Sicherheitsbewertung von Anonymisierungsverfahren im World Wide Web

Dogan Kesdogan (kesdogan@informatik.rwth-aachen.de)
Oliver Rattay (oliver.rattay@gmx.net)

RWTH Aachen, Lehrstuhl für Informatik 4
Ahornstr. 55, D-52074 Aachen

Abstract:
Anonymität im Internet ist nicht gegeben, da Verkehrsdaten (Quell- und Zieladressen) offen vorliegen. Ein Ansatz zur Gewährleistung von Anonymität sind die bekannten Mixe. Fast alle Vorschläge der Mixe für das Internet verzichten auf den Einsatz von Dummy-Nachrichten. Durch den Verzicht von Dummy-Nachrichten ist es verschiedenen Angriffsmethoden möglich, den Schutz zu kompromittieren. Der sogenannte *Ausschlussangriff* und zwei weitere Variationen vom Ausschlussangriff wurden bisher theoretisch untersucht. In dieser Arbeit werden sie erstmals mittels realer Daten bewertet und die Ergebnisse vorgestellt.

1 Einleitung

Heutige Netze schützen die Vermittlungsdaten (Adressen und Absender der Kommunikationspartner, momentaner Ort, etc.) nicht oder nur unzureichend. Aufgrund dieser sogenannten Verkehrsdaten können Profile der Kommunikationsteilnehmer erstellt werden (zum Beispiel wann ein Teilnehmer wie oft von wo mit wem kommuniziert)[1]. Zum Schutz der Verkehrsdaten sind sogenannte Anonymisierungstechniken bekannt und haben das Ziel die Erstellung solche Persönlichkeitsbilder zu verhindern.

Aber wie gut können Anonymisierungsverfahren die Erstellung solcher Persönlichkeitsbilder verhindern? Diese Frage ist insbesondere wesentlich, da praktische Implementierungen (wie zum Beispiel das Mix-Netz [Cha81]) nicht den theoretischen Modellen folgen. Diese „praktischen“ Varianten verzichten auf die Nutzung von sogenannten Dummy-Nachrichten.

Zur Bewertung der Anonymität von „praktischen“ Mixen wurde deshalb unter anderem der sogenannte Ausschlussangriff vorgeschlagen und entwickelt [KAP02]. Mit ihm ist es möglich die Stärke einer Anonymität im Internet zu quantifizieren. In den bisherigen Arbeiten wurden die Angriffe jedoch lediglich simuliert beziehungsweise analytisch be-

[1] Diese Feststellung beruht auf der Annahme, dass die Person eine wiederkehrendes Kommunikationsverhalten aufweist.

trachtet (vgl. [KAP02], [AKP03], [Dan03], [MD04] und [DS04]). In [Rat04] wurden die Angriffe erstmals unter realen Bedingungen untersucht.

Im nächsten Abschnitt werden die Anonymisierungsverfahren, die auf Mix-Technik basieren, vorgestellt. Der Grund dafür ist, dass im Zusammenhang mit Anonymität im Internet Mixe die größte Rolle spielen (siehe z.B. [JAP04], [FH04]). In Abschnitt 3 werden die aus der Literatur bekannten Angriffsalgorithmen vorgestellt. Das Experimentierumfeld und das verwendete Angriffswerkzeug wird in Abschnitt 4 dargestellt. In Abschnitt 5 werden dann einige Ergebnisse der Untersuchungen präsentiert und die Eignung der einzelnen Angriffsalgorithmen auf reale Daten bewertet.

2 Anonymität im Internet

Eine Möglichkeit die Verkehrsdaten (die Beziehung Sender und Empfänger) zu schützen ist, die Nachrichten nicht direkt, sondern über sog. Zwischenknoten (Mixe) zu schicken [Cha81]. Damit die Wege der Nachrichten weder anhand ihres äußeren Erscheinungsbildes (also ihre Länge und Codierung) noch anhand zeitlicher oder räumlicher Zusammenhänge verfolgt werden können, puffern die Mixe Nachrichten gleicher Länge von vielen Sendern, kodieren sie um und geben sie umsortiert aus. Das Umcodieren erfolgt durch Ver- und Entschlüsselung mittels eines Kryptosystems. Ein Mix muss darauf achten, dass er jede Nachricht nur einmal mixt, d.h. Nachrichtenwiederholungen vermeidet (siehe auch [PW85]).

Eine Methode, den Zeitpunkt des Sendens einer Nachricht zu verschleiern, ist Dummy-Traffic, d.h. das Senden bedeutungsloser Nachrichten, wenn keine bedeutungsvollen zu senden sind. Da in den zugrunde liegenden Untersuchungen kein Dummy-Traffic erzeugt worden ist, ist die Menge der Sender (die sogenannte Senderanonymitätsmenge) und die entsprechende Menge der Empfänger eines Schubes bekannt. Dies ist ein Informationsgewinn für den Angreifer und die Grundlage des Angriffs: Der Angreifer notiert sich bei jeder Kommunikation seines ausgewählten Opfers (hier z.B. Alice) die entsprechenden aufgerufenen Ziele. Die generelle Frage die in anderen Arbeiten bisher untersucht wurde ist, wie lange muss Alice über den Mix mit ihren Kommunikationspartner kommunizieren, bis der Informationsgewinn so groß ist, dass alle Kommunikationspartner von Alice bestimmt werden können.

Beispiel: Angreifer sieht zum Zeitpunkt t, dass Alice via Mix kommuniziert und schreibt sich die Menge der Ziele auf, hier $O_t = \{google.de, spiegel.de, gmx.net\}$. Der Angreifer weiss, dass Alice entweder mit google.de, spiegel.de oder mit gmx.net kommuniziert hat. Bei einmaliger Beobachtung kann der Angreifer nach dem Modell keine konkrete Aussage über das Ziel machen (da alle Ziele mit gleicher Wahrscheinlichkeit in Frage kommen). Im nächsten Abschnitt wird gezeigt, wie der Angreifer nach einer Anzahl von Beobachtungen doch noch Erfolg haben kann.

3 Angriffe

In diesem Abschnitt werden der Ausschlussangriff und zwei Variationen beziehungsweise Weiterentwicklungen vorgestellt. Da der Ausschlussangriff keine speziellen Anforderungen an die Implementierung des Anonymisierungsverfahrens stellt, wird der Mix, wie bereits erwähnt, im folgenden als Blackbox gesehen. Für den Angriff spielt es keine Rolle, ob das Anonymisierungsverfahren nur aus einem einzigen oder aus vielen hintereinander geschalteten Mix-Knoten besteht.

Der Angreifer erhält seine Angriffsinformationen, indem er die Ein- und Ausgangsleitungen des Mix-Systems beobachtet. Er erkennt hierdurch, welche Gruppe von Nutzern und welche Menge von Zielen im Zusammenhang stehen müssen. Der Angreifer kann sehen, ob der anzugreifende Nutzer an der beobachteten Kommunikation aktiv teilgenommen hat. Sollte dies nicht der Fall sein, trägt die Beobachtung nicht zur Aufhebung der Anonymität des Nutzers bei und kann somit verworfen werden.

3.1 Der Ausschlussangriff

Der Ausschlussangriff basiert auf dem Konzept der Schnittmengenangriffe und gehört somit zur Gruppe der kontextuellen Angriffe [Ray01]. Der Angriff besteht in seinem Grundgerüst aus zwei Phasen: der Lern- und der Ausschlussphase. In der Lernphase sucht der Algorithmus m disjunkte Beobachtungen. Nur so kann der Angreifer sicherstellen, dass für jedes Ziel auch eine dazugehörige Beobachtung vorliegt. Die in der Lernphase zusammengestellten Beobachtungen werden im folgenden auch als Basiselemente und die Menge aller als Basismenge bezeichnet. In der zweiten Phase werden mittels weiterer Beobachtungen die Elemente der Basismenge soweit verkleinert, bis die tatsächlichen Ziele des angegriffenen Opfers übrig bleiben.

Lernphase: In der Lernphase versucht der Angreifer für jedes Ziel seines Opfers eine passende Beobachtung zu erhalten. Der Ausschlussangriff geht davon aus, dass der Angreifer die Anzahl der von seinem Opfer kontaktierten Ziele m kennt[2]. Der Angreifer findet die Basiselemente dadurch, dass er Kommunikationen, an denen das ausgewählte Opfer aktiv teilgenommen hat, so lange beobachtet, bis ihm m paarweise verschiedene Beobachtungen $(O_1, \ldots, O_m)$ vorliegen, also $\forall i \neq j$ gilt $O_i \cap O_j = \varnothing$.

Ausschlussphase: Da es sich bei den Basiselementen lediglich um ausgewählte Beobachtungen handelt, kann jedes einzelne Element auch als Menge der Größe b von möglichen Zielen betrachtet werden. In der Ausschlussphase wird die Größe der einzelnen Elemente durch weitere Beobachtungen verkleinert. Der Algorithmus vergleicht in dieser Phase weitere Beobachtungen mit den Elementen der Basismenge. Die Basismenge kann immer dann verkleinert werden, wenn sich die neue Beobachtung O mit

[2] In [KAP02] wurde gezeigt, dass ein Angreifer mit Hilfe einer Statistik feststellen kann, ob er m korrekt abgeschätzt hat.

genau einem Basiselement O_i überschneidet, d.h. $O \cap O_i \neq \varnothing$ und $O \cap O_j = \varnothing$ für alle $j \neq i$. Diese Bedingung versichert dem Angreifer, dass O lediglich ein Ziel des angegriffenen Nutzers enthält. Das neue und kleinere Basiselement $O \cap O_i$ ersetzt das alte Element O_i und enthält immer noch das Ziel des Nutzers.

Die Ausschlussphase terminiert, sobald jedes der m Basiselemente $(O_1, \ldots, O_m)$ aus genau einem einzigen Ziel besteht. Die übrig gebliebenen Ziele sind die Ziele des angegriffenen Nutzers. Falls dem Angreifer nicht genügend Beobachtungen zur Verfügung stehen, werden unter Umständen nicht alle Ziele eindeutig identifiziert.

3.2 Der statistische Ausschlussangriff

In [Dan03] beschreibt George Danezis eine stark abgewandelte Variante des Ausschlussangriffs: den statistischen Ausschlussangriff. Im Gegensatz zum klassischen Ausschlussangriff liefert dieser Ansatz nicht immer korrekte Ergebnisse. Das Konzept lehnt sich an die Grundideen der Signalerkennung an. Der Angreifer versucht mit Hilfe der Beobachtungen eine Regelmäßigkeit beziehungsweise hier, die Ziele des Opfers zu erkennen.

Der statistische Ausschlussangriff, der in bisherigen Arbeiten lediglich simuliert (vgl. [MD04]) oder analytisch untersucht wurde (vgl. [DS04]), geht in seiner ursprünglichen Form davon aus, dass die Ziele vom Nutzer gleichverteilt kontaktiert werden. Bei längerer Beobachtung kann der Angreifer dann feststellen, dass die Ziele des Opfers öfters auftauchen[3]. Somit kann ein Angreifer die Häufigkeiten der einzelnen Ziele abzählen und erhält somit Kenntnisse über die von seinem angegriffenen Nutzer kontaktierten Ziele.

Ein Vorteil des Angriffs ist, dass keine Annahmen zur Zielmenge m benötigt werden. Das bedeutet, dass im Gegensatz zum klassischen Ausschlussangriff der Beobachtungszeitraum nicht zwangsläufig festgelegt werden muss. Ein Nachteil besteht darin, dass echte Nutzer ihre Ziele nicht gleichverteilt kontaktieren, so dass eine Normierung der Daten notwendig ist [Rat04].

3.3 Der kombinatorische Hitting-Set-Angriff

Der kombinatorische Hitting-Set-Angriff basiert, ebenso wie der Ausschlussangriff, auf der Tatsache, dass sich die Ziele des angegriffenen Nutzers mit jeder einzelnen durch den Angreifer gemachten Beobachtung überschneiden muss. Die kombinatorische Variante des Angriffs versucht durch Heuristiken den Suchraum aller möglichen Lösungen zu verkleinern und dadurch die korrekte Lösung schneller zu finden [KP04]. Dadurch ist es jedoch möglich, dass der Algorithmus die korrekte Lösung nicht findet. Als Heuristik wird ein Backtrackingverfahren verwendet.

Der Angriff setzt wie auch der klassische Ausschlussangriff voraus, dass die Anzahl der vom Opfer kontaktierten Ziele m bekannt beziehungsweise abgeschätzt werden kann.

[3] der Angreifer beobachtet nur Kommunikationen, an denen auch das Opfer teilgenommen hat

Während des Angriffs berechnet der Algorithmus in regelmäßigen Abständen die C wahrscheinlichsten Lösungskombinationen[4] $(L_1, \ldots, L_C)$ der Größe m, also $\forall\ 1 \leq i \leq C$ gilt $|L_i| = m$. Sollte sich lediglich ein Lösungskandidat L_x mit allen Beobachtungen $(O_1, \ldots, O_t)$ überschneiden, also $\forall\ 1 \leq i \leq t$ gilt $L_x \cap O_i \neq \varnothing$ und $\forall\ y \neq x$ gilt $L_y \cap O_j = \varnothing, \forall\ 1 \leq j \leq t$, ist es sehr wahrscheinlich, dass dieser die tatsächlichen Ziele des Opfers enthält.

4 Experimentierumfeld

Um die in Abschnitt 3 vorgestellten Angriffsalgorithmen unter realen Bedingungen zu testen und einen Vergleich der Ergebnisse zu den bisherigen simulierten Angriffen zu ermöglichen, mussten die Kommunikationen echter Nutzer beobachtet werden. Zu diesem Zweck wurden Logbücher eines Proxyservers als Ausgangsmaterial verwendet. In den Logbüchern sind alle World-Wide-Web-Zugriffe der angeschlossenen Rechner protokolliert.

Die für die Untersuchungen vorhandenen Daten stammen vom Proxyserver der RWTH Aachen, über den der größte Teil des Internetverkehrs der RWTH Aachen abgewickelt wird. Der Proxyserver wird von sehr unterschiedlichen Nutzern verwendet. Ein großer Teil der Nutzer wählen sich über eine Modemverbindung in das RWTH-Netz beziehungsweise das Internet ein. Zusätzlich wird der Server auch von einzelnen CIP-Pool-Terminals, Lehrstühlen (Zugang durch einen Router), Wohnheimen (ebenfalls Zugang durch Router oder eigene Proxyserver) und auch kompletten Nachbaruniversitäten (zum Beispiel Universität Köln) verwendet.

Während des Semesters erfolgen täglich knapp 2 Millionen Zugriffe aus dem RWTH-Netz auf den Proxyserver. Hinter einer protokollierten IP-Adressen stehen häufig mehrere Nutzer, die über einen eigenen Proxyserver auf den Server der RWTH Aachen zugreifen. Die Unterschiede zwischen den einzelnen Teilnehmern (protokollierten IP-Adressen) sind aus diesem Grund in den vorliegenden Daten, im Gegensatz zu Anonymisierungssystemen im Internet, besonders groß.

Um die vorliegenden Logbuchdaten einfach für die Untersuchungen verwenden zu können, mussten noch einige Aufbereitungsschritte durchgeführt werden. Dabei wurden im wesentlichen Informationen aus den Daten entfernt, die für die Angriffe nicht benötigt wurden. Die folgende Aufstellung zeigt die einzelnen durchgeführten Arbeitsschritte: entfernen unnötiger Informationen, entfernen fehlerhafter Einträge, Pseudonymisierung aller IP-Adressen und Zielsäuberung.

Neben dem Entfernen unnötiger Informationen (Zeitangaben, Dateigrößen und Angaben zu den Dateiformaten) und fehlerhaften Einträgen (z.B. durch Serverabsturz), wurden in einem weitere Schritt alle IP-Adressen durch Pseudonyme ersetzt.

Ein weiteres Problem bestand darin, dass eine große Anzahl unterschiedlicher Server-

[4] Mit Hilfe einer Häufigkeitsanalyse werden die Lösungskandidaten erzeugt, welche die am häufigsten aufgerufenen Zielen enthalten.

adressen die Nutzer auf die selben Webseiten führt. Die Adressen *www1.gmx.net* und *www2.gmx.net* führen beispielsweise beide zur selben Webseite. Insgesamt führt dieses Problem dazu, dass die Zielmenge sehr viel größer ist, als zunächst vermutet. Es ist durchaus realistisch, dass ein normaler Internetnutzer, der in einer Woche 50 verschiedene Webseiten aufgerufen hat, in den vorliegenden Logbüchern mit mehr als 500 Webservern kommuniziert hat. Um dieses Problem einzudämmen wurde eine sogenannte Zielsäuberung durchgeführt. Die Hauptaufgabe bestand darin, die Webserveradressen so zu vereinheitlichen, dass sie den durch den Nutzer aufgerufenen Zielen entsprechen. Eine genauere Beschreibung aller durchgeführten Arbeitsschritte kann in [Rat04] nachgelesen werden.

Um die Auswirkungen des Nutzerverhaltens auf die Güte der Anonymität beziehungsweise der Leistungsfähigkeit des Angriffsalgorithmus besser untersuchen zu können, hat sich eine Einteilung der Nutzer als sinnvoll erwiesen. Dazu wurden alle Nutzer des Proxyservers mit Hilfe einer zweidimensionalen Clusteranalyse in Gruppen eingeteilt. Das Verhalten der Nutzer wurde durch die Größen „Anzahl verschiedener Ziele“ und „Anzahl der Aufrufe“ bestimmt. Die Clusteranalyse wurde mehrfach angewendet, um eine Schrittweise Verfeinerung der interessanten Gruppen zu erzielen. Die Angriffe wurden dann auf wenige Gruppen beschränkt, um möglichst vergleichbare Ergebnisse zu den bisherigen Simulationen zu erreichen.

Um die Angriffe auf die Daten ausführen zu können, wurde im Rahmen der Untersuchungen ein modulares Angriffswerkzeug implementiert. Das Angriffsprogramm speichert alle wichtigen Zwischen- und Endergebnisse in einfach aufgebauten Textdateien. Dadurch war es möglich die Daten unkompliziert mit weiteren Programmen auszuwerten.

Der grundlegende Ablauf des Programms wird durch einzelne Phasen, die durch die einzelnen Module bearbeitete werden geprägt. Als Ausgangsposition liegen die pseudonymisierten und gesäuberten Logbuchdaten[5] vor. Um später überprüfen zu können, wie gut der Angriff funktioniert hat, müssen die tatsächlichen Ziele des Opfers in einem ersten Arbeitsschritt gespeichert werden. Danach werden aus den Logbuchdaten, gemäß vorgegebener Parameter, Anonymitätsmengen erstellt. Anschließend wird der ausgewählte Angriffsalgorithmus auf den erstellten Anonymitätsmengen ausgeführt und die Ergebnisse gespeichert. Abschließend werden die Ergebnisse des Angriffs mit den tatsächlichen Zielen verglichen und ausgewertet. Die folgende Auflistung enthält die wichtigsten Programmmodule mit einer kurzen Erklärung. Eine vollständige Beschreibung der einzelnen Module findet sich in [Rat04].

SPLIT-Modul: Dieser Programmteil wird dazu verwendet die vorliegenden Logbuchdaten auf den für den Angriff relevanten Teil zu reduzieren. Prinzipiell werden alle Einträge des aktuell angegriffenen Opfers in ein separates Logbuch übertragen. Hierdurch ist in den weiteren Schritten eine schnellere Bearbeitung möglich.

STATISTIC-Modul: Das STATISTIC-Modul erzeugt verschiedene Berichte aus den vorliegenden Logbuchdaten. Unter anderem wird eine Liste mit den tatsächlich aufgerufenen Zielen eines angegebenen Nutzers erstellt. Zusätzlich kann mit Hilfe dieses Moduls eine Clusteranalyse auf eine angegebene Nutzermenge ausgeführt werden.

[5] also nur noch ein Nutzer und das aufgerufene Ziel je Zeile

MIX-Modul: Das dritte Modul simuliert die Arbeit eines einfachen Mix-Anonymisierungsverfahrens gemäß dem in Abschnitt 2 vorgestellten Metamodells. Als Eingabe werden die vorliegenden Logbuchdaten verwendet. Als Ausgabe erzeugt das Modul die Beobachtungen des Angreifers. Sie enthalten eine Gruppe von Nutzern und eine Gruppe von Zielen, die durch die Nutzer innerhalb eines engeren Zeitraums aufgerufen worden sind. Diese Beobachtungen könnte auch ein echter Angreifer einfach durch abhören der Ein- und Ausgangsleitungen des Mix-Knotens erhalten.

Angriffsmodule: Das vierte Modul besteht je nach verwendetem Angriffsalgorithmus aus mehreren Untermodulen. Es existiert je ein Angriffsmodul für jeden in 3 vorgestellten Angriff. Das Angriffsmodul erhält als Eingabe die vom Mix-Modul bereitgestellten Beobachtungen und versucht die Ziele des angegebenen Nutzers eindeutig zu identifizieren. Mit Hilfe der vom STATISTIC-Modul berechneten Ziellisten, wird überprüft, ob die erkannten Ziele korrekt sind und wie viele Ziele nicht beziehungsweise falsch vom verwendeten Angriffsmodul erkannt worden sind.

5 Ergebnisse und Bewertung

Um die Frage nach der Eignung der verwendeten Angriffsalgorithmen mit realen Daten beantworten zu können wurden eine ganze Reihe von Untersuchungen durchgeführt. Durch verschiedene Angriffsszenarien und der Variation diverser Parameter, wurde die Vor- und Nachteile der Angriffsalgorithmen mit realen Daten herausgearbeitet.

Neben den Parametern des Anonymisierungsverfahrens (Größe der Anonymitätsmengen, Anzahl der Ziele aller Nutzer, ...) spielen auch einige Größen eine wichtige Rolle für die Stärke der Anonymität, die nicht direkt vom Anonymisierungsverfahren beeinflusst werden können. Zum Beispiel hat sich herausgestellt, dass das unmittelbare Nutzerverhalten einen sehr großen Einfluss auf den Erfolg des Angriffs und somit die Qualität der Anonymität hat.

5.1 Erkennungsraten der verschiedenen Angriffsalgorithmen

Eine interessante Frage, ist die Frage nach dem verwendeten Angriffsalgorithmus. Die Untersuchungen haben gezeigt, dass die Angriffsalgorithmen bei einer Beschränkung des Beobachtungszeitraums nicht immer alle Ziele des angegriffenen Nutzers erkennen. Je nach Angriffsszenario wird nur ein gewisser Prozentsatz der Ziele korrekt erkannt, der im folgenden auch als Erkennungsrate bezeichnet wird. In Abbildung 1 sind die Ergebnisse der verschiedenen Angriffsalgorithmen zusammenfassend dargestellt. Alle drei Angriffsvarianten wurden jeweils auf verschiedene Nutzertypen getestet, um die Vor- beziehungsweise Nachteile der einzelnen Algorithmen mit realen Daten herauszufinden.

Bei kleineren Anonymitätsmengen ist ein deutlicher Vorsprung des klassischen Ausschlussangriffs zu erkennen. Der Vorsprung der Erkennungsrate des Ausschlussangriffs verringert

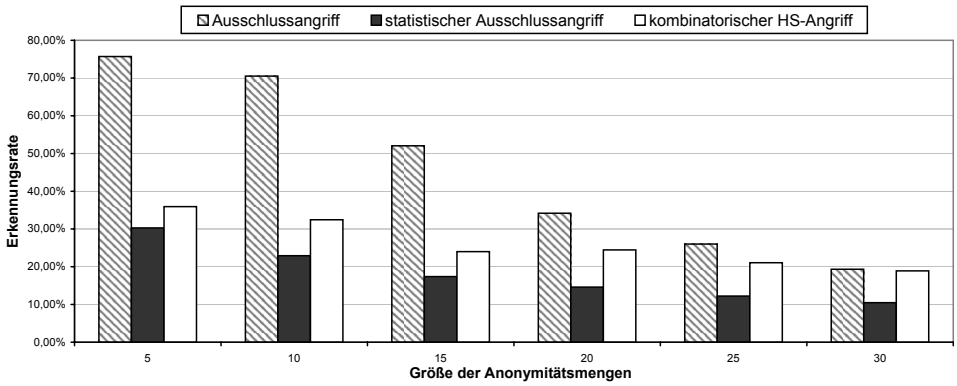

Abbildung 1: Erkennungsraten verschiedener Angriffsalgorithmen im Vergleich

sich mit zunehmender Batchgröße und ist gegenüber dem kombinatorischen Hitting-Set-Angriff ab einer Batchgröße von 30 kaum mehr erkennbar.

Die Erkennungsraten des statistischen Ausschlussangriffs mit realen Daten sind dagegen die schlechtesten der drei getesteten Algorithmen. Selbst bei geringen Batchgrößen werden maximal maximal 30% der Ziele erkannt. Dabei wurden die Eingabedaten des statistischen Ausschlussangriffs sogar über eine Gewichtung auf eine Gleichverteilung normiert, so dass die notwendigen Voraussetzungen bestmöglich erfüllt wurden. Ohne eine Normierung lagen die Werte für sämtliche Batchgrößen deutlich unter 5% (nicht in der Abbildung).

Die Ergebnisse des kombinatorischen Hitting-Set-Angriffs sind zwar besonders bei kleinen Batchgrößen deutlich schlechter als die des Ausschlussangriffs, liegen aber dennoch über den Ergebnissen des statistischen Ausschlussangriffs. Die Erkennungsrate fällt mit zunehmender Größe der Anonymitätsmengen nicht so stark ab, wie bei den beiden anderen Angriffsalgorithmen. Es ist anzunehmen, dass dieser Algorithmus für größere Anonymitätsmengen ($b > 30$) die besten Erkennungsraten liefert. Zusätzlich wurde im Laufe der Untersuchungen festgestellt, dass der Algorithmus bei größeren Anonymitätsmengen und längeren Beobachtungszeiträumen deutlich schneller als der klassische Ausschlussangriff arbeitet.

Insgesamt erscheint der klassische Ausschlussangriff für reale Angriffe und kleine Batchgrößen die beste Wahl zu sein. Für größere Anonymitätsmengen eignet sich der kombinatorische Hitting-Set-Angriff besser aufgrund seiner konstanteren Erkennungsrate und dem besseren Laufzeitverhalten.

5.2 Anteil der falsch erkannten Ziele

Ein großer Nachteil betrifft die Erkennung falscher Ziele (*false positives*). Der Anteil der falsch erkannten Ziele geht nicht in die Berechnung der Erkennungsrate mit ein. Sie de-

finierte sich lediglich aus dem Quotienten zwischen der Anzahl korrekt erkannter Ziele und der Anzahl aller Ziele des Opfers. Da der klassische Ausschlussangriff keine Ziele falsch erkennt, betrifft das Problem lediglich den statistischen Ausschlussangriff und den kombinatorischen Hitting-Set-Angriff.

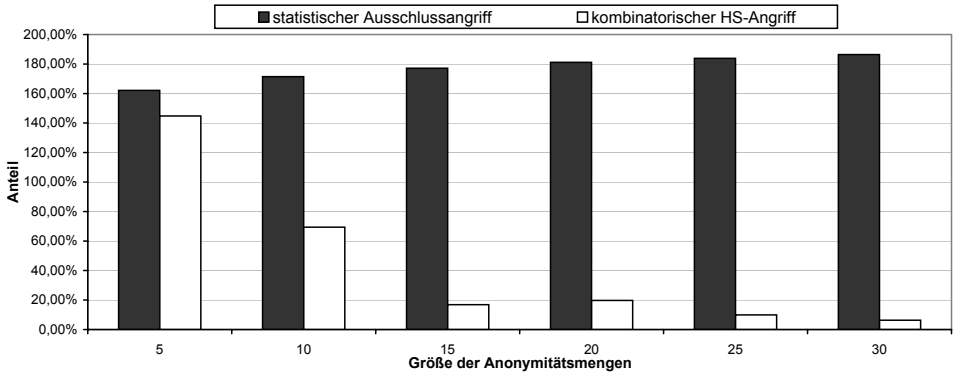

Abbildung 2: Anteil der falsch erkannten Ziele (*false positives*)

Um herauszufinden, wie groß der Anteil der falsch erkannten Ziele ist, wurden einige weitere Versuche durchgeführt. Abbildung 2 zeigt die Anteile der falsch erkannten Ziele. Als Bezugsgröße dient die Anzahl der tatsächlichen Ziele. Ein Anteil von 200% bedeutet zum Beispiel, dass die Anzahl falsch erkannter Ziele doppelt so groß ist, wie die Anzahl der tatsächlichen Ziele. Die Anzahl der falsch erkannten Ziele erhöht sich beim statistischen Ausschlussangriff mit zunehmender Batchgröße. Dieser Effekt ist damit zu erklären, dass mit zunehmender Batchgröße einige Ziele der anderen Nutzer (nicht vom Opfer selbst) häufiger vorkommen und durch den Algorithmus falsch erkannt werden.

Der Anteil der falsch erkannten Ziele beim kombinatorischen Hitting-Set-Angriff nimmt dagegen mit zunehmender Batchgröße ab und fällt insgesamt deutlich geringer aus, als beim statistischen Ausschlussangriff. Dieser Effekt ist darauf zurückzuführen, dass der kombinatorische Hitting-Set-Angriff bei nicht erfolgreicher Terminierung lediglich zwei Lösungskombinationen zurückgibt[6] und damit die Anzahl der falsch erkannten Ziele automatisch einschränkt. In den durchgeführten Versuchen hat sich dabei herausgestellt, dass die Lösungskandidaten meistens weniger als m mögliche Ziele enthielten. Außerdem kommt hinzu, dass der Algorithmus bei einigen Angriffen frühzeitig terminierte und die bis dahin kontaktierten m_t Ziele richtig identifiziert hat[7]. Dadurch hat der Algorithmus keine falschen Ziele erkannt, erreichte aber auch keine hundertprozentige Erkennungsrate. Mit zunehmender Batchgröße hat die Anzahl der frühzeitigen Terminierungen stetig zugenommen und aus diesem Grunde der Anteil der falsch erkannten Ziele abgenommen.

Durch die Versuche zeigt sich deutlich, dass der statistische Ausschlussangriff, zum einen aufgrund seiner schlechten Erkennungsrate und zum anderen wegen dem hohen Anteil falsch erkannter Ziele nur schlecht für reale Angriffe geeignet ist. Der kombinatorische

[6] Wenn der Algorithmus mehrere mögliche Lösungen findet, benötigt er noch weitere Beobachtungen.
[7] dabei ist $m_t < m$

Hitting-Set-Angriff verspricht dagegen besonders für große Anonymitätsmengen die besten Erfolgsaussichten.

5.3 Auswahl des Opfers

Ein weiterer interessanter Punkt ist die Frage nach der Qualität der Anonymität einzelnen Nutzerklassen. Es ist anzunehmen, dass nicht alle Nutzerklassen gleichermaßen vor Angriffen geschützt sind. Um die Unterschiede in der Qualität der Anonymität einzelnen Nutzergruppen zu erkennen, wurden mehrere Angriffe auf verschiedene Nutzertypen durchgeführt.

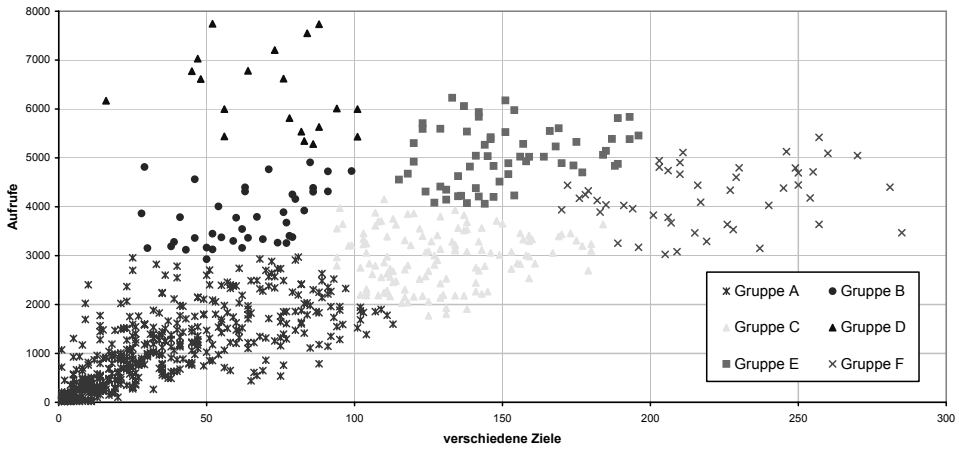

Abbildung 3: Clusteranalyse einer Teilgruppe von Nutzern

Abbildung 3 zeigt eine Clustereinteilung einer Nutzergruppe in sechs verschiedene Untergruppen. Abbildung 4 zeigt die erreichte Erkennungsrate des klassischen Ausschlussangriffs auf die einzelnen Nutzergruppen bei variierenden Batchgrößen[8].

Es ist eindeutig zu erkennen, dass der Ausschlussangriff für Nutzergruppen mit niedriger Zielmenge wesentlich bessere Ergebnisse liefert. Die höchste Erkennungsrate erzielt der Angriff mit der Nutzergruppe A, die maximal 120 verschiedenen Ziele kontaktieren. Besonders bei größeren Anonymitätsmengen, liegt die Erkennungsrate für diese Nutzergruppe deutlich höher als die anderer Nutzerklassen. Für die Nutzergruppen B (zwischen 3000 und 5000 Aufrufen) und D (zwischen 5000 und 8000 Aufrufen) verlaufen die Erkennungsraten für verschiedene Batchgrößen ungefähr gleich. Dies lässt die Vermutung zu, dass die absolute Ausprägung der Aufrufhäufigkeit keine so bedeutende Rolle spielt, wie die Zielanzahl des Opfers. Diese Annahme wird dadurch bestätigt, dass sich die Anonymität der Nutzer der Gruppe E schlechter aufheben lässt, als die der Nutzergruppe A. Die Relation zwischen Anzahl der Aufrufe und Anzahl der Ziele der Gruppe E entspricht

[8] Größe der Anonymitätsmengen

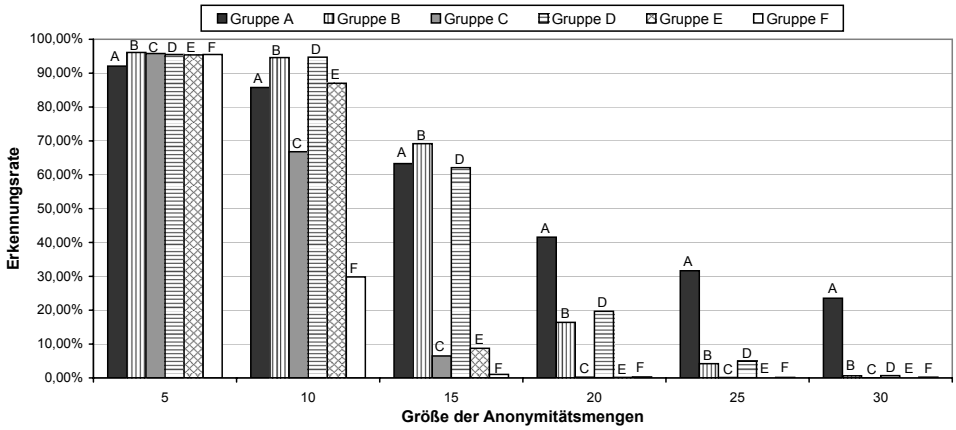

Abbildung 4: Einfluss der Opferauswahl auf die Erkennungsrate

ungefähr der, der Nutzer der Gruppe A.

Die Anonymität der Nutzergruppe F (bis zu 290 Ziele) lässt sich am wenigsten durch den Ausschlussangriff brechen. Schon ab einer Batchgröße von 10 liegt die Erkennungsrate knapp unter 30%. Wie auf Abbildung 3 zu erkennen ist, liegt diese Gruppe am rechten Rand. Dies bedeutet eine große Anzahl unterschiedlicher Ziele und eine bessere Anonymität. Die Nutzergruppe C hat bei gleichem Aufrufniveau wesentlich weniger Ziele und müsste leichter anzugreifen sein. Diese Vermutung wird durch Abbildung 4 bestätigt. Die Erkennungsrate liegt bei einer Batchgröße von 10 über 65%.

Die Ergebnisse zeigen, dass der Idealnutzer aus der Sicht des Angreifers eine sehr kleine Zielmenge hat und sich in einem Clusterdiagramm in der linken Hälfte befinden würde. Die Ergebnisse dieses Abschnitts lassen sich nur indirekt für die Sicherheit der Anonymisierungsverfahren im Internet nutzen. Letztendlich kann der Nutzer die Stärke der Anonymität durch sein eigenes Verhalten beeinflussen (z.B. durch den Einsatz von Dummy-Traffic). Weitere Ergebnisse und Bewertungen können in [Rat04] nachgelesen werden.

6 Schlussbemerkung

Ein wesentlicher Unterschied zwischen den bisher simulierten und den in [Rat04] durchgeführten Angriffen besteht in der Datengrundlage. Bei der Simulation wurden stochastische Verkehrsmodelle benutzt. Bei einem Vergleich der Angriffsergebnisse zeigt sich, dass echte Kommunikationsdaten teilweise unvorhersehbare und auch unerklärliche Kommunikationsmuster enthalten, die bei einer Simulation nicht berücksichtigt werden können.

Ein wichtiges Ergebnis ist, dass die Größe der Anonymitätsmengen eines Mix-Systems einen großen Einfluss auf die Durchführbarkeit der Angriffe und dadurch auch auf die Sicherheit der Anonymität hat. Während andere Schutzmaßnahmen oft nur schwer zu rea-

lisieren sind, ist eine Veränderung der Batchgröße bei jedem existierenden System leicht möglich. Außerdem hat sich gezeigt, dass der bereits 2002 vorgestellte Ausschlussangriff die Erwartungen für Angriffe auf reale Daten übertroffen hat, indem er besonders bei kleinen Anonymitätsmengen den erst ein Jahr später entwickelten und vorgestellten statistischen Ausschlussangriff übertroffen hat.

Die Frage nach dem geeigneten Angriffsalgorithmus richtet sich vor allem danach, welche Rahmenbedingungen durch das anzugreifende Anonymisierungsverfahren bereitgestellt werden. Mit den durchgeführten Angriffen wurde gezeigt, dass vor allem die Größe der Anonymitätsmengen darüber entscheidet, welcher Algorithmus besser geeignet ist.

Literatur

[AKP03] Dakshi Agrawal, Dogan Kesdogan und Stefan Penz. Probabilistic Treatment of MIXes to Hamper Traffic Analysis. In *Proceedings of the 2003 IEEE Symposium on Security and Privacy*, Mai 2003.

[Cha81] David L. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 24(2), Februar 1981.

[Dan03] George Danezis. Statistical Disclosure Attacks: Traffic Confirmation in Open Environments. In *Proceedings of Security and Privacy in the Age of Uncertainty, (SEC2003)*, Seiten 421–426, Athens, Mai 2003. IFIP TC11, Kluwer.

[DS04] George Danezis und Andrei Serjantov. Statistical Disclosure or Intersection Attacks on Anonymity Systems. In *Proceedings of 6th Information Hiding Workshop (IH 2004)*, LNCS, Toronto, Mai 2004.

[FH04] Tor: an anonymizing overlay network for TCP. Webseite, Freehaven.net, 2004.

[JAP04] JAP - Anonymity and Privacy. Webseite, Technische Universität Dresden, 2004.

[KAP02] Dogan Kesdogan, Dakshi Agrawal und Stefan Penz. Limits of Anonymity in Open Environments. In Fabien Petitcolas, Hrsg., *Proceedings of Information Hiding Workshop (IH 2002)*. Springer-Verlag, LNCS 2578, Oktober 2002.

[KP04] Dogan Kesdogan und Lexi Pimenidis. The Hitting Set Attack on Anonymity Protocols. In *Proceedings of 6th Information Hiding Workshop (IH 2004)*, LNCS, Toronto, Mai 2004.

[MD04] Nick Mathewson und Roger Dingledine. Practical Traffic Analysis: Extending and Resisting Statistical Disclosure. In *Proceedings of Privacy Enhancing Technologies workshop (PET 2004)*, LNCS, Mai 2004.

[PW85] Andreas Pfitzmann und Michael Waidner. Networks Without User Observability – Design Options. In *Proceedings of EUROCRYPT 1985*. Springer-Verlag, LNCS 219, 1985.

[Rat04] Oliver Rattay. Sicherheitsbewertung von Anonymisierungsverfahren im World Wide Web. Diplomarbeit, Lehrstuhl für Informatik 4, RWTH Aachen, Aachen, Deutschland, September 2004.

[Ray01] Jean-François Raymond. Traffic Analysis: Protocols, Attacks, Design Issues and Open Problems. In *Designing Privacy Enhancing Technologies*, Seiten 10–30, Berkeley, CA, USA, 2001. Springer-Verlag, LNCS 2009.