Anonyme und unbeobachtbare Kommunikation im Internet

Hannes Federrath
FU Berlin
feder@inf.fu-berlin.de

Stefan Köpsell
TU Dresden
sk13@inf.tu-dresden.de

Heinrich Langos
FU Berlin
langos@inf.fu-berlin.de

Abstract: Datenschutz im Internet kann nicht beim Schutz von Kommunikationsinhalten enden. Auch die äußeren Umstände einer Kommunikation, also wer wann mit wem kommuniziert, können schützenswerte personenbezogene Daten sein. Dieser Beitrag soll einen kurzen Überblick darüber geben, welche Systeme zum Schutz dieser Daten speziell beim Zugriff auf das WWW existieren, und wie das im BMWi-Projekt „AN.ON – Anonymität online“ entwickelte System deren Schutzniveau übersteigt. Ferner wird auf die bisherigen Erfahrungen im Probebetrieb und die noch zu lösenden Probleme eingegangen.

1 Motivation

Informationen, die eigentlich der Privatsphäre angehören und die durch Gesetze geschützt sind, werden in zunehmendem Maße ein Wirtschaftsgut. Persönlichkeits- und Interessenprofile werden gekauft und verkauft, um in einer zunehmend individualisierten Gesellschaft auch das Marketing von Produkten und Dienstleistungen individuell, also auf den einzelnen Kunden zugeschnitten, zu gestalten. Dies mag auf den ersten Blick für beide Seiten vorteilhaft sein. Probleme erwachsen aber aus den Datensammlungen, die dazu nötig sind, und aus der Art, wie diese Daten gesammelt werden. So kann zum Beispiel der Internet Service Provider sehr umfassende Daten sammeln, wenn er die Leitungen seiner Kunden beobachtet. Ohne aufwendige Maßnahmen kann er mitlesen, welche Online-Shops ein Nutzer besucht, für welche Nachrichten er sich interessiert und welche Meinung er (möglicherweise unter Pseudonym) in einem Forum äußert. Kommen dazu noch Informationen über den Gesundheitszustand, wird klar, warum die Privatsphäre auch im Internet des Schutzes bedarf, den sie in anderen Lebensbereichen genießt.

Juristisch ist der Schutz dieser Daten durch das Recht auf informationelle Selbstbestimmung geregelt. Die Durchsetzung eines solchen Rechtes muss aber auch durch technische Mittel unterstützt werden, ähnlich wie das Briefgeheimnis durch die Verwendung von Umschlägen und die Vertraulichkeit von E-Mails durch Verschlüsselung unterstützt wird.

Das bloße Verschlüsseln von Daten reicht jedoch nicht aus. Oft genug ist unabhängig vom Inhalt das Vorhandensein einer Kommunikationsbeziehung schon ein schützenswertes Datum. So kann beispielsweise das Abrufen von Webseiten mittels einer SSL-verschlüsselten Verbindung die genaue Webseite vor einem außenstehenden Beobachter verbergen, nicht

jedoch, von welchem Server diese Seite abgerufen wurde. Im Falle von Seiten, die vom Server einer AIDS- oder Drogenberatung abgerufen wurden, kann dies bereits eine Information sein, die man nicht in falschen Händen wissen möchte.

Dass Verbindungsdaten schützenswert sind, ist auch seitens des Gesetzgebers anerkannt worden. So enthält beispielsweise das Teledienstedatenschutzgesetz die Gebote, Dienste nach Möglichkeit auch anonym oder pseudonym nutzbar zu machen. und personenbezogene Verbindungsdaten, die nicht mehr für Abrechnung gebraucht werden, nach Ende der Nutzung zu löschen.

Die aktuelle Debatte über Vorratsspeicherung von Verbindungsdaten zur Verbrechensbekämpfung zeigt allerdings deutlich, dass diese Regelungen nicht unumstritten sind. Ob die Gefahren, die von einer so umfassenden Überwachung für einen demokratischen Rechtstaat ausgehen, nicht letztendlich größer sind als die Bedrohung durch Terroristen, bleibt eine offene Frage. Auch die Aufdeckung „normaler“ Verbrechen könnte durch Anonymisierungsdienste erschwert werden, und die gesellschaftliche Akzeptanz dieser Dienste hängt nicht zuletzt davon ab, wie ihr Nutzen und der mögliche Schaden gegeneinander abgewogen werden.

2 Vorhandene Lösungen

In den letzten Jahren entstanden mehrere Projekte und Dienste, die sich mit dem Anonymisieren von Internetverbindungen im World Wide Web befassen. Einige sind reine Forschungsprojekte, andere Dienste wurden bzw. werden kommerziell angeboten. Alle basieren darauf, die Nachrichten von und zu einem Webserver über Zwischenstationen (Proxies) zu leiten. In den sichersen Lösungen kommt dabei das Mix-Konzept [Chau81] zur Anwendung. Mixe sind spezielle Proxies, die asymmetrisch verschlüsselte Nachrichten empfangen, auf Wiederholungsangriffe (Replay) überprüfen, sammeln und entschlüsselt wieder ausgeben, so dass einem Beobachter die Zuordung von ein- und ausgehenden Nachrichten nicht bekannt wird. Mehrere Mixe unabhängiger Betreiber werden hintereinandergeschaltet, um auch eine Beobachtung durch die Mix-Betreiber zu verhindern. Solange wenigstens ein Mix einer Mix-Kette vertrauenswürdig ist, bleibt die Kommunikationsbeziehung zwischen Sender und Empfänger geschützt. Um zu verbergen, wann ein Sender Nachrichten sendet, werden ständig Scheinnachrichten gesendet (Dummy Traffic).

Anonymisierer müssen die folgenden bekannten Angriffe abwehren: **Verkettungsangriffe** führen direkt zu einer Zuordnung von eingehender und ausgehender Nachricht anhand von gleichen Inhalten, zeitnahem Auftreten oder auch nur gleichen Nachrichtenlängen. **Schnittmengenangriffe** sind durchführbar, wenn mehrere Nachrichten als zu einer Kommunikationsbeziehung gehörend erkannt werden können. Die durch wechselnde Nutzeraktivität schwankenden Mengen der möglichen Sender bzw. Empfänger werden durch Schnittmengenbildung verknüpft und immer weiter reduziert bis nur der tatsächliche Nutzer übrig bleibt. Bei **Flutungsangriffen** versucht der Angreifer, das System mit eigenen Nachrichten solange zu füllen, bis nur noch sehr wenige ihn interessierende Nachrichten von echten Absendern akzeptiert werden, die dann isoliert und deshalb nicht mehr

anonym sind. Wenn Zwischenstationen zu Angreifern werden und Informationen untereinander austauschen, spricht man von **gemeinschaftlichen Angriffen**.

Im Folgenden werden die bekanntesten Web-Anonymisierer kurz vorgestellt.

Anonymizer und Verwandte

Der bekannteste Web-Anonymisierer ist Anonymizer (http://www.anonymizer.com/). Dabei handelt es sich um einen sogenannten formularbasierten Proxy. Der Nutzer trägt auf der Webseite des Betreibers in einem Formular die URL ein, die er anonym abrufen möchte und lässt den Inhalt durch den Proxy holen und zu sich weiterleiten. Manche Proxies filtern noch Verweise auf aktive Inhalte sowie Cookies.

Der Anwender muss hierbei darauf vertrauen, dass der Proxy-Betreiber ihn nicht beobachtet und seine Aktivitäten protokolliert, da der Proxy genau erfährt, wann von welcher IP-Adresse welche Requests ausgehen. Ein Angreifer, der alle Kommunikation im Netz abhören kann, oder auch nur alle ein- und ausgehenden Verbindungen des Anonymisierers beobachtet, kann über die Inhalte der Nachrichten und deren Länge sowie die zeitlichen Korrelationen der ein- und ausgehenden Nachrichten verketten. Außerdem kann der Angreifer sofort alle Inhalte mitlesen, da keine Verschlüsselung verwendet wird. Zwar haben einige Anonymisierer nachgerüstet, indem sie neuerdings auch SSL-verschlüsselte Server betreiben, die Verkettbarkeit über die Nachrichtenlänge und zeitliche Korrelationen verhindert die Verbindungsverschlüsselung zwischen Benutzer und Anonymisierer jedoch nicht.

Crowds

Beim System Crowds [ReRu98] werden die Webzugriffe über zufällig ausgewählte Teilnehmer des Systems geleitet, bevor sie den Webserver erreichen. Bei jedem Teilnehmer wird die Anfrage, gesteuert von einem Zufallsprozess, mit einer Wahrscheinlichkeit p direkt an den Server geschickt bzw. mit $p - 1$ zu einem weiteren Teilnehmer.

Die Kommunikationsinhalte werden bei Crowds im Gegensatz zum Anonymizer zwischen den Nutzern verschlüsselt. Eine Verkettung über die Länge der Nachrichten und damit die Beobachtung ist jedoch nach wie vor möglich, wenn der Angreifer Verkehrsanalysen durchführt. Gegen Angriffe über die zeitliche Verkettung von eingehenden Nachrichten eines Knotens und deren Ausgabe wurden keine Schutzmaßnahmen vorgesehen.

Freedom

Die Software Freedom (http://www.freedom.net/) der kanadischen Firma Zero-Knowledge basiert auf der Idee, jene Funktionen eines Mixes zu implementieren, die der Verzögerung von Nachrichten nur wenig schaden, aber alle anderen Funktionen wegzulassen. Bei der Konzeption des Systems wurde also bewusst auf Sicherheit gegenüber sehr starken Angreifern verzichtet.

Letztendlich wurde bei Freedom der Performance der Vorzug gegeben, was sich der Benutzer mit einer hohen, aber nicht perfekten Sicherheit erkauft. So sind Flutungsangriffe und für sehr starke Angreifer auch Verkettungsangriffe möglich.

Das von Zero-Knowledge betriebene Freedom-Netzwerk, ein weltweiter Verbund von Freedom-Servern, sog. Anonymous Internet Proxies (AIP), wurde von 1999 bis 2001 kommerziell angeboten und im Herbst 2001 aus ökonomischen Gründen abgeschaltet.

Onion-Routing

Einen ähnlichen Ansatz wie Freedom verfolgte das Projekt Onion Routing [GoRS96] (http://www.onion-router.net/), indem auf die Funktionen eines Mix verzichtet wurde, die die Verzögerungszeit von Datenpaketen unvorhersagbar machen. Der zwiebelschalenartige Aufbau von Mix-Nachrichten gab dem Projekt den Namen.

Dummy Traffic wird nur zwischen den Onion-Routern erzeugt und bietet somit bei geringer Auslastung des Dienstes keinen (bzw. nur geringen) Schutz gegen Beobachtung, da die Enden eines Kommunikationskanals allein über die ausgetauschte Datenmenge verkettet werden können. Ferner ist eine Verkettung über die Länge der gesendeten Nachrichten möglich.

3 Das AN.ON-System

Im folgenden Abschnitt wird der technische Hintergrund des von uns entwickelten Systems (http://anon.inf.tu-dresden.de/) skizziert. Die Basis bilden die von David Chaum entwickelten Mixe [Chau81] unter Einbeziehung des Konzepts der symmetrischen Kanäle [PfPW89]. Als Verknüpfungsform der Mixe wurde die Kaskade gewählt, d.h. die Mixe werden in einer festen, nicht vom Nutzer bestimmbaren Reihenfolge durchlaufen. Die Kaskade besitzt gegenüber dem Mixnetz (freie Folge) eine geringere Komplexität, was zu höherer Sicherheit [BePS01] und geringerem Implementierungsaufwand führt.

Das Anonymisierungssystem AN.ON besteht aus drei Komponenten: einer Client-Software (genannt JAP), mehreren Anonymisierstationen (Mixe, die in einer Mix-Kaskade betrieben werden) und dem InfoService (Abbildung 1).

Über unseren Anonymisierungsdienst lassen sich proxy-fähige Dienste nutzen, d.h. die beim Benutzer zu installierende Software (JAP) implementiert eine Proxy-Schnittstelle, während hinter dem letzten Mix das entsprechende Gegenstück existiert. JAP sorgt außerdem für die Verschlüsselung der anonym zu übertragenden Daten und bereitet diese gemäß dem Protokoll des zugrundeliegenden Anonymisierungsdienstes auf. Ein Vorteil standardisierter Proxy-Schnittstellen ist, dass die Verarbeitung des Proxy-Protokolls auf erprobte und ausgereifte Komponenten ausgelagert werden kann, die oft noch zusätzliche Funktionalität (Zugriffskontrolle, Ressourcenbegrenzung, Caching etc.) bieten.

Um die Benutzung des Anonymisierungsdienstes zu erleichtern und dem Nutzer eine Rückmeldung über sein aktuelles Schutzniveau zu geben, wurde ein dritter Bestandteil

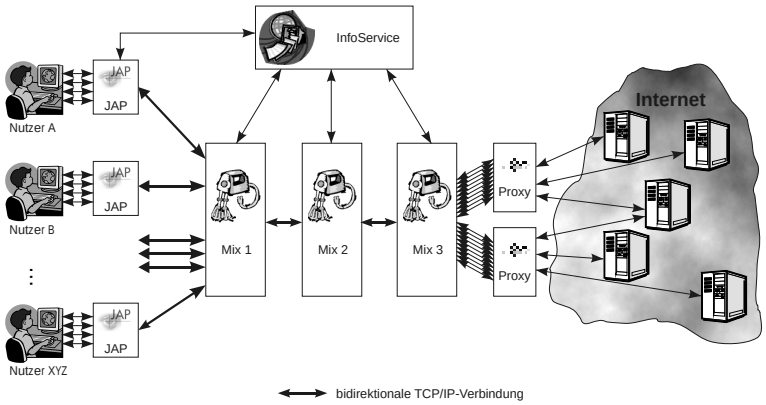

Abbildung 1: Architektur des Systems

in das Gesamtsystem aufgenommen – der sogenannte *InfoService*. Dieser ist mit einer Datenbank vergleichbar und hält abrufbar Informationen über die aktuell verfügbaren Mix-Kaskaden, deren Auslastung etc. bereit. Die Client-Komponente kann mit Hilfe der beim InfoService vorliegenden Daten dem Nutzer eine Vorstellung über seinen momentanen „Grad der Anonymität“ vermitteln.

Das System soll sowohl benutzbar als auch sicher sein. Dabei lässt sich keine generelle Priorität von Benutzbarkeit gegenüber Sicherheit festlegen. Ziel ist eine sinnvolle Verhältnismäßigkeit, da weder ein unbenutzbares aber sicheres noch ein unsicheres aber benutzbares System hilfreich sind. Benutzbarkeit aus Sicht des Endanwenders beinhaltet, dass Installation und Konfiguration der Client-Komponente einfach durchzuführen sind. Außerdem darf nicht das Gefühl aufkommen, dass die Nutzung des Anonymisierungsdienstes den Zugriff auf das Internet stark einschränkt. Dies bedeutet z.B., dass Durchsatz und Latenzzeit akzeptabel sind, was ein wichtiges Kriterium beim Entwurf des Systems ist.

Um möglichst vielen Menschen den Zugang zum Anonymisierungsdienst zu ermöglichen, ist es notwendig, dass die Client-Komponente auf vielen verschiedenen Hardware- und Betriebssystemplattformen ausgeführt werden kann. JAP wurde deshalb in der plattformunabhängigen Programmiersprache Java implementiert. Die Mixe wurden aus Performancegründen in C++ implementiert.

JAP und die Mixe kommunizieren über TCP/IP-Verbindungen. Die Verbindung zum eigentlichen Ziel (Kommunikationspartner, Internet-Dienst) wird durch die Proxies hergestellt. Zwischen den Mixen einer Kaskade besteht genau eine TCP/IP-Verbindung. Zwischen JAP und erstem Mix besteht pro JAP genau eine TCP/IP-Verbindung.

Die grundlegende Kommunikationseinheit zwischen JAP und den Mixen ist das *MixPaket*. Ein MixPaket gehört zu genau einem *MixKanal*. Ein MixKanal ist eine virtuelle Zusammenfassung mehrerer MixPakete. Im letzten Mix ist mit jedem MixKanal jeweils genau eine TCP/IP-Verbindung (mit einem Proxy) assoziiert, über die mehrere Verbindungen (z.B. HTTP-Requests) laufen können.

Ein MixPaket ist 998 Bytes groß. Die ersten vier Bytes des Pakets bilden eine Kanal-ID. Über diese erfolgt die Zuordnung des MixPakets zum MixKanal. Die Kanal-ID eines Kanals ändert sich von Mix zu Mix. Neue MixKanäle werden nur durch den JAP geöffnet. Der Beginn eines neuen MixKanals wird durch das Setzen eines Open-Flags signalisiert, das Ende durch ein Close-Flag. Das erste Paket eines MixKanals ist asymmetrisch gemäß dem Mix-Umkodierungsschema verschlüsselt (1024 Bit plain RSA) und enthält für jeden Mix jeweils einen symmetrischen Schlüssel. Alle weiteren Daten werden dann symmetrisch verschlüsselt (AES-128 im 128-Bit-OFB Modus).

Wird im Browser eine URL aufgerufen, so sendet der Browser den Request zunächst an den JAP. Dieser verschlüsselt die Anfrage für die Mix-Kaskade, sendet die Anfrage über den Anonymiserungsdienst und leitet die Antwort an den Browser zurück.

4 Praktische Erfahrungen

Wir haben unseren Anonymisierungsdienst erstmals im September 2000 der Öffentlichkeit zur kostenlosen Nutzung bereit gestellt. Mit einer Meldung auf dem Heise-News-Ticker [Heise01] vom Januar 2002 weckte der Dienst erstes öffentliches Interesse. In den folgenden Wochen stiegt die Nutzerzahl auf durchschnittlich 200–300 Nutzer pro Stunde. Im September 2001 benutzten durchschnittlich 500–600 Nutzer gleichzeitig den Anonymisierungsdienst. Ein Speicherausbau im Januar 2002 führte zu einer deutlichen Performancesteigerung des Dienstes, was wiederum zu einem Anstieg der Nutzerzahlen auf durchschnittlich 800–1000 Nutzer führte. Momentan werden über den Dienst ca. 4000 Web-Requests pro Minute abgewickelt. Dabei wird täglich ein Datenvolumen von ca. 90–100 GByte verarbeitet. Die Client-Software (JAP) wurde über 100 000 Mal von unserem Web-Server heruntergeladen. Auf Grund der Art des Dienstes ist es schwierig, eine Aussage darüber zu treffen, wie viele Menschen unseren Dienst regelmäßig nutzen. Wir schätzen, dass dies ca. 20 000 sind.

Während des Betriebs des Dienstes hat sich gezeigt, dass dieser auch angegriffen und missbraucht wird. Uns erreichen durschnittlich zwei Ersuchen pro Monat von Polizei bzw. Staatsanwaltschaft zur Herausgabe von Daten (die natürlich auf Grund der Art des Dienstes nicht vorhanden sind). In insgesamt zwei Fällen bestand Verdacht auf Beschaffung von kinderpornographischem Material. In der Mehrzahl der anderen Verdachtsfälle handelt es sich um Betrug. Dabei wurden mit Hilfe von gefälschten (bzw. fremden) Kreditkarten- oder Kontodaten Leistungen erschlichen.

Als Gegenmaßnahme bieten wir den Betreibern von Web-Foren, Gästebüchern etc. an, dass wir den Zugriff auf ihre Web-Seiten über den Anonymisierungsdienst sperren.

5 Ausblick

In Zukunft sollen weitere Bausteine implementiert werden, die zum sicheren Betrieb eines Mix basierten Anonymisierungsdienstes notwendig sind. Dies sind die Ticket-Methode [BeFK01] zur Verhinderung von Flutungsangriffen und die sog. Schübe, wobei die Schwierigkeit darin besteht, die in der theoretischen Welt getroffenen Annahmen über Quality of Service des zugrundeliegenden Transportnetzes und die darauf basierenden Konzepte in die Welt des Internet zu übertragen. Hier besteht nach wie vor Forschungsbedarf. Momentan werden Konzepte für die Kommerzialisierung unseres Dienstes entwickelt und implementiert. Ein Bezahlung der Anonymisierungsdienstleistung durch den Nutzer ist eine mögliche Option, um die Kosten des Netzwerkverkehrs abzudecken. Eine andere Option wäre die kostenlose Bereitstellung von solchen Anonymisierdienstleistungen durch den Staat, um das Datenschutzniveau seiner Bürger zu verbessern.

Literaturverzeichnis

[BeFK01] O. Berthold, H. Federrath, S. Köpsell. *Praktischer Schutz vor Flooding-Angriffen bei Chaumschen Mixen*. Patrick Horster (Hrsg.): Kommunikationssicherheit im Zeichen des Internet. DuD-Fachbeiträge, Vieweg, Wiesbaden, 2001, 235-249.

[BePS01] O. Berthold, A. Pfitzmann, R. Standtke. *The Disadvantages of Free MIX Routes and How to Overcome Them*. Designing Privacy enhancing Technologies, LNCS 2009, Springer-Verlag, Berlin 2001.

[Chau81] David Chaum. *Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms*. Communications of the ACM 24/2 (1981) 84–88.

[GoRS96] P. F. Syverson, D.M. Goldschlag, M.G. Reed. *Anonymous Connections and Onion Routing*. Proc. IEEE Symp. on Security and Privacy, Oakland, May, 1997.

[Heise01] Heise News Ticker. *TU-Software schützt vor Datenschnüfflern*. http://www.heise.de/newsticker/data/wst-10.01.01-000/default.shtml

[PfPW89] A. Pfitzmann, B. Pfitzmann, M. Waidner: Telefon-MIXe. *Schutz der Vermittlungsdaten für zwei 64-kbit/s-Duplexkanäle über den (2*64 + 16)-kbit/s-Teilnehmeranschluß*. Datenschutz und Datensicherung DuD /12 (1989) 605-622.

[ReRu98] M. K. Reiter and A. D. Rubin. *Crowds: Anonymity for Web Transactions*. ACM Transactions on Information and System Security. 1/1 (1998) 66–92.