaltsdaten bezeichne, die im Rahmen der Suchphase bei A angesammelt wurden — vereinfacht sei angenommen, dass stets gilt $c_i \subseteq c_{i+1}$. Weiterhin übermittelt der Nutzer seine Identität $id(N)$, d.h. Name, Straße, PLZ und Ort. Als Vorbedingung für die Transition t_1 muss gelten $\eta_1 \subseteq \mathcal{W}$ und $\eta_1 \neq \emptyset$, d.h. der Nutzer hat mindestens einen Artikel zum Kauf ausgewählt.

3.3 Bestellphase

▶ $(N, A, \{sid, id(N), w_1, \ldots, w_n, c_1\})$

In der Bestellphase p_2 ist der Nutzer N schließlich dem Anbieter namentlich bekannt. Das heißt, die Nutzungsdaten c_1 können an dieser Stelle mit der Nutzeridentität in einem Profil

zusammengefasst werden. Weiterhin werden die Bestellinformationen $\{w_1, \ldots, w_n\}$ verarbeitet und u.U. die Identität $id(N)$ überprüft. Die nächste Phase wird durch die Transition $t_2(N, A, \{sid, f, c_2\})$ eingeleitet, wobei f Finanzdaten des Nutzers seien wie bspw. Kontonummer und Bankverbindung oder Kreditkarteninformationen. Es muss folglich gelten: $f = \eta_2$.

3.4 Bezahlphase

► $(N, A, \{sid, f, c_2\})$

Auch in der Phase p_3 wird der Anbieter die ihm übermittelten Daten f prüfen und bei positiver Prüfung, d.h. Zahlungsfähigkeit des Nutzers, die bestellten Waren $\{w_1, \ldots, w_n\}$ versandfertig machen. Mit dem Initiieren der Bezahlphase ist die Kommunikation zwischen N und A beendet, d.h. ab diesem Zeitpunkt hat der Nutzer keine Möglichkeit mehr die Sitzung abzubrechen. Die letzte Phase p_4 leitet der Händler durch Benachrichtigung seines Paketdienstes $P \in \mathcal{A}$ ein, wobei er diesem die Adresse mitteilt, an der das Paket abgeholt werden kann, wohin es geliefert werden soll und evtl. weitere Daten d. Die Transition hierfür ist $t_3(A, P, \{id(A), id(N), d\})$.

3.5 Auslieferungsphase

► $(A, P, \{id(A), id(N), d\})$

Schließlich wird in der Auslieferungsphase p_4 das Paket mit den darin enthaltenen Waren $\{w_1, \ldots, w_n\}$ vom Paketdienst P beim Anbieter A abgeholt und an den Nutzer N ausgeliefert. Zuvor werden die bestellten Waren von A verpackt, sodass P i.Allg. der genaue Inhalt des Pakets nicht bekannt ist und nur N das Paket öffnen soll. Natürlich könnte P das Paket öffnen, dies kommt aber in der Praxis selten vor und könnte zudem von N entdeckt werden. Wir modelieren das Paket, genauer gesagt die darin eingeschlossenen Informationen x durch $enc_N(x)$. Die Ziel-Transition, d.h. die Übergabe der Waren von P an N, lässt sich dann beschreiben als $t_4(P, N, \{enc_N(w_1, \ldots, w_n)\})$.

3.6 Informationsverteilung

Der zuvor dargestellte Ablauf eines Einkaufsprozesses, trifft für die Mehrzahl der über das Internet abgewickelten Transaktionen zu. Im Gegensatz zu einem Einkauf in einem realen Kaufhaus sind nicht nur dem Nutzer alle Informationen des Einkaufs bekannt, sondern ebenfalls dem Anbieter A. Dieser erfährt Name und Adresse, die gekauften Waren, die Finanzinformationen des Nutzers und evtl. abgeleitete Informationen aus den vom Nutzer generierten *Clickstream*-Daten c_i. Im Kaufhaus hingegen würde ein Anbieter i.d.R. lediglich die vom Nutzer gekauften Waren kennen, sofern der Nutzer bar bezahlt.

Der Paketdienst P erfährt außer der Absender- und Lieferanschrift nichts über den Einkaufsvorgang. Er kann höchstens auf Grund des Absenders ungefähr die Art der Waren bestimmen, bspw. Bücher, CD, DVD, jedoch nicht deren genaue Titel. Damit ist es ihm nicht möglich, inhaltlichen Profilinformationen, wie bspw. Interessengebiete, zu erlangen.

Ziel dieser Arbeit ist es, den Einkauf im Internet hinsichtlich der Wissensverteilung soweit wie möglich dem Einkauf in einem realen Kaufhaus anzugleichen. Der zur Erreichung dieses Ziels gewählte pseudonyme Ansatz wird im Folgenden näher erläutert.

4 Pseudonymisierung des Interneteinkaufs

In diesem Abschnitt werden wir zeigen, wie die aus dem vorigen Abschnitt bekannten Phasen mit ihren Datenströmen pseudonymisiert werden können, sodass dem Händler die wahre Identität des Nutzers nicht bekannt wird. Das Neue an diesem Ansatz ist, dass auf vorhandene Infrastrukturen zurückgegriffen werden kann, insbesondere, dass keine weitere Partei hinzugefügt wird, die aktiv in den Geschäftsprozess eingreift.

In einer Phase vor dem eigentlichen Geschäftsprozess benötigen wir jedoch eine vertrauenswürdige dritte Partei T, die für die Ausgabe von Pseudonymen zuständig ist. Diese Partei ist aber ansonsten nicht direkt am Einkaufsprozess beteiligt, weshalb $T \notin \mathcal{A}$ gilt. T vergibt zertifizierte Pseudonyme $cert_T(ps(N))$ an Nutzer N, nachdem deren wahre Identität hinreichend überprüft wurde. Das heißt, T kennt die Zuordnung eines Pseudonyms $ps(N)$ zur wahren Nutzeridentität $id(N)$. Das Pseudonym $ps(N)$ dient dem Anbieter A als Sicherheit, falls in einer der Einkaufsphasen Probleme auftreten sollten, die einer namentlichen Identifizierung des Nutzers bedürfen. In diesem Bedarfsfall kann A sich zur Klärung an T wenden und ggf. die Aufdeckung des Pseudonyms $ps(N)$ beantragen.

Im Folgenden werden wiederum die vier Einkaufsphasen aus dem vorangegangen Abschnitt herangezogen, um den nunmehr pseudonymisierten Einkaufsprozess hinsichtlich der Informationsflüsse zu analysieren. Auf Unterschiede bzw. Gleichheiten zum nicht pseudonymisierten Einkauf wird in der jeweiligen Phase hingewiesen.

4.1 Start

Im Startzustand p'_0 findet wie zuvor die Initialisierung des Einkaufsvorgangs statt, die genau wie in Unterabschnitt 3.1 beschrieben abläuft. Die Start-Transition lautet dann ebenfalls $t'_0(N, A, \delta)$. Hier muss jedoch sichergestellt sein, dass δ keine Vereknüpfungsdaten wie Cookies enthält, sonst wäre eine Zuordnung verschiedener Pseudonyme zu einem Nutzer möglich.

4.2 Suchphase

▶ (N, A, δ)

Zu Beginn der Suchphase wird wiederum der Sitzung des Nutzers eine Identifikation sid durch den Anbieter zugeordnet. Nachdem der Nutzer sich in der Suchphase p'_1 zum Kauf der Waren $\{w_1, \ldots, w_n\}$, $w_i \in \mathcal{W}$ entschieden hat, muss er sich nach wie vor gegenüber dem Händler identifizieren. Im Gegensatz zur „alten“ Suchphase p_1, gibt der Nutzer statt seiner Identität $id(N)$ lediglich sein zertifiziertes Pseudonym $cert_T(ps(N))$ preis. Die neue Transition lautet dann $t'_1(N, A, \{sid, cert_T(ps(N)), w_1, \ldots, w_n, c_1\})$.

4.3 Bestellphase

▶ $(N, A, \{sid, cert_T(ps(N)), w_1, \ldots, w_n, c_1\})$

In der pseudonymisierten Bestellphase p'_2 ist der Nutzer im Gegensatz zur vorherigen Bestellphase p_2 nicht namentlich bekannt. Der Anbieter A kann dennoch die Identität des Pseudonyms $ps(N)$ prüfen, indem er das von T ausgestellte Zertifikat $cert_T(ps(N))$ überprüft. Ist diese Prüfung erfolgreich, hat A die Gewissheit, dass sich hinter dem Pseudonym $ps(N)$ eine identifizierbare Person verbirgt. Man beachte, dass zur Überprüfung

der Echtheit des präsentierten Pseudonyms die dritte Partei T nicht (direkt) in den Einkaufsprozess eingebunden werden muss, sondern eine *Offline*-Prüfung mit Hilfe ihres Prüfschlüssels und evtl. Sperrlisten ausreichend ist. Die ausgewählten Waren $\{w_1, \ldots, w_n\}$ werden von A wie bereits erwähnt über die Sitzungsidentifikation sid verwaltet bzw. dieser zugeordnet. Für den Übergang zur Bezahlphase müssen Finanzdaten f' übermittelt werden. Diese dürfen jedoch nicht die Identität $id(N)$ des Nutzers preisgeben. Denkbar sind hierfür elektronische *Pre-Paid*-Verfahren wie digitale Münzen[1] oder auch elektronische Lastschriftverfahren[9]. Im einfachsten Fall gilt $f' = \emptyset$. Dies ist bspw. dann der Fall, wenn der Nutzer sich entschieden hat, per Nachnahme zu zahlen. Die Transition für den Übergang in die Bezahlphase p'_3 ist dann $t'_2(N, A, \{sid, f', c_2\})$.

4.4 Bezahlphase

▶ $(N, A, \{sid, f', c_2\})$

Abhängig von der für die Bezahlphase p'_3 gewählten Zahlungsart werden zunächst die in f' übermittelten Finanzdaten überprüft und evtl. ein Zahlungsprotokoll zwischen Anbieter und Nutzer initiiert. Im Gegensatz zu p_3 ist für den Nutzer mit p'_3 und der Übermittlung der Finanzdaten die Sitzung noch nicht abgeschlossen. Denn noch ist die Lieferanschrift $id(N)$ des Nutzers nicht bekannt. Statt Name und Lieferadresse wie bisher an A zu übermitteln, schickt der Nutzer N die Daten stattdessen direkt an den Paketdienstleister P. Dies wird durch die Transition $t'_3(N, P, \{sid, id(N), id(A)\})$ beschrieben. Hierbei dient die Sitzungsidentifikation sid dem Paketdienst P für die spätere Zuordnung des Pakets zu einer Empfängeranschrift.

4.5 Auslieferungsphase

▶ $(N, P, \{sid, id(N), id(A)\})$

In der pseudonymisierten Auslieferungsphase p'_4 holt P ebenso wie in p_4 das Warenpaket bei A ab. Das vom Anbieter übergebene Paket ist in diesem Fall zunächst „nur“ mit der Sitzungsidentifikation sid versehen statt mit der Empfängeranschrift $id(N)$, da dies, neben dem Pseudonym, die einzige Bezugsinformation des Anbieters zur Bestellung ist. Der Paketdienstleister kann dann durch die ihm übermittelte Zuordnungsregel $sid \rightarrow id(N)$, das Paket dem richtigen Empfänger zuordnen. Man beachte, dass das Pseudonym hierfür dem Paketdienstleister nicht bekannt gemacht werden muss, d.h. er erfährt nicht die Zuordung $ps(N) \rightarrow id(N)$. Die Verwendung des Pseudonyms für die Paketzuordnung wäre auch nicht hinreichend. Denn der Pseudonymträger könnte mehrere Transaktionen unter dem gleichen Pseudonym bei A tätigen und für jede Transaktion eine andere Empfängeradresse angeben, bspw. wenn er Geschenke verschicken möchte. Die Zieltransition $t_4(P, N, \{enc_N(w_1, \ldots, w_n)\})$ wird unverändert aus p_4 übernommen. Denn nachdem P die Sitzungsidentifikation durch die Empfängeranschrift ersetzt hat, handelt es sich bei dem ausgelieferten Paket wieder um eine „gewöhnliche“ Warensendung.

4.6 Informationsverteilung

Im dargestellten pseudonymisierten Einkaufsprozess erlangt der Anbieter A weit weniger Wissen als zuvor. Er erfährt selbstverständlich weiterhin die Wareninformationen sowie *Clickstream*-Informationen c_i. Letztere können, wie eingangs bemerkt, dem Anbieter

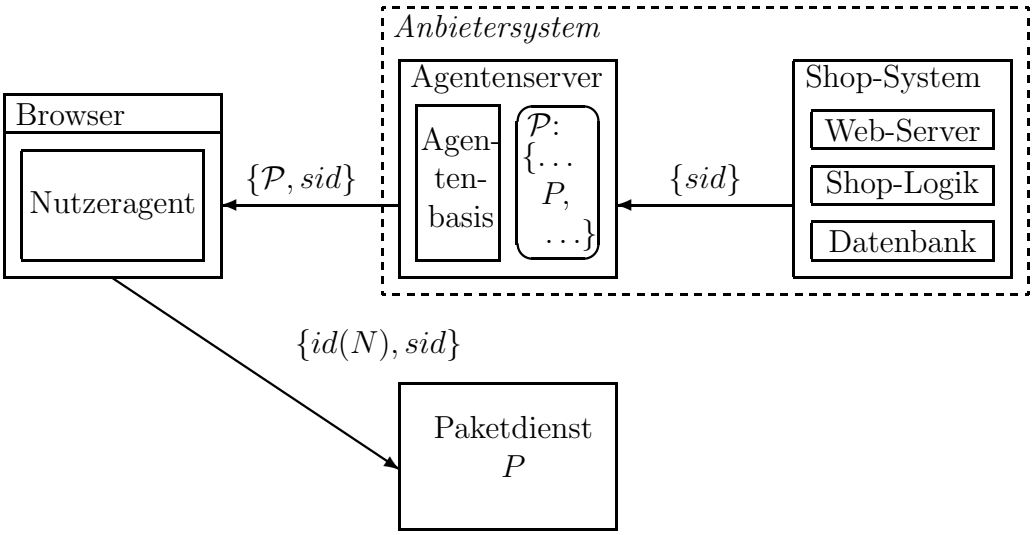

Abb. 1: Architektur des Pseudonym-Systems

Anhaltspunkte für die Gestaltung seines zukünftigen Warenangebots liefern. Durch die Übermittlung des Pseudonyms $ps(N)$ ist es ihm auch möglich, diese Profilinformationen einem bestimmten, jedoch nicht namentlich bekannten Kunden zuzuordnen. Die Finanzdaten f' können A *per se* keine personenbezogenen Daten liefern, da dies dem Wesen eines anonymen oder pseudonymen Zahlungssystems widersprechen würde.

Der Paketdienst P erfährt in der pseudonymen Variante neben den in Unterabschnitt 3.6 genannten Informationen noch zusätzlich die Sitzungsidentifikation *sid*. Da es sich hierbei aber um eine mehr oder weniger zufällige und einmalige Referenznummer handelt, kann diese Information hinsichtlich ihres Informationsgehalts für den Paketdienst vernachlässigt werden. Eine Sitzungsidentifikation wird i.d.R. keine Informationen über die gekauften Waren an P preisgeben, da sie bereits vor der Auswahl eines Artikels vom Anbieter festgelegt wird und sich üblicherweise Änderungen am Warenkorbinhalt nicht in *sid* widerspiegeln.

5 Architektur

Im Folgenden stellen wir eine Architektur für das zuvor beschriebene Pseudonym-System vor. Das System wurde im Projekt DASIT [5] entwickelt und bereits in einem Feldversuch praktisch erprobt. Die nachfolgend beschriebene Architektur ist in Abb. 1 schematisch dargestellt.

Den Kern des DASIT-Systems bilden auf Nutzerseite ein Agent in Form eines Java-Applets und ein dazu komplementärer Agentenserver auf Anbieterseite. Für den pseudonymen Einkauf dient der Nutzeragent hauptsächlich als vertrauenswürdige Anzeigekomponente, welche die vom Anbieter übermittelte Datenschutzpolitik $\mathcal{P}$ anzeigt sowie die zu erhebenden Daten entgegennimmt und an den jeweiligen Empfänger versendet. Der Programmcode des Agenten ist durch eine digitale Signatur geschützt, die von einem vertrauenswürdigen Dritten erstellt wird, bspw. von der Pseudonym-ausgebenden Stelle T aus Abschnitt 4. Dies dient dem Nutzer zur Erkennung von manipulierten Agenten,

die Daten erheben und diese an einen nicht genannten Dritten versenden könnten, statt an den behaupteten Empfänger. Durch einen manipulierten Agenten wäre es für einen betrügerischen Anbieter ein Leichtes, das Pseudonym des Nutzers aufzudecken, indem er durch den Agenten die Lieferanschrift statt an den Paketdienst an sich selbst schicken ließe.

Der Agentenserver dient zum einen dazu, den Agenten zu beherbergen und zum anderen der Anbindung des auf Anbieterseite vorhandenen Shop-Systems. Über den Server werden außerdem die beim Nutzer abgefragten Daten in die Shop-Datenbank eingespeist bzw. von dort abgefragt, falls der Nutzer sein Profil ergänzen oder abändern möchte. Auf dem Server stellt der Anbieter auch seine Datenschutzpolitik $\mathcal{P}$ ein, die u.a. beinhaltet, welche Profilinformationen zu welchem Zweck abgefragt werden und wer die Empfänger der Informationen sind. Dies schließt die Angabe des Paketdienstes ein, der die pseudonymen Warenpakete ausliefert. Die Datenschutzpolitik wird schließlich vom Nutzeragenten interpretiert.

Die Datentrennung zwischen Anbieter und Paketdienst hinsichtlich der Lieferanschrift wird über die Datenschutzpolitik festgelegt und durch den Agenten umgesetzt. Der Agent nimmt hierzu die vom Nutzer eingegebene Lieferanschrift $id(N)$ entgegen und verschickt diese über einen vertraulichen Kanal an eine URL des Paketdienstes P. Die URL entnimmt der Agent der Datenschutzpolitik $\mathcal{P}$ des Anbieters. Zur Sicherheit des Nutzers wird die genaue URL vom Agenten angezeigt, sodass der Nutzer überprüfen kann, ob die Daten tatsächlich an einen Paketdienst versendet werden. Dies setzt voraus, dass die URL zu einem dem Nutzer bekannten Paketdienst gehört.

6 Diskussion

Das in diesem Beitrag vorgestellte System setzt voraus, dass der Händler nicht mit dem Paketdienst zusammenarbeitet. Durch eine solche Koalition könnte sonst das Pseudonym des Nutzers leicht aufgedeckt werden. Es stellt sich also die Frage, was die beiden Parteien davon abhält zusammenzuarbeiten. Die erste, technische, Antwort lautet: Gar nichts! Durch Einführung zusätzlicher vertrauenswürdiger Parteien könnte eine Zusammenarbeit zwischen Parteien sicherlich erschwert werden, jedoch würde dies dem Ziel dieses Ansatzes bzgl. Einfachheit und ausschließlicher Nutzung von vorhandenen Infrastrukturen zu wider laufen.

Die zweite Antwort auf die gestellte Frage liefert deshalb weder technische noch organisatorische Argumente, sondern führt datenschutzrechtliche und wirtschaftlische Gründe an. Zunächst muss festgehalten werden, dass eine Koalition zwischen Händler und Paketdienst zum Zweck der Pseudonymaufdeckung einen Verstoß gegen europäisches Datenschutzrecht [4] sowie dessen Umsetzungen in nationale Datenschutzgesetze (bspw. das deutsche TDDSG [3]) bedeuten würde. Dies alleine hat in der Vergangenheit jedoch noch keine durchschlagende Wirkung erzielt. Das zweite Argument, der wirtschaftliche Faktor, ist deshalb auch gewichtiger. Wenn wir annehmen, dass ein Händler und ein Paketdienst eine Koalition bilden, so gehen beide hierdurch nicht nur ein strafrechtliches, sondern auch in wirtschaftlicher Hinsicht hohes Risiko ein. Denn sollte die Koalition bekannt werden, werden sich Kunden von dem Händler und dem Paketdienst zurückziehen. Somit wäre der wirtschaftliche Schaden sowie der Imageverlust für beide beträchtlich und sicher nicht

einfach wieder gutzumachen. Zudem stellt sich für den Paketdienst die Frage, welchen Vorteil er durch eine solche Koalition hätte. Eine Bestechung durch den Händler müsste für ihn das finanzielle Risiko aufwägen, dass sich durch das Bekanntwerden der Koalition ergeben würde. Denn es ist zu erwarten, dass sich (datenschutzbewusste) Kunden von ihm als Auslieferer zurückziehen würden. Das heißt, auch Kunden, die bei anderen Händlern pseudonym einkaufen, werden es sich nach Bekanntwerden einer solchen Koalition überlegen, ob sie ihre Waren noch über den betrügerischen Paketdienst ausliefern lassen. Dies wiederum hätte zur Folge, dass andere Händler, selbst wenn sie keine Koalitionsabsichten mit jenem Paketdienst hegen, den Auslieferer wechseln müssten, um das Vertrauen ihrer Kunden nicht zu verlieren.

Für den Käufer bestünde noch die Möglichkeit sich von Händler/Paketdienst-Koalitionen zu schützen, indem er sich seine bestellten Waren an ein Postfach liefern ließe. Dies setzt jedoch zum einen voraus, dass der Nutzer die Waren bereits elektronisch bezahlt hat (Nachnahme wäre in diesem Fall nicht möglich) und er zum anderen bereits ein solches Schließfach besitzt. Zudem bieten sich Postfächer nur für Waren bestimmter Größe an, verursachen zusätzlich Kosten und erfordern, dass der Nutzer das Postfach regelmäßig prüft, was i.d.R. unbequemer ist als eine Heimlieferung.

7 Verwandte Arbeiten

In der Literatur befassen sich zahlreiche Arbeiten mit anonymem oder pseudonymem Einkaufen im Internet. Es existieren eine Reihe von Vorschlägen, welche die Identität bzw. die Kommunikationsbeziehung des Kunden (Sender) vor dem Händler (Empfänger) oder auch vor dritten Parteien (Beobachtern) verbergen können [11, 10]. Elektronische Zahlungsysteme wie eCash [1] oder SET [9] wurden vorgeschlagen und umgesetzt, um anonyme bzw. pseudonyme Zahlungen im Internet zu ermöglichen. Alle diese Maßnahmen greifen jedoch nur dann, wenn es sich bei den einzukaufenden Gütern um immaterielle Waren handelt, die nicht physisch ausgeliefert werden müssen. Sobald eine physische Auslieferung notwendig wird, was heute i.d.R. der Fall ist, wird der Einsatz dieser Techniken nahezu sinnlos, da der Anbieter für die Auslieferung letztendlich doch Name und Lieferanschrift des Kunden erfährt. In [6] werden Informations-Zwischenhändler (*Infomediaries*) vorgeschlagen, welche im Namen des Kunden bei einem Händler Bestellungen durchführen. Ein solcher Zwischenhändler könnte dazu genutzt werden, die Ware zunächst geliefert zu bekommen, um sie dann an den tatsächlichen Kunden weiterzuverschicken. Diese Indirektion führt für den Kunden jedoch zu neuerlichen Versandkosten und u.U. auch zu einer Lieferverzögerung, da nicht sichtergestellt ist, dass der Zwischenhändler das Kundenpaket sofort weiterschickt.

Zusammenfassung

Wir haben einen praktikablen Lösungsansatz vorgestellt, der es möglich macht, dass Kunden im Internet einkaufen können ohne dabei personenbezogene Daten zu hinterlassen. Der Ansatz sucht dabei einen Interessensausgleich zwischen Händlern und Kunden. Erstere haben ein Interesse daran, ihre Kunden kennenzulernen, während letztere dies oft zu verhindern suchen. Der hier gewählte pseudonyme Ansatz ermöglicht beides, den Kunden bzw. dessen Interessen kennenzulernen, ohne ihn dabei namentlich erfassen zu müssen.

Die vorgestellte Lösung ist mit vertretbarem Aufwand bei den beteiligten Parteien in existierende Systeme integrierbar, ohne dass für den Betrieb neue, bisher nicht vorhandene Infrastrukturen geschaffen werden müssten.

Danksagung

Die Autoren danken Markus Schneider, Thomas Kunz sowie drei anonymen Gutachern für hilfreiche Hinweise und Kritik.

Literatur

[1] David Chaum. Privacy protected payments: Unconditional payer and/or payee untraceability. In *Smart Card 2000, Proceedings*. North Holland, 1989.

[2] Sebastian Clauß and Marit Köhntopp. Identity managements and its support of multilateral security. *Computer Networks – Special Issue on Electronic Business Systems*, (37), 2001.

[3] Bundesrepublik Deutschland. Gesetz über den Datenschutz bei Telediensten (Teledienstedatenschutzgesetz – TDDSG) vom 22. Juli 1997 (BGBl. I, S. 1870), zuletzt geändert durch Artikel 3 des Gesetzes über rechtliche Rahmenbedingungen für den elektronischen Geschäftsverkehr (Elektronischer Geschäftsverkehr-Gesetz – EGG) vom 14. Dezember 2001. BGBl. I, S. 3721, 2001.

[4] European Parliament and the Council of the European Union. Directive 95/46/ec of the european parliament and the council of 24 october 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Official Journal L 0281, October 1995.

[5] Rüdiger Grimm, Nils Löhndorf, and Philip Scholz. Datenschutz in Telediensten (DASIT) am Beispiel von Einkaufen und Bezahlen im Internet. *DuD, Datenschutz und Datensicherheit*, (5), May 1999.

[6] John Hagel and Jeffrey Rayport. The new infomediaries. *The McKinsey Quarterly*, (4), 1997.

[7] Donna L. Hoffman, Thomas P. Novak, and Marcos Peralta. Building consumer trust online. *Communications of the ACM*, 42(4), April 1999.

[8] Marit Köhntopp and Kristian Köhntopp. Datenspuren im Internet. *Computer und Recht (CR)*, (4), April 2000.

[9] Mastercard and VISA. SET – secure electronic transaction, specification version 1.0, books 1-3. http://www.setco.org/download/setbk_{1, 2, 3}.pdf, May 1997.

[10] Michael K. Reiter and Aviel D. Rubin. Crowds: Anonymity for web transactions. *ACM Transactions on Information and System Security*, 1(1), 1998.

[11] Paul F. Syverson, David M. Goldschlag, and Michael G. Reed. Anonymous connections and onion routing. In *Proceedings of 18th Symposium on Security and Privacy*. IEEE Press, May 1997.

[12] Sabine Wolters. Einkauf via Internet: Verbraucherschutz durch Datenschutz. *DuD, Datenschutz und Datensicherheit*, (5), May 1999.