Gottfried Wilhelm Leibniz Universität Hannover
Fachgebiet Distributed Computing
Distributed Computing Security Group

Bachelor-Arbeit
im Studiengang Informatik (B. Sc.)

Mechanismen zum Schutz der Privatsphäre in mobilen kontext-abhängigen Informationssystemen

Verfasser:	Philipp Tute
Erstprüfer:	Prof. Dr. rer. nat. M. Smith
Zweitprüferin:	Prof. Dr.-Ing. G. von Voigt
Betreuer:	M. Sc. B. Henne
Datum:	03. September 2010

Hannover, den 03.09.2010
Hiermit versichere ich, dass ich diese Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel verwandt habe.

Philipp Tute

Abstract

Mobile Endgeräte wie Smartphones und Laptops verfügen immer häufiger über schnelle Internetzugänge und GPS-Empfänger und laden ihre Benutzer ein, jederzeit an Web-2.0- oder standortbasierten Diensten teilzunehmen. Gerade soziale Netzwerke erfreuen sich zur Zeit großer Beliebtheit und erfahren einen weiteren Schub durch ihre allgegenwärtige Verfügbarkeit. Ihre Teilnehmer veröffentlichen Statusmitteilungen, Bilder und vieles mehr. Zudem bieten viele Angebote die Möglichkeit, Personen- oder Standortdaten zu Fotos und Dokumenten hinzuzufügen. In vielen Fällen geschieht dies jedoch auf Kosten der Privatsphäre der Nutzer. Dies liegt zum Teil an der lockeren Einstellung der Nutzer. Häufig bieten aktuelle mobile Informationssysteme jedoch nur wenige Möglichkeiten zum Schutz der eigenen Daten. Diese Arbeit gibt einen Überblick über den aktuellen Stand der Forschung zum Schutz der Privatsphäre in mobilen kontext-abhängigen Informationssystemen. Es werden ausgewählte Lösungsansätze vorgestellt und bewertet.

Inhaltsverzeichnis

Abbildungsverzeichnis

Abkürzungsverzeichnis

AID	Anonymous Identifier
API	Application Programming Interface
CBR	Case-Based Reasoning
IS	Identity Server
LBAC	Location Based Access Control
LBS	Location Based Service
MANet	Mobile Ad hoc Network
MITM	Man-In-The-Middle
OSNR	Obfuscated Social Network Routing
P2P	Peer-To-Peer
PKI	Public Key Infrastructure
QoS	Quality of Service, Dienstgüte
ROI	Region Of Interest
SN	Social Network, Soziales Netzwerk
SNS	Social Network Service
SSNR	Statisticulated Social Network Routing
TIS	Trusted Identification Service
TTP	Trusted Third Party

1 Einleitung

1.1 Mobile Informationssysteme

> „Every once in a while a revolutionary product comes along that changes everything.“
> – Steve Jobs

Moderne Informationssysteme haben ihren Nutzern vieles zu bieten. Sie werden durch ihre geringe Größe und gesteigerte Energieeffizienz zunehmend mobiler und bieten gleichzeitig immer mehr Funktionen bei stetiger Verbesserung der Nutzerfreundlichkeit. Angefangen bei Laptops und Netbooks über Tablet-PCs bis hin zu Handys und Smartphones steht dem Nutzer mittlerweile eine permanente Verbindung mit dem Internet zu Verfügung. Eingebaute GPS-Empfänger und die Möglichkeit, Standorte über Triangulation in Mobilfunk- oder WLAN-Netzen zu ermitteln, bieten darüber hinaus die Grundlage für eine allgegenwärtige Verwendung von standortbasierten Diensten. Aus der daraus resultierenden permanenten Verfügbarkeit von Informationen bieten sich sowohl für die Nutzer, als auch für die Anbieter verschiedenster Dienste völlig neue Möglichkeiten.
Vor allem soziale Netzwerke profitieren von jederzeit möglichem Internetzugang, indem ihre Nutzer Daten oder Informationen, wie Fotos, Statusupdates und Kurznachrichten, überall und zeitnah zu erwähnenswerten Geschehnissen veröffentlichen können. Dieser erweiterte Zugriff führt, gerade wenn er noch neu und aufregend ist, dazu, dass Anwender sich beinahe euphorisch auf die neuen Möglichkeiten stürzen. So vermeldet beispielsweise `facebook.com`, mit über 500 Millionen aktiven Nutzern Marktführer im Bereich sozialer Netzwerke, dass seine 150 Millionen mobilen Nutzer doppelt so aktiv sind, wie solche, die über einen herkömmlichen Computer auf die Seite zugreifen [17]. Es kann also davon ausgegangen werden, dass Nutzer, zumindest solange sie einer neuen Technologie noch nicht müde sind, diese gerne und viel verwenden.
Gleichzeitig können aber auch völlig neue Dienste entstehen. Kurzstreckennetzwer-

ke wie Bluetooth und WLAN ermöglichen eine neue Art von sozialen Netzwerken, die nicht mehr nur auf vorhandenen Verbindungen zwischen Personen basieren, sondern auch in Verbindung mit Standortdaten auf räumlicher Nähe basieren können. Es ist Nutzern somit möglich Freunde und Bekannte in ihrer Umgebung zu finden oder neue Bekanntschaften mit ähnlichen Interessen zu schließen. Anwendungen, die es erlauben Lokalitäten oder Veranstaltungen in der Nähe zu finden und zu bewerten, sind eine weitere Gruppe, die durch Standortbestimmung und mobiles Internet zunehmend an Bedeutung gewinnt. Gerade auf Smartphones, die der Besitzer im Normalfall immer bei sich trägt, helfen diese Dienste zum Beispiel, neue Restaurants in der Heimatstadt oder einen schnellen Imbiss auf Reisen zu entdecken. Zusätzlich bieten sie eine Plattform, auf der Empfehlungen und Bewertungen ausgetauscht werden können. Obwohl es ähnliche Dienste auch ohne mobilen Schwerpunkt gibt, beispielsweise gesammelt als Einträge in den Karten von Google Maps [28], kann eine zeitnahe Bewertung jedoch unter Umständen eine erhebliche Qualitätssteigerung bewirken. Wenn die Bewertung zeitnah abgegeben wird, sind die Eindrücke des Nutzers noch frisch und somit wesentlich genauer.
Auch die Anbieter verschiedener Dienste können von mobilen Informationssystemen profitieren. Vor allem für Werbung bietet sich die Chance, mögliche Kunden gezielter anzusprechen und maßgeschneiderte Angebote machen zu können. Ein lokales Café kann beispielsweise vorbeilaufende Kunden auf Aktionen oder Spezialitäten aufmerksam machen. Die Angaben eines Nutzers in einem sozialen Netzwerk sind ebenfalls ein starker Indikator für seine Vorlieben und Interessen und somit sehr erstrebenswerte Informationen für Anbieter von Werbung. Doch nicht nur die Werbung kann an den Nutzer angepasst werden. Durch die permanente Versorgung mit Informationen können Dienste individualisiert[1] und die Erfahrung für den Nutzer somit verbessert werden.

[1] z. B. auf Grund des aktuellen Standorts

1.2 Privatsphäre

> „Privacy bedeutet [...] mehr als ‚das Recht in Ruhe gelassen zu werden'[...]"
> – Rainer Kuhlen

Damit effektive Maßnahmen zum Schutz der Privatsphäre getroffen werden können, ist zunächst zu klären, was Privatsphäre überhaupt ist, damit ein umfassender Schutz sichergestellt werden kann. Der Schutz der Privatsphäre ist im deutschen Rechtswesen Teil des Persönlichkeitsrechts jedes Bürgers. Das (allgemeine) Persönlichkeitsrecht ist allerdings nicht explizit geregelt. Es setzt sich aus den Rechten auf Selbstbestimmung (Individualsphäre), auf Schutz der eigenen Gedanken und Gefühle (Intimsphäre), und auf Privatleben, also die Privatsphäre, zusammen. Während in der Antike und im Mittelalter sehr wenige Menschen, meist Adel oder Führungsschicht, den Luxus der Privatsphäre genießen konnten, weil beispielsweise Bedienstete häufig in gemeinsamen Räumen lebten, herrscht heutzutage ein starkes Verlangen nach ihrem Schutz. Auch die Privatsphäre umfasst mehrere Einzelaspekte. Weniger wichtig ist im Rahmen dieser Arbeit[2] die Unverletzlichkeit der eigenen Wohnung. Der entsprechende Artikel 13 des Grundgesetz beschäftigt sich hauptsächlich mit dem Eindringen in eine Wohnung.
Ein weiterer Aspekt der Privatsphäre, der gerade für die Informatik interessant ist, ist das Post- beziehungsweise das Fernmeldegeheimnis (Art. 10 GG). Es soll jedoch nur nebenbei Teil dieser Arbeit sein, die eigentliche Kommunikation zwischen Personen zu sichern. Hauptmerkmal dieser Arbeit ist der verbleibende Teil der Privatsphäre, also der Schutz von personenbezogenen Daten, häufig kurz als Datenschutz bezeichnet. Grundlage dieses Anspruches ist die informationelle Selbstbestimmung, also die Entscheidungsfreiheit eines Individuums über die Weitergabe und Verwendung von Daten, die seine Person betreffen. Der Schutz personenbezogener Daten ist nicht im Grundgesetz festgehalten, sondern wird auf Landesebene geregelt. Dabei werden mittlerweile alle persönlichen Daten als gleich wichtig für die Privatsphäre angesehen. Auf Grund der umfassenden Möglichkeiten, Daten zu sammeln und zu verknüpfen, können heute keine Informationen mehr als unbedeutend eingestuft werden. Im Rahmen des Datenschutzes werden dabei sowohl Nutzung, als auch Verarbeitung und Erhebung geregelt.
Das Internet als globale Plattform macht es jedoch nicht immer möglich, Daten den Vorschriften und Gesetzen eines Staates entsprechend zu schützen. Häufiges Beispiel

[2] und nur hier

für den lockeren Umgang mit privaten Daten stellen dabei die USA dar, in denen der Datenschutz kaum geregelt wird. Häufig ist es also so, dass Nutzer sich selbstständig um den Schutz ihrer Daten bemühen müssen, wenn dieser über einen minimalen Grad hinausgehen soll.

1.3 Motivation dieser Arbeit

> „Privatsphäre ist wie Sauerstoff - man schätzt sie erst, wenn sie fehlt."
> – John Emontspool

Immer wieder zeigt sich, dass Nutzer sich nur wenig oder gar nicht mit dem Schutz ihrer Daten auseinandersetzen [32, 40]. Häufig kommt es so zur Veröffentlichung privater Daten [54] oder sogar geheimer Informationen [50]. Doch auch, wenn ein Nutzer seine Daten aktiv schützen möchte, wird ihm oftmals keine Möglichkeit dazu geboten. Gerade standortbasierte Dienste wie Google Latitude [27] bieten nur wenig oder keinen Schutz für die Daten von Nutzern. Es lassen sich verschiedene Verhaltensmuster im Umgang mit privaten Daten erkennen. Die einen stürzen sich mit Freude auf neue Angebote wie soziale Netze und veröffentlichen ohne Vorbehalt ihre Daten, wohingegen andere diesen skeptisch gegenüberstehen und keine oder nur die nötigsten Informationen preisgeben. Doch auch vorsichtige Nutzer und Personen, die solche Dienste gar nicht verwenden, haben es jedoch immer wieder sehr schwer ihre Privatsphäre zu schützen. Denn selbst dann, wenn sie selber keine Informationen preisgeben, können andere, weniger vorsichtige Nutzer direkt (z. B. durch Fotos mit Tags) oder indirekt (durch öffentliche Nachrichten u.Ä.) sensible Daten preisgeben. Dies mag mit dem Vorsatz geschehen Schaden anzurichten. Häufig kann dies allerdings auf unbedachtes Handeln zurückgeführt werden, wenn ein Nutzer sich nicht im Klaren ist, ob die getaggte Person nicht Einwände gegen ihr Auftauchen hat. Bevor jedoch untersucht wird, wie und warum die Privatsphäre angegriffen werden kann, sollte zunächst festgestellt werden, warum Nutzer ihre Daten preisgeben oder schützen sollten.

Veröffentlichen privater Daten

Die einfachste Art etwas vor unerlaubten Blicken zu schützen ist es, diese Dinge für sich zu behalten. Es stellt sich also die Frage, warum jemand dem Internet Dinge anvertrauen sollte, die eigentlich schützenswert sind. Die Gründe hierfür sind vielfältig, lassen sich jedoch unter zwei wesentlichen Gesichtspunkten zusammenfassen:

mangelndes Interesse und Bequemlichkeit.
Häufig sind Nutzer sich nicht darüber im Klaren, was mit ihren Daten passiert und welche Auswirkungen dies für sie hat. Dies kann verschiedene Gründe haben. In den meisten Fällen ist es nur schwer möglich, bei einem Dienst einzusehen, wie die eigenen Daten geschützt oder weitergegeben werden. Da außerdem das Internet die erste Plattform ist, auf der Menschen, insbesondere nach dem Aufkommen von sozialen Netzwerken und Web 2.0, weltweit Informationen in Echtzeit zur Verfügung stellen und abrufen können, ist es leicht möglich die Gefahren für die Privatsphäre zu unterschätzen. Wenn vor den Zeiten des Internets Informationen veröffentlicht wurden, geschah dies meist in einem relativ kleinen Kreis, aus dem sich beispielsweise Fotos mit empfindlichem Inhalt leicht wieder entfernen ließen. Durch soziale Netzwerke ist nicht nur die Wahrscheinlichkeit für die Veröffentlichung solcher Bilder gestiegen, sie werden zugleich auch einer sehr viel größeren Menge zugänglich, wenn der Veröffentlichende nicht die geforderte Vorsicht walten lässt. Noch gravierender ist es, dass sich Informationen nur mit sehr großem Aufwand, man mag sogar behaupten gar nicht mehr, aus dem Internet entfernen lassen. Vielen Menschen, die sich mit der Technik hinter dem Internet und seinen Diensten nicht auseinandersetzen und sie lediglich als komfortables Werkzeug betrachten, sind diese Tatsachen wenig oder nicht bewusst, weshalb sie die Gefahren für ihre Privatsphäre schnell übersehen. Es ist auch leicht zu übersehen, inwiefern manche Daten überhaupt kritisch sein können.
Doch selbst Nutzer, die sich bewusst sind, was mit ihren Daten geschieht, wissen unter Umständen nicht, wie ihnen dies schaden kann. Denn gerade bei Diensten, die nur indirekt Informationen über ihren Nutzer erhalten, ist die Gefahr nicht einfach zu erkennen. Jemand, der anonyme standortbasierte Dienste nutzt, mag sich zwar der Tatsache bewusst sein, dass der Betreiber des Dienstes seinen Standort erfährt, jedoch nicht einsehen, wie seine Privatsphäre dadurch gefährdet werden kann.
Der zweite Gesichtspunkt betrifft sowohl informierte, als auch gleichgültige Nutzer. Generell kann davon ausgegangen werden, dass die meisten Menschen ihre Daten dann freigeben, wenn der daraus gewonnene Nutzen größer ist als die möglichen Folgen. Diese Entscheidung ist rein subjektiv und wird von dem persönlichen Verlangen nach Privatsphäre, aber auch von Fehleinschätzungen bestimmt. Mit weitreichender Verfügbarkeit und steigendem Umfang haben mobile Dienste einen großen gefühlten Nutzen und verleiten somit dazu, freizügiger mit persönlichen Daten umzugehen. So neigt eine Person möglicherweise dazu, einen Dienst zu nutzen, der ihr hilft ein Restaurant zu finden, wenn sie hungrig in einer fremden Stadt unterwegs ist und ein entsprechendes Programm sowieso auf ihrem Smartphone installiert ist.

Motivation für Angriffe

In vielen Fällen ist es nicht eindeutig, warum die versendeten Daten ein Risiko für die Privatsphäre darstellen. Wenn ein Nutzer sich nicht im Klaren darüber ist, wie seine Daten zu seinem Nachteil genutzt werden können, ist er möglicher Weise eher dazu geneigt sie preiszugeben. Deshalb sollen im Folgenden einige Gründe genannt werden, wodurch Gefährdungen entstehen können und wie persönliche Daten ausgenutzt werden können.
Nicht immer muss ein gezielter Angriff vorliegen, wenn die Privatsphäre eines Nutzers bedroht ist. Es ist häufiger so, dass ein Teilnehmer des Dienstes, meist der Anbieter selbst, gutmütig aber neugierig ist. Das heißt, dass er dem Nutzer nicht schaden, aber dennoch Informationen über ihn sammeln möchte. Der hauptsächliche Grund für dieses Verhalten ist Geld [18]. Werbeanbieter legen immer mehr Wert darauf ihre Anzeigen möglichst gut auf ihre Kunden abzustimmen und dementsprechend gut lassen sich Daten verkaufen, aus denen sie Vorlieben und Wünsche ihrer potentiellen Kunden erkennen können. Leider ist es dabei sehr schnell möglich, dass jemand im Besitz dieser Daten auf Dinge schließen kann, die ein negatives Licht auf den Nutzer werfen oder Geheimnisse enthüllen. Das Resultat eines solchen Schlusses ist dabei, unabhängig von seinem Wahrheitsgehalt, unter Umständen recht unangenehm für den Nutzer. So kann beispielsweise von häufig besuchten Standorten[3] oder von Statusnachrichten auf mögliche Interessen oder sogar Krankheiten geschlossen werden, die der Nutzer lieber geheim gehalten oder nur ausgewählten Personen zugänglich gemacht hätte.
Ein direkter Angriff auf die Privatsphäre liegt dann vor, wenn ein Angreifer gezielt Daten einer Person oder einer Gruppe sammelt, um deren Geheimnisse zu erfahren. So kann ein findiger Arbeitgeber zum Beispiel überprüfen, ob seine Mitarbeiter wirklich zum Arzt gehen, wenn sie sich krank melden, oder auf Geschäftsfahrten keine unerlaubten Abstecher machen, indem er sich Zugang zu den Standortdaten dieser verschafft. Aus sozialen Netzwerken könnte ein Stalker Adresse und Kontaktinformationen seines Opfers ermitteln um es verfolgen zu können [32, 57]. Viele solcher Angriffe lassen sich durch übliche Maßnahmen aus dem Bereich der Informationssicherheit[4] verhindern. Wenn dies jedoch nicht möglich ist, muss dafür gesorgt werden, dass die Daten von Nutzern so weit anonymisiert oder verschleiert sind, dass ein Angreifer keine Vorteile aus ihnen ziehen kann.

[3] oder Bildern mit Geo-Tags [33]
[4] z. B. Zugriffsschutz, Verschlüsselung

Ziel dieser Arbeit

Die zuvor beschriebenen Betrachtungen verdeutlichen, dass Maßnahmen entwickelt werden müssen um die Privatsphäre von Nutzern vor ihnen selbst und anderen zu schützen. Viele wissenschaftliche Arbeiten beschäftigen sich bereits mit dieser Aufgabe. Diese Arbeit stellt einige aktuelle Ansätze zum Schutz der Privatsphäre in mobilen Informationssystemen vor und bewertet sie. Ziel ist es, einen Überblick über den Stand der Forschung zu geben und bestehende wie zukünftige Probleme aufzuzeigen.

1.4 Inhalt und Aufbau dieser Arbeit

In Kapitel 2 wird eine Übersicht über verschiedene Arten von mobilen Informationssystemen gegeben. Die Dienstarten werden erläutert und Gefahren für die Privatsphäre der einzelnen Systeme dargestellt. Daraufhin werden nacheinander verschiedene Ansätze zu den einzelnen Dienstarten vorgestellt, die einen Überblick über die aktuelle Forschung geben. Jedes Kapitel wird mit einer Übersicht von Vor- und Nachteilen der vorgestellten Ansätze, sowie offener Probleme geschlossen. In Kapitel 3 werden Verfahren für die Verschleierung von Standorten und die Anonymisierung von Nutzern in verschiedenen Arten von standortbasierten Diensten vorgestellt. In Kapitel 4 werden soziale Netzwerke betrachtet. Es werden dabei sowohl herkömmliche Netzwerke mit einer mobilen Erweiterung, als auch Dienste, die speziell für die mobile Anwendung entwickelt wurden, betrachtet. Der Schwerpunkt liegt hierbei auf dem Veröffentlichen eigener Daten. In Kapitel 5 wird daraufhin auf Möglichkeiten zum Schutz gegen das Veröffentlichen eigener Daten durch Dritte eingegangen. Besonderer Fokus liegt hier auf dem Hochladen und Taggen von Fotos. Abschließend werden die Ergebnisse der einzelnen Kapitel in Teil 6 zusammengefasst und ein Fazit gezogen.

2 Klassifikation

Mobile Informationssysteme bieten und kombinieren verschiedene Dienste und offenbaren dementsprechend unterschiedliche Gefahren für die Privatsphäre ihrer Nutzer. Dieses Kapitel gibt einen Überblick über verschiedene Dienstarten und ihre Probleme. Abbildung 2.1 zeigt eine Einteilung von Diensten, die im Verlauf dieser Arbeit entstanden ist. Sie stellt jedoch nur eine mögliche Aufteilung dar, da viele Angebote und Forschungsprojekte die Grenzen zwischen den verwendeten Systemen stark verwischen. Die Einteilung basiert sowohl auf einer Zusammenfassung aktuell angebotener Dienste, als auch auf einem Überblick über aktuelle Forschungsgebiete.

Standortbasierte Dienste

Standortbasierte Dienste werden auf Grund von mittlerweile allgegenwärtiger Positionsbestimmung sehr häufig in neue oder vorhandene Dienste integriert. Ziel eines solchen Dienstes ist es, den Standort eines Nutzers zu erfassen und auszuwerten, um basierend auf diesem Informationen zu liefern oder Aufgaben auszuführen. Es lässt sich dabei zwischen zwei Verwendungszwecken des Standorts unterscheiden. Beide bringen eigene Anwendungen aber auch Gefahren mit sich. Im ersten Szenario kann ein Nutzer seinen Standort explizit bestimmen, um ihn dann zu veröffentlichen. Dies kann beispielsweise im Rahmen einer Webapplikation geschehen, in der andere Nutzer seinen Standort suchen können [27] oder von seiner Nähe in Kenntnis gesetzt werden, wenn sie dies wünschen. Bereits hier lässt sich eine Vermischung mit den Charakteristika eines sozialen Netzwerks erkennen, da häufig Freundeslisten gepflegt werden, die den Zugriff auf den Standort des Nutzers regeln. Dementsprechend muss in solchen Applikationen nicht nur auf den Schutz des Standortes vor Verfolgung, sondern auch auf den Schutz vor unberechtigtem Zugriff geachtet werden. Dabei kann ein Nutzer sowohl von einem externen Angreifer ausgespäht werden, als auch von dem Dienstanbieter selbst.
Die zweite Art von standortbasierten Diensten beschäftigt sich weniger mit dem Hochladen des Standortes, als viel mehr mit dem Herunterladen von Informationen.

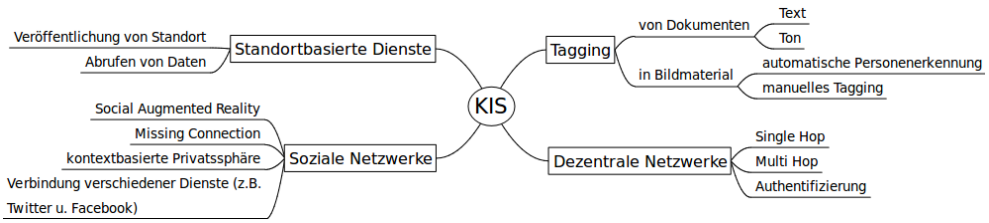

Abbildung 2.1: Übersicht über verschiedene kontextsensitive Informationssysteme

So gibt es Dienste wie Loopt Pulse [37], Qype [45] oder Google Maps [28], die ihren Nutzern Gastronomie- oder Veranstaltungsempfehlungen auf Grund ihres Aufenthaltsortes machen. Ein Angriff auf die Privatsphäre ist hier, setzt man einen sicheren Server und eine sichere Verbindung zu diesem voraus, alleine auf den Missbrauch der gesammelten Daten durch den Anbieter beschränkt. Zusätzlich erlauben es viele solcher Dienste einem Nutzer anonym zu bleiben oder ein Pseudonym zu verwenden. Durch die regelmäßige Übertragung von Standorten können jedoch Rückschlüsse auf regelmäßige Aufenthaltsorte eines Nutzers gezogen werden. Dadurch lassen sich Arbeitsplatz, Wohnort und durch diese die Identität des Ziels bestimmen. Es gilt also auch, die Identität eines Nutzers solcher Dienste zu schützen.
Der Standort eines Nutzers kann auf verschiedene Arten seine Privatsphäre gefährden. Schon wenn eine Person, der dies nicht vom Nutzer erlaubt wurde, Kenntnis über ihn hat, verletzt sie damit dessen Privatsphäre. Der im Endeffekt daraus entstehende Schaden hängt jedoch davon ab, was mit den Daten geschieht. Dieser Schaden kann genauso entstehen, wenn sich die Daten in der Hand einer berechtigten Person befinden. Deshalb legen Maßnahmen zum Schutz des Standortes, also der *Location Privacy*, nur wenig Wert auf den Zugriffsschutz. Sie konzentrieren sich stattdessen darauf, den versendeten Standort auf verschiedene Arten so zu manipulieren, dass eine genaue Bestimmung des Aufenthaltsortes eines Nutzers erschwert wird [2, 25]. Ziel dieser Ansätze ist es in den meisten Fällen, eine Fläche anstelle eines genauen Punktes zu erzeugen, in der sich ein Nutzer aufhalten kann.
Um die Identität eines Nutzers in einem standortbasierten Dienst zu schützen, ist es, gerade bei einer kontinuierlichen Übertragung von Messungen[1], notwendig Rückschlüsse auf seine Bewegung zu verhindern. Denn eine Spur, die ein Nutzer durch seine Anfragen hinterlässt, kann genauso wie eine Ansammlung einzelner Anfragen dazu führen, dass Rückschlüsse auf seinen Wohnsitz oder seine Arbeit möglich sind. Der häufigste Ansatz dies zu verhindern basiert darauf den Nutzer nur begrenzt zu

[1] z. B. GPS-Navigation

verfolgen. Dies kann durch eine zeitliche Grenze [30] oder auch durch die Aktivität anderer Nutzer in der Umgebung geregelt werden [6].
Eine Herausforderung, der sich alle bisherigen Ansätze stellen müssen, ist die Einhaltung einer gewissen Dienstgüte. Sucht ein Nutzer eine Tankstelle in seiner Nähe, muss er oder der Anonymisierungsdienst, den er verwendet, das Verhältnis zwischen Privatsphäre und Genauigkeit der Ergebnisse abwägen. Ein zu großer Fokus auf einen der beiden Aspekte kann die erhaltenen Ergebnisse oder die verwendete Verschleierung nutzlos machen. Ähnliches gilt für die Veröffentlichung des Standortes. Allerdings beeinflusst eine zu groß gewählte Privatsphärenebene eher die Qualität für andere Nutzer als für den Nutzer selbst. Bei aufwändigen Verfahren kommt zudem eine zeitliche Dimension hinzu, da große Verzögerungen entstehen können. Auch durch diese können manche Dienste erheblich an Qualität einbüßen.

Tagging

Mit steigender Verbreitung von Web-2.0-Diensten, die die Beteiligung ihrer Nutzer fordern, vermehrt sich auch das *Tagging*. Tagging bezeichnet das Hinzufügen von (Meta-)Informationen zu vorhandenen Daten. Ziel ist es, den Nutzern eines Angebots die Suche nach Inhalten zu erleichtern und diese mit zusätzlichen Daten anzureichern, indem andere Nutzer Schlüsselworte zu einem Dokument oder einer Website hinzufügen. In seiner ursprünglichen Form stellt Tagging kaum ein Sicherheitsrisiko dar, da das markierte Material im Allgemeinen vom Verfasser selbst veröffentlicht wurde. Mit dem Aufkommen von sozialen Netzwerken und Fotoportalen hat sich jedoch beispielsweise der Trend entwickelt, auch Personen auf Bildern und in Videos mit ihrem Namen zu kennzeichnen. Wenn die betroffene Person dem zustimmt, ist dies eine nützliche Funktion für ihre Freunde oder andere Interessierte. Ist sie sich jedoch des Taggings nicht bewusst, besteht auch hier ein Einbruch in ihre Privatsphäre. Denn ein Foto, das von einem Anderen hochgeladen und getaggt wurde, kann die Person zum Beispiel an einem Ort, an dem sie nicht beobachtet werden wollte, zeigen. Es kann sie auch in Begleitung einer Person zeigen, zu der sie ihre Verbindung geheim halten möchte. Es kann sie aber auch einfach in einer unpassenden Situation zeigen. Ein oft verwendetes Beispiel sind die Fotos von der letzten Feier, auf denen Gäste zu sehen sind, die das Feiern ein wenig übertrieben haben.
Ein wichtiger Unterscheidungspunkt für das Tagging von Bildern ist die Art und Weise, wie dies geschieht. Während ein Mensch von sich aus vielleicht nur wenige Bilder taggt, bei denen er dies für interessant und angemessen hält, gibt es Dienste, die automatisch Personen in Bildern erkennen und diese, beispielsweise basierend

auf dem Adressbuch des Nutzers, automatisch markieren [43]. Dieses automatische Tagging wird zu einem kritischen Problem, wenn Nutzer ihre Schnappschüsse von einem mobilem Gerät wie einem Smartphone aus an einen Webdienst schicken und sich danach nicht weiter mit ihnen beschäftigen. Eine allgegenwärtige Verfügbarkeit von Handykameras und mobilem Internet kann Anwender dazu verleiten besonders viele Fotos zu machen und diese unbedacht zu veröffentlichen. In Verbindung mit Geo-Tagging, also dem Anfügen des Ortes, an dem ein Foto entstanden ist, können dann sowohl andere Nutzer, als auch der Dienstbetreiber Aufenthaltsorte und Aktivitäten der gezeigten Personen ermitteln. Über viele solcher Fotos hinweg lassen sich so Personen verfolgen und Profile von ihnen erstellen [40].
Um Schutz vor diesen Gefahren zu gewährleisten, müssen Wege gefunden werden, das Tagging von Nutzern auf ihren Wunsch einzuschränken oder erst mit ihrer Zustimmung zuzulassen. Dabei existieren verschiedene Anforderungen bei verschiedenen Plattformen. So gilt es in sozialen Netzwerken dafür zu sorgen, dass Nutzer mitbestimmen können, was mit Bildern geschieht, auf denen sie zu sehen sind. Allerdings ist dies kein triviales Problem, da in der Regel mehrere Personen auf einem Foto abgebildet sind. Es ist also eine kollaborative Regelfindung für den Umgang mit solchen Bildern notwendig [49].
Maßnahmen zum Schutz bei automatischem Tagging sind komplexer umzusetzen. Durch den Abgleich mit anderen Bildern beziehungsweise dem Adressbuch des Hochladenden können Erkennungen durchgeführt werden, derer sich dieser nicht bewusst ist. Gerade das schnelle Veröffentlichen über ein mobiles Gerät veranlasst dazu, die Verwendung des Bildes nur kurz zu bedenken. Ansätze auf diesem Gebiet reichen von kryptografischen Verfahren, die bei der Suche nach einer Person ansetzen [47], über Schutz beim Erkennen der Gesichter bis hin zu Maßnahmen, die Gesichter vor einer Erkennung schützen[2]. Allerdings beschäftigen sich bisher nur wenige Arbeiten mit dem Tagging von Bildern. Ein klarer Schwerpunkt liegt stattdessen auf Gesichtserkennung in Überwachungssystemen oder auf der gezielten Suche von Personen.

Dezentrale Netzwerke

Mobile Geräte sind immer häufiger mit Kurzstreckenfunktechnologien wie Bluetooth oder WLAN ausgestattet. Mit Hilfe dieser Technologien können spontane *dezentrale Netzwerke* aufgebaut werden, die verschiedenen Zwecken dienen können. Die Reichweite solcher Netzwerke variiert dabei sowohl auf Grund der verwendeten Übertragungsform, als auch mit der Zahl der Knoten. In vielen Fällen wird Blue-

[2] so genannte *De-Identification*

tooth für den Aufbau eines solchen Netzes bevorzugt. Es schont den Akku im Vergleich zu WLAN und ist durch seine geringe Reichweite ein weniger geeignetes Ziel von Mithörern oder Man-in-the-Middle-Angriffen (MITM-Angriffen). Die Zahl der Knoten eines solchen Ad-hoc Netzwerkes wird maßgeblich durch seinen Zweck bestimmt. In Diensten, die dem Austausch von Kontaktinformationen oder dem Mitteilen der eigenen Anwesenheit[3] dienen, ist eine Single-Hop-Architektur ausreichend und zweckdienlich. In einem solchen Netz sind zwei Geräte oder eine Gruppe von Geräten direkt miteinander in Verbindung. In Multi-Hop-Systemen stehen die Teilnehmer nicht zwingend direkt miteinander in Kontakt. Es besteht vielmehr ein Netz aus vielen mobilen Geräten, die ähnlich einem klassischen gerouteten Netzwerk miteinander in Verbindung stehen (Mobile Ad hoc Network, MANet).
Diese Netze bieten eine Vielzahl von möglichen Angriffen. Der bereits erwähnte MITM-Angriff wird in einem Multi-Hop-Netzwerk erheblich erleichtert, da ein Angreifer die Übertragung zwischen zwei Geräten nicht mehr abfangen muss, sondern sich lediglich an dem Netz beteiligen muss. Befindet sich in einem Multi-Hop-Netzwerk erst einmal ein bösartiger Knoten, kann er Daten zurückhalten (*Black Hole*), nur gezielt weiterleiten (*Selective Forwarding*) oder verfälschen. Es gilt also in dezentralen Netzwerken, ob mobil oder nicht, einen vertrauenswürdigen und möglichst kurzen Pfad zu finden. Eine wichtige Aufgabe ist in diesem Rahmen die Authentifizierung der einzelnen Knoten, damit ein Pfad als sicher eingestuft werden kann.
Da viele mobile soziale Netze mit dem Hintergrund einer dezentralen Architektur entworfen werden und mobile dezentrale Netze momentan ihre Hauptanwendung in eben diesen Netzen finden, betrachtet diese Arbeit die beiden Themengebiete geschlossen in Kapitel 4.

Soziale Netzwerke

Soziale Netzwerke und insbesondere mobile soziale Netzwerke verbinden viele der bereits vorgestellten Dienste und fassen sie auf einer gemeinsamen Plattform zusammen. Social-Augmented-Reality-Anwendungen kombinieren so den Standort des Nutzers mit Tags anderer Nutzer um Informationen in ein Video der Umgebung einzublenden [51]. In einem solchen Fall sind lediglich Maßnahmen zum Schutz des Standortes eines Nutzers zu treffen, wenn auch mit einem besonders hohen Anspruch an die Genauigkeit des Ergebnisses. Ein Schlüsselattribut von sozialen Netzwerken stellen jedoch die Verbindungen zwischen ihren Nutzern dar. Diese, häufig

[3] z. B. bei Missing Connection Diensten, siehe Kapitel 4

als Freundschaften in entsprechenden Listen dargestellten, Verbindungen können ein kritisches Problem für die Privatsphäre darstellen. Denn selbst, wenn ein Nutzer seine Freundesliste nicht veröffentlicht, ist es doch oft möglich seine Verbindung zu anderen mindestens zum Teil zu rekonstruieren [7]. Eine Möglichkeit eine solche Verbindung aufzubauen, bieten so genannte Missing Connection Dienste. Solche Dienste haben das Ziel Personen, die sich zu einem vorherigen Zeitpunkt persönlich begegnet sind, die Kontaktaufnahme zu ermöglichen. Dieses Vorhaben wird durch mobile Geräte mit Standortbestimmung und Kurzstreckennetzwerken erheblich vereinfacht. Zusätzlich kann die Privatsphäre mit den richtigen Maßnahmen erheblich besser geschützt werden als zuvor. Denn in ihrer ursprünglichen Form haben solche Angebote es erfordert, dass eine Person ihr Anliegen auf einer öffentlichen Plattform, beispielsweise auf einem Message Board oder im Radio, bekannt macht. Durch den automatischen Austausch von Schlüsseln mit Personen in der Nähe und das Speichern des dazugehörigen Standortes, kann die Kontaktaufnahme zumindest zu Nutzern dieses Dienstes automatisiert und geschützt werden.

Das größte Problem von aktuellen sozialen Netzwerken ist jedoch ihre zentrale Architektur und die Ansammlung großer Datenmengen über ihre Nutzer. Werden verschiedene solcher Netzwerke kombiniert, steigt die Menge der vorhandenen Informationen sogar noch an. Dies kann auf verschiedene Arten zu Problemen führen. Zum einen stellt eine zentrale Verwaltung des Netzwerkes einen ausgezeichneten Angriffspunkt für Angreifer dar, die sich persönliche Daten von einem oder allen Nutzern des Netzwerkes aneignen möchten. Zum anderen kann der Anbieter des Dienstes selber meist auch auf diese Daten zugreifen. Eine mögliche Lösung dieser Probleme stellt das Ausweichen in dezentrale Architekturen dar, seien sie mobiler Natur oder in Form von klassischen Peer-to-peer (P2P) Netzwerken mit unbeweglichen Knoten.

Doch nicht nur Angreifer von außen sind eine Gefahr für die Sicherheit persönlicher Daten. Auch andere Nutzer eines sozialen Netzwerks können bei mangelnder Vorsicht des Anwenders Informationen erhalten, die nicht jedem offen stehen sollten. Doch nicht nur mangelnde Vorsicht ist ein Grund hierfür. Dadurch, dass ein Nutzer viele verschiedene Informationen und Daten, von Statusupdates bis zu Bildern, veröffentlicht, kann es insbesondere in Verbindung mit mobilem Internet schnell dazu kommen, dass die Übersicht verloren geht. Der Nutzer muss also dabei unterstützt werden, kontextbasierte Einstellungen für die Zugriffsrechte seiner Daten zu bestimmen. Die Zugriffsrechte in aktuellen sozialen Netzwerken lassen meist drei Arten von Zugriff zu. Private Daten können nur vom Nutzer selbst eingesehen werden, während öffentliche Daten oder teilweise öffentliche Daten von allen Nutzern

beziehungsweise von Freunden des Nutzers eingesehen werden können. Hierauf basierend können zwei Ansätze verfolgt werden, die den Schutz solcher Daten verbessern. Zum einen können die Entscheidungen, welche Daten öffentlich oder privat sind, vom System unterstützt oder teilautomatisiert werden. Dies hat nicht nur den Vorteil eines erhöhten Sicherheitsgefühls, es kann zudem die Bedienbarkeit eines entsprechenden Dienstes stark erhöhen, indem Nutzer sich weniger in Einstellungen für ihren Datenschutz aufhalten müssen. Zum anderen kann die Zugriffssteuerung verfeinert werden. So können zum Beispiel Regeln definiert werden, die Personengruppen oder die Tageszeit berücksichtigen. Auf diesem Weg könnte ein Nutzer seinen Standort in seinem Profil veröffentlichen, ihn für seine Arbeitskollegen jedoch nur in der Woche und während der Arbeitszeit zugänglich machen. Um optimalen Schutz zu erreichen, können beide Ansätze kombiniert werden [48].

3 Standortbasierte Dienste

3.1 Erläuterung und Gefahren

Standortbasierte Dienste (Location Based Services, LBS) und die Erweiterung vorhandener Dienste durch Standortdaten sind durch die Ausbreitung von mobilen Endgeräten (insbesondere Smartphones) mit der Möglichkeit zur Standortbestimmung (sowohl durch GPS als auch durch Funknetzwerke) in den Fokus von Wirtschaft und Forschung gerückt. Für den Endnutzer bieten solche Dienste viele interessante Möglichkeiten. So gibt es zum Beispiel eine Vielzahl von Angeboten, die es ermöglichen Personen in der Nähe zu finden [27, 36] und selber gefunden zu werden. Auch Dienste, die auf Grund des aktuellen Aufenthaltsortes eines Nutzers beispielsweise Veranstaltungs- und Restaurantvorschläge machen, sind durch permanent mögliche Standortbestimmung realisierbar geworden [37]. Durch die Allgegenwärtigkeit dieser Dienste sind allerdings auch massive Einbrüche in die Privatsphäre der Nutzer möglich. Ein neugieriger Arbeitgeber könnte verfolgen, ob seine Angestellten sich dort aufhalten, wo sie sollten. Er könnte durch regelmäßige Arztbesuche oder Besuche bei Spezialisten durch den Beobachteten sogar auf seinen Gesundheitszustand schließen. Immer wieder zeigen sich Menschen zudem von der Möglichkeit, dass Verbrecher diese Daten nutzen könnten, verängstigt [11].
Im Folgenden werden die wichtigsten Ansätze zum Schutz der Privatsphäre beim Umgang mit Standortdaten vorgestellt. Dabei wird zwischen zwei verschiedenen Grundideen unterschieden: dem Verschleiern des genauen Standorts und der Anonymisierung des Nutzers. Alle Ansätze haben andere Anwendungsgebiete und Probleme, welche genauer betrachtet werden. Abschließend wird zudem das Problem der Nutzerfreundlichkeit näher dargestellt. Ein Schwerpunkt liegt hier auf der intuitiven Bedienbarkeit und der Abstraktion der vorgestellten Verfahren, da es für einen Nutzer einfach sein muss, die richtigen Einstellungen zu treffen, damit er optimal geschützt wird.

3.2 Verschleiern des genauen Standorts

Häufig erlaubt es die Natur eines LBS nicht, die Identität des Nutzers geheim zu halten. Die Gründe hierfür können verschiedenster Art sein. Allen voran ist es möglich, dass der Nutzer explizit veröffentlichen möchte, wo er sich gerade befindet [27]. Aber auch, wenn der Standort zum Beispiel zur Autorisierung genutzt werden soll (*Location Based Access Control, LBAC*), kann es notwendig sein, die Identität eines Subjekts zu kennen.
Abbildung 3.1 gibt eine Übersicht über verschiedene Architekturen, die beim Verschleiern eines Standortes verwendet werden können. Es kann grob zwischen Verfahren, die auf dem Endgerät des Nutzers ablaufen, und Verfahren, welche eine Form von Middleware verwenden, unterschieden werden. Die feinere Klassifikation der Architekturen basiert auf [52].
Gerade LBAC stellt beim Verschleiern von Standorten sehr strenge Anforderungen an die erhobenen Daten. So müssen die Daten ausschließlich von dem erwarteten Nutzer stammen. Verfahren, die auf dem Vermischen von verschiedenen Daten basieren (wie k-Anonymität [25]), scheiden hier also aus. Zudem muss für eine erfolgreiche Autorisierung ein bestimmtes Maß an Genauigkeit erhalten bleiben. Nutzer müssen also Einbußen bei dem Grad der Verschleierung in Kauf nehmen. Wie groß diese Einbußen sind und wie sich ein solches System umsetzen lässt, wird in [3] behandelt. In [4] werden die verwendeten Mechanismen für die allgemeine Verwendung spezifiziert und nochmals verfeinert. Der verwendete Ansatz basiert dabei auf der einfachen Erkenntnis, dass es kaum möglich ist einen genauen Standort zu bestimmen. Stattdessen lässt sich auf Grund von Ungenauigkeiten nur ein Gebiet angeben, in dem sich der Nutzer befinden kann. Dieses Gebiet ist (zweidimensional betrachtet, die Höhe ist bei den meisten Standortangaben vernachlässigbar) kreisförmig. Dies ist darin begründet, dass die Messung des Standortes einen Punkt zurück gibt, um den der eigentlich Standort in jede Richtung, der Ungenauigkeit entsprechend, abweichen kann. Die Ungenauigkeit eines Standortes A_i lässt sich also durch den Radius r_i beschreiben, wie in Abbildung 3.2 (a) dargestellt.

Verschleierung durch Verändern der Fläche

Dementsprechend lässt sich ein gemessener Standort durch eine wie folgt definierte Fläche $A_i = (x_i, y_i, r_i) \subseteq \mathbb{R}^2$ darstellen, wobei (x_i, y_i) einen Punkt im verwendeten Koordinatensystem angibt und r_i die Abweichung in Form des Radius dieser Fläche darstellt. A_i muss dabei zwei Bedingungen erfüllen. Die Position des Nutzers muss

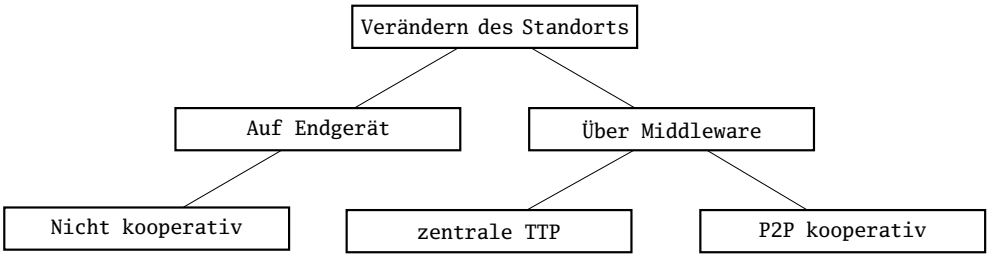

Abbildung 3.1: Grundlegende Architekturen für das Verschleiern eines Standortes

in ihr enthalten sein und muss mit gleicher Wahrscheinlichkeit an jedem Punkt in der Fläche liegen können. Zu diesem Standort lässt sich nun eine Relevanz $R_i = \frac{r_0^2}{r_i^2}$ bestimmen, wobei $R_i \in [0; 1]$. r_o stellt den Radius einer optimal bestimmten Fläche dar[1]. Diese Relevanz ist zunächst nur durch den Messfehler bestimmt. An dieser Stelle setzen die Entwürfe [3] und [4] an und lassen den Nutzer eine finale Relevanz R_f nach seinem Bedürfnis bestimmen. Welche Probleme für den Benutzer in der Bestimmung der Relevanz liegen, wird in 3.4 näher betrachtet.

Die Angabe der Relevanz bringt einige Vorteile mit sich. So ist ihre Verwendung nicht an den Kontext eines LBS gebunden. Eine vorgegebene Höchstgenauigkeit (bzw. Mindestfläche) erfüllt diesen Anspruch beispielsweise nicht, da je nach Verwendung unterschiedliche Mindestwerte notwendig sein können. So sind 100m in einem Fall mehr als ausreichend [2], während sie in einem anderen viel zu ungenau sind [3]. Wird dagegen ein auf Relevanz basiertes System verwendet, kann der Anbieter eines solchen, falls notwendig, einen optimalen oder ausreichenden Wert für r_0 festlegen. Im Gegensatz zu der Bestimmung von r_0 durch die Messgenauigkeit, ist dann ein $R_i > 1$ möglich, also eine Relevanz, die über dem Optimum liegt. Dies bedeutet, dass die Position eines Benutzers genauer bestimmt werden kann, als für den verwendeten Dienst notwendig. Mindestens genau so wichtig ist die Möglichkeit eine minimale Relevanz festzulegen, damit ein Dienst genutzt werden kann. Dies ist gerade bei Anwendungen wie LBAC von großer Bedeutung, wenn es notwendig ist ein gewisses Maß an Genauigkeit zu erhalten.

Wichtiger für den vorgestellten Ansatz ist jedoch die Tatsache, dass die gewünschte Relevanz nicht nur durch eine Änderung des Radius erreicht werden kann. Wenn der LBS es zulässt, kann so die Privatsphäre erheblich erhöht werden. In [4] wird ein System vorgestellt, das drei Arten von Operationen Unterscheidet:

[1] was optimal ist, hängt stark von der Anwendung und der verwendeten Technik ab
[2] z. B. Finden eines Restaurants
[3] z. B. LBAC

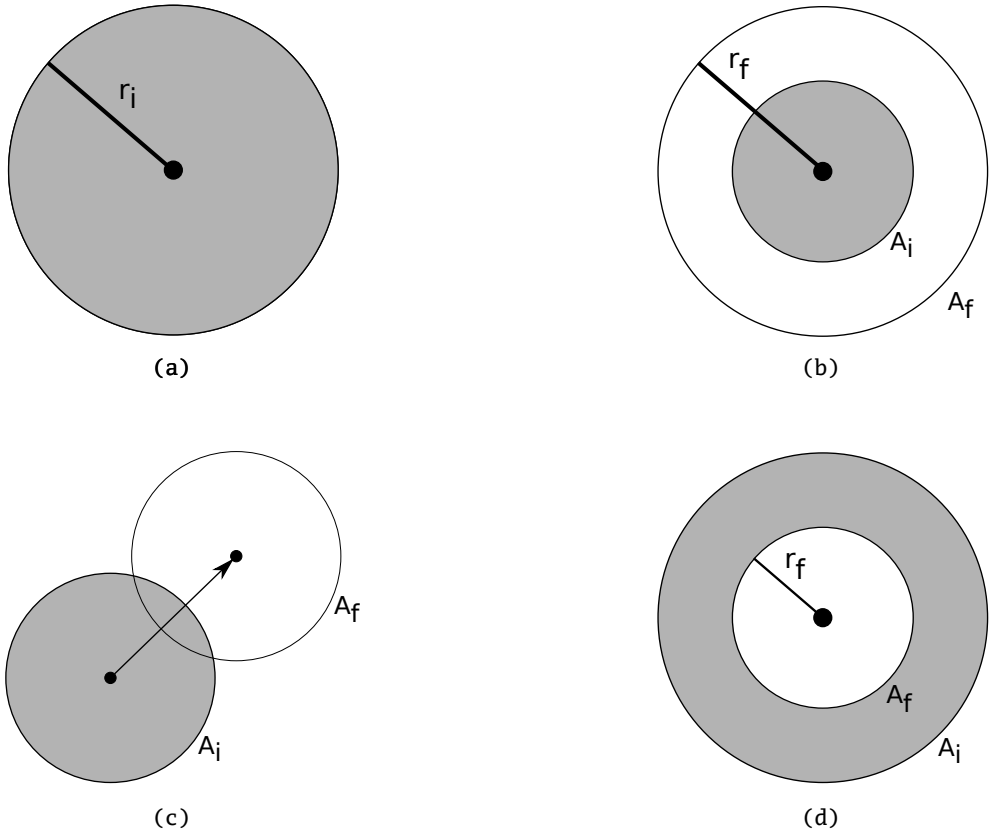

Abbildung 3.2: Fläche einer Messung und veränderte Flächen

- Vergrößern des Radius (*Enlarge*, E)
- Verkleinern des Radius (*Reduce*, R)
- Verschieben des Mittelpunkts (*Shift*, S)

Mit Hilfe dieser Operationen oder einer Kombination aus ihnen, dargestellt als Abfolge der Operationen [4], lässt sich eine Fläche A_f mit der gewünschten Relevanz R_f erzeugen. Das Vergrößern der Fläche (Abbildung 3.2 (b)), in der sich das Subjekt befinden kann, ist der naheliegendste Ansatz und wird in einer Vielzahl von Verfahren verwendet. Ihm liegt prinzipiell auch das später erläuterte Verfahren der Standort-k-Anonymität zugrunde. Der Ansatz folgt dem Gedanken, dass die Wahrscheinlichkeit, dass sich der Nutzer an einem Punkt in der Fläche aufhält, über die gesamte Fläche A_f gleichverteilt ist. Mit dem Vergrößern der Fläche sinkt damit auch die Wahrscheinlichkeit, den richtigen Ort auszuwählen. Einer entgegengesetzten (und prinzipiell kontraintuitiven) Idee folgt der R-Ansatz (Abbildung 3.2 (d)).

[4] z. B. ES für das Vergrößern des Radius gefolgt von einer Verschiebung

Zieht man die eben gemachten Überlegungen in Betracht, kann man zu dem Schluss kommen, dass durch eine kleine Fläche die Wahrscheinlichkeit eines Treffers steigen muss. Das ist allerdings nur zum Teil richtig. Das Verkleinern der Radius folgt der Überlegung, dass die wirkliche Position möglicherweise gar nicht mehr in der entstandenen Fläche enthalten ist. Denn während dies eine zwingende Bedingung für A_i ist, muss A_f sie nicht erfüllen. Durch eine Kombination mit der Shift-Operation kann die Wahrscheinlichkeit eines Ausschlusses der eigentlichen Position noch erhöht werden. S manipuliert als einziger der drei Ansätze nicht den Radius, sondern verändert x_i und y_i (Abbildung 3.2 (c)). Für diese Operation wird ein zufälliger Winkel θ gewählt und der Mittelpunkt der Fläche um die Distanz d verschoben, bis die gewünschte Relevanz erreicht ist. Dabei muss beachtet werden, dass $d < 2r_i$ gilt, damit die ursprüngliche und die entstandene Fläche nicht völlig unabhängig voneinander sind, sondern sich wenigstens minimal schneiden. Die erhöhte Privatsphäre resultiert hierbei sowohl aus dem Verkleinern des Schnittes beider Flächen, als auch aus der Tatsache, dass die Wahrscheinlichkeit der tatsächlichen Nutzerposition innerhalb der entstandenen Fläche nicht mehr gleichverteilt ist.
Auf Grund des zufällig gewählten Winkels und der damit kaum nachvollziehbaren Verschiebung ist ein Shift kaum von einem Angreifer rückgängig zu machen, stellt also einen sehr guten Schutz dar. Die folgenden Überlegungen zu möglichen Schwächen der verschiedenen Vorgehen lassen ihn deshalb als einzelne Operation außen vor. In Zusammenspiel mit den anderen Verfahren führt ein Shift jedoch zu einigen erwähnenswerten Effekten. E und R lassen sich zwar ebenfalls nur begrenzt entschleiern, ein Angreifer kann aber sogar ohne Kenntnis des verwendeten Operators eine Präferenz auf Grund der möglichen Veränderungen entwickeln. Diese Erkenntnis folgt, wenn man sich die möglichen Umkehr-Operationen der beiden Ansätze vor Augen führt. Um die Relevanz einer durch R erzeugten Fläche zu erhöhen, liegt es nahe, den Radius dieser zu vergrößern, um die Ursprungsgröße möglichst genau anzunähern. Entsprechend bietet es sich an, eine vergrößerte Fläche wiederum zu verkleinern. Zieht man nun allerdings eine Kombination aus Shift und Enlarge in Betracht (SE oder ES), ist es möglich, dass A_i, also die Originalfläche, nicht komplett von der erzeugten Fläche A_f eingeschlossen, sondern lediglich von ihr geschnitten wird. In einem solchen Fall bietet es sich für einen Angreifer an, r_f nicht zu verringern, sondern den Radius ebenfalls zu vergrößern[5]. Ein kleiner Radius würde hier die Schnittfläche zwischen A_i und der entstehenden Fläche und somit die Wahrscheinlichkeit die tatsächliche Position zu finden verringern. Wird der Radius hingegen vergrößert, wird ein größerer Teil der ursprünglichen Fläche überdeckt.

[5] Für RS oder SR ändert sich das vorgehen nicht

Kennt ein Angreifer die verwendete Operation, hilft ihm dies nur begrenzt. Ist A_f auffällig klein, so kann er auf die Verwendung von R schließen und den Radius für eine Rekonstruktion vergrößern. Um einen durch Vergrößern verschleierten Standort zu bestimmen, ist jedoch zudem das Wissen notwendig, ob eine Verschiebung durchgeführt wurde, damit die passende Reaktion gewählt werden kann. Die Vergrößerung stellt in Verbindung mit einem (möglichen) Shift also einen stärkeren Schutz dar. Damit ein Angreifer nicht sicher wissen kann, welche Operation zur Verschleierung des Standorts verwendet wurde, kann diese zufällig aus einer vorher definierten Menge gewählt werden. Dies kann jedoch zu einer leichteren Präferenzfindung für den Angreifer beitragen, wenn man die eben gemachten Überlegungen heranzieht. Denn in zwei von drei Fällen (E oder R mit Schnitt der Ursprungsfläche) führt ein Vergrößern des Radius zu einer erhöhten Relevanz, während nur im Falle des vollständigen Einschlusses der Originalfläche durch die vergrößerte Fläche A_f eine Verkleinerung ratsam ist. Doch selbst, wenn ein Angreifer weiß, wie A_f erzeugt wurde, kann er diesen Vorgang nur bedingt rückgängig machen. Ohne Kenntnis von A_i, bei der eine Entschleierung nicht mehr notwendig wäre, kann ein Angreifer nicht wissen, wie stark er den Radius verändern muss um eine höhere Relevanz zu erzielen. Ist die Änderung zu gering, geht potentielle Relevanz verloren. Ist sie zu groß, kann es sogar sein, dass sich die Relevanz im Vergleich zu A_f verschlechtert. Kennt der Angreifer jedoch die Technik, mit der der ursprüngliche Standort ermittelt wurde, und die ungefähren Umstände, unter denen dies geschah[6], könnte er abschätzen, wie genau die ursprüngliche Messung war und daraus einen Rahmen für die Vergrößerung beziehungsweise Verkleinerung des Radius ermitteln. Das Verfahren der Standortverschleierung durch eine Kombination von Operationen bietet also unter der Voraussetzung, dass es korrekt verwendet wird, einiges an Schutz. Dabei ist es wichtig zu beachten, dass das Verfahren nur für eine einmalige Standortbestimmung zu verwenden ist. Es ist nicht geeignet um Bewegungen zu verfolgen. Der Versuch dies zu tun, würde zwei Probleme aufwerfen. Erstens würde die Messung nur mit begrenzter Genauigkeit die zurückgelegte Strecke wiedergeben, da durch zufällige Shifts Sprünge entstehen können. Zweitens wäre es einem Angreifer dennoch sehr einfach möglich Verhaltensmuster und gegebenenfalls sogar den Standort des Nutzers zu einem bestimmten Zeitpunkt nachträglich zu errechnen, indem er, je nach zeitlicher Auflösung der Messungen, die Schnittpunkte der ermittelten Flächen bestimmt.

Das Verfahren gleicht seine leichten Einbußen im Bereich der Sicherheit durch einige Vorteile für den Nutzer aus. So kann bei diesem Ansatz vollständig auf Midd-

[6] z. B. Aufenthalt in Gebäude

leware verzichtet werden. Während viele Ansätze darauf aufbauen, dass eine Verschleierung oder Anonymisierung von einem vertrauenswürdigen Dritten durchgeführt wird, können die unbearbeiteten Rohdaten permanent in den Händen des Nutzers bleiben. Die relativ einfachen Verschleierungsoperationen können auf Grund des geringen Rechenaufwands problemlos direkt auf den meisten heute erhältlichen Smartphones und vergleichbaren Geräten ausgeführt werden. Das bedeutet nicht nur, dass die Daten nicht weitergegeben werden müssen und somit sehr viel besser vor, vor allem, ehrlichen, aber neugierigen Drittanbietern geschützt sind, es heißt auch, dass ein LBS möglicherweise nicht einmal mitbekommt, dass er mit veränderten Daten arbeitet. Die geringe Unterscheidbarkeit der veränderten und einer gemessenen Fläche trägt zusätzlich dazu bei. Der Dienstanbieter muss also nicht einmal damit rechnen verschleierte Standortdaten zu erhalten. Das selbe gilt natürlich auch für einen Angreifer, der in diesem Fall gar nicht erst versucht die ursprüngliche Position zu bestimmen, wodurch die Privatsphäre des Nutzers besser geschützt ist, als es mit jeder Verschleierung möglich wäre. Ein weiterer Vorteil bietet sich durch das Durchführen der Veränderungen auf dem eigenen Gerät. Der Nutzer kann ohne Mehraufwand die zu erreichende Relevanz R_f für eine beliebige Zahl an Diensten oder Kontakten selber wählen. Er kann beispielsweise guten Freunden genauere Positionsangaben schicken als flüchtigen Bekannten oder einem Restaurantfinder. Schlussendlich ist zu bemerken, dass die Verschleierung vollständig unabhängig von anderen Nutzern durchgeführt werden kann. Im Gegensatz zu Standort-k-Anonymität und ähnlichen Verfahren kann sie also auch in Bereichen mit wenigen Nutzern ohne Einbußen verwendet werden. Die Ansätze [3] bzw. [4] stellen ein gutes Beispiel für eine der drei grundlegenden Architekturen dar [52]. Sie beschreiben einen nicht kooperativen Weg um den eigenen Standort zu schützen.

Standort-k-Anonymität

Aus dem Bereich der Trusted-Third-Party(TTP)-Ansätze, welche die zweite mögliche Architektur darstellen, stammt die Standort-k-Anonymität. Auf einer TTP basierende Verfahren bieten auf Grund der Vielzahl verfügbarer Informationen bessere Möglichkeiten zum Schutz der Privatsphäre eines Nutzers. Sie bringen gleichzeitig allerdings auch Nachteile mit sich. Allen voran steht der vertrauenswürdige Dritte selbst. Es stellt nicht nur einen erheblichen Mehraufwand dar, eine Infrastruktur für die Anonymisierung aufzubauen. Damit diese vertrauenswürdig ist, sollte sie zudem von einem unabhängigen Anbieter zur Verfügung gestellt werden. Denn einem LBS-Anbieter, dem man nur seine verschleierten Daten anbieten möchte, möchte man

wahrscheinlich auch *nur* diese anbieten.
k-Anonymität ist ein Verfahren, das hauptsächlich zum Anonymisieren von sensiblen Daten, wie z. B. medizinischen, verwendet wird [20]. Es basiert darauf, dass veröffentlichte Datensätze mindestens k verschiedenen Individuen zugeordnet werden können und somit nicht festgestellt werden kann, zu wem die Daten ursprünglich gehörten. Standort-k-Anonymität basiert auf dem selben Ansatz. Allerdings soll hier nicht der Nutzer anonymisiert werden[7], sondern sein Standort, wodurch sich die genannte Definition leicht ändert. Bei der Standort-k-Anonymität soll die Identität des Nutzers erkennbar bleiben, sein Standort jedoch soll nicht von $k - 1$ anderen Standorten zu unterscheiden sein [25]. Eine Vielzahl von Arbeiten [52, 20, 15] hat sich sowohl mit der k-Anonymität, als auch mit Standort-k-Anonymität befasst und dabei verschiedene Schwerpunkte betrachtet. Im Folgenden soll eine Auswahl von Prinzipien und Verfahren im Bezug auf Standort-k-Anonymität als Beispiel einer TTP-Architektur vorgestellt werden.
Um die Voraussetzung der k-Anonymität zu erfüllen, muss ein Algorithmus aus der Menge aller bekannten Standortdaten mindestens k Standorte auswählen und aus diesen eine Fläche generieren, in der alle gewählten Koordinaten enthalten, aber nicht mehr erkennbar sind. Wenn diese Daten jedoch willkürlich gewählt werden, ergeben sich starke Probleme für die Qualität des Dienstes (Quality of Service, QoS). Denn je weiter die einzelnen Standorte voneinander getrennt sind, desto größer wird die resultierende Fläche ausfallen. Das gleiche kann passieren, wenn k bei zu geringer Nutzerdichte zu groß gewählt wird. Um dies zu verhindern sollten zwei Maßnahmen getroffen werden. Zuerst sollte k entsprechend des Verlangens nach Privatsphäre eines Nutzers gewählt werden. Zudem sollte eine zusätzliche maximale räumliche Ausdehnung festgelegt werden, die nicht überschritten werden darf. Da es bei ungünstiger Wahl dieser Parameter[8] zu sehr langen Wartezeiten kommen kann, die die QoS weiter verringern und manche Dienste gegebenenfalls sogar unbenutzbar machen können, sollte zudem eine maximale zeitliche Ausdehnung festgelegt werden, nach deren Ablauf die Nachricht verworfen und eine neue Anfrage gestellt werden muss. Sind diese Voraussetzungen getroffen, gilt es k zusammenhängende Nachrichten zu finden und aus ihnen eine Fläche zu erzeugen. Die Schwierigkeit bei diesem Vorgehen liegt vor allem darin die kleinste Fläche zu finden, die mindestens k Nachrichten einschließt, dabei jedoch die Parameter jeder einzelnen Nachricht beachtet und deren maximale Ausdehnung nicht überschreitet. Das größte Problem erwächst

[7] Dies ist zwar möglich, zunächst soll jedoch nur die Verschleierung des Standortes betrachtet werden
[8] Für weitere Probleme bei der Parameterwahl siehe 3.4 *Problem der Nutzerfreundlichkeit*

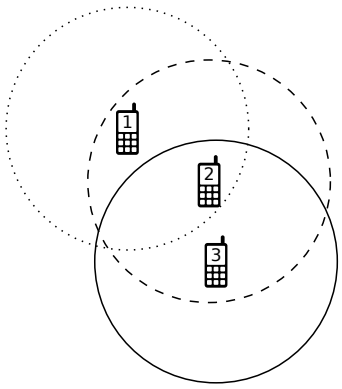

Abbildung 3.3: Drei mobile Geräte mit ihrer maximalen verschleierten Flächen.

dabei aus der Tatsache, dass nicht vorhersehbar ist, wie groß die maximalen Flächen und minimalen k der einzelnen Nachrichten sind. Um diese Aufgabe dennoch erfüllen zu können, wird in [25] ein middlewarebasierter Algorithmus vorgestellt, der eine solche Fläche findet und es dabei zulässt, dass jeder Nutzer seine eigenen Parameter festlegt. Dieser *Clique Cloak Algorithmus* soll hier exemplarisch die Funktionsweise von Standort-k-Anonymität darstellen.

Verschleierung durch Verändern der Fläche

In dem vorgestellten Ansatz treten Nutzer nicht direkt mit einem LBS in Kontakt. Die Kommunikation findet über einen *Anonymisierungsserver* statt. Damit ein Angreifer Standortdaten nicht schon bei ihrer Übertragung abfangen kann, findet diese mit gängigen Methoden verschlüsselt statt. Auf Grund der relativ geringen ausgetauschten Datenmenge sollte eine asynchrone Verschlüsselungsmethode ausreichend sein, da zum Beispiel ein SSL Handshake mehr Zeit kosten könnte, als durch die Verwendung eines solchen Verfahrens gewonnen würde. Nach dem Entschlüsseln einer Nachricht muss zunächst sichergestellt werden, dass Identifizierungsmerkmale wie eine IP-Adresse oder eine Gerätekennung entfernt werden, damit eine Verfolgung über diese Daten nicht mehr möglich ist. Daraufhin wird die Nachricht an eine sogenannte *Message Pertubation Engine* auf dem Server übergeben, die sie verarbeitet oder, falls notwendig, zunächst zwischenspeichert. Ankommende Nachrichten werden dazu in einer einfachen FIFO-Warteschlange abgelegt und nacheinander bearbeitet. Hier setzt der eigentliche Algorithmus zum Finden einer passenden verschleiernden Fläche an. Auf dem Server sind alle Standorte in einem ungerichteten Graphen $G(S, E)$ organisiert, in den die aktuell zu verarbeitende Nachricht eingefügt wird.

Jede Nachricht m ist ein Knoten dieses Graphen. Aus den in ihr spezifizierten maximalen räumlichen Ausdehnungen wird ihr zudem ein räumlicher Rahmen berechnet. Eine Kante zwischen zwei Knoten existiert genau dann, wenn sich beide Nachrichten in dem Rahmen der jeweils anderen Nachricht befinden. Durch diese Bedingung ist sichergestellt, dass alle benachbarten Knoten zum Erzeugen der verschleierten Fläche verwendet werden können, ohne dass die Grenzen der Ausdehnung überschritten werden. Siehe dazu auch Abbildung 3.3. Die Geräte 1 und 2 sowie Geräte 2 und 3 in der Abbildung können eine gemeinsame Fläche bilden. Zwischen den Geräten 1 und 3 liegt eine zu große Distanz. Um die gesuchte Fläche zu erzeugen, wird nur die Menge U aller Nachbarn vom m betrachtet, die ein k kleiner oder gleich dem der aktuellen Nachricht besitzen. Nachrichten, die ein größeres k benötigen, können auf Grund der Natur der durchgeführten Suche nicht mit anonymisiert werden. Nachrichten, die nicht mit m benachbart sind, können nicht zur Anonymisierung von m verwendet werden, da die räumliche Distanz zwischen beiden Standorten zu groß ist, um die Bedingungen wenigstens einer Nachricht zu erfüllen. Nachrichten auf die einer dieser beiden Punkte zutrifft, können somit von der weiteren Suche ausgeschlossen werden. Aus den Nachrichten in U muss nun eine Clique gefunden werden, die mindestens k Elemente (inklusive m) enthält. Da jede Kante bedeutet, dass die verbundenen Knoten eine gemeinsame Fläche erzeugen können, ist in dieser Clique, falls sie existiert, sichergestellt, dass alle enthaltenen Nachrichten in den Grenzen aller anderer Nachrichten liegen und somit auch die erzeugte Fläche diese Bedingung erfüllt. U wird bei der Suche nach einer Clique so lange durchlaufen und ungeeignete Nachrichten entfernt, bis sich keine Änderungen mehr ergeben. Ungeeignet sind Nachrichten dann, wenn sie weniger als $k-1$ Nachbarn in U haben. Mit diesen Nachrichten kann keine Clique mit mindestens k Elementen gefunden werden, da sie selber mit mindestens $k-1$ und der aktuellen Nachricht verbunden sein müssten. Sie werden deshalb aus U entfernt. Bei den nun verbleibenden Nachrichten ist sowohl sichergestellt, dass sie die räumlichen Bedingungen erfüllen, als auch, dass sie über genügend Nachbarn verfügen, um eine k-elementige Clique formen zu können. Mit einem beliebigen Algorithmus werden nun so lange Cliquen in U gesucht, bis eine Clique gefunden wird, die mindestens $k-1$ Knoten und m enthält. Existiert keine ausreichend große Clique in der die Ursprungsnachricht enthalten ist, verbleiben alle Nachrichten im Graphen. Nachrichten, die länger als erlaubt auf dem Server lagen, werden verworfen und das Verfahren mit der nächsten Nachricht wiederholt. Wurde eine Clique gefunden, werden die Standorte in den enthaltenen Nachrichten anonymisiert, indem eine Fläche erzeugt wird, die die Standorte aller Nachrichten umfasst. Die gefundenen Nachrichten werden dann in zufälliger Rei-

henfolge an den LBS weitergeleitet um Rückschlüsse aus ihren Ankunftszeiten zu vermeiden. Schlussendlich werden die verwendeten und abgelaufene Nachrichten aus dem Graphen entfernt und die nächste Nachricht kann verarbeitet werden.
Die Suche nach geeigneten Nachrichten beziehungsweise einer entsprechenden Clique ist ein kritischer Abschnitt in diesem Verfahren. Werden die Rahmen von Nachrichten zu groß gewählt, muss ein sehr großer Teilgraph durchlaufen werden, da viele Nachrichten zu m benachbart sind. Entsprechend können recht große Verzögerungen entstehen, die je nach Art des angesprochenen Dienstes mehr oder weniger kritisch sind, auf jeden Fall aber eine Verschlechterung der Dienstgüte bedeuten. Ebenso kann ein zu groß gewähltes k dazu führen, dass es aufwändiger wird, eine geeignet große Clique zu finden, und die Erfolgschancen stark sinken. Dieses Problem kann jedoch durch die Angabe eines maximalen k nur begrenzt kontrolliert werden, da die Grenze für k stark von der Zahl der aktuell im System befindlichen Nachrichten abhängt[9]. Die benötigte Mindestzahl von Nutzern stellt eine der Hauptschwierigkeiten dieses und ähnlicher Verfahren dar. Um diese Probleme abzuschwächen, wurden schon in [25] erste Ansätze vorgestellt.
Um die Verweildauer von Nachrichten im System zu verringern und so die QoS erheblich zu erhöhen, ist es wünschenswert möglichst viele Nachrichten gleichzeitig zu anonymisieren. Eine Möglichkeit dies zu erreichen, ist das Parallelisieren der Suche in unabhängigen Teilgraphen. Während dies theoretisch durchaus möglich ist, müsste überprüft werden, unter welchen Bedingungen genügend Nachrichten und Teilgraphen vorhanden sind, damit sich auch in der praktischen Umsetzung ein Vorteil aus der parallelen Verarbeitung ergibt. Ein praktisch wahrscheinlich effektiverer Ansatz setzt bei der Wahl des verwendeten k an. Während der vorgestellte ursprüngliche Clique Cloak Algorithmus das k der aktuell verarbeiteten Nachricht verwendet, können möglicherweise mehr Nachrichten gleichzeitig erfasst werden, wenn der größte k-Wert in der Nachbarschaft von m verwendet wird. Da dies jedoch auch zu Verzögerungen führen kann, wenn das größte k wiederum zu groß ist, ist es sinnvoll zunächst alle Nachbarn in die Suche nach einer Clique einzubeziehen (ursprünglich wurden Nachrichten mit größerem k als m ignoriert) [52] und beginnend bei dem höchsten k nach einer Clique zu suchen. Wenn keine ausreichend große Clique gefunden wird, wird das nächst niedrigere k verwendet, bis entweder ein passender Teilbaum gefunden wurde oder k kleiner ist als von der ursprünglichen Nachricht angegeben. Dieses Vorgehen führt zwar möglicherweise zu einer Erhöhung der Zahl benötigter Suchen, kann aber gerade bei Nachrichten mit kleinem k die Verweildauer einiger Nachrichten gleichzeitig verringern. Jedoch ist auch dieser Ansatz stark

[9] Der Clique Cloak Algorithmus ist laut [25] performant bis zu einem k von ca. 10

abhängig von der Zahl der Nutzer des Systems. Damit Nachrichten schneller in den Graphen eingepflegt werden können, kann es bei wenigen vorhandenen Nachrichten hilfreich sein diese zunächst zu sammeln und die eigentliche Suche erst auszuführen, wenn der Graph im entsprechenden Umfeld eine gewisse Dichte erreicht hat. Es sollte hierbei allerdings bedacht werden, dass Nachrichten eine maximale Verweildauer besitzen, und ein Anonymisierungsversuch durchgeführt werden, bevor diese verfallen.

Dezentrale Architektur für k-Anonymität

Nach der benötigten Mindestnutzerzahl stellt der zentrale Server, der die Nutzerdaten verarbeitet, einen weiteren großen Schwachpunkt des Verfahrens dar. Ein einzelner Server wirkt sich nicht nur als Flaschenhals aus, der bei zu vielen Nachrichten die Dienstgüte aller Teilnehmer verschlechtern kann. Er stellt zudem auch einen Single Point of Failure dar. Wenn der zentrale Server ausfällt, legt dies den gesamten Dienst lahm. Gravierender noch ist es, wenn der Server Ziel eines Angriffes wird. Wenn dieser Erfolg hat, kann der Angreifer die Standortdaten aller mit dem System verbundenen Nutzer ermitteln. Das selbe gilt für den Fall, dass der Betreiber der Middleware nicht so vertrauenswürdig ist wie angenommen. Um diese Probleme zu beseitigen, wird in [59] ein verteiltes System zur k-Anonymisierung vorgestellt. Anstatt einem zentralen werden hier mehrere Server[10] eingesetzt, die jeweils nur eine Teilmenge aller Standorte kennen. Jeder Nutzer teilt seine Standortinformationen genau einem dieser *Location Broker* seiner Wahl mit. Ein weiterer Server vermittelt zwischen den einzelnen Brokern und ermittelt, ohne Nutzerdaten zu kennen, ob sich mindestens k Nutzer in der Nähe des Anfragenden befinden. Eine dritte Serverart bietet Informationen über die vorhandenen Location Broker an. Einige Server sind kombinierbar, es ist allerdings darauf zu achten, dass Nutzerdaten und Standorte nur auf dem eigens dafür vorgesehenen Location Broker vorhanden sind. Dieser muss also alleine betrieben werden. Möchte ein Nutzer seinen Standort verschleiern, muss er zunächst einen geeigneten Location Broker finden. Dazu lädt er die gesamte Liste verfügbarer Broker herunter, um seinen Standort nicht an den entsprechenden Server senden zu müssen, und meldet sich bei einem geeigneten Server an. Anders als bei dem zuvor vorgestellten Ansatz kann ein Nutzer die verschleierte Fläche selbstständig berechnen ohne die Standorte anderer Nutzer kennen zu müssen. Dazu wählt er zunächst eine Fläche aus, dessen Größe er für geeignet hält, um k andere Nutzer zu umfassen. Daraufhin fragt er bei den für diese Fläche zuständigen Location Brokern

[10] z. B. einer für jeden Mobilfunkanbieter

an, wie viele Nutzer für diese Fläche registriert sind[11]. Er erhält die Ergebnisse dieser Anfrage in verschlüsselter Form. Dies geschieht, damit kein Nutzer weiß, wie viele andere Nutzer sich um ihn herum aufhalten, da dies als möglicher Angriffspunkt genutzt werden könnte. Die Verschlüsselung muss es dabei ermöglichen die Summe der verschlüsselten Zahlen zu bilden, ohne dass diese entschlüsselt werden[12]. Die immer noch verschlüsselte Summe wird nun an einen weiteren Server geschickt, der sie entschlüsselt und dem anfragenden Nutzer mitteilt, ob das Ergebnis seinen Ansprüchen genügt, damit kein Broker die tatsächliche Zahl der Nutzer in der Umgebung kennen kann. Befinden sich nicht genug andere Teilnehmer in der gewählten Fläche, kann der Nutzer sie vergrößern und eine neue Anfrage stellen. Dieses Verfahren entfernt effektiv das Problem des alleinigen Servers. Fällt beispielsweise ein Location Broker aus, kann ein Nutzer problemlos auf einen benachbarten Server ausweichen. Allerdings ist ein erheblicher Mehraufwand im Aufbau der Infrastruktur notwendig, der unter Umständen sogar zwischen verschiedenen Anbietern koordiniert werden muss. So ist es vor allem notwendig ein Protokoll zu finden, auf dessen Basis sich die Server einzelner Anbieter verständigen und Standorte sowie Statusmeldungen austauschen können. Zudem ist beim Erzeugen der verschleierten Fläche zu beachten, wie diese erzeugt wird. Wird die Fläche durch einfaches Vergrößern erzeugt, ist sie nicht nur relativ einfach ihrer ursprünglichen Größe anzunähern. Es müsste außerdem überprüft werden, ob nicht das in [4] vorgestellte Verfahren effektivere Ergebnisse liefert, die vor allem ohne Infrastruktur berechnet werden könnten. Eine Kombination beider Verfahren wäre jedoch durchaus denkbar.

Die dritte grundlegende Architektur nach [52] stellt die Peer-to-Peer-kooperative Verschleierung dar. Sie vertraut darauf, dass Nutzer untereinander Daten austauschen um ihre Standorte für einen LBS zu verschleiern. Verfahren dieser Art setzen allerdings voraus, dass alle Nutzer sich gegenseitig vertrauen, und machen es einem Angreifer entsprechend einfach Standorte zu ermitteln. Sie seien hier nur der Vollständigkeit halber erwähnt.

3.3 Anonymisierung und Bewegung

Nicht immer muss die Identität eines Nutzers bekannt sein, um Dienste nutzen zu können. So ist es zum Finden eines Restaurants beispielsweise nicht zwingend not-

[11] Selbst, wenn die Anfrage über mehrere Broker stattfindet, muss der Nutzer nur an einem registriert sein

[12] siehe [59] für Beispiele

wendig persönliche Daten (außer dem Standort natürlich) preiszugeben. Fordert der verwendete standortbasierte Dienst (LBS) jedoch eine Identifikation, kann es ausreichend sein ein Pseudonym anzugeben. Im ersten Fall erweist sich die Verfolgung eines Nutzers als recht schwierig. Einzelne Anfragen sind einer Person maximal über die Ursprungsadresse, die sich durchaus ändern kann, oder eine Gerätekennung, die vor dem Versenden einer Nachricht gegebenen Falls entfernt oder zur Verschleierung geändert werden kann, möglich und damit leicht zu umgehen. Wird ein Pseudonym verwendet[13], kann dies jedoch eine Verfolgung sehr einfach machen. Vor allem, wenn regelmäßig das selbe Pseudonym verwendet wird, kann die Identität des Nutzers ohne Probleme ermittelt werden. Dies kann zum Beispiel geschehen, indem ein Pseudonym verfolgt wird und häufige Aufenthaltsorte bestimmt werden. Hält sich eine Person den Großteil eines Tages am selben Ort auf, kann davon ausgegangen werden, dass sie dort arbeitet. Ihr nächtlicher Standort ist mit großer Wahrscheinlichkeit ihr Zuhause. Gerade, wenn ein Angreifer gezielt nach einer Person sucht, sind ihm diese Standorte mit großer Sicherheit bekannt und es ist ihm möglich sein Ziel aus allen Pseudonymen zu erkennen. Es ist möglich, dass ein Nutzer sein Pseudonym für jede Anfrage zufällig generiert. Dennoch lässt sich möglicherweise seine Identität aufdecken. Kennt ein Angreifer Ursprung und Ziel der Bewegung einer Person, kann er er über die An- und gegebenenfalls Abmeldung von Pseudonymen auf dem Weg zwischen den beiden Punkten Rückschlüsse auf deren Verwender ziehen. Um die Identität eines Nutzers zu schützen, muss also seine Bewegung verschleiert werden. Im Folgenden werden zwei Ansätze vorgestellt, die dies mit unterschiedlichen Hintergedanken zu erreichen versuchen.
Beide Ansätze versuchen dazu, ähnlich der k-Anonymität, den Nutzer in der Menge untertauchen zu lassen. Sie sind jedoch mit verschiedenen Anforderungen entstanden. Während der erste Ansatz, ursprünglich vorgestellt in [6], versucht die Bewegung eines Nutzers zwischen der Verwendung zweier Dienste zu verschleiern, ist das zweite Verfahren darauf ausgelegt die Privatsphäre beim expliziten Aufzeichnen von Bewegungen zu schützen [30].

Mix Zones

In [6] ist ein Verfahren beschrieben, das auf so genannten *Mix Zones* aufbaut, in denen Nutzer nicht voneinander zu unterscheiden sind. Obwohl dieser Ansatz schon 2003 veröffentlich wurde, soll er hier grundlegend erläutert werden, da auch neue-

[13] Wenn man so möchte, kann auch die Adresse eines Gerätes als Pseudonym angesehen werden. Eine Verwendung ist also durchaus wahrscheinlich

re Verfahren das Prinzip der Mix Zones verwenden [21]. Ziel des Ansatzes ist es[14] einem Nutzer den Wechsel zwischen verschiedenen Standorten und den dort angebotenen Diensten zu ermöglichen, ohne dass seine Bewegungen verfolgt werden können. Für seine Verwendung wird eine gewissen Infrastruktur vorausgesetzt, die vor allem vertrauenswürdig sein muss. Es wird eine Middleware angenommen, bei der sich Nutzer und Dienste gleichermaßen anmelden. Ein Dienst gibt dabei zudem eine Region Of Interest (ROI), also einen räumlichen Bereich, in dem ihr Angebot genutzt werden kann, an. So kann ein Nutzer, sollte er dies wünschen, zum Beispiel über aktuelle Sonderangebote informiert werden, wenn er sich in der Nähe eines Elektronikmarktes aufhält. Um sein Interesse bekannt zu machen, kann er sich bei der Middleware für einen Dienst oder, je nach Implementierung und Vorlieben, eine Gruppe von Diensten anmelden. Betritt er nun die ROI einer entsprechenden Applikation oder verlässt sie, wird dies mit einem Pseudonym bei dem Dienst bekanntgemacht. Das Pseudonym dient dazu rein als Antwortadresse für den Dienst und wird für jeden Eintritt in eine ROI zufällig generiert, damit eine Verfolgung über dieses nicht möglich ist. Es ist allerdings ohne weiteres möglich eine Person trotz verschiedener Pseudonyme alleine durch ihre Bewegung, in diesem Fall die An-/Abmeldung bei verschiedenen Diensten, zu verfolgen [31]. Aus diesem Grund werden Mix Zones eingeführt. Dies sind die Regionen zwischen verschiedenen Diensten. Genauer ausgedrückt sind Mix Zones die Flächen maximaler Größe, in denen kein Nutzer bei einem Dienst angemeldet ist. Ähnlich einer Fußgängerzone können Nutzer hier nach dem Verlassen eines Ladens (bzw. einer ROI) in der Menge untertauchen, bis sie den nächsten Laden betreten. Der Nutzer wird versteckt, indem die Standorte von Personen in der Mix Zone nicht gespeichert beziehungsweise weitergegeben werden. Somit kann jeder Nutzer, der sich an einem Dienst anmeldet, aus der Menge der Nutzer in dem gemischten Bereich stammen. Es lässt sich nicht mehr eindeutig feststellen, welcher Nutzer genau diese Aktion durchgeführt hat. Um dies realisieren zu können, muss die Zone gewissen Ansprüchen genügen. Eine wichtige Eigenschaft ist die Größe des gemischten Bereichs. Die Middleware führt periodische Updates der Standorte ihrer Nutzer durch. Ist die Mix Zone so groß gewählt, dass ein Nutzer sie nicht innerhalb eines Zeitschlitzes durchqueren kann, könnte ein Angreifer einen Nutzer, der sich an einem neuen Dienst anmeldet, identifizieren, indem er ausschließt, dass andere Nutzer rechtzeitig zu diesem Standort gelangen konnten. Dies ist vor allem dann möglich, wenn die Mix Zone nur eine geringe Nutzerzahl aufweist. Ein zweiter wichtiger Punkt ist also die aktive Nutzerzahl. Wie im echten Leben ist es einfacher in einer großen Nutzermenge unerkannt zu bleiben. Je größer die Men-

[14] in diesem Fall

ge von Nutzern ist, desto schwerer ist es, ähnlich der k-Anonymität, die Bewegung eines Nutzers durch die Mix Zone zu verfolgen. Um zu verhindern, dass eine Verfolgung besser möglich ist als gewünscht, kann dem Nutzer ein Überblick über die durchschnittliche Zahl anderer Personen in der Mix Zone gegeben werden. Anhand dieser Werte kann dieser dann entscheiden, ob er sich an seinem aktuellen Standort für Dienste anmelden möchte. Eine weitere Möglichkeit für den Nutzer wäre es, sich bei der Middleware anzumelden und eine minimale Zonengröße anzugeben. Wird diese Größe unterschritten, werden keine Standortupdates an Applikationen weitergeben. Wie geeignet diese Einstellungsmöglichkeiten sind, wird in 3.4 untersucht.
Das gravierendste Problem dieses Ansatzes liegt, vor allem bei geringen Nutzerzahlen, in der Art, wie Menschen sich bewegen. Betreten beispielsweise zwei Personen eine Mix Zone in entgegengesetzter Richtung, kann davon ausgegangen werden, dass sie im Allgemeinen nicht mitten in der Bewegung umkehren. Verlassen also kurz darauf zwei Personen die Zone in ebenfalls entgegengesetzten Richtungen, sind mit relativ großer Sicherheit Aussagen über ihre Herkünfte und damit eine Verfolgung möglich. Um die Privatsphäre der Nutzer effektiv schützen zu können, ist also nicht nur eine ausreichend große Zahl von Nutzern, sondern auch ein Minimum an Aktivität und Bewegung dieser gefordert, um genug Rauschen zu erzeugen, das die Verfolgung eines einzelnen Ziels ausreichend erschwert.

Verschleiern von Bewegung mit Minimalgenauigkeit

Ein Verfahren, das versucht auf schwach besuchte Gebiete einzugehen, stammt aus der Straßenvermessung und Verkehrsüberwachung und ist in [30] dargestellt. Ziel und besondere Herausforderung dieses Entwurfes war es, sowohl stark, als auch schwach befahrene Straßen vermessen zu können und dabei sowohl den Schutz der Privatsphäre der Vermesser als auch eine Mindestgenauigkeit[15] zu gewährleisten. In einem solchen Szenario ist es nur schwer möglich Standorte zu verzerren oder zu verschleiern, da sonst möglicherweise einzelne Straßen nicht mehr voneinander unterschieden werden können. Ebenso soll das Straßennetz möglichst vollständig und inklusive selten benutzter Straßen erfasst werden, wodurch eine gewisse zeitliche Auflösung notwendig wird. Ein Verfahren ähnlich den Mix Zones, bei dem unter Umständen gewartet werden muss, bis sich eine gewisse Zahl anderer Teilnehmer in der Nähe befindet, damit ein Standort veröffentlicht wird, ist demnach ebenfalls ungeeignet.
Das folgende Verfahren kombiniert die Unsicherheit, die Verfahren wie k-Anonymität

[15] räumlich wie zeitlich

erzeugen, mit einer maximalen zeitlichen Genauigkeit. Diese wird durch den Konfusions-Timeout beschrieben. Diese Zeit gibt an, wie lange ein Fahrzeug maximal verfolgt werden darf. Sie kann von jedem Nutzer nach seinen Vorlieben gewählt werden und beeinflusst seine Privatsphäre maßgebend. Je größer dieser Zeitrahmen gewählt wird, desto länger wird die Bewegung aufgezeichnet und desto wahrscheinlicher wird eine Verfolgung. Als eine weitere Metrik für das Maß der Privatsphäre wird Konfusion verwendet. Die Konfusion entspricht der Entropie eines beliebigen Abschnittes der verfolgten Strecke. Sie ist demnach definiert als $H = -\sum(p_i * \log(p_i))$. p_i gibt dabei die Wahrscheinlichkeit an, dass die Standortdaten i zum verfolgten Fahrzeug gehören. Diese wird sowohl, wiederum ähnlich der k-Anonymität, durch die Zahl anderer Fahrzeuge in der Nähe als auch durch die Entfernung zum letzten Messpunkt des gesuchten Fahrzeugs bestimmt. Auch das Mindestmaß an Konfusion kann vom Nutzer gewählt werden.
Basierend auf diesen beiden Parametern kann eine vertrauenswürdige Middleware nun entscheiden, ob der Standort eines Fahrzeugs veröffentlicht werden darf. Dazu wird zunächst überprüft, ob das Fahrzeug länger, als durch seinen Konfusions-Timeout erlaubt, verfolgt wurde. Ist dies nicht der Fall, kann sein Standort ohne weiteres veröffentlicht werden. Wurde die zeitliche Grenze jedoch überschritten, muss geprüft werden, ob genügend andere Fahrzeuge in der Nähe sind, um die Minimalkonfusion zu erreichen. Wird die Grenze hierzu überschritten, muss abschließend geprüft werden, ob nach Entfernen aller anderen nicht zu verwendenden Standorte immernoch genug Konfusion erreicht werden kann. Ist dies der Fall, können alle Standorte, deren Bedingungen erfüllt werden, veröffentlicht werden. Zusätzlich wird die Zeit seit der letzten Konfusion für alle Fahrzeuge, die ihre Konfusionsgrenze überschritten haben auf den aktuellen Zeitpunkt gesetzt, da nun die Zeit seit der letzten Konfusion von neuem beginnt.
Durch die erzwungenen Aussetzer der Standortbestimmung bietet dieser Ansatz auch in Gebieten mit geringer Nutzerdichte ein gutes Maß an Privatsphäre und deckt dabei, nach Angaben seiner Entwickler und abhängig der verwendeten Parameter, bis zu 95% der befahrenen Straßen ab. Der weitreichende Einsatz eines solchen Verfahrens könnte dank der immer größeren Verbreitung von Geräten mit eingebauten GPS Empfängern beispielsweise ein crowdbasiertes Verkehrsfrühwarnsystem schaffen, an dem Nutzer teilnehmen und dennoch ihre Privatsphäre wahren können.
Doch auch dieser Ansatz schützt nur begrenzt vor gezielten Angriffen. Gerade, wenn Ursprung und Ziel einer Person bekannt sind, kann ein Angreifer ermitteln, wann sein Ziel sich zwischen beiden Punkten bewegt hat und unter Umständen sogar, welchen Weg er gewählt hat. Es empfiehlt sich also im Allgemeinen Dienste, die die

Verfolgung der eigenen Bewegung ermöglichen, nur dann zu nutzen, wenn man bereit ist, zumindest grob verfolgt werden zu können.

3.4 Problem der Nutzerfreundlichkeit

Die vorgestellten Maßnahmen zum Schutz der Privatsphäre bieten eine Vielzahl von Möglichkeiten für den Nutzer seinen bevorzugten Grad an Verschleierung zu wählen. Bei vielen dieser Möglichkeiten stellt sich jedoch die Frage, wie intuitiv sie zu benutzen sind. Dies ist ein wichtiges Bewertungskriterium, da es maßgeblich die korrekte Verwendung eines Dienstes mitbestimmt. Nutzer können auf vielfältige Arten, mit Absicht oder aufgrund fälschlicher Annahmen, dazu verleitet werden, suboptimale Entscheidungen in ihren Einstellungen zu treffen oder einen Dienst aufgrund zu komplizierter Einstellungen gar nicht erst zu nutzen. Gerade, wenn ein Maß für Privatsphäre zu abstrakt gewählt ist, kann ein Nutzer schnell Parameter wählen, die seinen Standort nicht in dem Umfang schützen, den er sich erhofft. Im Folgenden soll untersucht werden, wo die Schwächen der zuvor erläuterten Ansätze liegen, die solche Fehler provozieren können. Nachfolgend sollen Maßnahmen aufgezeigt werden, mit deren Hilfe die Wahl der richtigen Einstellungen für den Nutzer erleichtert werden können.
Alle bisher erläuterten Ansätze und Verfahren, die Wert darauf legen, dem Nutzer eine Wahlmöglichkeit für die Stärke seiner Privatsphäre zu geben, teilen sich ein gemeinsames Problem. Ihre Möglichkeiten für Nutzereinstellungen sind sehr abstrakt. In [4] wird dem Nutzer die Wahl einer Relevanz eingeräumt. Es ist jedoch fraglich, wie einfach dieser den Wert der Relevanz auf die Realität übertragen kann. Einzuschätzen, mit welchem Wert für diese Relevanz die Privatsphäre im gewünschten Maße geschützt ist, kann gerade bei Zahlen zwischen 0 und 1 schwer fallen und damit einhergehend einen Mangel an Vertrauen in diese Wahl und damit in das Verfahren hervorrufen. Ein ähnlicher, auf Relevanz basierender, Parameter wird als Ergänzung in [3] vorgestellt. Es wird vorgeschlagen einen Nutzer nicht die Relevanz direkt, sondern die Abschwächung der Relevanz des gemessenen Standortes, wählen zu lassen. Diese ist definiert als $\lambda = \frac{R_f}{R_i}$. Die Abschwächung kann von einem Nutzer also in Prozent angegeben werden. Dies erhöht zwar unter Umständen das Verständnis für die getroffene Wahl, macht das Abschätzen eines geeigneten Wertes jedoch nur begrenzt einfacher, da das Resultat, also die Größe und Verschiebung der entstehenden Fläche, immer noch sehr abstrakt beschrieben ist. Das Ergebnis der Verschleierung lässt sich auch so nur begrenzt ablesen. Eine sehr viel verständlichere

Methode, um die Genauigkeit zu bestimmen bietet sich dabei schon durch die Natur des Ansatzes an. In Verbindung mit Kenntnis der Umgebung, beispielsweise durch offene Kartendaten von OpenStreetMap [41], ließe sich die Umgebung der aktuellen Messung ermitteln und in die Verschleierung einbeziehen. Ein Nutzer könnte dann wählen, auf welcher Ebene er erkannt wird. Beispiele hierfür sind „Stadt“, „Stadtteil“ oder „Straßenblock“. Entsprechend könnte dann die gemessene Fläche verzerrt und verschoben werden, bis sie das gewählte Gebiet abdeckt.
Eine andere Wahl, die aber ähnliche Probleme mit sich bringt, ist dem Nutzer bei dem vorgestellten Ansatz der k-Anonymität gegeben [25]. Festzulegen, mit wie vielen anderen Personen ein Nutzer vermischt werden möchte, gibt ihm ein besseres Gefühl für das Verhältnis seiner Entscheidung. Allerdings ist es immernoch schwer abzuschätzen, welcher Wert ausreichend groß ist, und wo der Unterschied zwischen dem Ergebnis zweier ähnlicher Werte ist [58]. Ebenfalls in [24] wird eine mögliche Entschärfung des Problems angegeben, indem der Nutzer seine Privatsphäre in Prozent angibt. Der zugehörige k-Wert kann dann als $k = (1-\frac{P}{100})^{-1}$ berechnet werden, da die Wahrscheinlichkeit in k Standortdaten gefunden zu werden mit $\frac{1}{k}$ gegeben ist, wenn dem Angreifer nur diese Standorte bekannt sind. P ist dabei die Angabe des Nutzers. Allerdings hat diese Lösung die selben Probleme. So fällt es schwer sich 50% Privatsphäre vorzustellen. Eine weitere zu definierende Größe des Ansatzes ist die maximale räumliche und zeitliche Ausdehnung der erzeugten Fläche. Wie in 3.2 beschrieben, kann eine zu große oder zu kleine[16] Wahl die Suche nach einer geeigneten Clique erheblich erschweren und verlangsamen. Eine Möglichkeit dieses Problem einzuschränken, ist es dem Nutzer Vorschläge für die größte erlaubte Fläche zu machen. Diese können in einer verständlichen und vorstellbaren Form, zum Beispiel „Stadt“ oder „Stadtteil“, angeboten werden. Die maximal sinnvolle zeitliche Ausdehnung ist jedoch von Dienst zu Dienst unterschiedlich und dem Nutzer nicht zwingend bewusst. Um dieses Problem in den Griff zu bekommen, kann dieser Wert nicht von dem Nutzer, sondern von einem LBS direkt angegeben werden, der vorher beispielsweise eine optimale Größe empirisch bestimmt hat.
Ein weiteres Problem, das sich nicht aus zu hoher Abstraktion, sondern mangelnder Information für den Nutzer ergibt, zeigt sich bei dem, in [6] vorgestellten, Ansatz der Mix Zones. Es wird vorgeschlagen, dass der Nutzer eine minimale Zahl anderer Teilnehmer angibt, die in einer solchen Zone vorhanden sein muss, bevor er sich bei einem LBS anmeldet. Auch hier kann nicht klar gesagt werden, was ein realistischer Wert ist, ohne sich vorher eingehend mit dem Gebiet, in dem man sich befindet, und dem Dienst, den man verwenden möchte, auseinandergesetzt zu haben. So kann

[16] bei zu großem k

es passieren, dass eine durchschnittliche Nutzerzahl in einer Region durchaus erreichbar, in einem anderen Gebiet jedoch unrealistisch hoch ist. Es ist beispielsweise davon auszugehen, dass in einem Stadtzentrum sehr viel mehr potentielle Nutzer zu finden sind, als dies ein einem Randbezirk der Fall ist. Ein Nutzer müsste also je nach seinem aktuellen Aufenthaltsort einen neuen Wert wählen, um optimale Ergebnisse zu erzielen. Es ist fraglich, ob, beziehungsweise in welchem Rahmen, Nutzer dazu bereit sind.

Ein Lösungsansatz

Ein weiteres wünschenswertes Ziel ist es, einen Ansatz zu entwickeln, der es dem Nutzer ermöglicht nicht zwischen einmaligen und kontinuierlichen Standortdiensten unterscheiden zu müssen. Im Folgenden wird ein Verfahren vorgestellt, das viele der vorgestellten Probleme ausräumt und zudem für die Verschleierung sowohl eines einzelnen Standortes als auch eines Bewegungspfades geeignet ist [58]. Das Verfahren basiert grundsätzlich auf dem Prinzip der k-Anonymität. Es bildet eine verschleierte Fläche jedoch nicht nur anhand der Nutzerzahl in dieser. Da Personen, die beispielsweise einen LBS von ihrem Büro aus benutzen, oft Anfragen von der selben Position aus durchführen, ist es wahrscheinlicher, dass eine Anfrage in entsprechenden Gebieten von ihnen stammt. Dieser Zustand wird als *dominante Anwesenheit (engl.: dominant presence)* bezeichnet. Dominante Anwesenheit kann die Anonymität eines Standortes stark verzerren, sodass ein vermeintlich ausreichender Wert von k zu einem sehr viel schwächeren Schutz als erhofft führt. Anstatt also auf die aktuelle Zahl von Nutzern zu vertrauen, wird zusätzlich die Vergangenheit des betroffenen Bereiches und der Nutzer hinzugezogen. Dies erfordert natürlich erst recht eine vertrauenswürdige Middleware, da Standorte nicht nur verarbeitet, sondern auch gespeichert werden müssen. Mit diesen Daten wird die Beliebtheit eines Bereiches berechnet. Dazu wird zunächst die Entropie der Nutzerdaten in einem räumlichen Bereich R berechnet als $E(R) = -\sum_{i=1}^{m} \frac{n_i}{N} \log \frac{n_i}{N}$. m gibt dabei die Zahl der Nutzer in R, n_i die Zahl der Fußabdrücke des Nutzers i in R und N die Zahl aller Fußabdrücke in R an. Die Beliebtheit eines Bereiches berechnet sich dann als $P(R) = 2^{E(R)}$ mit $1 < P(R) \leq m$. Sie ist maximal, wenn alle Nutzer die gleiche Anzahl an Fußabdrücken hinterlassen haben, und entsprechend minimal, wenn jeder Nutzer nur einen Fußabdruck hat und alle restlichen Daten einem dominanten Nutzer gehören. Dadurch und durch m als maximalen Wert für $P(R)$ ist sichergestellt, dass viel besuchte Gebiete eine hohe Beliebtheit haben, die abgeschwächt wird, wenn

sich dominante Nutzer in diesen aufhalten. Sendet ein Nutzer nun eine Anfrage an einen LBS, muss die Middleware ein Gebiet minimaler Größe[17] finden, das den Nutzer einschließt und seiner zuvor definierten mindesten Beliebtheit entspricht. Diese Beliebtheitsgrenze könnte von einem Nutzer wie in den meisten Verfahren als Zahl angegeben werden. Es wird jedoch vorgeschlagen, dass ein Nutzer einen öffentlichen Bereich angibt, der seiner Vorstellung von Beliebtheit entspricht. Das könnte beispielsweise ein Einkaufzentrum oder eine öffentliche Fläche, wie ein Park sein. Auf diese Art kann ein Nutzer sich gut vorstellen, welche Auswahl er trifft, da er sich nicht auf eine abstrakte Zahl verlassen muss, sondern Erfahrungen aus seinem täglichen Leben zu Rate ziehen kann.
Auf die bisher dargestellte Weise liefert das Verfahren aus [58] einen guten Schutz für einzelne Anfragen. Wenn sich der Nutzer jedoch bewegt und dabei kontinuierliche Anfragen an einen LBS stellt, kann ein Angreifer ihn möglicherweise identifizieren, indem er einen Nutzer sucht, der in allen veröffentlichten Messungen zu finden ist. Um dies zu verhindern, muss die beschriebene Ermittlung einer verschleiernden Fläche leicht modifiziert werden. Dazu wird die Zahl der Nutzer, die zum Bestimmen der Beliebtheit herangezogen werden, eingeschränkt. Eine Folge T von Standortdaten ist genau dann zum Darstellen des Pfades des Nutzers geeignet, wenn $P_S(R_i) \geq P(R)$ für alle $R_i \in T$ gilt und der Nutzer in allen R_i eingeschlossen ist. $P_S(R)$ ist die Beliebtheit eines Bereiches R, wenn nur die Nutzer in der Menge S betrachtet werden. In diesem Fall enthält S die Nutzer, die in allen R_i zu finden sind. Durch diese Einschränkung ist sichergestellt, dass nicht nur ein Nutzer in allen Teilstücken des Weges vorhanden ist und somit der gesamte Weg der geforderten Beliebtheit entspricht.
Dieser Ansatz vereint viele Vorteile in sich. So muss sich der Nutzer nur eines Anonymisierungsdienstes bedienen und braucht nicht zwischen Einzel- und Folgeanfragen zu unterscheiden. Zudem kann er ein vertrautes Umfeld als öffentlichen Bereich wählen, wodurch sowohl seine Vorstellung einer guten Wahl als auch sein Vertrauen in diese Wahl gestärkt wird. Doch auch dieser Ansatz zeigt das häufige Problem der Mindestnutzerzahl. Auch hier ist es möglich, dass die gewünschte Beliebtheit auf Grund mangelnder registrierter Nutzer nicht erreicht wird.

[17] um eine gewisse QoS zu erhalten

3.5 Zusammenfassung

Die vorangegangenen Betrachtungen haben gezeigt, dass auf vielfältige Arten versucht wird den Schutz von Standortdaten zu sichern. Die vorgestellten Ansätze unterscheiden sich dabei sowohl in den verwendeten Techniken, als auch in ihren Schwerpunkten. Es hat sich gezeigt, dass es kein Verfahren gibt, das für jedes Einsatzgebiet und unter allen Voraussetzungen geeignet ist. So gibt es Verfahren, die keine spezielle Infrastruktur benötigen und unabhängig von anderen Nutzern funktionieren, dafür aber nur begrenzten Schutz bieten [4]. Es gibt Ansätze, die darauf vertrauen Nutzer untereinander zu vermischen, sodass diese nicht mehr zu unterscheiden sind. Diese Ansätze benötigen jedoch meist eine gewisse Nutzerbasis [59] oder Infrastruktur mit genügend Rechenleistung um effektiv arbeiten zu können [25, 58]. Zudem sind viele Entwürfe entweder auf einzelne Abfragen [4, 25] oder auf kontinuierlichen Datenaustausch [30] ausgelegt. Den meisten Verfahren haftet dabei das Problem mangelnder Nutzerfreundlichkeit an. Es gibt jedoch Bestrebungen auch diese zu beseitigen [58]. In Tabelle 3.1 wird noch einmal ein Überblick über die Vor- und Nachteile der vorgestellten Verfahren gegeben.
Da die Privatsphäre bei standortbasierten Systemen ein sehr aktuelles Thema ist, ist damit zu rechnen, dass es noch einige Veröffentlichungen auf diesem Gebiet geben wird. Schwerpunkte könnten dabei auf der Nutzerfreundlichkeit, Effizienz und Robustheit, also dem Schutz gegen Angriffe, der Verfahren liegen. Dabei ist es auf Grund der vielen möglichen Anwendungsgebiete unwahrscheinlich, dass es eine allumfassende Lösung geben wird. Es gilt jetzt, wie auch in Zukunft, die Vor- und Nachteile einzelner Verfahren abzuwägen und sich für ein Verfahren oder eine Kombination aus mehreren zu entscheiden.
Es ist jedoch ein grundlegend wichtiger Schritt unerlässlich, um die Verbreitung von schützenden Maßnahmen bei standortbasierten Diensten zu fördern. Um eine weitreichende Abdeckung von Diensten auch zwischen verschiedenen Anbietern zu ermöglichen, müssen grundlegende Protokolle und Standards entwickelt werden, die den Datenaustausch zwischen der Middleware verschiedener Verschleierungstechniken und vor allem auch mit den LBS selbst sicherstellen. Zur Zeit gehen die meisten Dienste nicht davon aus verschleierte Daten zu empfangen. Ein Hauptgrund hierfür kann es sein, dass es noch kein einheitliches Verfahren gibt, um diese zu verarbeiten. Momentan ist es dadurch den LBS-Anbietern selbst überlassen eine Verschleierung anzubieten oder nicht. Zudem muss sichergestellt sein, dass die Middleware, die von den meisten Ansätzen als vertrauenswürdig vorausgesetzt wird, diesem Anspruch tatsächlich genügt. Denn gerade für Drittanbieter einer solchen bietet es sich an, ge-

Tabelle 3.1: Übersicht über die vorgestellten Verfahren

Verfahren	Vorteile	Nachteile
Verzerren der Messung [4]	keine Middleware; geringer Rechenaufwand; unabhängig von LBS; unabhängig von anderen Nutzern	u.U. leicht rückgängig zu machen; ungenau und anfällig bei Bewegung
k-Anonymität [25]	starker Schutz; gleichzeitige Anonymisierung möglich; dezentrale Verwendung möglich [59]	hoher Aufwand: Suche nach k Nachbarn, Infrastruktur; minimale Nutzerzahl notwendig; zentraler Server u.U. Flaschenhals
Mix Zones [6]	verschleiert Bewegung; verbindet Dienste miteinander	Verfolgung möglich; minimale Nutzerzahl benötigt; Infrastruktur notwendig
Konfusion [30]	verschleiert Bewegung; geringe Mindestnutzerzahl; hohe QoS	Verfolgung möglich, wenn auch schwer; Infrastruktur;
Beliebtheit [58]	Bewegung und Einzeldaten verschleiert; nutzerfreundlich; guter Schutz	Infrastruktur; Rechenaufwand; minimale Nutzerzahl benötigt

sammelte Daten beispielsweise an Werbeunternehmen zu verkaufen, welche solche dann für das Erstellen personalisierter Werbung nutzen können. Die Entwicklung von Schutzmaßnahmen für Standortdienste ist also, hoffentlich, noch lange nicht an ihrem Ende angelangt.

Weitere Arbeiten

Alle Arbeiten, die sich mit dem Schutz von Standortdaten auseinandersetzen, zu berücksichtigen, ist im Rahmen dieser Arbeit nicht möglich. Deshalb ist im Folgenden eine Auswahl weiterer Arbeiten aufgeführt, die sich ebenfalls mit diesem Thema auseinandersetzen.

- *On the Anonymity of Home/Work Location Pairs* [26]
  Ein Angriffsszenario, bei dem Wohnort, Arbeitsplatz und Identität von Nutzern aus Standortdaten bestimmt werden.
- *On the Optimal Placement of Mix Zones* [22]
  Freudiger et al. greifen das Problem von Größe und Standort der vorgestellten

Mix Zones auf. Sie zeigen, dass eine geschickte Wahl der Zonen den gebotenen Schutz stark erhöht.

- *A survey of computational location privacy* [34]
  Die Arbeit gibt einen Überblick über den Standpunkt, den Nutzer zum Schutz ihres Standortes einnehmen. Des Weiteren werden Gefahren und Schutzmaßnahmen zusammengefasst und erläutert.

4 Mobile Soziale Netzwerke

4.1 Erläuterung und Gefahren

Soziale Netzwerke (Social Network Services, SNS) haben mit dem Aufkommen des Web 2.0 stetig an Beliebtheit gewonnen. SNS können in Zweck und Umfang stark variieren. Manche Dienste legen einen Schwerpunkt auf professionelle Kontakte [46], andere gestalten sich als reines Unterhaltungsportal [16]. Es ist also notwendig zu definieren, was ein soziales Netzwerk ausmacht. Boyd und Ellison [10] haben solche Netzwerke definiert als „*[...]web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system.*“ SNS sind also Systeme, die es einem Nutzer erlauben ein Profil anzulegen und Verbindungen zu anderen Nutzern aufzubauen. Viele Dienste wie `facebook.com` erlauben es ihren Nutzern zudem Nachrichten zu verschicken, Kommentare (z. B. zu einem Bild oder auf einer Pinnwand) zu hinterlassen oder Daten, wie Bilder oder Dokumente, zu veröffentlichen.

Mit den zahlreichen Möglichkeiten von sozialen Netzwerken, entstehen allerdings auch zahlreiche Probleme. Bei einem SNS sammeln Nutzer eine große Menge an persönlichen Daten in einem System. Viele dieser Daten sind dabei öffentlich einsehbar. Doch auch nicht oder teilweise[1] veröffentlichte Daten können von unerwünschten Personen, zum Beispiel vom Betreiber oder von Angreifern, eingesehen werden. Zudem können Verbindungen und Freundschaften zu Personen erkannt werden, die ein Nutzer unter Umständen geheim halten möchte. Werden diese Verbindungen aufgedeckt kann dies nicht nur persönliche Folgen haben[2], sondern sich auch auf das Berufsleben einer oder mehrerer Personen auswirken[3].

Das folgende Kapitel stellt gezielte Verfahren vor, die die Daten und Kontakte der

[1] z. B. nur für Freunde
[2] z. B. bei Ehebruch
[3] z. B. bei Verbindungen zu einer Konkurrenzfirma

Nutzer eines SNS schützen sollen. Dazu werden Möglichkeiten für die Geheimhaltung von Verbindungen evaluiert, Maßnahmen zum Erstellen effektiver Rechteverteilungen erläutert und dezentrale Ansätze vorgestellt, die einen zentralen Dienstbetreiber unnötig machen. Es wird zudem auf die Bedeutung von mobilen Geräten für die Entwicklung von sozialen Netzwerken und Schutzmechanismen eingegangen.

4.2 Freundschaftsbeziehungen

Das Abbilden von Beziehungen im echten Leben auf Freundeslisten in einem Onlinedienst stellt ein wesentliches Kernelement von sozialen Netzwerken dar. Gleichzeitig liefert es aber auch viele Möglichkeiten für andere in die Privatsphäre eines Nutzers einzudringen. Erlaubt es ein Nutzer beispielsweise nicht, dass seine Daten für Werbung ausgewertet werden, kann ein Werbeanbieter die Vorlieben seiner Freunde heranziehen. In der Annahme, dass befreundete Personen ähnliche Interessen haben, kann er so ein Profil für seine Anzeigen erstellen. Existierende soziale Netzwerke wie `facebook.com` erlauben es ihren Nutzern häufig ihre Freundeslisten geheim zu halten, wenn sie dies wünschen. Gründe hierfür gehen über die Auswertung dieser Verbindungen hinaus. Ein Nutzer möchte sich unter Umständen nicht nur vor der Auswertung seiner Daten und Verbindungen schützen. Es kann auch vorkommen, dass er sich Komplikationen mit anderen Bekannten sparen möchte. Diese können zum Beispiel aus einer Beziehung zu unbeliebten Dritten oder dem Ausschlagen einer Freundschaft entstehen. In letzterem Fall könnte es passieren, dass sich die abgelehnte Person beleidigt fühlt, wenn sie bereits bestehende Freundschaften zu Personen, die sie für weniger beliebt hält, erkennt.
Zur Zeit gibt es wenige Maßnahmen, die über das Verstecken der Freundesliste hinaus gehen. Doch selbst, wenn ein Nutzer seine Beziehungen geheim hält, lassen sich Rückschlüsse auf seine Bekanntschaften ziehen. Denn Freundschaften können in den meisten Fällen als symmetrisch angesehen werden. Wenn Alice mit Bob befreundet ist, ist Bob üblicherweise auch mit Alice befreundet. In einer Untersuchung konnten so 89% der Freundschaftsbeziehungen eines sozialen Netzwerks ermittelt werden [7]. In der selben Arbeit werden zwei Verfahren vorgestellt um die Beziehungen eines Nutzers effektiv geheim zu halten. Unterschieden wird hier zwischen einem klassischen SNS, bei dem Daten zentral gespeichert werden, und einer verteilten Architektur, da sich letzteres als erheblich komplizierter erweist.
Der Vorteil eines zentralen SNS besteht in diesem Fall darin, dass die Freundesliste eines Nutzers problemlos bei jeder Anfrage neu erzeugt werden kann. Dadurch, dass

dabei für jeden Nutzer geprüft werden kann, ob dieser seine Verbindungen geheim halten möchte, kann diese Liste gefiltert werden. Nutzer können entsprechend ausgeblendet werden. Weist der Dienst genügend Nutzer auf, kann problemlos angegeben werden, wie viele Nutzer dabei versteckt bleiben. Ist das Netzwerk jedoch relativ klein, kann es möglich sein die versteckten Nutzer mit großer Wahrscheinlichkeit zu identifizieren. In diesem Fall sollte nur die Zahl der einsehbaren Freunde angegeben werden.

In einem verteilten System sind die Verbindungen zwischen Nutzern und ihre Präferenzen nicht zentral gespeichert. Sie können damit nur mit großem Aufwand, wenn überhaupt, auf die Einstellungen jedes Teilnehmers geprüft werden. Modellieren lassen sich die Verbindungen zum Beispiel mit Hilfe von FOAF [53]. Hierbei werden die Freunde einer Person in einem XML-Dokument festgehalten. Ziel und gleichzeitig Nachteil einer solchen Speicherung ist die Maschinenlesbarkeit eines solchen Dokumentes. Deshalb werden nach [7] nur Verbindungen als Klartext dargestellt, die wirklich öffentlich sind. Um anonyme Verbindungen darzustellen, wird eine Verschlüsselung verwendet. Möchte ein Nutzer A seine Verbindung zu einer anderen Person B geheim halten, speichert er seine verschlüsselte ID $E_B(A)$. E_k ist eine Funktion, die einen Wert mit Hilfe von Schlüssel k verschlüsselt. Ob E_k symmetrisch oder asymmetrisch verschlüsselt, wird nicht näher erläutert. Eine symmetrische Verschlüsselung scheint jedoch ausreichend zu sein. In diesem Fall speichert A also seine eigene ID, die von B mit seinem geheimen Schlüssel verschlüsselt wurde. B speichert entsprechend $E_A(B)$, wenn die Freundschaft symmetrisch ist. Durch die Verwendung solcher Werte kann das FOAF-Dokument beider Nutzer veröffentlicht werden, ohne dass die Verbindung zwischen ihnen bekannt wird. Damit kann das Dokument zudem in einem Cache gespeichert werden, damit die Verbindungen zu anderen Nutzern auch eingesehen werden können, wenn der entsprechende Nutzer nicht erreichbar ist. Indem A seine ID an B schickt, kann sie allerdings immernoch ihre Freundschaft mit B bestätigen. Dieser muss zur Kontrolle lediglich $E_B(A)$ mit seinem Schlüssel entschlüsseln und das Ergebnis mit ID_A vergleichen. Da B einen geheimen Schlüssel verwendet, um $E_B(A)$ zu erzeugen, kann er sich zudem sicher sein, dass er seine Freundschaft zu A bestätigt hat. Ein bösartiger Nutzer M kann mit $E_B(A)$ keine Freundschaft zu B vortäuschen. Er müsste vielmehr in Besitz von $E_B(M)$ sein, der nur von B ausgestellt werden kann. Dies ist vor allem wichtig, wenn Daten nur an Freunde herausgegeben werden sollen, damit M sich diese nicht mit falscher Identität erschleicht.

Routing über Freundschaftsbeziehungen

Die Verwendung von Freundeslisten kann in mobilen Systemen über die SNS-üblichen Zwecke hinausgehen. Ein Ansatz, der gerade für spontane dezentrale Netzwerke geeignet ist, ist das Heranziehen solcher Listen für das Routing in opportunistischen Netzwerken. Opportunistisches Routing zeichnet sich dadurch aus, dass es keinerlei Kenntnis über die Topologie eines Netzes erfordert. Es kann deshalb mit mobilen Geräten und Kurzstreckennetzwerken wie Bluetooth verwendet werden, ohne dass Nutzer etwas davon mitbekommen. Größter Nachteil eines solchen Vorgehens sind jedoch die mitunter langen Übertragungszeiten von bis zu mehreren Stunden. Jedoch sind spezielle Routingalgorithmen notwendig, die auf den wechselhaften Aufbau eines solchen Netzwerks abgestimmt sind. Ein möglicher Weg eine Nachricht zuzustellen stellt das so genannte *Epidemic Routing* dar [56]. Bei diesem Verfahren überträgt ein Gerät Nachrichten an alle anderen Geräte in seiner Nähe. Es betreibt also Flooding. Im Rahmen eines mobilen SNS ist dieses Verfahren jedoch nur sehr begrenzt geeignet. Zwar kann eine möglichst schnelle Zustellung der Nachrichten erreicht werden. Gleichzeitig ist das Verfahren jedoch auch sehr verschwenderisch im Bezug auf die Zahl der gemachten Übertragungen und den Batterieverbrauch der beteiligten Geräte. Einen gezielten Routingpfad im Voraus zu finden ist jedoch ebenfalls nicht möglich, da nicht abzusehen ist, welche Geräte sich wann begegnen. In [42] wird vorgeschlagen, die Freundschaftsverbindungen eines Nutzers heranzuziehen und so ein teilweise gezieltes Routing möglich zu machen. Grundlage des Ansatzes ist die Annahme, dass Mitglieder eines SN eine größere Wahrscheinlichkeit haben sich untereinander zu treffen[4], als zufällige fremde Personen. Dementsprechend kann eine hohe Wahrscheinlichkeit für eine schnelle und vollständige Übertragung angenommen werden. Allerdings muss der Sender einer Nachricht eine Liste seiner Freunde mit jeder Nachricht verschicken, damit jeder Knoten potentielle weitere Knoten erkennen kann. Es kommt also zu einem Broadcast der Netzwerkstruktur, der unter Umständen leicht belauscht werden kann. Zudem können anonyme Teilnehmer an dem Netzwerk unter Umständen durch bekannte Verbindungen de-anonymisiert werden.

Um dem entgegen zu wirken, sind zwei Maßnahmen von Parris et al. vorgeschlagen worden [42], die sich gegebenenfalls auch kombinieren lassen. Im Falle des Statisticulated Social Network Routing (SSNR), wird die Freundesliste des Absenders verändert. Dazu werden bei jeder Nachricht zufällige Einträge aus ihr entfernt oder hinzugefügt. Der Nutzer kann, wenn dies gewünscht ist, den Grad der Veränderung

[4] z. B. bei gemeinsamen Aktivitäten

festlegen. Es ist jedoch fraglich, ob dieses Vorgehen praktisch sinnvoll ist. Im Allgemeinen ist nicht davon auszugehen, dass ein Nutzer zwischen zwei Nachrichten seine Vorlieben verändert. Durch das Verändern der Kontaktliste kann ein Angreifer, der wenige oder nur eine Nachricht belauscht, nicht mit Sicherheit feststellen, wie die vollständige und korrekte Freundesliste aussieht. Um dies tun zu können, müssen sehr viele Nachrichten mitgeschnitten und verglichen werden. Auf Grund der wechselhaften Struktur eines opportunistischen Netzwerks und dem Verwenden von Netzwerken mit geringer Reichweite ist dies allerdings nur mit erheblichem Aufwand möglich.

Das Obfuscated Social Network Routing (OSNR) vermeidet es die Freundschaften des Absenders in Klartext zu versenden. Stattdessen wird ein Bloom-Filter verwendet, um diese zu speichern. Ein Bloom-Filter ist eine Datenstruktur mit fester Länge, die eine wahrscheinlichkeitsbasierte Suche nach in ihr gespeicherten Elementen ermöglicht. Dabei kann es mit geringer Wahrscheinlichkeit zu False Positives kommen, bei denen Elemente als vorhanden angezeigt werden, obwohl sie es nicht sind. False Negatives sind jedoch nicht möglich. In einem solchen Bloom-Filter werden die IDs[5] der Freunde eines Nutzers zusammen mit einem Salt um Rainbowtable-Angriffe zu vermeiden abgelegt. Die so gespeicherten IDs sind nicht mehr zu identifizieren. Es kann lediglich überprüft werden, ob ein Nutzer mit dem Absender befreundet ist. Für die Auswahl von Knoten für die Nachrichtenübertragung ist dies völlig ausreichend. Allerdings kann ein Angreifer durch einen Brute-Force-Angriff die Struktur der Freundesliste bestimmem. Zudem kann er gezielt nach bestimmten Personen suchen. Durch die Natur eines Bloom-Filters muss er jedoch mit False Positives rechnen und kann nicht davon ausgehen, dass seine Ergebnisse garantiert korrekt sind.

Kombiniert man SSNR und OSNR, kann diese Unsicherheit stark vergrößert werden. Das Selbe gilt allerdings auch für Knoten, die nicht bösartig sind. Es kann somit vorkommen, dass Nachrichten an Personen weitergeleitet werden, die nicht in der Freundesliste des Senders stehen. Da die Liste durch OSNR jedoch verschleiert wurde, entsteht so keine große Gefahr durch diese Knoten. Durch das Versenden an zusätzliche Knoten kann jedoch die Übertragungszeit in manchen Fällen reduziert werden. Ein weiterer Vorteil besteht in der konstanten Größe eines Bloom-Filters. Bei reinem SSNR kann die Freundesliste vor allem dann, wenn sie grundsätzlich sehr umfangreich ist, zu enormem Overhead bei jeder Übertragung führen. Ein Bloom-Filter dagegen hat bei jedem Füllstand die selbe, vorher festgelegte, Größe. Diese sollte jedoch sinnvoll gewählt sein, da die Wahrscheinlichkeit für False Positives

[5] z. B. MAC-Adresse, IMEI oder Nutzer-ID

sonst schnell sehr hoch werden kann. In diesem Fall wäre das Verfahren nur begrenzt effektiv, da es dem Epidemic Routing sehr nahe kommt.

4.3 Dezentrale Soziale Netzwerke

Klassische soziale Netzwerke besitzen schon aufgrund ihres Auftretens als Webapplikationen eine zentrale Struktur. Sämtliche Nutzerdaten werden auf Servern des Anbieters gespeichert und verwaltet. Dies ermöglicht zwar den einfachen und gewohnten Zugriff über ein Webinterface, stellt aber gleichzeitig auch eine große potentielle Gefahr für die Daten der Nutzer dar. Sowohl der Betreiber selbst, als auch ein Angreifer, der sich Zugang zu diesen Servern verschafft, kann unabhängig von allen Einstellungen der Nutzer sämtliche Daten einsehen. Mobile Geräte und P2P-Systeme ermöglichen hingegen den Aufbau von spontanen und dezentralen Netzwerken, bei denen alle Daten nur bei dem Nutzer selbst oder bei Personen, denen er vertraut, gespeichert werden. Solche Systeme erlauben zudem eine höhere Flexibilität der angebotenen Dienste und machen neue Anwendungen möglich.
Dezentrale SNS bringen jedoch auch neue Sicherheitsprobleme mit sich. Vor allem, wenn ein Angreifer gezielt Informationen über eine Person sammeln möchte, bietet die verteilte Struktur einfache Möglichkeiten hierfür. Der einfachste Weg, an solche Informationen zu gelangen, ist die Teilnahme an einem solchen Dienst und das Mitschreiben der gewünschten Daten. Die Übertragung muss also vor Mithörern und nicht vertrauenswürdigen Teilnehmern geschützt werden. Dies kann teilweise durch kryptografische Verfahren geschehen. Jedoch ergeben sich gerade bei P2P-Architekturen erhebliche Sicherheitsprobleme bei dem Austausch gemeinsamer Schlüssel. Zudem muss die Identität der Knoten in einem solchen Netzwerk sichergestellt werden, um Manipulationen vermeiden zu können.

Physikalische Begegnungen als Vertrauensgrundlage

Ein Vorteil mobiler Geräte ist die Möglichkeit Vertrauen durch räumliche Nähe aufzubauen. Wenn Personen sich persönlich gegenüber stehen, können sie sicher sein, dass sie mit dem richtigen Gegenüber kommunizieren. Sie können dann zum Beispiel Schlüssel für eine spätere verschlüsselte Kommunikation austauschen. Da durch diese Schlüssel eine physikalische anstelle einer digitalen Identität bestätigt wird, kann auf aufwändige Strukturen wie ein Web of Trust oder eine *Public-Key-Infrastructure*

(PKI) verzichtet werden. Ein Verfahren zum Austausch von Public Keys[6] zwischen mehreren Geräten wird in [35] spezifiziert. Die maximale Zahl von Teilnehmern ist dabei durch die Verwendung von Bluetooth begrenzt, welches maximal acht aktive Geräte in einem Netzwerk zulässt. Um die Identität der Teilnehmer bei einer späteren Verbindung sicherzustellen, tauschen diese selbstsignierte Zertifikate inklusive öffentlichem Schlüssel aus. Da das Verfahren lediglich Schutz vor Manipulation, jedoch nicht vor Lauschangriffen bietet, sollten nur Schlüssel eines asymmetrischen Verfahrens oder andere öffentliche Daten getauscht werden. Keinesfalls sollten symmetrische Schlüssel oder gar private Schlüssel über eine solche Verbindung getauscht werden. Damit kein Angreifer sich das Vertrauen der anderen Teilnehmer erschleicht, indem er seinen Schlüssel den übertragenen hinzufügt[7], geben alle Teilnehmer zunächst die Größe ihrer Gruppe auf ihren Geräten an. Jeder Teilnehmer darf dann genau einen Datensatz an alle anderen schicken. Die Zahl der empfangenen Nachrichten wird daraufhin mit der angegebenen Gruppengröße abgeglichen. Um eine Manipulation durch einen Man-In-The-Middle-Angriff zu verhindern wird zudem ein Hash der empfangenen Daten gebildet. Dieser wird auf allen Geräten in einer leicht vergleichbaren Form, zum Beispiel als Kombination von Farben, angezeigt und von den Nutzern kontrolliert. Wurde eine Nachricht verändert oder ausgetauscht, fällt dies mindestens dem Absender der ursprünglichen Nachricht auf.

Implementierung eines sicheren dezentralen SNS

Viele Arbeiten zum Schutz der Privatsphäre in mobilen sozialen Netzwerken beschäftigen sich nicht mit Lösungen für einzelne Probleme, sondern stellen vollständige soziale Netzwerke vor. Deshalb soll im Folgenden ein solches Netzwerk vorgestellt werden, bei dem Wert auf eine vollständig verteilte Struktur und starken Schutz gelegt wurde. *Safebook* [12] wurde von Cutillo et al. entwickelt. Dabei wurde Wert auf drei Schwerpunkte gelegt. Der Dienst soll es seinen Nutzern ermöglichen alle ihre Daten und ihre Kommunikation, bis hin zu ihrer Teilnahme an dem SN, zu verstecken. Es soll zudem gewährleisten, dass Integrität sowohl von Daten, als auch von Identitäten, gegeben ist. Es soll also sichergestellt sein, dass jeder Nutzer wirklich der ist, für den er sich ausgibt. Dies ist ein Hauptproblem in dezentralen Netzwerken, da eine zentrale Kontrolle fehlt. Der dritte Schwerpunkt betrifft die Verfügbarkeit des Dienstes. Gerade, wenn ein SNS für produktive Zwecke[8] eingesetzt wird, stellt eine

[6] oder beliebigen anderen Daten
[7] die Teilnehmer könnten dann annehmen bereits mit ihm in Kontakt gestanden zu haben
[8] z. B. bei Kollaborationen verschiedener Forschungseinrichtungen

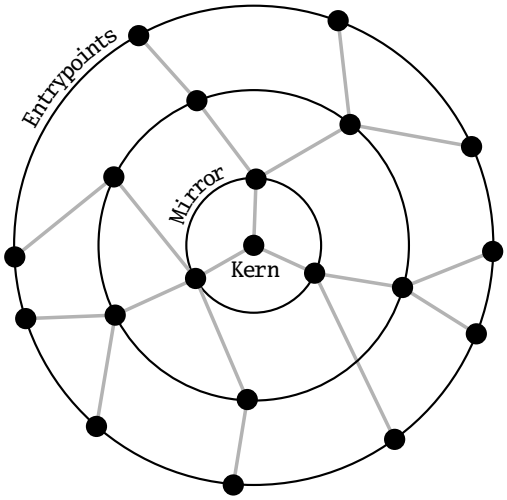

Abbildung 4.1: Eine Matrjoschka mit drei Kreisen

verteilte Struktur ein großes Hindernis dar. Es muss sichergestellt sein, dass die Daten eines Nutzers auch dann verfügbar sind, wenn dieser gerade nicht erreichbar ist. Auch *Safebook* kommt jedoch nicht vollständig ohne einen vertrauenswürdigen Dritten (TTP) aus. Zur Kontrolle der Identität eines Nutzers wird ein so genannter *Trusted Identification Service (TIS)* eingesetzt. Dessen Aufgabe besteht darin, die Identität eines neuen Nutzers zu überprüfen. Dies geschieht wie zum Beispiel bei der Ausstellung eines Zertifikats in einer PKI, indem der Nutzer seine Identität bei einer entsprechenden Institution bestätigen lässt. Der TIS kennt allerdings nicht mehr als die Identität der Nutzer. Er weiß somit lediglich über deren Teilnahme an dem SN Bescheid.

Möchte sich ein Nutzer B bei *Safebook* anmelden, muss er zunächst von einem anderen Nutzer A eingeladen werden. Dadurch ist ein Mindestmaß von Vertrauen bei allen Nutzern gegeben. Hat der neue Nutzer dann seine Identität beim TIS bestätigt, schickt er zwei öffentliche Schlüssel P_b^+ und I_b^+ an diesen. Diese Schlüssel gehören zu jeweils einem Schlüsselpaar, wie es bei asynchronen Kryptographieverfahren üblich ist. Sie werden von B selbst erzeugt. Der TIS berechnet nun das Pseudonym P_b und die Knoten-ID I_b und schickt die von ihm signierten Zertifikate $Cert(P_b, P_b^+)$ und $Cert(I_b, I_b^+)$ an B. Die Zertifikate bescheinigen, dass P_b und I_b aus den entsprechenden öffentlichen Schlüsseln erzeugt wurden. Somit ist die Echtheit der beiden Werte sichergestellt. Die Knoten-ID wird bei einer Übertragung innerhalb des SN verwendet, um Ende-zu-Ende-Integrität und Sicherheit durch Signatur beziehungsweise Verschlüsselung zu gewährleisten. Das Pseudonym wird verwendet, um das selbe bei jedem Hop der Übertragung zu gewährleisten.

Die Umgebung eines Nutzers ist nach einem Matrjoschkaprinzip aufgebaut. Ähnlich den russischen Puppen, umgibt sich ein Teilnehmer mit konzentrischen Kreisen vertrauenswürdiger Knoten. In Abbildung 4.1 ist eine solche Matrjoschka mit drei Ebenen dargestellt. Der Nutzer selbst bildet dabei den Kern des Systems. Knoten im ersten Kreis um ihn herum werden als *Mirrors* bezeichnet, da die Daten des Nutzers auf diesen Knoten gespiegelt werden, um sie so jederzeit verfügbar zu machen. Je nach Einstufung dieser Daten werden sie dabei verschlüsselt (nur begrenzt zugängliche Daten) oder im Klartext (öffentliche Daten) gespeichert. Private Daten werden nicht auf die Mirrors übertragen. Welches Verfahren für die Verschlüsselung eingesetzt wird, ist in [12] nicht näher beschrieben. Der innere Kreis wird von Knoten gebildet, denen der Nutzer sowohl in dem SN als auch im echten Leben vertraut. Der äußere Kreis der Matrjoschka besteht aus den so genannten *Entrypoints*. Nur über diese Knoten kann der Kern angesprochen werden. Möchte ein anderer Teilnehmer Kontakt mit dem Nutzer Kontakt aufnehmen, muss er dies über die Entrypoints tun, welche die Nachricht dann rekursiv bis zu einem Mirror weiterleiten. Dabei muss nicht jeder Kreis der Matrjoschka angesprochen werden. Es können Pfade existieren, die mache Kreise zwischen Entrypoints und Mirrors überspringen. Kommt eine Anfrage bei einem Mirror an, leitet er die angeforderten Daten an die nächste Ebene weiter und der Pfad wird zurückverfolgt. Dabei kann kein Knoten zwischen einer neu generierten und einer weitergeleiteten Anfrage unterscheiden. Der Ursprung bleibt also unbekannt und wird so ebenfalls geschützt.
Die Mirrors können auf Grund dieser Architektur Daten nicht nur an autorisierte Nutzer weitergeben. Die Zugriffskontrolle erfolgt deshalb über die Verschlüsselung. Nur wenn der Anfragende bereits zuvor Kontakt mit dem Nutzer hatte und im Besitz des entsprechenden Schlüssels ist, kann er die Daten entschlüsseln. Der Nutzer kann verschiedene Vertrauensebenen erstellen, indem er Daten mit unterschiedlichen Schlüsseln sichert und diese an die entsprechenden Personen weitergibt.
Safebook ist ein System, dass sehr starken Wert auf den Schutz seiner Nutzer legt. Dementsprechend aufwändig sind die verwendeten Verfahren und Strukturen. Die hohe Aktivität in aktuellen SNS zeigt jedoch, dass viele Nutzer sich mindestens im Moment nur wenig Sorgen um den Schutz ihrer Daten machen. Der zusätzlich nötige Aufwand, allen voran die Notwendigkeit einer Einladung und das verifizieren der eigenen Identität, kann dementsprechend viele Nutzer abschrecken. Eine weitreichende Verbreitung eines solchen Systems ist also zur Zeit nicht zu erwarten. Dennoch bietet *Safebook* einige interessante Ansätze für den Schutz der Nutzer in zentralen und dezentralen sozialen Netzwerken.

4.4 Räumliche Nähe

Ein großer Vorteil von mobilen gegenüber herkömmlichen sozialen Netzwerken ist die Möglichkeit auf räumliche Nähe einzugehen. Herkömmliche soziale Netzwerke können allerdings mit einem mobilen Interface um eine ähnliche Funktionalität erweitert werden. Dabei können zwei Arten von Anwendungen unterschieden werden. Bei dem ersten Anwendungsfall kann ein Nutzer auf Freunde und Bekannte in seiner Nähe hingewiesen werden. Im zweiten Fall kann ein Nutzer neue Menschen kennenlernen. In beiden Fällen muss er seine Anwesenheit an einem Standort bekannt machen. Tut er dies über die Bekanntgabe seines Standorts auf einem Server, kann ein entsprechender Dienst als standortbasiert angesehen werden. Es können dann Schutzmaßnahmen wie in Kapitel 3 *Standortbasierte Dienste* beschrieben verwendet werden. Er kann sich aber auch über ein Kurzstreckennetzwerk wie WLAN direkt bei anderen Geräten bekannt machen. Die Gefahren eines solchen Ansatzes sollen im Folgenden behandelt werden.

Veröffentlichen der eigenen ID

Um die eigene Anwesenheit bekannt zu machen, muss ein Nutzer Daten von sich preisgeben. Ein einfacher Weg hierfür ist die Bekanntgabe der eigenen ID. Mit ihrer Hilfe können anwesende Nutzer dann Informationen über den Nutzer in dem eigentlichen sozialen Netzwerk abrufen. Je nach Art des verwendeten SNS können dabei verschiedene Gefahren entstehen. Wie bei jedem Dienst, der standortgebunden ist, kann der Nutzer über diese ID verfolgt werden. Da jedoch ein Netzwerk mit geringer Reichweite verwendet wird, muss ein Angreifer sich mit dem Ziel bewegen, wodurch der eigentliche Angriff sinnlos wird. Gravierender ist das Aufzeichnen der ID für eine spätere Verwendung. Wenn die ID nicht weiter überprüft wird, kann sich ein Angreifer zu einem späteren Zeitpunkt und an einem anderen Ort für den Nutzer ausgeben. Er kann so dessen Anwesenheit vortäuschen oder Informationen von anderen Nutzern anfordern. In [5] wird deshalb die Verwendung eines einmalig gültigen *Anonymous Identifier (AID)* vorgeschlagen.
Die AID ist eine Nonce, die von einem vertrauenswürdigen Server[9] generiert wird. Ein Nutzer A fordert diese einzeln oder zu mehreren von diesem Identity Server (IS) an. Sie wird dann beispielsweise als Hash der Nutzer-ID in Verbindung mit einem zufälligen Salt erzeugt und dem Nutzer zugeordnet. Ziel ist es, zu verhindern, dass

[9] z. B. betrieben vom Anbieter des SNS

die AID auf die tatsächliche ID des Nutzers abgebildet werden kann. A veröffentlicht dann seine Anwesenheit ohne seine ID bekannt zu geben. Möchte ein Nutzer B mit ihm Kontakt aufnehmen, kann A ihr eine AID zukommen lassen. Über diese kann B Informationen vom IS abrufen. Tut sie dies oder läuft die festgelegte Lebensdauer der AID ab, entfernt der Server die AID und sie verliert ihre Gültigkeit. Der Nutzer kann gegebenenfalls festlegen, welche Informationen übertragen werden. Im ursprünglichen Ansatz von Beach et al. ist festgelegt, dass der IS keine Daten preisgibt, durch die ein Nutzer direkt identifiziert werden kann [5].

Missed Connections

Dienste, bei denen sich Nutzer kennen lernen, müssen nicht zwingend in Echtzeit ablaufen. Missed Connection Dienste sind ein solcher Fall. Bei ihnen können Nutzer nachträglich herausfinden, wer sich zu einem bestimmten Zeitpunkt in ihrer Nähe befunden hat. Ein bekanntes Beispiel hierfür sind Radiosendungen, in denen Personen anrufen und beispielsweise „die hübsche junge Frau im roten Kleid auf dem Konzert gestern Abend“ um eine Kontaktaufnahme bitten können. Ein weiteres Beispiel sind Internetforen oder Messageboards, in denen eine entsprechende Anfrage gestellt werden kann. Diese Verfahren bieten jedoch nur eine geringe Erfolgschance[10]. Sie fordern allerdings vom Nutzer, dass dieser einige möglicherweise problematische Daten preisgibt. Er muss nicht nur eine Kontaktmöglichkeit angeben, die von Angreifern zum Beispiel für Scherze oder Spam missbraucht werden kann. Er gibt vor allem öffentlich bekannt, wann er sich wo aufgehalten hat und mit wem er Kontakt hatte oder haben möchte. Manweiler et al. haben in [38] ein System für Begegnungen als Vertrauensgrundlage vorgestellt, das diese Probleme beseitigen sollen. Wie schon bei *Safebook* wird eine vollständige Implementierung eines Missed Connections Dienstes vorgestellt, die jedoch gute Schutzmaßnahmen für die Kommunikation zweier Personen aufweist. Deshalb sollen die Grundzüge des Verfahrens hier vorgestellt werden.

Das gesamte Verfahren basiert auf dem Austausch und der Verwendung von gemeinsamen symmetrischen Schlüsseln. Diese werden periodisch zufällig erzeugt und über Bluetooth an andere Geräte in der Nähe gesendet. Ein Nutzer speichert alle gesendeten und empfangenen Schlüssel mit Ort und Zeit des Empfangs. Später lädt er dann Hashes der gesammelten Schlüssel[11] auf einen zentralen Server. Ein Nutzer kann

[10] die gemeinte Person muss zufällig die Nachricht mitbekommen und sich angesprochen fühlen

[11] einen Hash pro Schlüssel

nur einen Teil des Hashes mit einer Länge seiner Wahl auf den Server laden um k-Anonymität für diesen Schlüssel zu erreichen. Dabei entsteht jedoch eine größere Menge von empfangenen Nachrichten im späteren Verlauf des Austauschs. Nachdem er seine Hashes mit dem Server synchronisiert hat, kann der Nutzer eine Nachricht auf den Server laden. Er wählt dazu den Schlüssel, der zu Ort und Zeit, an der er die gewünschte Person getroffen hat, passt. Mit diesem Schlüssel verschlüsselt er die Nachricht und markiert sie mit dem zugehörigen Hash. Außerdem wird eine Nonce zu der Nachricht hinzugefügt, um Replay-Angriffe zu vermeiden. Der Server leitet die Nachricht dann an alle Nutzer weiter, die den selben Hash hochgeladen haben. Hat der Nutzer den Hash gekürzt, leitet der Server die Nachricht entsprechend an alle in Frage kommenden Empfänger weiter. Dadurch bleibt auch vor dem Server verborgen, wer die Zielperson ist. Gleichzeitig erhält der Nutzer jedoch auch mehr Nachrichten, die nicht an ihn gerichtet sind, da er von mehr Hashes betroffen sein kann. Empfängt ein Nutzer eine solche Nachricht, versucht er sie mit dem passenden Schlüssel zu entschlüsseln. Dies gelingt also nur Personen, die wirklich zur richtigen Zeit am richtigen Ort waren. Alle anderen müssen die Nachricht verwerfen. Es bleibt jedoch immernoch den Nutzern selbst überlassen die gemeinte Person zu identifizieren, wenn mehrere Personen gleichzeitig anwesend waren. Dies geschieht im Laufe des Gesprächs mit Fragen wie „Was hattest du an?“ und stellt somit eine große Schwachstelle des Verfahrens dar.

Um zu verhindern, dass nachzuvollziehen ist, welche Nutzer eine Unterhaltung führen, müssen gesonderte Maßnahmen getroffen werden. Denn ein Angreifer oder ein neugieriger Betreiber könnten über die Zahl der getauschten Nachrichten Rückschlüsse auf ein Gespräch ziehen. Unterhalten sich zwei Teilnehmer, tauschen sie mit großer Wahrscheinlichkeit eine ähnliche Zahl von Nachrichten aus. Damit solche Muster nicht erkannt werden, wird eine feste Zahl von Nachrichten in jeder Konversation vorgeschrieben. Nutzer können diese Zahl selbst wählen, allerdings nicht für jede Konversation. Tauschen Teilnehmer weniger Nachrichten aus, verschicken sie Dummys, also zufällige Nachrichten. Diese sind mit einem zufälligen Schlüssel verschlüsselt, damit der Empfänger sie als solche erkennt. Mehr Nachrichten als die gewählte Anzahl dürfen nicht ausgetauscht werden. Es sollte möglichst bald auf ein anderes Kommunikationsmedium umgestiegen werden. Lange Unterhaltungen belasten zudem andere Nutzer mit für sie unwichtigen Nachrichten.

Damit ein Angreifer keine falschen Nachrichten in die Konversation einschiebt, wird vorgeschlagen die Abstände zwischen den Nachrichten festzulegen. Gegebenenfalls müssen Nachrichten dann verzögert oder ebenfalls Dummys geschickt werden. Allerdings kommt es dadurch zu erhöhten Verzögerungen während der Kommunikation

beziehungsweise zu erhöhtem Overhead für andere Teilnehmer.
Viele der Schutzmaßnahmen, die in dem Verfahren vorgestellt werden, können auf andere Systeme und Anwendungsbereiche übertragen werden. Vor allem die anonyme Kommunikation durch die Verwendung gekürzter Schlüsselhashes kann auf andere Systeme übertragen werden. Die in diesem Abschnitt und in Abschnitt 4.3 vorgestellten Maßnahmen stellen gute Möglichkeiten dar, um vorherige Begegnungen auf verschiedene Arten als Vertrauensbasis verwenden zu können.

4.5 Regeln zur Freigabe persönlicher Daten

Soziale Netzwerke sind eine beliebte Plattform, um Daten zu veröffentlichen. Nutzer können beispielsweise anderen ihre Aktivitäten mitteilen, Bilder veröffentlichen oder ihren Standort anzeigen lassen. Je größer die Datenmenge eines Nutzers ist, desto schwerer fällt es jedoch, Regeln für den Zugriff auf sie aufzustellen. Viele Systeme bieten zudem nur eine relativ grobe Unterscheidung von Rechten. Werden allerdings mehr Einstellungsmöglichkeiten angeboten, kann es schnell dazu kommen, dass Nutzer den Überblick oder die Lust am Erstellen von Regeln verlieren. Gerade, wenn Nutzer Daten über mobile Geräte veröffentlichen, beschäftigen sie sich möglicher Weise nur sehr kurz oder gar nicht mit deren Schutz. In [48] wird ein Verfahren vorgestellt, das komplexe Regeln erlaubt und den Nutzer beim Aufstellen dieser unterstützt.
Der Ansatz schlägt eine Semantic Web Umgebung wie OWL für die Modellierung von Regeln vor. Solche Regeln können beliebig komplex werden und unter Umständen zwischen verschiedenen SNS ausgetauscht werden. Zusätzlich können Templates erstellt werden, um Nutzern das Definieren neuer Regeln zu erleichtern. Regeln können verschiedene Kriterien aufweisen. Sie können den Zugriff für bestimmte Gruppen von Personen, zu bestimmten Tageszeiten oder an bestimmten Orten regeln. Sadeh et al. schlagen außerdem vor den Nutzer bei Zugriffsversuchen auf seine Daten zu benachrichtigen. So soll der Nutzer Vertrauen in seine Regeln gewinnen und potentielle Angreifer abgeschreckt werden. Werden allerdings zu häufig Daten eines Nutzers angefragt, kann dies schnell zu einer Belästigung werden. Ein möglicherweise geeigneter Ansatz ist die Einsicht vergangener Zugriffe. Ein Nutzer kann einsehen, wann ein Zugriff erlaubt und wann er verweigert wurde. Aufgrund dieser Entscheidungen kann er dann seine Regeln verfeinern. Experimente im Rahmen der Arbeit haben jedoch gezeigt, dass Nutzer Regeln nur bis zu einem gewissen Punkt verfeinern können. Über diesen Punkt hinaus sind sie nicht in der Lage einen verbesserten

Schutz zu erzielen. Um den Nutzer zu unterstützen, werden dessen Regeln deshalb mit Hilfe von *Case-Based Reasoning (CBR)* weiter verfeinert. Dazu wird das Prinzip der k nächsten Nachbarn angewandt. Bei jeder Anfrage wird diese mit den k ähnlichsten vergangenen Anfragen verglichen. Die von der Mehrheit dieser Nachbarn getroffene Entscheidung wird für den aktuellen Fall übernommen. Damit der Nutzer sich nicht übergangen fühlt, kann darauf verzichtet werden den Fall automatisch zu entscheiden. Stattdessen können dem Nutzer Vorschläge für die Verfeinerung seiner Regeln gemacht werden.

4.6 Zusammenfassung

Mobile soziale Netzwerke werfen viele verschiedene Probleme auf. Dementsprechend zahlreich sind die verschiedenen Schwerpunkte von aktuellen Lösungsansätzen. Es wurden Verfahren für die Geheimhaltung von Freundschaftsverbindungen, das Aufstellen von Zugriffsregeln und den Schutz der eigenen Identität vorgestellt. Außerdem wurden Maßnahmen zum Aufbau und Schutz von dezentralen sozialen Netzwerken vorgestellt. Viele aktuelle Arbeiten auf diesem Gebiet teilen sich jedoch ein gemeinsames Problem. Ein großer Teil dieser Arbeiten stellt vollständige soziale Netzwerke mit gewissen Schwerpunkten vor. Aufgrund der aktuellen Beliebtheit vorhandener SNS ist jedoch nicht damit zu rechnen, dass Nutzer in großen Zahlen auf einen neues Netzwerk umsteigen, um besseren Datenschutz zu erfahren. Die meisten dieser neuen Systeme liefern zudem individuelle Probleme, die Nutzer von einem Wechsel abhalten können. Das in dieser Arbeit erwähnte *Safebook* erschwert beispielsweise die Anmeldung neuer Nutzer erheblich. Dadurch könnten viele neugierige Nutzer abgeschreckt werden.
Einige Ansätze, wie das Verwenden von OWL, um Regeln zu definieren, bieten jedoch die Möglichkeit bestehende Netzwerke integriert zu werden. Zur Zeit sind solche allgemeinen Ansätze jedoch nur sehr selten zu finden. Auf diesem Gebiet besteht also noch einiges an Forschungsbedarf.

Weitere Arbeiten

Im Folgenden sind einige Arbeiten aufgeführt, die im Rahmen dieser Arbeit keine Erwähnung gefunden haben.

- *Safebook: Feasibility of transitive cooperation for privacy on a decentralized social network* [13]
  Die Entwickler von *Safebook* untersuchen in dieser Arbeit die Durchführbarkeit des vorgestellten SNS. Sie betrachten dabei Verfügbarkeit, Verzögerungen bei der Übertragung und Schutz der Nutzer in dem Dienst. Der Schwerpunkt der Arbeit liegt allerdings auf der Performance.

- *The privacy jungle: On the market for data protection in social networks* [9]
  Bonneau und Preibusch untersuchen die Struktur des aktuellen Marktes für soziale Netzwerke. Sie erläutern zudem den Stellenwert von Privatsphäre in aktuellen Diensten.

- *Myths and fallacies of personally identifiable information* [39]
  Ein kurzer Artikel über Probleme beim Anonymisieren von Daten und die Identifikation von Personen über scheinbar ungefährliche Informationen.

5 Tagging

5.1 Erläuterung und Gefahren

Tagging bezeichnet das Anreichern von Daten mit zusätzlichen Informationen. Viele moderne Kameras und Mobiltelefone mit Kamera bieten beispielsweise die Möglichkeit den Aufnahmeort eines Bildes in einem so genannten Geo-Tag festzuhalten. Fotodienste wie Flickr [19] erlauben außerdem das Hinzufügen von Tags, die den Inhalt des Bildes beschreiben. Soziale Netzwerke bieten darüber hinaus häufig die Möglichkeit Personen in Bildern zu markieren. Manche Dienste führen dazu sogar automatische Erkennung durch [43]. Während die bisher vorgestellten Dienste hauptsächlich eine Gefahr für die Daten des Nutzers selbst beinhalten, stellt Tagging zunehmend eine Bedrohung für die Privatsphäre anderer Nutzer dar. Personen, die auf Bildern markiert werden, können aus verschiedenen Gründen Einwände dagegen haben. Eine Untersuchung [8] hat ergeben, dass viele Nutzer sich dabei kaum Sorgen wegen unbekannten Personen machen. Viel mehr wollen sie häufig vermeiden, dass bestimmte Gruppen aus ihrem Umfeld Bilder zu sehen bekommen. Solche Gruppen können zum Beispiel die eigene Familie, Freunde oder den Arbeitgeber und Kollegen umfassen. Die Hauptsorge der Nutzer besteht dabei darin, bei etwas Unangenehmem oder Verbotenem fotografiert und vor Allem veröffentlicht zu werden. Nutzer befürchten, dadurch Schaden für ihr Ansehen in bestimmten Gruppen zu erleiden oder ihre Karriere zu gefährden. Ein häufig geäußertes Beispiel hierfür stellt der Konsum von Alkohol oder anderen Drogen dar. Weniger Sorgen machen Nutzer sich dagegen wegen ihres Standorts [23, 8, 1].

Tagging von Texten und Dokumenten hat sich mit dem Aufkommen des Web 2.0 stark verbreitet. Da die getaggten Daten im Allgemeinen explizit vom Nutzer veröffentlicht wurden, um von anderen eingesehen zu werden, stellt diese Form des Tagging jedoch kaum eine Gefahr dar. Sie sei hier nur der Vollständigkeit halber erwähnt.

Bisher existieren nur wenige Arbeiten, die sich mit Tagging und Privatsphäre ausein-

andersetzen. Die wenigen vorhandenen Ansätze sollen im Folgenden erläutert werden.

5.2 Geo-Tagging

Für Standortdaten, die beim Geo-Tagging an ein Bild angefügt werden, gelten die bereits im Kapitel Standortbasierte Dienste beschriebenen Gefahren und Maßnahmen. Darüber hinaus können allerdings noch weitere Gefahren hinzukommen. Das Hauptproblem beim Geo-Tagging stellt dabei die geringe Aufklärung der Nutzer dar. Manche Geräte hängen den Standort automatisch an Fotos, wenn dies nicht explizit deaktiviert wurde. Dadurch wissen viele Nutzer nicht, dass der Aufnahmeort in ihren Bildern gespeichert wird [23]. Zudem unterschätzen manche Nutzer die Genauigkeit der verwendeten Standortdaten.
Eine weitere unterschätzte Gefahr stellen die Fotos selbst dar. Sie können Angreifern helfen ein Ziel für, zum Beispiel, einen Einbruch zu wählen. Ein Angreifer kann über die *APIs (Application Programming Interfaces)* von Diensten wie Twitter [55] gezielt Daten mit Geo-Tags sammeln [32, 44]. Wenn unter den gefundenen Daten Fotos mit Bildern aus der Wohnung eines Nutzers sind, kann der Angreifer nach wertvollen Gegenständen Ausschau halten. Vor allem Dienste, bei denen Nutzer etwas verkaufen oder tauschen, stellen hier ein besonders geeignetes Ziel für das Sammeln von Daten dar. Die hier verwendeten Bilder geben oft Auskunft über den Besitz eines Nutzers. Unter Umständen gibt dieser sogar Zeiten an, während denen er erreichbar ist. Ein Einbrecher kann so nicht nur Wert der möglichen Beute, sondern auch eine geeignete Zeit für seinen Einbruch bestimmen. Alternativ können Dienste wie Google Maps [28] und Street View [29] verwendet werden, um die Nachbarschaft und das Grundstück des Nutzers einzuschätzen.
Abgesehen von Schutzmaßnahmen, wie sie im Kapitel 3 *Standortbasierte Dienste* beschrieben wurden, lassen sich nur wenige weitere Maßnahmen treffen. Friedland und Sommer sehen als Hauptziel für den Schutz der Nutzer ihre Aufklärung [23]. Die Nutzer müssen sowohl über das Vorhandensein von Geo-Tagging, als auch seine Folgen aufgeklärt werden.

5.3 Tagging in Bildern

Das Markieren von Personen in Bildern ist vor allem in sozialen Netzwerken wie `facebook.com` sehr beliebt. Es bietet die Möglichkeit Bilder mit den Profilen der

darin erscheinenden Nutzer zu verbinden. Auf diese Weise können andere problemlos alle Fotos einsehen, auf denen ein Nutzer auftaucht. Der Nutzer selbst möchte dies aber unter Umständen verhindern. Dabei entstehen verschiedene Spannungen. Entfernt ein Nutzer sein Tag von einem Bild, wird das Bild nicht gelöscht. Es kann also immernoch von anderen gefunden werden. Gerade, wenn das Bild von einem Freund des Nutzers stammt, kann es mit großer Wahrscheinlichkeit trotzdem gefunden werden. Einem erneuten Tagging steht damit nichts entgegen. Zudem ist dieses Bild dann für keinen Nutzer mehr mit dem Profil verbunden. Häufig wollen Nutzer Fotos allerdings nur vor einigen Personen oder Gruppen verstecken [8]. Des Weiteren könnte sich ein uneinsichtiger Tagger persönlich angegriffen fühlen, wenn seine Tags entfernt werden.

Gezielter Ausschluss

Besmer und Lipford stellen deshalb in [8] das Verfahren *Restricting Others* vor. Ziel ist es, mit dem Besitzer des Bildes in Kontakt zu treten und gezielt Nutzern den Zugriff auf das Bild zu verweigern. Dazu kann der Nutzer eine Nachricht an den Besitzer schicken, in dem er um den Ausschluss bestimmter Nutzer bittet. Damit der Nutzer sicher sein kann, dass der Besitzer des Bildes diese Anfrage nicht gegen ihn verwendet[1], werden die Namen der Vetroffenen nicht an den Besitzer gesendet. Während über die Maßnahme verhandelt wird, bleibt der Nutzer ungetaggt, damit er in der Zwischenzeit nicht bereits mit dem Bild in Verbindung gebracht wird. Lehnt der Besitzer die Anfrage ab, kann der Getaggte ebenfalls entscheiden, das Tagging vollständig entfernen.
Stimmt der Besitzer der Bitte zu, wird den ausgewählten Nutzern automatisch der Zugriff auf das Bild verweigert. Somit bleibt die endgültige Kontrolle über das Bild bei seinem Besitzer. Der getaggte Nutzer hat aber dennoch die Möglichkeit, seine Interessen gezielt durchzusetzen. Allerdings muss er dazu genau wissen, welche anderen Nutzer er ausschließen möchte. Dabei kann es schnell vorkommen, dass einzelne Nutzer vergessen werden. Ebenso kann es passieren, dass ein neuer Nutzer zu dem Netzwerk hinzukommt, der Getaggte aber vergisst, dass es Bilder gibt, die dieser nicht sehen soll. Sollen Bilder beispielsweise vor Arbeitskollegen geheim gehalten werden, kann ein neuer Mitarbeiter unter Umständen diese Bilder dennoch einsehen. Das gesamte Verfahren könnte damit vergebens gewesen sein. Um solche und ähnliche Probleme zu vermeiden, wird in [8] vorgeschlagen, Algorithmen einzusetzen, die

[1] z. B. zur Erpressung

den Nutzer bei seinen Entscheidungen unterstützen. Genaue Angaben hierzu werden jedoch nicht gemacht.

Kollaboratives Verwalten von Inhalten

Während das zuvor vorgestellte Verfahren expliziten Wert darauf legt, die Rechte an dem Bild beim Urheber zu belassen, wird in [49] ein anderer Ansatz gewählt. Hauptanliegen der Arbeit ist es, einen simplen Mechanismus zu entwickeln, mit dessen Hilfe Entscheidungen über die Freigabe von gemeinsamen Dokumenten und Bildern getroffen werden können. Eine große Herausforderung ist es dabei, Regeln für ein Bild[2] zu finden, die allen Betroffenen zusagen. Dies ist ohne enormen Kommunikationsaufwand der Nutzer nur schwer zu erreichen. Da Nutzer jedoch nicht vom Aufwand einer solchen Einigung abgeschreckt werden sollen, wird versucht eine Mehrheitsentscheidung zu finden. Die Probleme einer solchen Entscheidung sind ein wichtiges Thema in der Volkswirtschaftslehre. Dementsprechend bedient sich der Ansatz eines Prinzips aus dieser. Für das Treffen einer Entscheidung wird die Verwendung der Clarke-Steuer [14] vorgeschlagen.
Die Nutzer geben den Wert an, den sie in verschiedenen Einstellungen[3] sehen. Dazu müssen sie Punkte auf die einzelnen Möglichkeiten verteilen. Diese Punkte kann ein Nutzer verdienen, indem er Mitbesitzer für die von ihm veröffentlichten Daten bestimmt. Somit ist gleichzeitig ein Anreiz für die Verwendung des Verfahren gegeben. Die Punkte, die die Nutzer auf die verfügbaren Optionen verteilt haben, werden für jede Option summiert. Die Option mit dem höchsten Ergebnis wird als Gruppenpräferenz ausgewählt. Die entsprechenden Einstellungen werden daraufhin auf das Bild angewendet.
Damit die Nutzer fair und ehrlich abstimmen, wird nun die Clarke-Steuer eingesetzt. Durch sie werden Nutzer, die besonders viele Punkte auf den Sieger gesetzt haben, besteuert und müssen entsprechend viele Punkte abgeben. So wird sichergestellt, dass Nutzer mit vielen Punkten diese nicht vollständig einsetzen, sondern entsprechend ihren Präferenzen einen geeigneten Einsatz abwägen. Die Nutzer müssen jedoch in der Lage sein den Wert der einzelnen Optionen zu bestimmen. Dies ist gerade in Relation zu anderen Nutzern nur begrenzt möglich. So kann es schnell passieren, dass ein Nutzer im Vergleich zu den Mitbietern übermäßig viele oder zu wenige Punkte ausgibt.

[2] das Verfahren ist nicht auf Bilder begrenzt; im Rahmen dieser Arbeit wird allerdings nur von Bildern ausgegangen
[3] z. B. öffentlich, privat, nur für Freunde einsehbar

Ein weiteres Problem stellt die mögliche Menge an gemeinsamen Daten dar. Unter Umständen muss ein Nutzer sehr häufig das Verfahren durchlaufen. Dabei kann es vorkommen, dass er sehr oft ähnliche Vorlieben angibt. Es kann dadurch zu mangelndem Interesse und einer geringeren Anwendung des Verfahrens kommen. Um dem entgegen zu wirken, schlagen die Entwickler des Verfahrens vor das Verfahren teilweise zu automatisieren. Dazu werden Bilder anhand vorhandener Tags mit vorherigen verglichen. Das System schlägt dem Nutzer dann vor, die Regel anzuwenden, die bei dem ähnlichsten Bild verwendet wurde. Stimmen alle Nutzer dieser Regel zu, wird sie angewendet. Es wird nur noch abgestimmt, wenn ein Nutzer dem Vorschlag nicht zustimmt.

5.4 Zusammenfassung

Geo-Tagging und das Tagging von Personen ermöglichen massive Eingriffe in die Privatsphäre von betroffenen Nutzern. Dennoch gibt es bisher nur wenige Maßnahmen, um Nutzer zu schützen. Im Fall von Geo-Tagging liegt das Problem dabei hauptsächlich an der Sache selbst. Eine der wenigen möglichen Schutzmaßnahmen ist die Aufklärung der Nutzer über Möglichkeiten und Gefahren.
Beim Tagging von Personen bieten sich jedoch noch viele Möglichkeiten den aktuellen Schutz zu verbessern. Die vorgestellten Verfahren gehen bereits in die richtige Richtung. Beide weisen jedoch noch individuelle Schwächen auf. Vor allem konzentrieren sich beide Verfahren auf Tagging innerhalb eines sozialen Netzwerks. In den meisten Fällen werden Nutzer hier informiert, wenn sie auf einem Bild vermerkt werden. Dienste wie Flickr [19] oder Picasa [43] erlauben es allerdings, auch Nutzer einzutragen[4], die nicht bei dem entsprechenden Dienst angemeldet sind. Solche Tags können genutzt werden, um Bilder einer Person zu finden, die sich dessen nicht bewusst ist. Vor allem in Kombination mit Geo-Tags und gegebenenfalls Tags, die den Inhalt des Bildes beschreiben, können Tags so eine enorme Gefahr darstellen. Dies ist ein Gebiet, auf dem noch einige Arbeit möglich und notwendig ist.

[4] als externes Tag oder direkt im Bild

6 Fazit

In dieser Arbeit wurden verschiedene Gefahren für die Privatsphäre in mobilen Informationssystemen aufgezeigt und ein Überblick über eine Auswahl von Lösungsansätze für diese Probleme gegeben. Es wurden Verfahren zur Verschleierung des genauen Standortes eines Nutzers vorgestellt. Dabei wurde zwischen Verfahren, die mit dem Standort eines einzelnen Nutzers arbeiten, und solchen, die k-Anonymität oder ähnliche Prinzipien umsetzen, unterschieden. Zudem wurden Maßnahmen erläutert, die die Identität von Nutzern in standortbasierten Diensten schützen sollen. Es wurde auch auf Probleme eingegangen, die bei der Wahl von Parametern durch die Nutzer der einzelnen Verfahren auftreten können.
Des Weiteren wurden Maßnahmen zum Schutz verschiedener Aspekte der Privatsphäre in mobilen sozialen Netzwerken dargestellt. Es wurden Verfahren zum Schutz der Beziehungen zwischen Nutzern dargestellt, Möglichkeiten zum Entfernen eines allwissenden zentralen Servers aufgezeigt und die Nutzung von mobilen Geräten zum Vertrauensaufbau durch physikalische Begegnungen vorgestellt. Außerdem wurde ein Verfahren zur Unterstützung der Nutzer beim Aufstellen von Zugriffsregeln für ihre Daten erläutert.
Darüber hinaus wurden Gefahren beim Anreichern von Daten mit Standort- oder Personendaten aufgezeigt. Es wurden erste Lösungsansätze vorgestellt.

Möglichkeiten für weitere Forschung

Der Schutz von Standortdaten wird von vielen Arbeiten thematisiert. Dementsprechend existieren viele und sichere Verfahren auf diesem Gebiet. Viele dieser Verfahren sind jedoch für spezielle Anwendungen entworfen worden oder liefern Daten, wie zum Beispiel geographische Flächen an Stelle eines einzelnen Punktes, die von vielen Diensten nicht verarbeitet werden können. Ein wichtiger Schritt zu sicheren standortbasierten Diensten ist deshalb die Entwicklung einer Schnittstelle zum Austausch von Standortdaten, die verschleierte Standorte vorsieht. Zur Zeit sind Nutzer

darauf angewiesen auf die Anbieter entsprechender Dienste zu vertrauen. Eine solche Schnittstelle kann unabhängige Verschleierungsdienste oder -applikationen ermöglichen. Ein weiterer wichtiger Schritt ist die Aufklärung der Nutzer über die Möglichkeiten, die sich aus ihren Standortdaten für Angreifer ergeben.
Ähnliche Probleme stellen sich im Bezug auf den Datenschutz in sozialen Netzwerken. Viele der Schutzmaßnahmen, die in dieser Arbeit vorgestellt wurden, stammen aus vollständigen Implementierungen sozialer Netzwerke. Oftmals sind die vorgestellten Maßnahmen zudem mit einem Mehraufwand für die Nutzer verbunden. Es ist also notwendig Verfahren zu entwickeln, die einfach in vorhandene soziale Netzwerke integriert werden können und die leicht zu bedienen sind. Dezentrale Netzwerke und Dienste, die die Möglichkeiten von mobilen Endgeräten voll auslasten, sind eine viel versprechende Entwicklung. Auch sie können und müssen durch eine einfachere Benutzbarkeit stark verbessert werden.
Tagging ist ein Gebiet, auf dem bisher nur sehr wenig Forschung betrieben wurde. Während dies bei Geo-Tagging wohl in den begrenzten Möglichkeiten begründet liegt, ist das Tagging von Personen bisher kaum in den Fokus der Forschung gerückt. In diesem Bereich sind noch viele Aspekte, allen voran der Schutz der Nutzer selbst, offen. Durch die starke Beliebtheit von Smartphones mit eingebauter Kamera und schnellem mobilen Internet gewinnt dieses Thema jedoch zunehmend an Bedeutung. Zudem ist dies ein Gebiet, auf dem Nutzer sich selbst kaum durch vorsichtiges Verhalten schützen können. Von allen in dieser Arbeit vorgestellten Gefahren verdient das Tagging zur Zeit wohl die meiste Aufmerksamkeit.

Literaturverzeichnis

[1] Shane Ahern, Dean Eckles, Nathan Good, Simon King, Mor Naaman, and Rahul Nair. Over-Exposed? Privacy Patterns and Considerations in Online and Mobile Photo Sharing. pages 357–366, 2007.

[2] C. A. Ardagna, M. Cremonini, E. Damiani, and P. Samarati. Location Privacy Protection Through Obfuscation-Based Techniques. *Ifip International Federation For Information Processing*, pages 47–60, 2007.

[3] C. A. Ardagna, M. Cremonini, E. Damiani, S. De Capitani Di Vimercati, and A. P. Samarati. middleware architecture for integrating privacy preferences and location accuracy. *In Proc. of IFIP SEC 2007, Sandton, South Africa, May*, 2007.

[4] Claudio Ardagna, Marco Cremonini, Sabrina De Capitani di Vimercati, and Pierangela Samarati. An Obfuscation-Based Approach for Protecting Location Privacy. *IEEE Transactions on Dependable and Secure Computing*, (May 2009):1–16, 2009.

[5] Aaron Beach, Mike Gartrell, and Richard Han. Solutions to Security and Privacy Issues in Mobile Social Networking. *2009 International Conference on Computational Science and Engineering*, pages 1036–1042, August 2009.

[6] A. R. Beresford and F. Stajano. Location privacy in pervasive computing. *IEEE Pervasive Computing*, 2(1):46–55, 2003.

[7] Preibusch Beresford. Privacy-Preserving Friendship Relations for Mobile Social Networking. *w3.org*, 2009.

[8] Andrew Besmer and Heather Lipford. Moving Beyond Untagging: Photo Privacy in a Tagged World. *Interfaces*, pages 1563–1572, 2010.

[9] Joseph Bonneau and S Preibusch. The privacy jungle: On the market for data protection in social networks. *The Eighth Workshop on the Economics of*, pages 1–45, 2009.

[10] Danah M. Boyd and Nicole B. Ellison. Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*, 13(1):210–230, October 2008.

[11] cnet News - Twitter user says vacation tweets led to burglary. http://news.cnet.com/8301-1009_3-10260183-83.html, 08 Juni 2009. Abruf: 11.08.2010.

[12] Leucio Cutillo, Refik Molva, and Thorsten Strufe. Safebook: A privacy-preserving online social network leveraging on real-life trust. *IEEE Communications Magazine*, 47(12):94–101, December 2009.

[13] Leucio Antonio Cutillo, Refik Molva, and Thorsten Strufe. Safebook: Feasibility of transitive cooperation for privacy on a decentralized social network. *2009 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks & Workshops*, (217141):1–6, June 2009.

[14] Harold Demsetz, Property Rights, and See Paul Samuelson. MULTIPART PRICING OF PUBLIC GOODS Edward H. Clarke. *Star*, 7, 1964.

[15] Rinku Dewri, Indrajit Ray, Indrakshi Ray, and Darrell Whitley. On the Optimal Selection of k in the k-Anonymity Problem. 00:2008–2010, 2008.

[16] Facebook. http://www.facebook.com. Abruf: 23.08.2010.

[17] Facebook - Statistik. http://www.facebook.com/press/info.php?statistics. Abruf: 11.08.2010.

[18] Facebook-Datenschutzrichtlinien. http://www.facebook.com/policy.php. Stand: 19.08.2010.

[19] Flickr. http://www.flickr.com/. Abruf: 31.08.2010.

[20] S. Foresti. k- Anonymity.

[21] Julien Freudiger, Maxim Raya, Mark Felegyhazi, Panos Papadimitratos, and Jean-Pierre Hubaux. Mix-zones for location privacy in vehicular networks. 2007.

[22] Julien Freudiger, Reza Shokri, and Jean-pierre Hubaux. On the Optimal Placement of Mix Zones. pages 216–234, 2009.

[23] Gerald Friedland and Robin Sommer. Cybercasing the Joint: On the Privacy Implications of Geo-Tagging. In *Proceedings of the Fifth USENIX Workshop*

on Hot Topics in Security (HotSec 10), 2010.

[24] B. Gedik. Location Privacy in Mobile Systems: A Personalized Anonymization Model. *25th IEEE International Conference on Distributed Computing Systems (ICDCS'05)*, pages 620–629, 2005.

[25] B. Gedik. Protecting Location Privacy with Personalized k-Anonymity: Architecture and Algorithms. *IEEE Transactions on Mobile Computing*, 7(1):1–18, January 2008.

[26] Philippe Golle and Kurt Partridge. On the Anonymity of Home/Work Location Pairs. pages 390–397, 2009.

[27] Google Latitude. http://www.google.com/intl/en_us/mobile/latitude/. Abruf: 11.08.2010.

[28] Google Maps. http://maps.google.com/. Abruf: 11.08.2010.

[29] Google Maps mit Street View. http://maps.google.de/intl/de/help/maps/streetview/. Abruf: 02.09.2010.

[30] M. Gruteser, Baik Hoh, Hui Xiong, and Ansaf Alrabady. Preserving privacy in gps traces via uncertainty-aware path cloaking. *Proceedings of the 14th ACM conference on Computer and communications security*, pages 161 – 171, 2007.

[31] Marco Gruteser and Baik Hoh. On the Anonymity of Periodic Location Samples. pages 179–192, 2005.

[32] I Can Stalk You. http://www.icanstalku.com/. Abruf: 19.08.2010.

[33] Ben Jackson. Locational privacy and wholesale surveillance via photo services, 2010. The 8th Hackers On Planet Earth Conference.

[34] John Krumm. A survey of computational location privacy. *Personal and Ubiquitous Computing*, 13(6):391–399, October 2008.

[35] Y. H. Lin, Ahren Studer, H. C. Hsiao, and J. M. McCune. Spate: Small-group pki-less authenticated trust establishment. *Proceedings of the*, pages 1–14, 2009.

[36] Loopt. http://www.loopt.com/loopt. Abruf: 11.08.2010.

[37] Loopt Pulse. http://www.loopt.com/looptpulse. Abruf: 11.08.2010.

[38] Justin Manweiler, Ryan Scudellari, and Landon P. Cox. SMILE: Encounter-Based Trust for Mobile Social Services. *Encounter*, pages 246–255, 2009.

[39] Arvind Narayanan and Vitaly Shmatikov. Myths and fallacies of personally identifiable information. *Communications of the ACM*, 53(6):24, June 2010.

[40] nytimes.com - Web Photos That Reveal Secrets, Like Where You Live. http://www.nytimes.com/2010/08/12/technology/personaltech/12basics.html?_r=1&scp=1&sq=geo-tagging&st=cse, 11. August 2010. Abruf: 19.08.2010.

[41] OpenStreetMap. http://www.openstreetmap.org/. Abruf: 11.08.2010.

[42] Iain Parris, Greg Bigwood, and Tristan Henderson. Privacy-enhanced social network routing in opportunistic networks. *2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pages 624–629, March 2010.

[43] Picasa Web Albums. http://picasa.google.com/intl/en_us/features-nametags.html. Abruf: 23.08.2010.

[44] Please Rob Me. http://pleaserobme.com/. Abruf: 02.09.2010.

[45] Qype. http://www.qype.com. Abruf: 11.08.2010.

[46] ResearchGATE. http://www.researchgate.net. Abruf: 23.08.2010.

[47] Ahmad-reza Sadeghi, Thomas Schneider, and Immo Wehrenberg. Efficient Privacy-Preserving Face Recognition. 2009.

[48] Norman Sadeh, Jason Hong, Lorrie Cranor, Ian Fette, Patrick Kelley, Madhu Prabaker, and Jinghai Rao. Understanding and capturing people's privacy policies in a mobile social networking application. *Personal and Ubiquitous Computing*, 13(6):401–412, October 2008.

[49] Anna Cinzia Squicciarini, Mohamed Shehab, and Federica Paci. Collective privacy management in social networks. *Proceedings of the 18th international conference on World wide web - WWW '09*, page 521, 2009.

[50] tagesschau.de - Das kurze, aber erfolgreiche Leben der Robin S. http://www.tagesschau.de/ausland/facebookfake100.html, 02. August 2010. Abruf: 11.08.2010.

[51] Tagwhat. http://www.tagwhat.com/. Abruf: 24.08.2010.

[52] Hassan Takabi, James B. D. Joshi, and Hassan A. Karimi. A Collaborative k-Anonymity Approach for Location Privacy in Location-Based Services. *Information Sciences*, 2009.

[53] The Friend of a Friend (FOAF) project. http://www.foaf-project.org/. Abruf: 02.09.2010.

[54] thinq.co.uk - 100 million Facebook pages leaked on torrent site. http://www.thinq.co.uk/2010/7/28/100-million-facebook-pages-leaked-torrent-site/, 28. July 2010. Abruf: 11.08.2010.

[55] Twitter. http://twitter.com/. Abruf: 01.09.2010.

[56] Amin Vahdat and David Becker. Epidemic Routing for Partially-Connected Ad Hoc Networks. *Science*.

[57] Paul Vet. Geo-tagging: Opting in to Total Surveillance, 2010. The 8th Hackers On Planet Earth Conference.

[58] Toby Xu and Ying Cai. Feeling-based location privacy protection for location-based services. *Proceedings of the 16th ACM conference on Computer and communications security - CCS '09*, page 348, 2009.

[59] Ge Zhong and Urs Hengartner. A distributed k-anonymity protocol for location privacy. *2009 IEEE International Conference on Pervasive Computing and Communications*, pages 1–10, March 2009.