

Klassifizierung von Anonymisierungstechniken

Dogan Kesdogan^{1*} · Roland Büschkes²

¹o.tel.o communications GmbH & Co
Abt. Unternehmenssicherheit
D-51063 Köln
Dogan.Kesdogan@o-tel-o.de

²Lehrstuhl für Informatik IV
RWTH Aachen
D-52056 Aachen
roland@i4.informatik.rwth-aachen.de

Zusammenfassung

Mit der fortschreitenden Vernetzung von Rechner- und Kommunikationssystemen gewinnen datenschutzfreundliche Technologien zunehmend an Bedeutung. In der aktuellen Literatur werden verschiedene Techniken diskutiert, die insbesondere auch die Anonymisierung der Nutzer ermöglichen und deren Unbeobachtbarkeit sicherstellen. Für den Nutzer, der solche Techniken anwenden will, ist es wichtig, die verschiedenen vorgeschlagenen Techniken im Hinblick auf ihre Sicherheit und Leistungsfähigkeit bewerten und vergleichen zu können. In dieser Arbeit wird die bisher auf dem Gebiet existierende modelltheoretische Welt erweitert und Klassifizierungsgrößen vorgeschlagen, welche die geforderte Einordnung der Techniken ermöglichen. Die exemplarische Anwendung dieser Größen auf aktuell diskutierte Anonymisierungstechniken wird dazu genutzt, einen Überblick über den aktuellen Forschungsstand auf dem Gebiet zu geben.

1 Einleitung

Die zunehmende Verschmelzung der Computer- und Telekommunikationsindustrie wird es zukünftigen Nutzern ermöglichen, immer und überall auf eine Vielzahl von Diensten zugreifen zu können. Millionen von Menschen, ob zu Hause oder unterwegs, werden die Möglichkeit haben, die digitalen Netze für die unterschiedlichsten Arten von Anwendungen zu nutzen. Beispiele sind die Telearbeit und die elektronische Abwicklung von Geschäften (ECommerce). Offen bleibt jedoch die Frage, ob und unter welchen Umständen die Nutzer diese Dienste in Anspruch nehmen werden.

Eine entscheidende Voraussetzung für die Akzeptanz der neuen Dienste ist, daß sie eine mit ihren konventionellen Pendanten vergleichbare Mindestfunktionalität erbringen müssen. Unter dieser Voraussetzung ist die Vertraulichkeit (Privacy) eine wesentliche Anforderung. Es muß bei Bedarf garantiert sein, daß Informationen unbeobachtbar und/oder anonym gesendet

* Die Arbeit von D. Kesdogan wurde von der Gottlieb Daimler- und Karl Benz-Stiftung gefördert.

werden können. Dies gilt im besonderen Maße für offene Systemumgebungen wie das Internet. [Pfit90] konkretisiert dies als technische Schutzanforderung wie folgt:

Sender und/oder Empfänger von Nachrichten sollen voreinander anonym bleiben können, und Unbeteiligte (inkl. Netzbetreiber) sollen nicht in der Lage sein, sie zu beobachten.

Zur Erfüllung dieser technischen Schutzanforderung werden in der aktuellen Literatur verschiedene Techniken diskutiert. In dieser Arbeit werden Klassifizierungsgrößen für Techniken zur Sicherstellung der Anonymität und Unbeobachtbarkeit (im folgenden kurz Anonymisierungstechniken genannt) vorgeschlagen, die den Vergleich und die Einordnung der Techniken im Hinblick auf ihre Sicherheit und ihre Leistungsfähigkeit ermöglichen. Grundlage hierzu ist die Erweiterung der bisherigen Modellwelt auf diesem Gebiet um die Klassen des probabilistischen und praktischen Schutzes. Die konkreten Verfahren werden anschließend anhand der definierten Klassifizierungsgrößen bewertet und damit verbunden ein aktueller Überblick über das Gebiet der Anonymisierung und Unbeobachtbarkeit gegeben. Die Arbeit schließt mit einer Zusammenfassung und einem Ausblick auf zukünftige Arbeiten auf diesem Gebiet.

2 Klassifizierung von Anonymisierungstechniken

Die in dieser Arbeit vorgeschlagene Klassifizierung der Anonymisierungstechniken lehnt sich an die in der Kryptographie verwendeten Modellwelten an. In der Kryptographie wird die Sicherheit in drei Modellwelten untersucht und bewiesen. Dies sind die informationstheoretische, die komplexitätstheoretische und die praktische Modellwelt. Die Kryptographie verwendet bei der Erweiterung des Modells der informationstheoretischen zur komplexitätstheoretischen Sicherheit die Komplexitätstheorie (insbesondere die Zahlentheorie). Im Bereich der Anonymität und Unbeobachtbarkeit ist bisher nur die informationstheoretische Modellwelt bekannt [Pfit90][Pfit98]. Diese läßt sich analog zu den kryptographischen Modellen erweitern, was insbesondere zur Einführung der probabilistischen Modellwelt auf der Basis der Warteschlangentheorie [KEB98] führt.

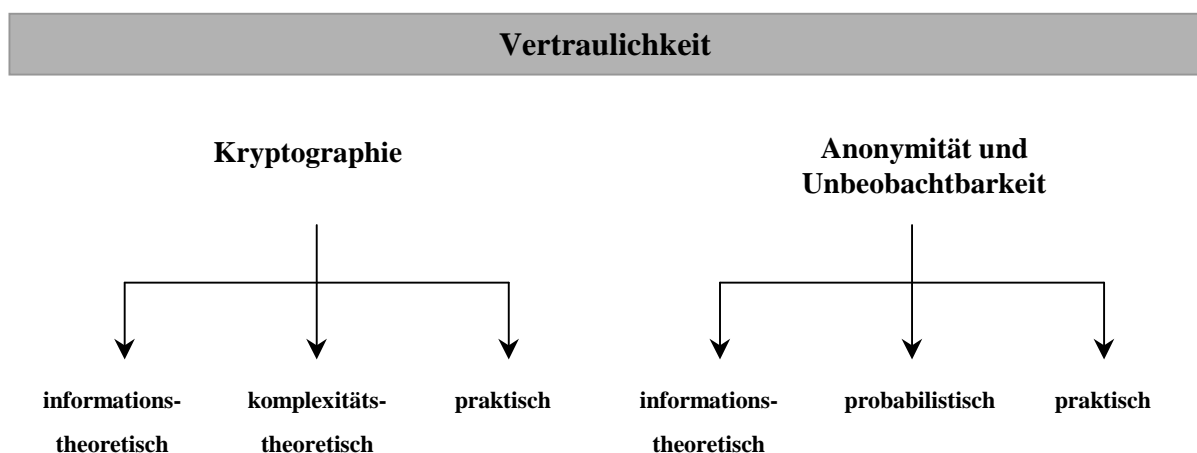


Abb. 1: Modellwelten in der Vertraulichkeit

Darüberhinaus existiert auch hier die Klasse der praktisch sicheren Techniken. Insgesamt ergibt sich das in Abb.1 dargestellte Bild.

In den nachfolgenden Abschnitten wird die informationstheoretische Modellwelt gemäß dem obigen Schema erweitert und damit das erste Klassifizierungsmerkmal eingeführt, nämlich die Einordnung einer Anonymisierungstechnik in die entsprechende Sicherheits- bzw. Schutzklasse.

2.1 Schutzklassen

Beim Schutz sogenannter Verkehrsdaten (z.B. Sender- und/oder Empfängeridentität, benutzer Dienst etc.) steht die Tarnung prinzipiell beobachtbarer Aktionen im Vordergrund. Verfahren, die diese Tarnung innerhalb des Netzes gewährleisten, existieren in großer Anzahl. Sie basieren in der Regel jedoch auf einem der folgenden drei *Grundverfahren*:

1. Implizite Adressierung und Verteilung [FL75][Karg77],
2. MIX [Chau81], und
3. DC-Netz [Chau88].

Das erste Verfahren geht auf Farber, Larson und Karger zurück und die weiteren auf D. Chaum. Die Sicherheit dieser Verfahren wurde von A. Pfitzmann untersucht und entsprechend erweitert [Pfit90]. Weiterhin wurden diese Verfahren von A. Pfitzmann in der gleichen Arbeit in eine informationstheoretische Modellwelt eingebettet.

In der neueren Literatur werden weitere Verfahren vorgeschlagen [FKK96][GT96][RR97][SGR97][SM96], die beweisbar nicht in der in [Pfit90] vorgeschlagenen Modellwelt liegen. Der von diesen Techniken geleistete Schutz kann somit nicht im Rahmen der informationstheoretischen Modellwelt gemessen werden, obwohl ein gewisses Maß an Schutz gewährleistet wird. Daher werden in dieser Arbeit weitere Modellwelten eingeführt, die nicht den perfekten Schutz als Ziel haben. Als Leitbild dienen die in der Kryptographie gebräuchlichen Sicherheitsmodelle. In Analogie zu diesen Modellen bieten sich hinsichtlich des Schutzes der Kommunikationsbeziehung die folgenden drei Schutzmodelle an:

1. Perfekter Schutz,
2. Probabilistischer Schutz, und
3. Praktischer Schutz.

Diese Schutzmodelle werden im folgenden definiert. Die in der Literatur diskutierten Verfahren lassen sich damit gemäß Abb.2 gliedern. Die in Abb.2 fett umrandeten Bereiche kennzeichnen neue Erweiterungen der bisherigen Theorie.

Das MIX-Verfahren wird in [Chau81], Stop-and-Go-MIXe in [KEB98], Onion Routing in [GRS96][SGR97], NDM in [FKK96], Babel-MIXe in [GT96], Remailer in [Cott][Scha96][SM96] und Crowds in [RR97] vorgestellt. Auf die einzelnen Verfahren wird zum

Teil später noch genauer eingegangen. Wie Abb.2 deutlich macht, gehen die verschiedenen Modellwelten von unterschiedlichen Annahmen bzgl. der Fähigkeiten eines Angreifers aus.

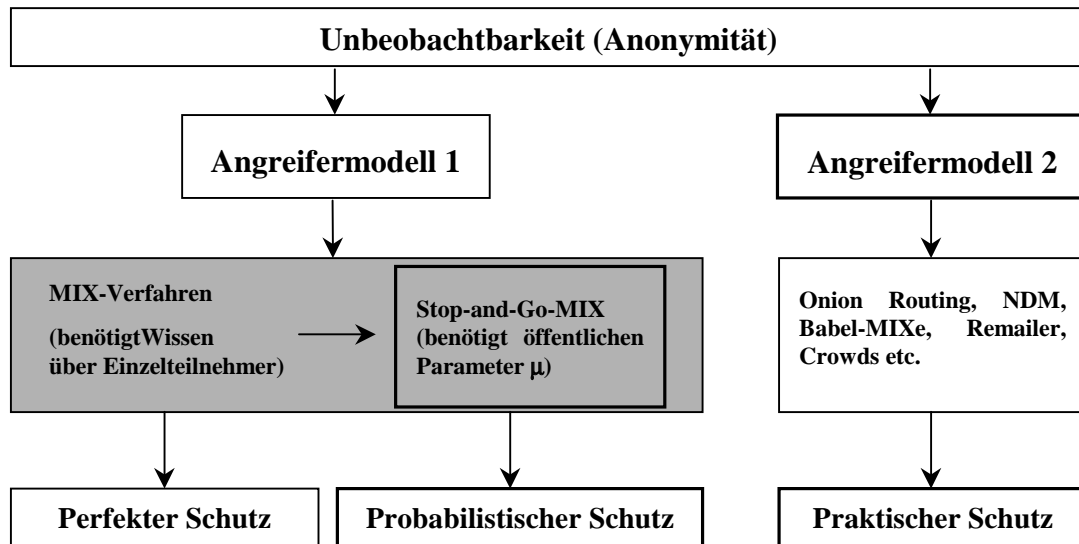


Abb. 2: Erweiterung der Modellwelt

Nachfolgend wird daher zunächst auf die verschiedenen Angreifermodelle eingegangen.

2.2 Angreifermodelle

Grundlage jeder Sicherheitsbetrachtung ist die genaue Festlegung der Fähigkeiten des Angreifers, gegen den man sich schützen möchte. Die dieser Arbeit zugrundeliegenden Angreifermodelle unterscheiden sich bzgl. der Mächtigkeit des jeweiligen Angreifers. Sie werden wie folgt definiert:

Angreifermodell 1: Omnipräsenter Angreifer

Der omnipräsente Angreifer besitzt die folgenden Fähigkeiten:

- eingesetztes Verfahren ist dem Angreifer genau bekannt,
- kann alle Wege (Leitungen) der gesendeten Nachricht gleichzeitig abhören,
- kontrolliert alle benutzten Vermittlungsrechner und kann deshalb beliebig viele Nachrichten in das System selbst einspielen, löschen und modifizieren,
- hat unbegrenzten Speicherplatz und Kommunikationskapazität sowie beliebig kurze Reaktionszeiten, und
- kann **nicht** die benutzten Verschlüsselungsverfahren brechen.

Die Definition des omnipräsenten Angreifers erfolgt unabhängig von der tatsächlichen Netzstruktur, d.h. unabhängig von Größe und räumlicher Ausbreitung des Netzes. Für offene Kommunikationsumgebungen wie das Internet ist dies sicherlich keine realistische Annahme. Daher bietet sich als realistischeres Modell für solche Umgebungen eine netzabhängige Definition des Angreifermodells an:

Angreifermodell 2: Teilweise präsenter Angreifer

Der teilweise präsente Angreifer besitzt die folgenden Fähigkeiten:

- eingesetztes Verfahren ist dem Angreifer genau bekannt,
- kann nur teilweise Wege (Leitungen) der gesendeten Nachricht gleichzeitig abhören,
- kontrolliert Teile der benutzten Vermittlungsrechner und kann deshalb eingeschränkt Nachrichten in das System selber einspielen, löschen und modifizieren,
- hat unbegrenzten Speicherplatz und Kommunikationskapazität sowie beliebig kurze Reaktionszeiten, und
- kann **nicht** die benutzten Verschlüsselungsverfahren brechen.

Die Angreifermodelle selber werden nicht als Klassifizierungsmerkmal herangezogen, da sie durch das jeweiligen Schutzmodell impliziert werden.

3 Perfekte, probabilistische und praktische Unbeobachtbarkeit

In diesem Abschnitt werden die Beziehungen der drei vorgeschlagenen Schutzmodelle zueinander (siehe Abb.2) und die Schutzmodelle selbst vorgestellt.

Anonymisierungstechniken haben zum Ziel, die folgenden personenbezogenen Informationen zu verbergen:

- Die Kommunikation soll gegenüber Unbeteiligten weitgehend *unbeobachtbar* sein.
- Gegenüber Beteiligten soll die Kommunikation im allgemeinen *anonym* erfolgen.
- Verkehrereignisse sollen *unverkettbar* sein.

Diese Begriffe werden in [Pfit93] und [Pfit98] präzisiert. Die Unbeobachtbarkeit wird dabei wie folgt definiert:

Def. 1: Unbeobachtbar, perfekt unbeobachtbar

Ein Ereignis E heißt **unbeobachtbar** bezüglich eines Angreifers A , wenn die Wahrscheinlichkeit des Auftretens von E nach jeder für A möglichen Beobachtung B sowohl echt größer 0 als auch echt kleiner 1 ist. Für A gilt für alle B : $0 < P(E|B) < 1$.

Das Ereignis E heißt **perfekt unbeobachtbar** bezüglich eines Angreifers A , wenn die Wahrscheinlichkeit des Auftretens von E vor und nach jeder für A möglichen Beobachtung B gleich ist, d.h. für alle B : $P(E) = P(E|B)$.

Gemäß Shannon [Shan49][Hell77] bedeutet *perfekter Schutz*, daß alle möglichen Beobachtungen durch den Angreifer diesem keinen Informationsgewinn bringen. Es handelt sich dabei also um eine Maximalforderung. Zur Realisierung dieser Maximalforderung ist eine Umgestaltung der heute existierenden Netze notwendig [PPW88]. Darüber hinaus wird insbesondere authentisches Wissen bzgl. der Identität der Teilnehmer benötigt, da immer

mehrere Teilnehmer zusammenarbeiten müssen, um Anonymität und Unbeobachtbarkeit zu gewährleisten. Diese Anforderungen lassen sich in heutigen, offenen Kommunikationsumgebungen wie dem Internet nur schwer erfüllen.

Sollen die Schutzanforderungen jedoch gemäß der heutigen Gestalt der Netze definiert werden, um diese ohne umfangreiche und kostenintensive Modifikationen und ohne die Notwendigkeit des Teilnehmerwissens weiterbenutzen zu können, so führt dies zu einem schwächeren Sicherheitsmodell, dem Modell des *probabilistischen Schutzes*:

Def. 2: Probabilistisch unbeobachtbar

Ein Ereignis E heißt probabilistisch **unbeobachtbar** bezüglich eines Angreifers A , wenn die Wahrscheinlichkeit des Auftretens von E nach jeder für A möglichen Beobachtung B echt größer 0 ist, mit Wahrscheinlichkeit $1-\alpha$ auch echt kleiner 1 ist und wenn folgende Bedingungen bezüglich α erfüllt sind:

- Die a-posteriori-Wahrscheinlichkeit α ist gleich der a-priori-Wahrscheinlichkeit (d.h. unabhängig vom Angreifer).
- Es existiert ein Systemparameter μ , dessen lineare Änderung α exponentiell gegen 0 konvergieren läßt.
- Für A gilt für alle B : $0 < P(E/B) < 1$ mit Wahrscheinlichkeit $1-\alpha$ und $P(E/B) = 1$ mit Wahrscheinlichkeit α .

Def. 2 erfüllt offensichtlich die Anforderungen einer komplexitätstheoretischen Sicherheit. Wenn ein Algorithmus bekannt ist, der probabilistische Sicherheit gewährleistet (z.B. mit $\alpha=10^{-20}$), dann besteht ein erfolgreicher Angriff darin, diesen bis zum Erfolg immer zu wiederholen. Da die Unsicherheit α unabhängig von der Anzahl der Versuche immer konstant bleibt, enden die Versuche im Erwartungswert erst nach $5 \cdot 10^{19}$ Schritten. Können diese Versuche in polynomieller Zeit verwirklicht werden, dann gehört der erfolgreiche Angriff sicherlich in die Klasse NP.

Die Modellwelt des praktischen Schutzes ergibt sich, indem man vom Angreifermodell 1 zum Angreifermodell 2 wechselt, d.h. die Annahme eines omnipräsenten Angreifers fallen läßt und durch das realistischere Modell eines teilweise präsenten Angreifers ersetzt. Zu beachten ist hierbei, daß die tatsächlichen Fähigkeiten des teilweise präsenten Angreifers in ihrer konkreten Ausprägung von Fall zu Fall schwanken können. Dies bedeutet aber auch, daß dieser Klasse zuzuordnende Anonymisierungstechniken eventuell gegen die eine konkrete Ausprägung des Angreifermodells Schutz bieten, gegen eine andere konkrete Ausprägung jedoch nicht. Damit ist für die Anonymisierungstechniken dieser Klasse keine Ordnung definiert. Vielmehr ergibt sich eine entsprechende Halbordnung.

Dieses indeterministische Angreifermodell bedarf auch einer entsprechenden indeterministischen Schutzanforderung, die zur Definition *des praktischen Schutzes* führt:

Def. 3: Praktisch unbeobachtbar

Ein Ereignis E heißt praktisch **unbeobachtbar** bezüglich des teilweise präsenten Angreifers A , wenn die Wahrscheinlichkeit des Auftretens von E nach jeder für A

möglichen Beobachtung B echt größer 0 ist und mit Wahrscheinlichkeit $1-\alpha$ auch echt kleiner 1 ist. Für A gilt für alle B : $0 < P(E/B) \leq 1$.

Damit sind die drei wesentlichen Schutzklassen definiert und die formale Einordnung existierender Techniken im Hinblick auf ihre Sicherheit ist möglich. Im nachfolgenden Abschnitt werden zusätzlich Größen zur Beschreibung der Leistungsfähigkeit eingeführt.

4 Wirkungsgrad und Zeiteffizienz

Das Erreichen einer Anonymisierung der Kommunikationsteilnehmer und der Schutz ihrer Kommunikationsbeziehung ist in der Regel mit dem Versenden von echten Nachrichten und Scheinnachrichten verbunden. Generell ist es für Anonymisierungstechniken daher von Bedeutung, trotz des jeweiligen Angreifermodells eine hohe Nachrichtenübertragungsrate zu gewährleisten. In der Regel müssen zum Erreichen des jeweiligen Schutzziels $n > 1$ Pakete von n verschiedenen Teilnehmern erzeugt werden. Befinden sich unter diesen n Paketen k echte Nachrichtenpakete und m Scheinnachrichten ($m := n - k$), so soll das Verhältnis $\eta := k/n$ als Wirkungsgrad¹ bezeichnet werden. η kann nie größer als 1 ($\eta \leq 1$) sein, da die Anzahl der ausgetauschten „echten“ Nachrichtenpakete nie größer sein kann als die Gesamtzahl der ausgetauschten Nachrichtenpakete. Wird eine echte Nachricht zum Erreichen der Anonymisierung wiederholt gesendet (z.B. über Zwischenknoten), so werden diese zusätzlichen Übertragungen ebenfalls wie Scheinnachrichten behandelt.

Offensichtlich ist es erstrebenswert, ein Anonymisierungsverfahren mit hohem Wirkungsgrad ($\eta \approx 1$) zu entwerfen. Angenommen, es gäbe solch ein Verfahren. Das würde bedeuten, daß bei der Anwendung des Verfahrens eine Anonymitätsmenge existiert, in der alle Teilnehmer etwas zu versenden haben. Da jedoch nicht immer alle n Teilnehmer etwas zu versenden haben, muß das Verfahren entsprechend lange warten, d.h. das Zeitintervall muß so groß gewählt werden, daß die n Teilnehmer mit hoher Wahrscheinlichkeit etwas zu versenden haben. Dieser Zeitaufwand kann als Zeiteffizienz τ gemessen werden. Weiteren Einfluß auf die Zeiteffizienz hat die Umleitung von Nachrichten über zusätzliche Knoten, da hier ebenfalls Zeit eingeübt wird.

Da die verschiedenen Verfahren unterschiedliche Techniken einsetzen, muß zur Berechnung der beiden Kenngrößen auf ein vereinfachtes, einheitliches Modell zurückgegriffen werden. Dieses Modell geht davon aus, daß jeder Teilnehmer direkt und zeitgleich mit jedem anderen Teilnehmer kommunizieren kann (vollvermaschtes Netz). Die konkrete Netzwerktopologie, die gegebenenfalls eine Optimierung ermöglicht, bleibt dabei unberücksichtigt. Das unmittelbare Senden der Nachricht vom Sender zum Empfänger wird damit als Normfall betrachtet. Die beiden Kenngrößen definieren demzufolge den Mehraufwand bzgl. der Bandbreite sowie der Zeit und werden wie folgt berechnet:

¹ In der Physik wird das Verhältnis der abgegebenen Leistung an einer Maschine zur zugeführten Leistung als Wirkungsgrad η definiert. Dies ist sinnvoll, da jede Maschine eine größere Leistung aufnimmt, als sie abgibt. Die Analogie zur Informationsrate bei den Anonymisierungstechniken ist offensichtlich.

$$\eta = \frac{\text{Anzahl richtige Nachrichten}}{\text{Gesamtanzahl Nachrichten}}, \quad \tau = \frac{\text{Zeit zum direkten Senden einer Nachricht}}{\text{Zeit zum Senden mittels konkretem Verfahren}}.$$

Mit dem Wirkungsgrad und der Zeiteffizienz werden zwei Größen definiert, die eine Klassifizierung der verschiedenen Anonymisierungstechniken gemäß ihrer Leistungsfähigkeit ermöglichen. Zusammen mit den drei definierten Schutzmodellen ergibt sich ein Klassifizierungsschema, das eine Ordnung und einen Vergleich der existierenden Techniken ermöglicht und damit das Arbeitsgebiet der Anonymisierungstechniken weiter strukturiert.

5 Anonymisierungstechniken und ihre Klassifizierung

Im folgenden werden exemplarisch verschiedene existierende Anonymisierungstechniken gemäß dem oben definierten Schema klassifiziert und bewertet. Der Wirkungsgrad und die Zeiteffizienz werden dabei lediglich angegeben. Auf eine detaillierte Herleitung wird verzichtet. Sie beruht aber auf den jeweils angegebenen Arbeiten und kann mittels dieser einfach nachvollzogen werden.

5.1 Perfekter Schutz

Ziel der Verfahren dieser Klasse ist es, die Information darüber, wer wann von wo und wie lange mit wem kommuniziert hat, perfekt zu schützen.

1. *Verteilung*: Bei der Verteilung [FL75][Karg77][Pfit90] sendet der Sender eine Nachricht nicht nur an den eigentlichen Empfänger, sondern an eine Gruppe von Empfängern (Anonymitätsmenge). Die Adressierung des eigentlichen Empfängers geschieht durch eine implizite Adresse.
2. *DC-Netze*: Ein DC-Netz [Chau88] sendet für jedes Nutzbit auf dem unsicheren Netz n Schlüsselbits. Diese werden mit der XOR-Funktion summiert, so daß das Nutzbit perfekt in die n Schlüsselbits der n verschiedenen Teilnehmer eingebettet ist. Für dieses überlagernde Senden benötigen die beteiligten Stationen paarweise gemeinsame geheime Schlüssel, die so lang sind, wie die zu versendende Nachricht, und nur einmal verwendet werden dürfen. Zunächst werden in jeder Station lokal die vorhandenen Schlüssel und evtl. die zu sendende Nachricht überlagert (XOR-Funktion). Danach werden global die lokalen Ergebnisse überlagert. Wenn nun genau eine Station eine Nachricht x sendet, entspricht die verteilte Gesamtsumme dieser Nachricht x , da ja alle Schlüsselwörter genau zweimal addiert werden. Falls von den n Teilnehmern mehrere in der gleichen Periode eine Nachricht senden möchten, kommt es zu einer erkennbaren Kollision, die entsprechend aufgelöst werden kann².
3. *MIXe*: Das MIX-Verfahren hat das Ziel, die Nachrichtenvermittlung vom Sender bis zum Empfänger durch den Einsatz von Zwischenknoten, sogenannte MIX-Stationen, unverfolgbar zu machen. Dazu sammelt eine MIX-Station genügend viele Datenpakete von ge-

² In der Analyse wird dazu Slotted ALOHA benutzt. Der Gesamtverkehr des Netzes ist dabei ein Poissonstrom mit Parameter G und die mittlere Anzahl von Sendeversuchen ist gleich e^G .

nügend vielen Absendern und gibt sie so verändert wieder aus, daß ein Außenstehender ein Eingangspaket nicht zu einem Ausgangspaket in Bezug setzen kann. In der hier betrachteten Variante sammelt der MIX dazu n Datenpakete und gibt diese dann in einem Schub aus. Dies kann entweder synchron oder asynchron geschehen. Im synchronen Fall sendet jeder Teilnehmer in jedem Taktzyklus eine (Schein-)Nachricht. Im asynchronen Fall ist das System nicht getaktet. Ein Loop-Back garantiert dabei die sichere Bildung der Anonymitätsmenge. Für weitere Protokolldetails sei auf [Chau81] und [Pfit98] verwiesen.

	Direktes Senden	Verteilung	DC-Netz	MIX-Kaskade (synchron)	MIX (asynchron)
Angreifermodell	-	1	1	1	1
Sender-Anonymisierung	Nein	Nein	Ja	Ja	Ja
Empfänger-Anonymisierung	Nein	Ja	Ja	Ja	Ja
Wirkungsgrad η	1	$\frac{1}{n}$	$\frac{1}{n \cdot (n-1) \cdot e^G}$	$\frac{k}{n \cdot (N+1) + N \cdot (2 \cdot n)}$	$\frac{1}{(N+1) + 2 \cdot N}$
Zeiteffizienz τ	1	1	$\frac{1}{e^G}$	$\frac{1}{(N+1) + 2 \cdot N}$	$\frac{1}{(N+1) + N\mu_s \frac{n-1}{2\lambda} + 2 \cdot N}$
Spontane Kommunikation	-	Nein	Nein	Nein	Ja

Tabelle 1: Perfekter Schutz

Tabelle 1 ordnet die Verfahren gemäß dem oben vorgestellten Klassifizierungsschema. Zusätzlich wird noch angegeben, ob der zugrundeliegende Gruppenbildungsmechanismus eine spontane Kommunikation zuläßt. n bezeichnet die Zahl der Empfänger (Verteilung) bzw. die Anzahl der Knoten im DC-Netz. Bei den MIX-Verfahren wird davon ausgegangen, daß N MIXe benutzt werden und jeder MIX mit der Rate μ_s senden kann und mit der Rate λ empfängt. m bezeichnet dabei die Anzahl der von einem Knoten generierten Scheinnachrichten. Als Referenz wird zusätzlich das direkte, nicht anonyme Senden angegeben.

5.2 Probabilistischer Schutz

Der Klasse der probabilistischen Schutz bietenden Verfahren ist zur Zeit nur ein einziges Verfahren zuzuordnen, die Stop-and-Go-MIXe [KEB98]. Ein SG-MIX arbeitet grundsätzlich wie ein konventioneller MIX, sammelt jedoch keine feste Anzahl von Nachrichten. Ein Sender A wählt zum anonymen Versenden einer Nachricht n verschiedene SG-MIXe aus. Für jeden dieser Knoten i berechnet er ein Zeitfenster $(TS^{\min}, TS^{\max})_i$ und eine zufällige Verzögerungszeit T_i gemäß einer Exponentialverteilung mit geeignetem Parameter μ_w . Diese Information wird dem Paket hinzugefügt, bevor es mit dem öffentlichen Schlüssel des Knotens verschlüsselt wird. Der SG-MIX i entschlüsselt das Paket und entnimmt das Zeitfenster (TS^{\min}, TS^{\max}) . Sollte der Ankunftszeitpunkt des Pakets nicht innerhalb die-

ses Zeitfensters liegen, so verwirft er das Paket. Ansonsten leitet er das Paket nach T_1 Zeiteinheiten an den nächsten SG-MIX bzw. den Empfänger weiter.

SG-MIX	
Angreifermodell	1
Sender-Anonymisierung	ja
Empfänger-Anonymisierung	ja
Wirkungsgrad η	$\frac{1}{N+1}$
Zeiteffizienz τ	$\frac{1}{(N+1) + N\mu_s \frac{1}{\mu_w}}$
Spontane Kommunikation	ja

Tabelle 2: Probabilistischer Schutz

Tabelle 2 ordnet das Verfahren gemäß dem oben vorgestellten Klassifizierungsschema. Es wird davon ausgegangen, daß N SG-MIXe benutzt werden und jeder SG-MIX mit der Rate μ_s sendet.

5.3 Praktischer Schutz

Zu der Klasse der praktischen Schutz bietenden Verfahren gehören die meisten in der jüngeren Literatur vorgeschlagenen Verfahren.

1. *NDM*: Bei der NDM-Methode [FKK96] wählt der Sender unabhängig und zufällig N spezielle Zwischenknoten (Security Agents, SA) aus, über die seine Nachricht geleitet werden soll. Er verschlüsselt das zu versendende Paket mit den entsprechenden öffentlichen Schlüsseln der SAs. Weiterhin lassen sich hier die von den MIXen bekannten Techniken zur Replay-Erkennung, zur indeterministischen Verschlüsselung und zur einheitlichen Nachrichtenlänge anwenden.
2. *Crowds*: Crowds [RR97] ermöglicht es einer Gruppe von Nutzern anonym Web-Seiten abzurufen. Dazu wird die Anfrage eines Nutzers vom aktuellen Zwischenknoten entweder an den Web-Server (mit Wahrscheinlichkeit $(1 - p_f)$) oder aber an einen weiteren Zwischenknoten (mit Wahrscheinlichkeit p_f) weitergeleitet.

Insbesondere in der Klasse des praktischen Schutzes gibt es noch weitere Verfahren, die hier nicht vorgestellt wurden. Diese können analog bewertet und eingeordnet werden.

	NDM	Crowds
Angreifermodell	2	2
Sender-Anonymisierung	ja	ja
Empfänger-Anonymisierung	ja	ja
Wirkungsgrad η	$\frac{1}{N+1}$	$\frac{1-p_f}{2-p_f}$
Zeiteffizienz τ	$\frac{1}{N+1}$	$\frac{1-p_f}{2-p_f}$
Spontane Kommunikation	ja	nein ³

Tabelle 3: Praktischer Schutz

6 Ausblick

In dieser Arbeit wurde ein Klassifizierungsschema vorgestellt, das die Beurteilung und den Vergleich von Anonymisierungstechniken hinsichtlich ihrer Sicherheit und Leistungsfähigkeit ermöglicht. Die Sicherheitsklassifikation beruht auf der Erweiterung der bisherigen informationstheoretischen Modellwelt um die probabilistische und die praktische Modellwelt. Zur Beurteilung der Leistungsfähigkeit wurden die beiden Parameter Wirkungsgrad und Zeiteffizienz eingeführt.

In zukünftigen Arbeiten soll das Klassifikationsschema zu einer allgemeinen Theorie der Anonymisierungstechniken erweitert werden, die insbesondere die gemeinsamen Grundelemente der Verfahren (Bildung der Anonymitätsmenge, Umkodierung, Adreßumsetzung etc.) formal und systemunabhängig beschreibt.

Literatur

- [Chau81] Chaum, D.: Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms. In: Comm. ACM, Vol. 24, No. 2, Februar 1981, S. 84-88.
- [Chau88] Chaum, D.: The Dining Cryptographers Problem: Unconditional Sender and Recipient Untraceability. In: Journal of Cryptology 1/1, 1988, S. 65-75.
- [Cott] Cottrell, L.: Mixmaster & Remailer Attacks (<http://obscura.com/~loki/remailer-essay.html>).
- [FP97] Federrath, H., Pfitzmann, A.: Bausteine mehrseitiger Sicherheit. In: Mehrseitige Sicherheit in der Kommunikationstechnik, Addison-Wesley, Bonn, 1997, S. 83-104.

³ Das Gruppenmanagement macht eine explizite Anmeldung erforderlich.

- [GT96] Gülçü, C., Tsudik, G.: Mixing Email with Babel. In: Proc. Symposium on Network and Distributed System Security, San Diego, IEEE Computer Society Press, 1996.
- [GRS96] Goldschlag, D.M., Reed, M.G., Syverson, P.F.: Hiding Routing Information. In: Information Hiding, Springer-Verlag LNCS 1174, 1996, S. 137-150.
- [Hell77] Hellman, M.E.: An extension of the Shannon Theory Approach to Cryptography. In: IEEE Transactions on Information Theory, Band IT-23, No. 3, Mai 1977.
- [FKK96] Fasbender, A., Kesdogan, D., Kubitz, O.: Variable and Scalable Security: Protection of Location Information in Mobile IP. In: Proc. VTC'96, Atlanta, 1996.
- [FL75] Farber, D.J., Larson, K.C.: Network Security Via Dynamic Process Renaming. In: Fourth Data Communications Symposium, Quebec City, Canada, 1975.
- [Karg77] Karger, P.A.: Non-Discretionary Access Control for Decentralized Computing Systems. Master Thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, Report MIT/LCS/TR-179, 1977.
- [KEB98] Kesdogan, D., Egner, J., Büschkes, R.: Stop-And-Go-MIXes Providing Probabilistic Anonymity in an Open System. Erscheint in: Proc. Second Workshop on Information Hiding (IHW98), LNCS (Springer-Verlag).
- [PPW88] Pfitzmann, A., Pfitzmann, B., Waidner, M.: Datenschutz garantierende offene Kommunikationsnetze. In: Informatik-Spektrum 11/3, 1988, S. 118-142.
- [Pfit90] Pfitzmann, A.: Dienstintegrierende Kommunikationsnetze mit teilnehmerüberprüfbarem Datenschutz. In: IFB 234, Springer-Verlag, Heidelberg, 1990.
- [Pfit93] Pfitzmann, A.: Technischer Datenschutz in öffentlichen Funknetzen. In: Datenschutz und Datensicherheit (DuD), August 1993, S. 451-463.
- [Pfit98] Pfitzmann, A.: Datensicherheit und Kryptographie. Vorlesungsskript TU Dresden, WS 1997/98.
- [RR97] Reiter, M.K., Rubin, A.D.: Crowds: Anonymity for Web Transactions. DIMACS Technical Report 97-15, <http://www.research.att.com/projects/crowds/>, April 1997.
- [Scha96] Schaarschmidt, S.: Anonyme Remailer im Internet. Großer Beleg am Institut für Theoretische Informatik der TU Dresden, Januar 1996.
- [Shan49] Shannon, C.E.: Communication theory of secrecy systems. In: Bell System Technical Journal, 28, 1949, S. 656-715.
- [SGR97] Syverson, P.F., Goldschlag, P.F., Reed, M. G.: Anonymous Connections and Onion Routing. In: Proc. 1997 IEEE Symposium on Security and Privacy, Mai 1997.
- [SM96] Strassmann, P. A., Marlow, W.: Risk-Free Access Into The Global Information Infrastructure Via Anonymous Re-Mailers. In: American Programmer, Vol. 9, Issue 5, 1996, <http://www.strassmann.com/pubs/anon-remail.html>.