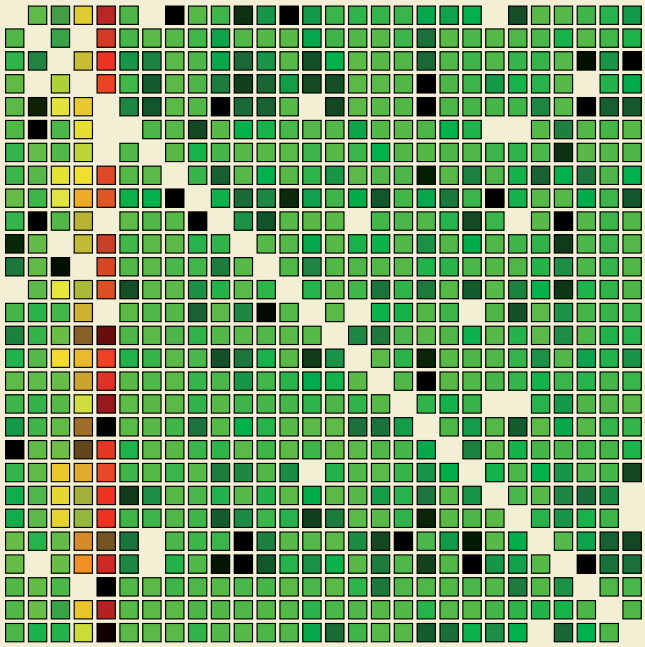
Daniel Kraft

Verteilte Zugangskontrolle in offenen Ad-hoc-Netzen



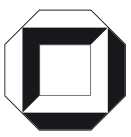
universitätsverlag karlsruhe

Daniel Kraft

Verteilte Zugangskontrolle in offenen Ad-hoc-Netzen

Verteilte Zugangskontrolle in offenen Ad-hoc-Netzen

von
Daniel Kraft



universitätsverlag karlsruhe

Dissertation, Universität Karlsruhe (TH)
Fakultät für Informatik, 2006

Impressum

Universitätsverlag Karlsruhe
c/o Universitätsbibliothek
Straße am Forum 2
D-76131 Karlsruhe
www.uvka.de

Universitätsverlag Karlsruhe 2007
Print on Demand

ISBN: 978-3-86644-121-7

Verteilte Zugangskontrolle in offenen Ad-hoc-Netzen

Zur Erlangung des akademischen Grades eines

Doktors der Ingenieurwissenschaften

von der Fakultät für Informatik
der Universität Karlsruhe (TH)
genehmigte

Dissertation

von
Daniel Kraft
(geb. Müller)
aus Freudenstadt

Tag der mündlichen Prüfung: 22. Juni 2006
Erster Gutachter: Prof. em. Dr. Dr. h.c. mult. Dr.-Ing. E.h. G. Krüger
Zweiter Gutachter: Prof. Dr.-Ing. R. Vollmar

Dank

Zum Gelingen dieser Arbeit haben viele Menschen beigetragen, bei denen ich mich an dieser Stelle dafür herzlich bedanken möchte.

An erster Stelle danke ich ganz besonders meinem Doktorvater Prof. Dr. Dr. h. c. mult. Dr.-Ing. E. h. mult. Gerhard Krüger, der mir im Rahmen meiner Tätigkeit an seinem Lehrstuhl die Möglichkeit zur Erstellung dieser Dissertation gab. Er sorgte für ausgezeichnete Arbeitsbedingungen und ermöglichte mir beispielsweise auch die Vorstellung meiner Arbeiten auf internationalen Konferenzen. Seinen erfahrenen Rat und seine Unterstützung ließ er mir auch über seine Emeritierung hinaus zuteil werden und widmete seinem letzten Promovenden damit einen Teil seines wohlverdienten Ruhestandes.

Herrn Prof. Dr. Roland Vollmar möchte ich herzlich für die bereitwillige Übernahme des Korreferats, sein Interesse und den angenehmen Kontakt im Rahmen der Begutachtung der Arbeit danken.

Meine Kollegen am Institut für Telematik der Universität Karlsruhe, an dem ich als wissenschaftlicher Mitarbeiter tätig war, sowie Frau Prof. Dr. Martina Zitterbart, welche die Institutsleitung nach der Emeritierung Prof. Krügers übernahm, sorgten stets für ein fachlich kompetentes Umfeld und eine freundliche und konstruktive Arbeitsatmosphäre. In besonderer Erinnerung bleiben mir unter Anderem die fruchtbare und rundum angenehme engere Zusammenarbeit mit Dr. Roland Bless und Dr. Frank Pählke und die vielen anregenden Diskussionen und Gespräche mit Dr. Marc Bechler, Verena Kahmann und Dr. Bernhard Thurm. Sehr hilfreich war auch die stetige und kompetente Unterstützung durch die Mitarbeiter aus Verwaltung und Technik; ganz besonders möchte ich dabei Monika Joram für ihr Engagement danken.

Herrn Prof. Dr. Günter Schäfer, der mich auch für die Arbeit am Institut gewonnen hatte, Herrn Dr. Stefan Dresler sowie Herrn Prof. Dr. Klaus Wehrle gebührt besonderer Dank für die fortgesetzte Unterstützung bei der Fertigstellung meiner Arbeit auch über ihren jeweiligen eigenen Weggang vom Institut hinaus. Durch fachliche Diskussion und konstruktive Kritik am probegelesenen Manuskript haben sie wichtige Beiträge dazu geleistet.

Dankend erwähnt werden sollen auch die unzähligen Autoren freier Software, welche eine unverzichtbare Grundlage für die mit der Dissertation verbundenen praktischen Arbeiten gelegt haben. Unter Anderem betrifft dies das Betriebssystem GNU/Linux, das Textsatzsystem LaTeX, die Simulationsumgebung OMNeT++, das Schaltkreisanalysesystem Gnucap, die Visualisierungswerkzeuge Gnuplot und Graphviz und die Skriptsprache Perl.

An der herausragenden letzten Stelle möchte ich mich endlich herzlich bei meiner Familie bedanken, die allzu oft und lange auf ihren Ehemann und Vater verzichten musste, der sich in seiner Arbeit vergraben hatte. Trotzdem fand ich bei meiner Frau Annette steten Glauben an den Erfolg und in ihr eine begeisterte erste Leserin, was mich sehr bestärkte. Einen ganz wesentlichen Anteil am Gelingen haben auch meine Schwiegereltern, welche mir durch ihren persönlichen Einsatz viel Freiraum verschafften, sowie meine Eltern, die mir – neben vielem Anderen – in den zeitkritischsten Phasen eine ideale Arbeitsumgebung boten. Allen gilt mein aufrichtiger Dank.

Karlsruhe, im Februar 2007 *Daniel Kraft*

Inhaltsverzeichnis

Kapitel 1

Einleitung

Mobilität spielt in der heutigen Telekommunikation eine zentrale Rolle. Zur Vernetzung mobiler Geräte kommen derzeit sowohl bei der Sprach- als auch bei der zunehmend aufkommenden Datenkommunikation hauptsächlich infrastrukturbasierte Netze zum Einsatz, bei denen die mobilen Geräte der Netzteilnehmer drahtlos mit ortsfest installierten Zugangsstationen kommunizieren, welche den Übergang zu einem leitungsgebundenen Netz bilden. Beispiele dafür sind Mobiltelefonnetze (GSM, UMTS) oder auch Internet-Zugänge über drahtlose lokale Netze. Im Unterschied zu solchen Infrastrukturnetzen setzen *Ad-hoc-Netze* keine ortsfest installierten Komponenten voraus: Die Teilnehmer kommunizieren mit Hilfe der drahtlosen Übertragungstechnik direkt miteinander. Netzknoten, die sich außerhalb ihrer gegenseitigen Sendereichweite befinden, sind darauf angewiesen, dass dazwischen liegende Knoten Nachrichten für sie weiterleiten, also die Funktion von Zwischensystemen (Routern) in herkömmlichen Netzen übernehmen. Ein wesentlicher Vorteil von Ad-hoc-Netzen ist die Unabhängigkeit vom Vorhandensein bzw. der Funktionsfähigkeit einer vorgegebenen Infrastruktur sowie von den Bedingungen, die der Betreiber einer Infrastruktur für deren Nutzung festlegt. Neben Ad-hoc-Netzen mit vorgegebener Teilnehmermenge, wie sie z. B. firmenintern oder bei der Katastrophenhilfe vorteilhaft eingesetzt werden können, bilden eine allgemeinere Ausprägung die *offenen Ad-hoc-Netze*, bei denen jederzeit neue, vorher unbekannte Teilnehmer beitreten können.

1.1 Problemstellung

Da alle Dienste innerhalb von Ad-hoc-Netzen (inklusive der Weiterleitung von Nachrichten) allein von den Netzteilnehmern erbracht werden, können Ad-hoc-Netze nur dann funktionieren, wenn zumindest ein großer Teil der Teilnehmerschaft sich auch aktiv beteiligt und die eigenen Ressourcen zum Nutzen der Allgemeinheit einbringt. Alle Komponenten von Ad-hoc-Netzen sind in der Regel Eigentum der Teilnehmer. Bei diesen besteht deshalb berechtigtes Interesse daran, das Netz vor unerwünschter Benutzung durch Individuen zu schützen, die zwar gerne Netzdienste nutzen, aber zwecks Schonung der eigenen Ressourcen nicht zu deren Erbringung beitragen wollen. Analog zu Infrastrukturnetzen, wo der Betreiber der Infrastruktur in der Regel die Benutzung und Beeinträchtigung durch nicht zahlende Nutznießer verhindern möchte, ist deshalb auch in Ad-hoc-Netzen die Durchführung einer *Zugangskontrolle* sinnvoll. Dabei ergeben sich allerdings einige besondere Probleme: Während in Infrastrukturnetzen die Zugangskontrolle von besonderen, gut geschützten und stets verfügbaren Komponenten durchgeführt wird, stehen solche spezialisierten Komponenten in Ad-hoc-Netzen nicht zur Verfügung, da aufgrund der hochdynamischen Netztopologie, der Störanfälligkeit der drahtlosen Übertragung und der Begrenztheit der Energievorräte mobiler Geräte eine ständige Verfügbarkeit

und Erreichbarkeit bestimmter Knoten nicht garantiert werden kann. Insbesondere in offenen Ad-hoc-Netzen stellt sich zudem die Frage, wodurch einzelne Knoten legitimiert werden könnten, derart übergeordnete Rollen zu übernehmen.

Einheitliche Sicherheitsrichtlinien, wie sie in Infrastrukturnetzen vom Betreiber festgesetzt werden, um zu bestimmen, wer unter welchen Bedingungen Zugang erhält, sind in offenen Ad-hoc-Netzen, in denen jeder Knoten eine eigene administrative Domäne darstellt, natürlicherweise nicht vorhanden und auch nicht durchsetzbar.

1.2 Zielsetzung

Ziel dieser Arbeit ist die Entwicklung eines Zugangskontrollsystems für offene Ad-hoc-Netze, welches vollkommen *verteilt* realisiert wird, also ohne ausgezeichnete Komponenten auskommt: Alle Teilnehmer sollen darin gleichberechtigt dieselbe Funktion erfüllen. Die Zugangskontrolle wird damit grundsätzlich von jedem Teilnehmer für jeden anderen durchgeführt, und statt der sonst bei Zugangskontrollsystemen gebräuchlichen einmaligen Berechtigungsprüfung bei der Anmeldung an einem vorgegebenen Zugangspunkt erfolgt die Prüfung hier bei jeder Inanspruchnahme von Diensten; die Zugangskontrolle ähnelt damit einer Zugriffskontrolle. Das Kriterium für die Zugangsentscheidung soll das Verhalten der Knoten selbst sein, welches von ihren Nachbarn zu beobachten und zu bewerten ist: Zugang erhalten solche Teilnehmer, die selbst zur Erbringung der Netzwerkdienste beitragen. Nutznießer, die nichts beitragen, werden dagegen ausgeschlossen. Die Zugangskontrolle muss dabei so erfolgen, dass einerseits die Offenheit eines Netzes, also die Möglichkeit des Beitritts neuer Teilnehmer, möglichst wenig beeinträchtigt wird, andererseits aber Teilnehmer, die wegen unfairen Verhaltens ausgeschlossen worden sind, nicht einfach durch einen Wechsel ihrer Identität wieder vollen Zugang erhalten. Voraussetzung für die Durchführung einer Zugangskontrolle ist eine Möglichkeit zur authentischen Identifikation von Kommunikationspartnern und Nachrichten, denn einerseits müssen Verhaltensbeobachtungen den verursachenden Netzknoten zugeordnet werden können und andererseits muss sichergestellt werden, dass Ausschlussmaßnahmen gegen sich unkooperativ verhaltende Knoten auch tatsächlich nur genau diese treffen. Da die in Infrastrukturnetzen gebräuchlichen Verfahren für Authentisierung bzw. Schlüsselverwaltung (z. B. Kerberos, Public-Key-Infrastrukturen) ständig verfügbare zentrale Komponenten voraussetzen, sind sie aus den oben genannten Gründen in Ad-hoc-Netzen nicht ohne Weiteres einsetzbar. Ein Ziel dieser Arbeit ist es deshalb auch, verteilte Dienste für diese Zwecke zur Verfügung zu stellen.

1.3 Lösungsansatz

Um Einschätzungen über die Kooperationswilligkeit benachbarter Knoten und damit die Voraussetzung für die Zugangsentscheidung zu gewinnen, sieht das hier vorgeschlagene Konzept die Beobachtung des von den Nachbarn erzeugten und dank der Rundrufcharakteristik drahtloser Übertragungstechniken auch mithörbaren Netzverkehrs vor. Beobachtetes korrektes Verhalten, beispielsweise bei der Weiterleitung von Datenpaketen oder bei der Erbringung anderer Dienstleistungen, wird dabei positiv gewertet, falsches Verhalten negativ.

Die sich ergebende Gesamtwertung wird anhand einer zweidimensionalen Bewertungsmetrik errechnet, die neben positiven und negativen Einschätzungen auch die Unsicherheit aufgrund fehlender Beobachtungen ausdrücken kann. Durch die Wertung sowohl positiver als auch negativer Beobachtungen werden Nachteile anderer Entwürfe vermieden, bei denen jeweils entweder beobachtetes Fehl-

verhalten zum Ausschluss führt oder eine Art „Bezahlung“ für erbrachte Dienste geleistet wird. Unvermeidliche Ungenauigkeiten der Beobachtung führen beim ersten Ansatz leicht zu fälschlichem Ausschluss, und durch Identitätswechsel kann ein Konto an negativen Beobachtungen annulliert werden. Der zweite Ansatz erfordert eine Verrechnung der erarbeiteten Beträge, für die meist besondere, regelmäßig verfügbare Abrechnungsknoten vorausgesetzt werden.

Die eigentliche Zugangskontrolle wird aufgrund der ermittelten Einschätzungen von jedem Knoten zunächst im Wesentlichen auf der Ebene der Paketweiterleitung umgesetzt, indem ausschließlich Pakete weitergeleitet werden, deren Quelle als kooperativ eingeschätzt wird. Sofern bei einem weiterleitenden Knoten keine Einschätzung bezüglich der Quelle eines Pakets vorhanden ist, kann eine Einschätzung der eigenen Nachbarn eingeholt werden, und unter Umständen kann auch die Quelle selbst eine Art Zertifikat eines für den Weiterleitenden akzeptablen Befürworters liefern. Falls auch auf diese Weise keine positive Einschätzung gewonnen werden kann, wird die Dienstleistung verweigert. Für dem Netz neu beitretende Knoten, über die zunächst keine Einschätzungen vorhanden sind, existiert ein Bürgschaftsverfahren, durch welches ein etablierter Knoten einen Teil seiner eigenen positiven Einschätzung auf einen Kandidaten übertragen kann.

Zur Authentisierung von Netzknoten und Nachrichten wird Public-Key-Kryptographie verwendet. Als Teilnehmeridentitäten werden dabei die öffentlichen Schlüssel selbst eingesetzt, sodass keine Sicherung der Zuordnung zwischen Identitäten und Schlüsseln (etwa durch Zertifikate) erforderlich ist. Sichere Nachrichtenauthentisierung kann durch digitale Signaturen erreicht werden. Um die passenden Schlüssel zur Verifikation zu identifizieren und zu beschaffen oder auch zur direkten Identifikation von Teilnehmern in Situationen, wo keine aufwändige Sicherung nötig ist, dient ein Schlüsselverwaltungsverfahren, das öffentliche Schlüssel bei Bedarf schnell dort zur Verfügung stellt, wo sie benötigt werden.

Um zu ermitteln, unter welchen Randbedingungen (etwa Netzgröße, -dichte und -belastung) das Verfahren effektiv und effizient arbeitet, wurde es in einer Simulationsumgebung implementiert. Simuliert werden unterschiedliche Ad-hoc-Netze mit verschiedenen Modellen für die Benutzermobilität, um einerseits Vergleichbarkeit mit verwandten Arbeiten zu erreichen und andererseits realitätsnahe Ergebnisse zu erhalten. Die Funktion der einzelnen Komponenten des entwickelten Systems und die Wirksamkeit des Zugangskontrollmechanismus im Ganzen wird auch anhand der Reaktion auf simulierte Missbrauchsversuche überprüft. Zusätzlich zur Simulation wird die Sicherheit der in dieser Arbeit entworfenen Zugangskontrollmechanismen und -protokolle gegen Umgehung und Missbrauch analytisch untersucht.

Insgesamt liefert der vorgeschlagene Ansatz ein neuartiges Konzept für eine Zugangskontrolle in offenen Netzen beschränkter Größe, welches vollkommen ohne die in Ad-hoc-Netzen nicht realisierbaren hochverfügbaren Komponenten auskommt. Im Unterschied zu Ansätzen, bei denen zentrale Komponenten zur Verbesserung der Verfügbarkeit auf Untermengen der Teilnehmerknoten verteilt werden, wird durch die vollständige Gleichwertigkeit aller Knoten im vorgeschlagenen Konzept zusätzlich die eingangs angesprochene Problematik der Legitimation hervorgehobener – auch verteilt realisierter – Funktionsträger vollständig umgangen.

1.4 Gliederung

Im Anschluss an diese Einleitung werden in Kapitel 2 zunächst Grundlagen zum Aufbau und zur Funktion von Kommunikationsnetzen, zur Mobilkommunikation und zur Netzwerksicherheit behandelt. Nach der Beschreibung von Besonderheiten bei Ad-hoc-Netzen werden außerdem existierende

Ansätze für die Zugangskontrolle sowie die zugrundeliegende Authentisierung in Ad-hoc-Netzen vorgestellt.

Kapitel 3 ist der Vorstellung des in dieser Arbeit entwickelten Konzepts für die Zugangskontrolle in Ad-hoc-Netzen gewidmet. Nach einer Zusammenfassung der Ziele und Voraussetzungen erfolgt dies zunächst in Form eines Überblicks über alle Bestandteile und deren Zusammenwirken, bevor detaillierter auf einzelne Verfahren und Abläufe eingegangen wird.

In Kapitel 4 wird zunächst die Umsetzung des entwickelten Konzepts in Form einer Simulation behandelt, wobei die verwendeten Modelle sowie einige Details der Implementierung beschrieben werden. Die mit Hilfe der Simulation gefundenen Ergebnisse werden im Anschluss daran vorgestellt und diskutiert.

Kapitel 5 schließt die Arbeit mit einer Zusammenfassung der Ergebnisse und einem Ausblick auf weiterführende Arbeiten ab.

Kapitel 2

Grundlagen und Stand der Technik

In diesem Kapitel werden sowohl Grundlagen eingeführt, die zum Verständnis der folgenden Ausführungen erforderlich sind, als auch einige existierende Arbeiten vorgestellt, die Ansätze und Voraussetzungen für die Zugangskontrolle in Ad-hoc-Netzen liefern.

Zu Beginn des Kapitels werden allgemeine Grundlagen zum Aufbau und zur Funktion von Kommunikationsnetzen behandelt, bevor anschließend auf Besonderheiten der Mobilkommunikation eingegangen und exemplarisch der Standard IEEE 802.11 für drahtlose lokale Netze, der auch zum Aufbau von Ad-hoc-Netzen verwendet werden kann, detaillierter vorgestellt wird. Es folgen Grundbegriffe der Netzwerksicherheit, wichtige kryptographische Algorithmen und gebräuchliche Verfahren für Authentisierung und Schlüsselverwaltung. Weiterhin wird eine Metrik zur Bewertung von Vertrauen und Authentizität vorgestellt. Schließlich werden, nachdem noch auf Besonderheiten von Ad-hoc-Netzen bezüglich Netzwerkfunktionalität und Netzwerksicherheit eingegangen wurde, existierende Ansätze für Authentisierung, Schlüsselverwaltung und Zugangskontrolle in Ad-hoc-Netzen beschrieben.

2.1 Architektur von Telekommunikationssystemen

2.1.1 Schichten und Dienste

Durch die Vielzahl sehr unterschiedlich gearteter Teilaufgaben, die bei der Übermittlung von Information zwischen räumlich getrennten Parteien zu erledigen sind (und auf die in den folgenden Abschnitten noch etwas genauer eingegangen wird), weisen Telekommunikationssysteme eine recht hohe Komplexität auf. Um diese Komplexität beherrschbar zu machen, werden sie modular aufgebaut, wobei sich insbesondere eine Strukturierung in mehrere horizontale *Schichten* bewährt hat.

Jede Schicht übernimmt dabei eine klar abgegrenzte Aufgabe. Sie erbringt für die nächsthöhere Schicht einen bestimmten *Dienst*, d. h. sie stellt einen bestimmten Umfang an Funktionalität zur Verfügung. Bei der Erbringung dieses Dienstes greift sie selbst auf den Dienst der nächsttieferen Schicht zurück, den sie damit also um ihren eigenen Beitrag aufwertet bzw. von dem sie abstrahiert. Jede Schicht – bis auf die jeweilige oberste und die unterste – besitzt damit genau zwei *vertikale Schnittstellen* (pro Instanz) zu den benachbarten Schichten. Die oberste Schicht stellt die eigentliche Anwendung dar, die den gesamten Telekommunikationsdienst in Anspruch nimmt. Die unterste Schicht besteht aus einem physikalischen Medium, über welches Signale zur Überbrückung der räumlichen Distanz übertragen werden können.

Bei der Erbringung ihres Dienstes kommunizieren die verschiedenen räumlich getrennten Instanzen einer Schicht unter Verwendung des von der nächsttieferen Schicht angebotenen, geringerwertigen Kommunikationsdienstes miteinander. Die Regeln für diese *horizontale* Kommunikation zwischen Instanzen derselben Schicht werden als *Protokolle* bezeichnet.

Abbildung 2.1 zeigt als Beispiel drei kommunizierende Systeme bei Verwendung einer Schichtenarchitektur mit fünf Schichten. Die beiden äußeren Systeme beinhalten Instanzen aller Schichten und werden als *Endsysteme* bezeichnet, da sie Endpunkte von Kommunikationsbeziehungen auf der Anwendungsschicht darstellen können. Das mittlere ist ein *Zwischen-* oder *Vermittlungssystem*, das zur Verbindung zweier Streckenabschnitte mit getrennten physikalischen Medien dient, aber selbst keinen Kommunikationsendpunkt darstellen kann.

Der Vorteil der Schichtenarchitektur bei Telekommunikationssystemen gegenüber einem monolithischen Aufbau ist ihre Modularität, welche die unabhängige Entwicklung und Wartung einzelner Schichten erlaubt und Wiederverwendbarkeit bzw. Flexibilität bezüglich ihrer Kombination bietet. Gegenüber diesen Vorteilen fallen die Nachteile des insgesamt etwas höheren Implementierungs- und manchmal auch Laufzeitaufwands meist nicht ins Gewicht.

2.1.2 Paketvermittelte Netze

Physikalische Medien – die immer die unterste Schicht von Telekommunikationssystemen bilden – verbinden jeweils eine bestimmte Menge von Systemen. Oft sind es nur genau zwei Systeme, die durch eine Punkt-zu-Punkt-Strecke verbunden werden; bei mehr als zwei Systemen hat das Medium in der Regel Rundrufcharakteristik, d. h. die von einer Station ausgesandten Signale werden von allen anderen Stationen gleichermaßen (wenn auch unter Umständen mit unterschiedlicher Signalstärke und zeitlicher Verzögerung) empfangen. In beiden Fällen ist die Menge der Empfänger einer ausgesandten Nachricht bereits durch die physikalischen Gegebenheiten festgelegt.

Um beliebige Punkt-zu-Punkt-Kommunikationsbeziehungen innerhalb einer größeren, nicht durch die Notwendigkeit des Vorhandensein eines gemeinsamen physikalischen Mediums beschränkten Teilnehmermenge zu ermöglichen, ist eine *Vermittlungsfunktion* erforderlich, die jeweils an Stellen, an denen auf dem Weg zwischen zwei beliebigen Kommunikationspartnern ein physikalisches Medium endet und ein neues beginnt, einsetzt, um die Unterbrechung zu überbrücken. Durch die Vermittlungsfunktion wird ein Telekommunikationssystem zu einem *Vermittlungsnetzwerk*.

Neben *leitungsvermittelten Netzen* wie beispielsweise dem herkömmlichen Telefonnetz, bei denen durch die (auf Schicht 0 wirkende) Vermittlungsfunktion eine (zumindest bezüglich der zur Signalübertragung verwendeten Eigenschaften) durchgehende physikalische Verbindung zwischen Kommunikationspartnern geschaltet wird, werden zur Datenkommunikation heute meist *paketvermittelte Netze* eingesetzt. Bei ihnen werden Sender und Empfänger entkoppelt, indem die Vermittlungsfunktion die auf einem Medium empfangenen, in mit einer Adressangabe versehene *Pakete* begrenzter Größe unterteilten Nachrichten[1] zwischenspeichert und in der Folge auf einem anderen Medium wieder aussendet.

In Telekommunikationssystemen, die auf paketvermittelten Netzen basieren, können von den höheren Schichten sowohl *verbindungslose* als auch *verbindungsorientierte* Dienste angeboten werden.

[1] Das Wort „Nachricht“ soll auch im Folgenden für inhaltlich zusammenhängende Informationseinheiten verwendet werden, die aber nicht unbedingt „in einem Stück“ übertragen werden müssen. Ein „Paket“ ist dagegen zumindest auf der betrachteten Schicht immer zusammenhängend. Wenn – wie insbesondere in Kapitel 3 häufig der Fall – davon ausgegangen werden kann, dass jede Nachricht auch in einem einzelnen Paket übertragen wird, sind die Begriffe nahezu gleichbedeutend.

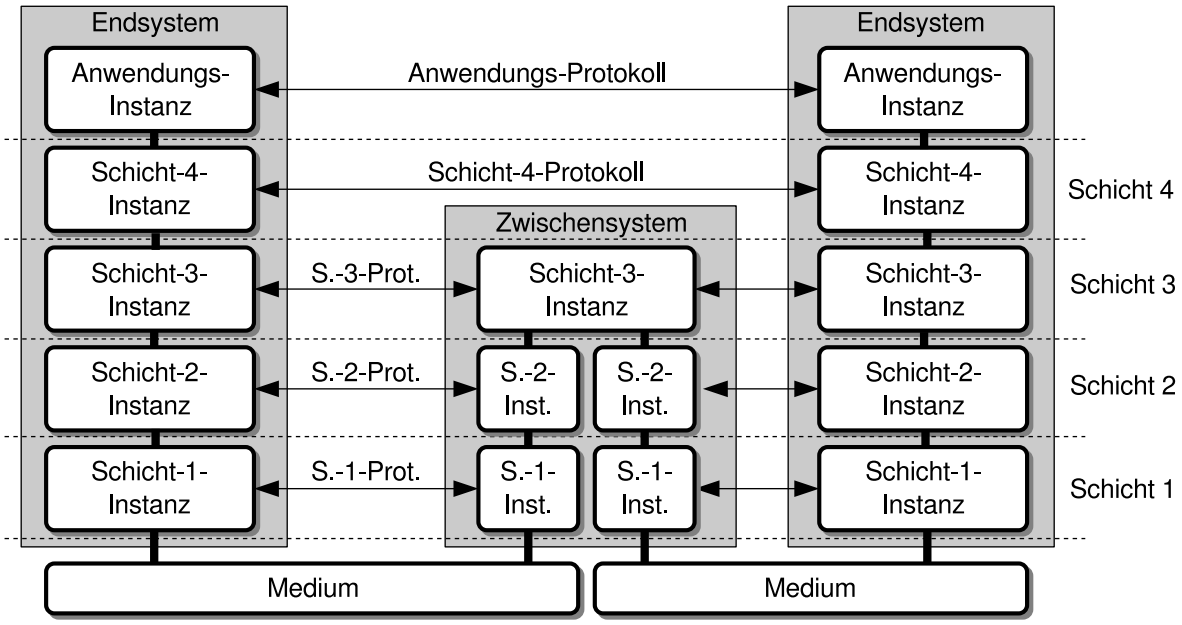


Abbildung 2.1: Schichtenarchitektur kommunizierender Systeme (Beispiel)

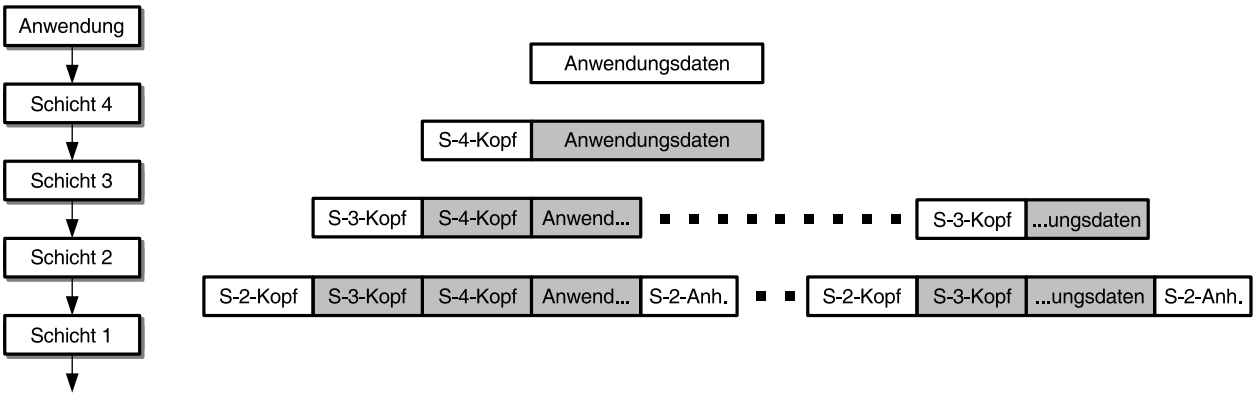


Abbildung 2.2: Ergänzung eines Anwendungsdatagramms um protokollspezifische Zusatzinformation beim Durchlaufen der Schichten des Kommunikationssystem. Die neue Information ist jeweils weiß unterlegt, grau unterlegt sind die von der jeweiligen Schicht transparent als Nutzdaten weitergereichten Teile.

Beim verbindungslosen Dienst müssen die zu übertragenden Daten von der darüberliegenden Schicht in kleineren Einheiten (so genannten *Datagrammen*) und jeweils zusammen mit einer Empfänger-Adresse übergeben werden. Dem Empfänger wird zu den erhaltenen Datagrammen auch jeweils die Absender-Adresse mitgeteilt. Beim verbindungsorientierten Dienst gibt die höhere Schicht dagegen einmalig die Anweisung zum Aufbau einer *(virtuellen) Verbindung* zum Adressaten, und bei der anschließenden Datenübertragung brauchen keine Adressen mehr übergeben zu werden. Die Verbindung wird schließlich auf Anweisung der höheren Schicht (oder unter Umständen im Fehlerfall auch von der erbringenden Schicht selbst) wieder abgebaut.

In jeder Schicht eines Kommunikationssystems müssen deren Instanzen auf den End- und gegebenenfalls Zwischensystemen zur Erbringung der jeweiligen Dienste neben den zu übermittelnden Nutzdaten in der Regel zusätzliche Information austauschen (wie z. B. Adressen). Diese protokollspezifische Zusatzinformation wird in der Regel in Form eines *Protokollkopfs* (engl. Header) den Nutzdaten vorangestellt; seltener werden Protokollinformationen auch noch hinten an die Nutzdaten angehängt. Abbildung 2.2 zeigt beispielhaft, wie ein von einer Anwendung an den Kommunikationsdienst der Schicht 4 übergebenes Datagramm beim Durchlaufen der Schichten des Kommunikationssystem von oben nach unten jeweils um protokollspezifische Zusatzinformation ergänzt wird. Von Schicht 3 wird die ihr übergebene Übertragungseinheit außerdem aufgeteilt, um eine durch Schicht 2 vorgegebene Längenbeschränkung einzuhalten; jeder der beiden Teile erhält einen eigenen Protokollkopf. Schicht 2 ergänzt neben einem Protokollkopf außerdem einen Anhang, bevor das Resultat an Schicht 1 zur Übertragung über das Medium übergeben wird. Auf eventuellen Zwischensystemen und dem Zielsystem werden die jeweils hinzugefügten Zusatzinformationen von jeder durchlaufenen Schicht wieder entfernt (und aufgeteilte Einheiten wieder zusammengefügt), bevor die Nutzdaten an die nächsthöhere Schicht übergeben werden.

2.1.3 Referenzarchitekturen

Um die Vorteile der Modularität in Schichten untergliederter Kommunikationssysteme voll ausnutzen zu können, ist eine internationale Standardisierung der vertikalen Schnittstellen zwischen den Schichten sinnvoll. Idealerweise erhält man so ein offenes System, bei dem Schichten unterschiedlicher interner Funktionsweise relativ frei kombiniert werden können.

Eine zusätzliche Standardisierung von Protokollen, also der horizontalen Schnittstellen zwischen den Instanzen derselben Schicht in verschiedenen Systemen, ermöglicht die Kommunikation zwischen

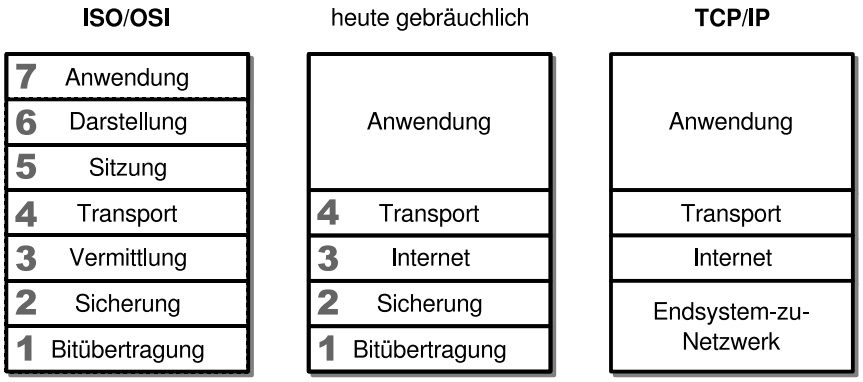


Abbildung 2.3: Referenzarchitekturen im Vergleich

den Systemen auch dann, wenn unterschiedliche Implementierungen für die Instanzen der Schicht verwendet werden. Erst damit werden so große offene Netze wie das Internet ermöglicht, in dem beliebige Systeme verschiedener Hersteller problemlos kommunizieren können.

Ein wichtiger Schritt bei der Standardisierung der Schnittstellen in Kommunikationssystemen war das so genannte *ISO/OSI-Basisreferenzmodell für die Kommunikation in offenen Systemen* (Open Systems Interconnection Basic Reference Modell) [ISO94], das erstmals 1984 von der *International Organization for Standardization* (ISO) spezifiziert wurde. Es definiert die in Abbildung 2.3 links gezeigten sieben Schichten mit deren genauen Aufgaben und Schnittstellen untereinander:

1. Die *Bitübertragungsschicht* sorgt für die ungesicherte Übertragung von Daten in Form unstrukturierter Bitfolgen als Signale über ein physikalisches Medium zwischen zwei Systemen, die durch dieses Medium direkt verbunden, also netztopologisch benachbart sind. Innerhalb dieser Schicht sind die Eigenschaften der verwendeten Medien und die einzusetzenden Codierverfahren festgelegt.

2. Die *Sicherungsschicht* bietet eine bezüglich ihrer Zuverlässigkeit abgesicherte Datenübertragung zwischen benachbarten Systemen an: Bitfolgen werden in so genannte Rahmen unterteilt, und Fehler bei der Übertragung der Rahmen werden (z. B. anhand zugefügter Prüfsummen) erkannt und (z. B. durch Wiederholung einzelner Rahmen) behandelt. Außerdem wird durch *Flusskontrolle* dafür gesorgt, dass der Empfänger nicht durch zu schnelle Datenübertragung überfordert wird. Falls ein geteiltes Medium verwendet wird, ist die Kontrolle des Zugriffs verschiedener Systeme auf das Medium ebenfalls Aufgabe der Sicherungsschicht.

   Ein Reihe sehr gebräuchlicher Standards für konkrete Implementierungen der Bitübertragungs- und Sicherungsschicht unter Einsatz unterschiedlicher Übertragungstechniken für lokale Netze ist die Serie 802.x der IEEE. Dort wird die Sicherungsschicht weiter unterteilt in die zwei Unterschichten *Logical Link Control* (LLC) und *Medium Access Control* (MAC): Die LLC-Schicht ist bei allen 802.x-Standards gleich, die MAC-Schicht enthält die veränderlichen Bestandteile, unter Anderem die Medienzugriffskontrolle.

3. Die *Vermittlungs-* oder *Netzwerkschicht* enthält die im vorigen Abschnitt beschriebene Vermittlungsfunktion, durch die einzelne Verbindungsstrecken der Sicherungsschicht zu Vermittlungsnetzen verknüpft werden. Eine wichtige hierbei zu erfüllende Aufgabe ist die Bestimmung günstiger Wege durch das Netz zwischen beliebigen Endsystemen. Bei paketvermittelten Netzen müssen dann weiterhin Pakete entlang solcher Wege weitergeleitet werden. Unter Umständen sind dabei unterschiedliche Eigenschaften der vernetzten Teilstrecken auszugleichen, indem z. B. Paketgrößen durch Aufteilen von Paketen (die sog. Fragmentierung) verän-

dert werden. Außerdem sollte mittels *Staukontrolle* dafür gesorgt werden, dass Stausituationen, die entstehen können, wenn zu viele Pakete zu einem Zeitpunkt einen bestimmten Netzbereich durchqueren sollen, erkannt und behoben werden.

Der Name „Netzwerkschicht“ könnte suggerieren, dass erst mit Einführung dieser Schicht (beim Aufbau „von unten“ her) von einem „Netzwerk“ gesprochen werden kann. Tatsächlich bezeichnet man aber oft auch schon den auf Schicht 2 erreichten Zusammenschluss mehrerer Endsysteme ohne Vermittlungsfunktion als Netzwerk, beispielsweise im Fall so genannter lokaler Netze (*Local Area Networks*, LANs) auf Basis der oben genannten IEEE-Standards der 802.x-Serie. Bei solcher Bezeichnung verbindet dann die Vermittlungsfunktion mehrere Netze zu einem so genannten Inter-Netz, wofür das Internet ein Beispiel ist.

4. Die *Transportschicht* erweitert die von der Vermittlungsschicht angebotene Endsystem-zu-Endsystem-Kommunikation um die Möglichkeit, einzelne Anwendungen auf den Endsystemen zu adressieren. Die Instanzen höherer Schichten gehören damit jeweils schon zu bestimmten Anwendungen, weshalb die Schichten 5 bis 7 auch als *anwendungsorientierte* Schichten im Unterschied zu den *transportorientierten* Schichten 1 bis 4 bezeichnet werden.

   Zur Ergänzung der bereits auf tieferen Schichten durchgeführten und sich damit jeweils nur auf einzelne Teilstrecken beziehenden Maßnahmen dieser Art führt die Transportschicht Fehlererkennung und -behebung sowie Flusskontrolle auf Ende-zu-Ende-Basis durch.

5. Die *Sitzungs-* oder auch *Kommunikationssteuerungsschicht* dient zur Strukturierung des Datenaustauschs zwischen Anwendungen. So kann z. B. ein Senderecht verwaltet werden, und es können Synchronisationspunkte gesetzt werden, zu denen im Fehlerfall zurückgesprungen werden kann.

6. Die *Darstellungsschicht* dient zum Abgleich unterschiedlicher Datenformate auf den kommunizierenden Systemen mit Hilfe einer standardisierten Übertragungssyntax.

7. Die *Anwendungsschicht* enthält die eigentliche Anwendung, die das Kommunikationssystem nutzt. In der ISO/OSI-Standardisierung werden zur Unterstützung des Anwendungsentwicklers eine Reihe von Funktionsblöcken aus Bereichen wie elektronischer Post, Dateiübertragung und entferntem Prozeduraufruf vorgegeben.

Die von der ISO in Zusammenhang mit dem Basisreferenzmodell standardisierten Protokolle und die Dienste der Schichten 5 bis 7 konnten sich in der Praxis nicht durchsetzen, was unter Anderem wohl in ihrer hohen Komplexität begründet liegt. Die vorgestellte Schichtenstruktur kommunizierender Systeme spielt aber noch immer eine Rolle als strukturelle Grundlage beim Entwurf und der Beschreibung von Telekommunikationssystemen.

Bereits vor der Standardisierung des ISO/OSI-Basisreferenzmodells entwickelten sich die Vorläufer des heutigen Internet auf Basis einer eigenen Schichtenarchitektur und zugehöriger Dienste und Protokolle. Dieses nach seinen Vermittlungs- und Transportschichtprotokollen auch TCP/IP-Modell genannte Referenzmodell unterscheidet nur vier Schichten, stimmt also nicht mit dem ISO/OSI-Modell überein. Es lässt sich aber ungefähr auf dieses abbilden. Abbildung 2.3 stellt beide Modelle gegenüber (ganz links ISO/OSI, ganz rechts TCP/IP). Die Dienste und Protokolle des TCP/IP-Modells und damit des Internets werden von der *Internet Society* (ISOC) standardisiert, einer offenen internationalen Gesellschaft. Die Standardisierungsdokumente werden innerhalb der untergeordneten *Internet Engineering Task Force* (IETF) unter Beteiligung aller interessierten Parteien entwickelt.

In der Netzwerkschicht kommt im TCP/IP-Modell das *Internet-Protokoll* (IP) zum Einsatz, welches durch Hilfsprotokolle für Fehlermeldung und -diagnose (*Internet Control Message Protocol*, ICMP),

Verwaltung von Kommunikationsgruppen (*Internet Group Management Protocol*, IGMP) und Abbildung zwischen IP-Adressen und Adressen der Sicherungsschicht (*Address Resolution Protocol*, ARP) ergänzt wird. Das Internet-Protokoll liefert einen verbindungslosen und unzuverlässigen Dienst: Pakete können – beispielsweise in Überlastsituationen – ohne Rückmeldung an den Absender verworfen werden.

In der Transportschicht werden sowohl ein verbindungsorientierter als auch ein verbindungsloser Dienst angeboten. Ersterer wird mit Hilfe des *Transmission Control Protocol* (TCP) erbracht, letzterer verwendet das *User Datagram Protocol* (UDP). TCP transportiert einen kontinuierlichen Datenstrom zuverlässig zur Zielanwendung, indem es ihn selbst in Pakete unterteilt und Paketverluste oder andere Übertragungsfehler durch Übertragungswiederholungen repariert. UDP dagegen bietet nur einen unzuverlässigen Datagramm-Dienst, und damit – abgesehen von der Möglichkeit der Adressierung von Anwendungen anstelle von Endsystemen – im Wesentlichen dieselbe Funktionalität wie IP selbst.

Alles oberhalb der Transportschicht zählt im TCP/IP-Modell zur Anwendungsschicht. Die Funktionen der Sitzungs- und Darstellungsschicht müssen – falls benötigt – von der jeweiligen Anwendung selbst übernommen werden. Es existiert eine große Zahl standardisierter Protokolle, die teilweise wiederum aufeinander aufbauen, so dass sich dann auch eine weitere, für den jeweiligen Anwendungsfall spezialisierte Schichtung ergibt. Zwei der bekanntesten Protokolle der Anwendungsschicht sind das Hypertext Transfer Protocol (HTTP) zur Übertragung von Dokumenten des WWW und das Simple Mail Transfer Protocol (SMTP) zur Übertragung elektronischer Post (E-Mail).

Die Bitübertragungs- und die Sicherungsschicht wurden im TCP/IP-Modell ursprünglich nicht klar definiert; es wurde lediglich erwähnt, dass die Endsysteme mit dem Netz verbunden sein müssen, so dass Pakete übertragen werden können [CeKa74, Post81]. In neueren Internet-Standards wird ebenfalls ISO/OSI-Terminologie verwendet (beispielsweise [Simp94]), so dass sich das in Abbildung 2.3 in der Mitte dargestellte Referenzmodell ergibt, das heutzutage meist zugrundegelegt wird.

2.1.4 Wegfindung und Weiterleitung in paketvermittelten Netzen

Zwei zentrale Funktionen der Vermittlungsschicht paketvermittelter Netze sind die Bestimmung von Wegen zwischen kommunikationswilligen Endsystemen über eine Reihe von Zwischensystemen hinweg sowie die Weiterleitung von Paketen entlang dieser Wege. Beide Funktionen werden oft ungenau ohne Unterschied mit dem englischen Wort *Routing* bezeichnet – die korrekte Bezeichnung unterscheidet *Wegfindung* (engl. Routing) und *Weiterleitung* (engl. Forwarding).

In der Vermittlungsschicht des Internet ist genau genommen nur die Weiterleitung implementiert, während die Wegfindung durch einen unabhängigen Prozess (bezeichnet als Routing-Daemon) erfolgt. Aus der von ihr ermittelten Weginformation wird die zur Weiterleitung durch das eigene System benötigte Information (im Folgenden als Weiterleitungsinformation bezeichnet) extrahiert und in eine als Schnittstelle fungierende Datenbasis einspeist, wo sie von der Vermittlungsschicht bei der Weiterleitung einzelner Pakete ausgelesen werden kann.

Die Weiterleitungsinformation gibt der Vermittlungsschicht an, über welche Netzwerkschnittstelle, letztendlich also über welches physikalische Medium, ein gegebenes Paket weitergeleitet werden muss, damit es – eventuell nach weiteren Weiterleitungsschritten – ein bestimmtes Ziel erreicht. Die Auswahl der Netzwerkschnittstelle erfolgt anhand der Adresse des Zielsystems. Da in großen Netzen wie dem Internet nicht jedes Endsystem eine Liste aller übrigen erreichbaren Endsysteme halten kann, um diesen die jeweils zu verwendende Netzwerkschnittstelle zuordnen zu können, müssen die dabei verwendeten Adressen hierarchisch aufgebaut sein, so dass auf einfache Weise große Adressbereiche, die sich bezüglich der lokalen Wegewahl nicht unterscheiden, zusammengefasst und einer

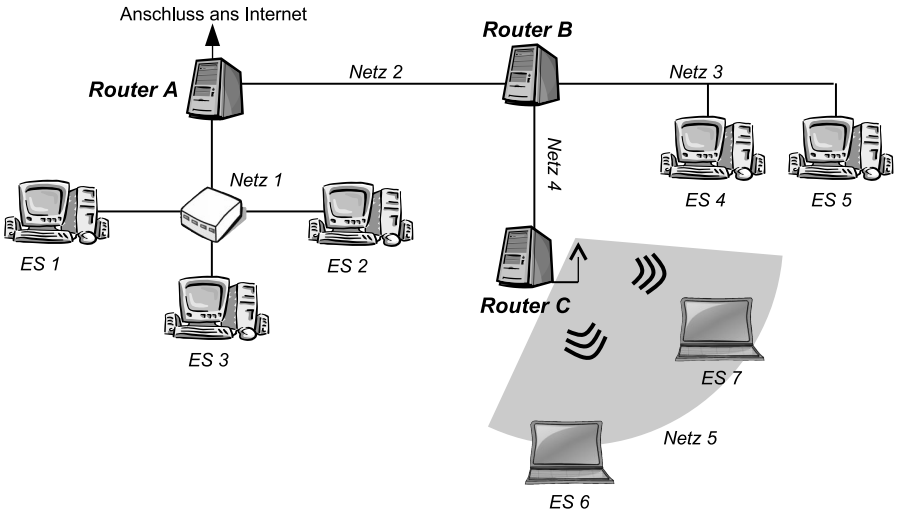


Abbildung 2.4: Beispielnetz bestehend aus fünf einzelnen lokalen Netzen, die über Router (Knoten mit Vermittlungsfunktion) verknüpft werden

Netzwerkschnittstelle zugeordnet werden können. Tatsächlich besteht auf den im Internet verwendeten Adressen eine solche Hierarchie, die anhand administrativer Zuständigkeiten und anhand der Netztopologie aufgebaut ist.

Bei der Einbindung mobiler Endsysteme in das Internet entsteht wegen der topologischen Strukturierung des Adressraums das Problem, dass Endsysteme, die sich zu unterschiedlichen Zeitpunkten an verschiedenen Stellen der Netztopologie befinden, keine permanente und immer topologisch richtig bleibende Adresse besitzen können. Als Lösungsansatz dafür würde von der IETF *Mobile IP* entwickelt [Perk02]; dabei verwaltet ein Platzhaltersystem (der Heimatagent) in einem ortsfesten Heimatnetz die Abbildung von einer stets gleichbleibenden, dem Adressbereich des Heimatznetzes entnommenen Adresse des mobilen Systems auf eine zusätzliche, sich je nach Aufenthaltsort verändernde Adresse aus dem jeweiligen lokalen Adressbereich. In nicht mit dem Internet verbundenen Ad-hoc-Netzen werden dagegen in der Regel keine weltweit eindeutigen Adressen benötigt, und auch eine topologische Strukturierung ist nicht erforderlich. In Abschnitt 2.4.3 wird noch näher auf Besonderheiten der Netzwerkschicht in Ad-hoc-Netzen eingegangen.

Abbildung 2.4 zeigt ein Beispiel für ein aus fünf Schicht-2-Netzen (LANs) bestehendes Inter-Netz. Die Schicht-2-Netze werden über Knoten mit Vermittlungsfunktion – so genannte *Router* – verknüpft und besitzen teilweise unterschiedliche Architekturen: Während die Netze 2 und 4 Punkt-zu-Punkt-Strecken sind, verwendet Netz 3 ein geteiltes Medium, an das alle Knoten angeschlossen sind. Bei Netz 1 hat das Medium eine sternförmige Struktur, eine besondere Komponente mit einer Art Vermittlungsfunktion auf Schicht 2 (üblicherweise als *Switch* bezeichnet) sorgt dafür, dass jeder der an sie angeschlossenen Knoten mit jedem anderen kommunizieren kann. Netz 5 schließlich verwendet drahtlose Übertragung. Ein Paket von Endsystem ES 2 an ES 7 wird auf Schicht 3 von den Routern A, B und C in dieser Reihenfolge weitergeleitet.

Die für die Weiterleitung benötigte topologieabhänge Weiterleitungsinformation wird bei so kleinen Netzen wie dem aus Abbildung 2.4 in der Regel statisch „von Hand“ konfiguriert. In größeren Netzen wird sie durch ein Wegfindungsverfahren automatisch ermittelt und ständig aktualisiert. Dazu verwendet die Wegfindungsinstanz (der Routing-Daemon) ein *Wegfindungsprotokoll* (auch Routing-Protokoll), in dessen Rahmen sie Informationen mit anderen Wegfindungsinstanzen austauscht; die ausgetauschten Nachrichten werden häufig als *Routing-Updates* bezeichnet. In Gebrauch sind zwei verschiedene Typen von Wegfindungsverfahren:

- Bei *Link-State-Verfahren* flutet jeder Router regelmäßig ein so genanntes Link-State-Paket (d. h. er sendet es an alle Router des Netzwerks), das alle seine Nachbarn angibt, also diejenigen Router, zu denen er momentan eine direkte Verbindung auf der darunterliegenden Schicht hat; Letzteres wird meist regelmäßig durch kurze Testnachrichten kontrolliert. Anhand der empfangenen Link-State-Pakete bestimmt jeder Router die Netztopologie in Form eines Graphen und berechnet darin die kürzesten Wege zu allen anderen Routern, etwa mit Hilfe des Dijkstra-Algorithmus [Dijk59].

  Das Ergebnis hängt nicht von Berechnungen anderer Router ab, da die Link-State-Pakete nur sicheres Wissen enthalten, im Unterschied zum unten beschriebenen Distanz-Vektor-Verfahren, dessen Update-Nachrichten abgeleitete Information tragen. Die Fehlerdiagnose ist bei Link-State-Verfahren einfach, da die Link-State-Pakete unverändert durchs Netz wandern. Link-State-Verfahren reagieren rasch auf Topologieänderungen und Knotenausfälle. Die Größe der Link-State-Pakete hängt nicht von der Netzgröße ab, die von den Routern zu speichernde Information ist aber recht umfangreich.

- Bei *Distanz-Vektor-Verfahren* pflegt jeder Router eine Liste, in der er die Distanz zu jedem möglichen Ziel speichert, sowie den jeweiligen Nachbar-Router, über den der Pfad der angegebenen Länge führt. (Diese zweite Angabe kann man als Richtungsvektor auffassen, daher der Name des Verfahrens.) Anfangs enthält die Liste nur den Router selbst mit Entfernung Null, alle übrigen Einträge erhalten die Entfernung „unendlich“ bzw. sind nicht vorhanden, was Ersterem gleichgesetzt wird. Am Anfang und periodisch oder bei jeder Änderung an seiner Liste teilt ein Router seine gesamte Liste den benachbarten Routern mit. Bei Erhalt einer solchen Mitteilung von einem Nachbarn wird überprüft, ob eine der darin angegebenen Entfernungen zuzüglich der Distanz zum Nachbarn (häufig einfach pauschal als 1 gerechnet) niedriger ist als die selbst gespeicherte. Falls ja, wird der Eintrag auf die neue Entfernung und den zugehörigen Nachbarn geändert.

  Distanz-Vektor-Verfahren haben den Vorteil, einfach implementierbar und wenig rechenaufwändig zu sein. Durch die schrittweise erfolgende Verbreitung von Änderungen können aber Informationen unterschiedlicher Aktualität gleichzeitig im Netz existieren, was zur (vorübergehenden) Bildung von Schleifen in der Wege-Information führen kann [CRKGLA89]. Einzelne Knoten haben nie eine Gesamtsicht auf die Netztopologie und können veraltete Information nicht erkennen; diese wird deshalb noch so lange weiter verbreitet, bis die enthaltene Distanz, die bei jeder Weitergabe erhöht wird, eine festgelegte Obergrenze erreicht, bei der das Ziel als unerreichbar gewertet wird (dies ist das so genannte Count-to-Infinity-Problem). Es gibt Verfahren, um diese Probleme zu entschärfen.

Auf Besonderheiten bei der Wegfindung in Ad-hoc-Netzen wird in Abschnitt 2.4.3 eingegangen.

2.2 Mobilkommunikation

Mobilität spielt in der modernen Telekommunikation eine immer größer werdende Rolle. Neben der mittlerweile nahezu überall und jederzeit gegebenen Möglichkeit zur Sprachkommunikation mit entfernten Gesprächspartnern entsteht dabei zunehmend auch ein Bedarf für Datenkommunikation, z. B. zum Austausch von schriftlichen Mitteilungen, Bildern oder anderen Dokumenten zwischen mobilen Geräten. Bei der Unterstützung von Mobilität in Kommunikationssystemen treten neuartige Problemstellungen auf, auf die im Folgenden kurz eingegangen werden soll.

2.2.1 Begriffe

Mobilität ist zunächst eine Eigenschaft eines (menschlichen oder auch in einer rechnergestützten Anwendung bestehenden) Nutzers von Telekommunikationssystemen, der die Dienste des Systems unabhängig von seinem Aufenthaltsort in gleicher Weise nutzen möchte. Dies kann z. B. durch die Bereitstellung von (selbst ortsfesten) Dienstzugangspunkten an den jeweiligen Aufenthaltsorten des Nutzers ermöglicht werden. Mit Hilfe mobiler Geräte können andererseits aber auch mobile Dienstzugangspunkte angeboten werden, die funktionsfähig bleiben, während die Geräte ihren Aufenthaltsort verändern.

Unabhängig davon, ob Kommunikationssysteme mobile Nutzer unterstützen oder nicht, lassen sich die verwendeten Kommunikationsmedien in zwei Kategorien einteilen. Man spricht von *leitungsgebundener* Kommunikation, wenn Leitungen zur Übertragung von Signalen verwendet werden, und von *drahtloser* Kommunikation bei der Übertragung durch den freien Raum, meist mittels elektromagnetischer Wellen. Im Ausgleich für die Vorteile der drahtlosen Kommunikationstechnik, durch die Kommunikation unabhängig vom Vorhandensein von Leitungen ermöglicht wird, muss als Nachteil in Kauf genommen werden, dass drahtlose Kommunikation sehr viel störanfälliger als leitungsgebundene ist.

Grundsätzlich sind für alle Kombinationen von mobil versus ortsfest bzw. drahtlos versus leitungsgebunden sinnvolle Anwendungen denkbar. Für diese Arbeit sind aber hauptsächlich Szenarien mit mobilen Geräten und drahtloser Kommunikationstechnik relevant, also der Fall, der sich am meisten von „herkömmlichen“ Kommunikationssystemen unterscheidet.

Vorherrschend bei der Vernetzung mobiler Geräte sind derzeit infrastrukturbasierte Netze, bei denen die Geräte der Netzteilnehmer drahtlos mit ortsfest installierten Zugangsstationen kommunizieren, welche den Übergang zu einem leitungsgebundenen Netz bilden. Beispiele dafür sind Mobiltelefonnetze (GSM, UMTS) oder auch Internet-Zugänge über drahtlose lokale Netze (z. B. IEEE 802.11, siehe Abschnitt 2.2.4). Im Unterschied zu solchen Infrastrukturnetzen setzen *Ad-hoc-Netze* keine ortsfest installierten Komponenten voraus: Die Teilnehmer kommunizieren mit Hilfe drahtloser Übertragungstechnik direkt miteinander; Netzknoten, die sich außerhalb ihrer gegenseitigen Sendereichweite befinden, sind dabei darauf angewiesen, dass dazwischen liegende Knoten Nachrichten für sie weiterleiten, also die Funktionalität von Zwischensystemen (Routern) in herkömmlichen Netzen übernehmen (siehe Beispiel in Abbildung 2.5). Die Knoten stellen damit sozusagen selbst die benötigte und in diesem Fall mobile Infrastruktur. Ein wesentlicher Vorteil von Ad-hoc-Netzen ist die Unabhängigkeit vom Vorhandensein bzw. der Funktionsfähigkeit einer vorgegebenen Infrastruktur sowie von den Bedingungen, die der Betreiber einer solchen Infrastruktur für deren Nutzung festlegt. Sie können beispielsweise auch dort eingesetzt werden, wo diejenige Partei, die den Betrieb einer ortsfest installierten Infrastruktur legitimieren müsste, kein Interesse daran hat, dies zu tun.

2.2.2 Eigenschaften mobiler und drahtloser Kommunikation

Drahtlose Kommunikation unterscheidet sich von leitungsgebundener hauptsächlich in den folgenden Eigenschaften:

- Die verfügbare Bandbreite ist erheblich geringer als bei leitungsgebundener Kommunikation, was auch daran liegt, dass das Medium immer unter allen Nutzern geteilt werden muss. Ein „Raummultiplex“ wie bei Leitungen ist nicht möglich.
- Die Fehlerrate ist auf drahtlosen Übertragungsstrecken wesentlich höher, was dadurch begründet ist, dass elektromagnetische Wellen aus in der Regel reichlich vorhandenen anderen Quellen

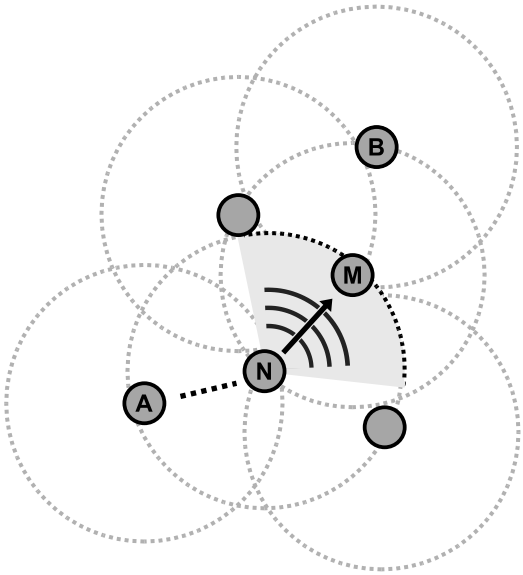


Abbildung 2.5: Weiterleitung im Ad-hoc-Netz: Ein von Knoten A an Knoten B zu übertragendes Paket wird von den Knoten N und M weitergeleitet. Die Übertragungsreichweite ist durch gestrichelte Kreise um die potentiellen Sender angedeutet.

als Störungen wirken. Die höhere Fehlerrate wirkt sich in der Praxis z. B. bei Staukontrollverfahren wie demjenigen des Transportprotokolls TCP negativ aus, welches Paketverluste dann fälschlich als Anzeichen für die Überlastung von Zwischensystemen wertet.

- Die Verzögerung bei der Übertragung von Paketen bzw. die Schwankungsbreite der Verzögerung ist höher, was teilweise durch Fehlerkontroll- und -korrekturverfahren verursacht wird. Auch hierdurch können unerwünschte Wechselwirkungen mit Staukontrollmechanismen oder sogar mit Anwendungsprotokollen oder Benutzern entstehen, wenn diese knappe Annahmen zu Übertragungszeiten machen.

- Übertragungsstrecken fallen relativ häufig aus, entweder, weil die Endpunkte ihre gegenseitige Sendereichweite verlassen, oder weil die elektromagnetischen Wellen durch Hindernisse abgeschattet werden. Aufgrund physikalischer Gegebenheiten im Raum zwischen den Endpunkten oder verschiedener Sendeleistungen der Endpunkte können auch asymmetrische Strecken auftreten, bei denen nur in eine Richtung Daten übertragen werden können.

- Der einfache Zugang zum Medium ohne die Notwendigkeit einer Verkabelung bedeutet auch, dass jede Kommunikation innerhalb der Sendereichweite mitgehört werden kann und auch Eingriffe in die Kommunikation zwischen anderen relativ leicht möglich sind. Dies erleichtert einige Angriffe und erfordert deshalb besondere Sicherheitsmaßnahmen.

Durch die Mobilität der Teilnehmer ergeben sich weitere Unterschiede:

- Die hohe Dynamik der Netztopologie stellt höhere Anforderungen an die unteren Schichten bis zur Netzwerkschicht.

- Auch die Menge der an einem Netz teilnehmenden Knoten kann sich sehr viel dynamischer entwickeln als bei leitungsgebundenen bzw. ortsfesten Netzen, einerseits weil die Teilnahme an den Aufenthalt in einem bestimmten räumlichen Gebiet gebunden ist, welches betreten und verlassen werden kann, und andererseits, weil mobile Geräte häufiger zu Energiesparzwecken abgeschaltet werden als ortsfeste, die meist mit dem Stromnetz verbunden sind.

- An Geräte und Verfahren entsteht die Anforderung, dass sie möglichst weltweit einsetzbar sein sollen, und zwar möglichst lizenzfrei, also ohne jeweils erst eine besondere Nutzungserlaubnis erwerben zu müssen. Die Abhängigkeit von nationalen Regelungen zur Nutzung von Frequenzbereichen erschwert die Schaffung weltweit nutzbarer Verfahren für drahtlose Übertragung; effektiv stehen nur wenige geeignete Frequenzbereiche frei zur Verfügung, und es dürfen dort nur geringe Sendeleistungen verwendet werden.

Je nach Leistungsfähigkeit der verwendeten Endgeräte können weitere Einschränkungen entstehen:

- Mobile Geräte haben nur einen beschränkten Energievorrat, sofern sie nicht an ihrer jeweiligen Aufenthaltsort an eine Stromversorgung angeschlossen werden. Verwendete Verfahren müssen deshalb Energie sparend arbeiten, was einerseits durch Einsparung von Rechenoperationen, wesentlich mehr noch aber durch das Verkürzen von Sendevorgängen erreicht werden kann.
- Eine Folge der Energieknappheit ist, dass mobile Geräte meist mit schwächeren Prozessoren ausgestattet sind, die nur eine beschränkte Rechenleistung zur Verfügung stellen.
- Auch die Speicherkapazität mobiler Geräte, sowohl bezogen auf flüchtigen als auch auf nicht flüchtigen Speicher, ist geringer, um Platz und Energie zu sparen.

2.2.3 Drahtlose Übertragungstechnik

Gegenstand dieses Abschnitts sind verschiedene Ausprägungen drahtloser Übertragungstechnik, welche meist die Grundlage für die Kommunikation zwischen den Knoten von Ad-hoc-Netzen darstellen. Bei den für die Datenübertragung verwendeten elektromagnetischen Wellen kann zunächst bezüglich des Frequenzbereichs grundsätzlich zwischen Radiowellen und Lichtwellen unterschieden werden.

Während die Übertragung von Radiowellen stark reglementiert ist, gibt es keine Beschränkungen für den Einsatz von Lichtwellen. Meist wird infrarotes Licht verwendet. Es gibt viele Störquellen wie z. B. Tageslicht oder Wärmequellen, deshalb ist die Reichweite bei ungerichteter Übertragung sehr beschränkt. Bei gerichteter Übertragung können zwar größere Entfernungen überbrückt werden, aber die Notwendigkeit einer genauen Ausrichtung auf einen bestimmten Empfänger macht das Verfahren dann für den Aufbau von Netzen schon deshalb kaum nutzbar, da so nicht mit mehreren anderen Teilnehmern gleichzeitig kommuniziert werden kann. In jedem Fall werden Lichtwellen durch Hindernisse absorbiert oder reflektiert, so dass die Kommunikation nur bei Sichtverbindung (in gewissen Grenzen auch indirekt über Reflektion) möglich ist. Besonders sinnvoll genutzt werden kann Infrarot-Übertragung dann, wenn die einfache Abschirmbarkeit erwünscht ist, z. B. zur gezielten Übertragung zwischen zwei Geräten, wenn ein für Dritte schwer manipulierbarer Kanal benötigt wird.

Radiowellen durchdringen bzw. umlaufen Hindernisse wesentlich besser (was allerdings auch für die zum Teil breitbandigen Abstrahlungen von Störquellen gilt) und erlauben auch eine ungerichtete Übertragung über Entfernungen, die für die Vernetzung mobiler Teilnehmer interessant sind. Für die Übertragung zum Aufbau lokaler Netze werden deshalb auch hauptsächlich Radiowellen verwendet.

Zum Aufbau von Ad-hoc-Netzen bietet sich die Verwendung standardisierter Techniken für drahtlose lokale Netze an (Wireless Local Area Networks, WLANs), die jederzeit ohne weitere Vorbereitungen zur Kommunikation mit in der näheren Umgebung (z. B. bis zu 200 m Entfernung) befindlichen Geräten genutzt werden können. Infrastruktur ist nicht zwingend erforderlich, und bei Verwendung lizenzfreier Frequenzbereiche fallen keine Kosten an. Es gibt mittlerweile eine ganze Reihe von Standards für drahtlose lokale Netze. IEEE 802.11 und Bluetooth sind die beiden, die sich bisher der besten Akzeptanz erfreuen:

- Der IEEE-Standard 802.11 [IEE99b] für drahtlose lokale Netze knüpft an die etablierte Reihe 802.x von Standards für andere Verfahren zum Aufbau lokaler Netze an, deren am häufigsten eingesetzter Vertreter wahrscheinlich 802.3 (besser bekannt unter dem Namen „Ethernet“) ist. Alle diese Standards bieten eine einheitliche Schnittstelle nach oben hin an und können problemlos zusammen mit dem Internet-Protokoll verwendet werden.

  IEEE 802.11 erlaubt den Aufbau infrastrukturbasierter Netze, bei denen die Kommunikation stets über fest installierte Basisstationen erfolgt, die meist gleichzeitig einen Übergang zu leitungsgebundenen Netzen und letztendlich dem Internet darstellen. Es kann aber auch in einem Ad-hoc-Modus betrieben werden, in dem auf Infrastruktur verzichtet wird und die Teilnehmer direkt miteinander kommunizieren.

  Die ursprüngliche Version des Standards sah eine Datenrate von 1 Mbit/s sowie optional auch 2 Mbit/s vor; inzwischen gibt es aber neuere Varianten mit wesentlich höheren Übertragungsraten. Die Reichweite kann bei der geringsten Datenrate im Freien bis zu etwa 500 m betragen, innerhalb von Gebäuden bis zu 250 m. Höhere Datenraten sind allerdings nur über kürzere Entfernungen erreichbar (z. B. bis zu 250 m im Freien und 50 m in Gebäuden bei 11 Mbit/s).

- Das Bluetooth-Verfahren [Blu01a, Blu01b] wurde insbesondere für Zwecke der einfachen Nahbereichsvernetzung als Ersatz für Kabelverbindungen z. B. zwischen kleinen mobilen (Mobiltelefone, PDAs) oder auch fest installierten (Fernseher, Videorecorder usw.) Geräten entwickelt. Die Adapter sollten möglichst preisgünstig sein, um sie in möglichst viele Geräte integrieren zu können.

  Bluetooth realisiert ausschließlich Ad-hoc-Netze. Diese haben zunächst sehr spezielle Charakteristika: Ein so genanntes Pikonetz kann nur bis zu acht aktive Knoten enthalten, von denen einer als Leitstation ausgezeichnet ist. Daten können direkt nur zwischen dieser Leitstation und jeweils einer der übrigen Stationen übertragen werden. Alle weiteren Funktionen müssen von höheren Schichten erbracht werden. Die Bluetooth-Spezifikation verwendet eine eigene Schichtenarchitektur, die viele spezielle Funktionsblöcke für bestimmte Anwendungen wie z. B. Sprachübertragung, Telefonie, Modememulation oder Datenabgleich bei persönlichen digitalen Assistenten (PDAs) enthält. Die Verwendung des Internet-Protokolls über Bluetooth ist möglich, erfordert aber spezielle Adaptionsmechanismen. Mittlerweile wurde auch bei der IEEE eine Arbeitsgruppe mit der Bezeichnung 802.15 gegründet, welche sich an der Standardisierung von Bluetooth beteiligt und Wechselwirkungen mit 802.11 betrachtet.

  Die Sendeleistung ist entsprechend der vorgesehenen Einsatzszenarien geringer als bei 802.11. Dabei sind drei verschiedene Leistungsklassen vorgesehen, deren typische Reichweiten mit 10, 20 und 100 m angegeben sind. Die Datenrate beträgt knapp 1 Mbit/s.

Die Verfahren IEEE 802.11 und Bluetooth haben mittlerweile eine recht große Verbreitung gefunden. Adapter für IEEE 802.11 sind in vielen mobilen Rechnern bereits eingebaut oder problemlos nachrüstbar, und viele Mobiltelefone, PDAs und andere Kleingeräte sind mit Bluetooth-Schnittstellen ausgestattet. Damit ist bei Verwendung dieser Verfahren zum Aufbau von Ad-hoc-Netzen eine recht hohe Wahrscheinlichkeit gegeben, andere potentielle Netzteilnehmer mit kompatibler Übertragungstechnik vorzufinden und damit eine gewisse Teilnehmerdichte erreichen zu können.

Auf IEEE 802.11 wird in Abschnitt 2.2.4 noch näher eingegangen. Nähere Informationen zu anderen Verfahren sind z. B. in [Schi03] zu finden.

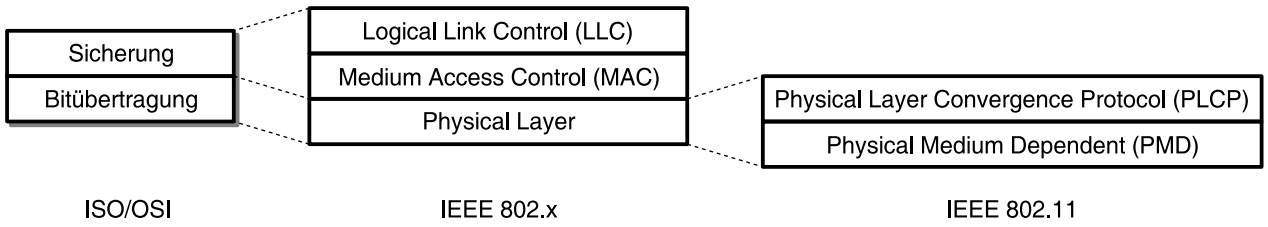


Abbildung 2.6: Schichtenarchitektur bei IEEE 802.11

2.2.4 Drahtlose lokale Netze nach IEEE 802.11

Durch die folgende exemplarische Vorstellung eines sehr gebräuchlichen Verfahrens für drahtlose lokale Netze soll ein Eindruck von den Eigenschaften und von Möglichkeiten, diese Eigenschaften bei der Realisierung eines Zugangskontrollsystems auf den darüberliegenden Schichten zu nutzen, vermittelt werden.

2.2.4.1 Architektur

Die Standards für unterschiedliche Ausprägungen lokaler Netze der Serie 802.x der IEEE spezifizieren jeweils die beiden untersten Schichten des ISO/OSI-Referenzmodells (siehe Abschnitt 2.1.3), also die Bitübertragungs- und die Sicherungsschicht. Dabei wird die Sicherungsschicht nochmals in zwei Unterschichten mit den Bezeichnungen *Logical Link Control* (LLC) und *Medium Access Control* (MAC) aufgeteilt (siehe Abbildung 2.6 Mitte).

Die an das jeweilige physikalische Übertragungsverfahren angepasste MAC-Schicht übernimmt insbesondere die Kontrolle des Zugriffs auf geteilte Übertragungsmedien sowie die Flusskontrolle (plus Fragmentierung und bei 802.11 auch Verschlüsselung – eine Funktion, die bei den älteren, kabelgebundenen lokalen Netzen nicht vorgesehen ist). Die allen Ausprägungen gemeinsame LLC-Schicht (spezifiziert als 802.2) verbirgt die Unterschiede der verschiedenen MAC-Schichten und bietet damit eine einheitliche Schnittstelle nach oben zur Vermittlungsschicht. Sie bietet außerdem die Möglichkeit, die durch die MAC-Schichten realisierten unzuverlässigen Datagramm-Dienste durch die Verwendung von Sequenznummern und Empfangsbestätigungen zuverlässig und verbindungsorientiert zu machen..

Da der Standard 802.11 für drahtlose lokale Netze drei unterschiedliche Übertragungsverfahren erlaubt, ist die Bitübertragungsschicht ebenfalls nochmals in zwei Unterschichten unterteilt (siehe Abbildung 2.6 rechts): Die untere wird mit *Physical Medium Dependent* (PMD) bezeichnet und beinhaltet die von den jeweiligen Übertragungsverfahren abhängigen Teile wie Modulation und Kodierung. Die obere, bezeichnet als *Physical Layer Convergence Protocol* (PLCP), stellt wiederum eine gemeinsame Schnittstelle nach oben zur MAC-Schicht dar und beinhaltet außerdem gemeinsame Funktionen wie die Erzeugung des Signals zur Überwachung des Mediums.

Wie bereits erwähnt können lokale Netze nach 802.11 sowohl im Infrastruktur- als auch im Ad-hoc-Modus betrieben werden. Im Infrastruktur-Modus kommunizieren die mobilen Endgeräte auf physikalischer Ebene jeweils nur mit einem bestimmten der normalerweise fest installierten *Zugangspunkte (Access Points)*. Der Zugangspunkt regelt auch den Medienzugriff innerhalb des so genannten *Basic Service Set* (BSS) – dieser Begriff fasst einen Zugangspunkt und alle mit ihm verbundenen Endgeräte zusammen. Mehrere BSS werden durch ein (im Standard nur bezüglich seiner Dienste, nicht aber bezüglich seines genauen Aufbaus spezifiziertes) *Distribution System* miteinander und evtl. auch mit anderen Netzen verbunden, so dass insgesamt ein größeres logisches Netz entsteht, welches auch als

Extended Service Set bezeichnet wird. Da Infrastrukturnetze für die vorliegende Arbeit nicht relevant sind, wird im Folgenden nicht näher auf den Betrieb von 802.11 im Infrastruktur-Modus eingegangen (siehe dazu z. B. [Schi03]).

Im Ad-hoc-Modus wird die Gesamtheit der Endgeräte, die mit demselben Übertragungsverfahren im selben Frequenzbereich arbeiten und sich innerhalb ihrer gegenseitigen Reichweite befinden, als *Independent Basic Service Set* (IBSS) bezeichnet. Dieser Begriff ist damit allerdings nicht besonders klar spezifiziert: Zwei Endgeräte, die sich außerhalb ihrer gegenseitigen Sendereichweite befinden, können eigentlich nicht im selben IBSS sein; wenn der Abstand allerdings durch weitere Endgeräte überbrückt wird, ist unklar, wo nun die Grenzen des IBSS sein sollen. Ein Verfahren zur Weiterleitung von Paketen über mehrere Knoten hinweg wird vom Standard aber nicht vorgegeben; diese Funktionalität muss also gegebenenfalls von der Vermittlungsschicht erbracht werden, d. h. zumindest ein Teil der Endgeräte muss als Vermittlungssystem arbeiten.

2.2.4.2 Bitübertragung

IEEE 802.11 unterstützt sowohl Infrarot- als auch Funkübertragung, wobei für letztere außerdem zwei verschiedene Bandspreizungsverfahren angeboten werden: *Frequency Hopping Spread Spectrum* (FHSS) und *Direct Sequence Spread Spectrum* (DSSS). Die Bandspreizverfahren dienen dazu, die Signale eines Übertragungskanals über einen größeren Frequenzbereich zu verteilen, so dass sie weniger empfindlich gegen schmalbandige Störungen werden. Der zur Verfügung stehende Frequenzbereich wird dabei von mehreren Übertragungskanälen gleichzeitig genutzt, die sich aber aufgrund unterschiedlicher Parametrisierung des Bandspreizverfahrens gegenseitig nur unwesentlich stören. Auf die genaue Funktionsweise von FHSS und DSSS und ihren Einsatz bei IEEE 802.11 soll hier nicht näher eingegangen werden (siehe dazu z. B. [Schi03]). Das in der Praxis am häufigsten eingesetzte Übertragungsverfahren ist die Funkübertragung mit DSSS.

Die Funkübertragungsverfahren arbeiten im lizenzfreien 2,4-GHz-Band. Die dort zur Verfügung stehende Bandbreite und die maximale erlaubte Sendeleistung sind allerdings von nationalen Regulierungen abhängig und beispielsweise für Nordamerika, Europa und Japan jeweils etwas unterschiedlich. Mikrowellengeräte, schnurlose Telefone (DECT) und Bluetooth arbeiten im selben Frequenzband, sind also potentielle Störungsquellen.

Für alle Übertragungsverfahren sind im ursprünglichen Standard die zwei unterschiedlichen Übertragungsraten 1 Mbit/s und 2 Mbit/s vorgesehen, wobei die erste immer unterstützt werden muss und die zweite optional ist. Für die Koordination der Endgeräte wichtige Steuerungsinformation wird immer mit der niedrigsten Datenrate von 1 Mbit/s übertragen. Damit soll erreicht werden, dass auch einfachere und billigere Geräte stets mit allen anderen Geräten kommunizieren können.

Neben dem transparenten Übertragungskanal stellt die Bitübertragungsschicht der MAC-Schicht außerdem ein Signal zur Verfügung, das angibt, ob das Medium zur Zeit frei ist (*Clear Channel Assessment*, CCA).

Neuere Varianten von IEEE 802.11 sind die Standards 802.11b, 802.11a und 802.11g (die in dieser Reihenfolge fertiggestellt wurden, auch wenn die zugehörigen Arbeitsgruppen in alphabetischer Reihenfolge gegründet wurden); noch in Entwicklung befindet sich IEEE 802.11n. Sie spezifizieren jeweils eine neue Bitübertragungsschicht mit höheren Übertragungsraten. Die MAC-Schicht bleibt jeweils gleich wie beim ursprünglichen Standard.

- IEEE 802.11b [IEE99c, IEE01] arbeitet ebenfalls im 2,4-GHz-Bereich und bietet zusätzliche optionale Datenraten von 11 und 5,5 Mbit/s an, die einerseits durch verbesserte Kodierungsverfahren und andererseits durch Einsparungen bei den Paketformaten ermöglicht werden. Wie

schon beim ursprünglichen Standard wird Steuerungsinformation immer mit der niedrigsten Datenrate übertragen; diese Forderung beschränkt auch die tatsächlich erreichbare maximale Nettodatenrate auf ca. 6 Mbit/s.

- IEEE 802.11a [IEE99a] ist nicht wie IEEE 802.11b eine interoperable Erweiterung des IEEE-802.11-Standards, sondern verwendet in der Bitübertragungsschicht sowohl andere Frequenzbereiche als auch andere Modulations- und Kodierungsverfahren, so dass Geräte mit IEEE-802.11a-Schnittstelle nicht mit solchen mit IEEE-802.11(b)-Schnittstelle kommunizieren können.

  Die verwendeten Frequenzbereiche liegen im 5-GHz-Band und sind ebenfalls lizenzfrei nutzbar. Die genauen Frequenzbereiche sind wieder regional etwas unterschiedlich, und nur ein Teil davon ist wirklich weltweit nutzbar. Das 5-GHz-Band wird bisher wenig von Massenmarkt-Geräten benutzt, so dass aus dieser Richtung weniger Störungen zu erwarten sind. Allerdings arbeiten Radargeräte und manche Satellitenübertragungen in diesem Bereich, deshalb ist insbesondere in Europa bei Überschreitung einer bestimmten Sendeleistung (50 mW) die Implementierung zusätzlicher Funktionen zur Anpassung der Übertragungsleistung sowie zur dynamischen Auswahl freier Frequenzbänder [IEE03a] vorgeschrieben.

  IEEE 802.11a setzt das so genannte OFDM-Verfahren (*Orthogonal Frequency Division Multiplex*) ein. Dabei wird statt einer einzigen Trägerfrequenz eine ganze Reihe benachbarter Träger gleichzeitig verwendet (48 bei IEEE 802.11a), auf welche die Bits des zu übertragenden Datenstroms sozusagen aufgeteilt werden. Damit kann die Schrittfrequenz der Übertragung vermindert werden, was die Empfindlichkeit gegenüber Störungen, die durch die Ausbreitung der Signale auf unterschiedlichen Wegen und die sich daraus ergebenden Laufzeitunterschiede entstehen, verringert.

  Unter idealen Bedingungen können auf der Bitübertragungsschicht Datenraten bis zu 54 Mbit/s erreicht werden. Allerdings durchdringen die Signale wegen der kürzeren Wellenlänge feste Hindernisse wesentlich schlechter. Die maximale Datenrate ist deshalb praktisch nur bei direkter Sichtverbindung erreichbar.

- IEEE 802.11g [IEE03b] ist wieder eine zu IEEE 802.11b interoperable Erweiterung, die dieselben Frequenzbereiche im 2,4-GHz-Band und für die dort bereits vorhandenen Datenraten auch dieselben Modulations- und Kodierungsverfahren verwendet. Unter Verwendung von OFDM können aber außerdem höhere Datenraten bis 54 Mbit/s erreicht werden. Als Vorteil gegenüber IEEE 802.11a können die besseren Ausbreitungseigenschaften des 2,4-GHz-Bandes gesehen werden, ein Nachteil ist dagegen, dass weiterhin nur drei überschneidungsfreie Kanäle zur Verfügung stehen.

- IEEE 802.11n soll, wenn der Standard fertiggestellt ist, nochmals höhere Datenraten erreichen, nämlich bis zu etwa 540 Mbit/s. Die Kompatibilität zu IEEE 802.11b und IEEE 802.11g soll dabei gewahrt bleiben; es werden auch dieselben Frequenzbereiche genutzt. Ein wesentlicher Faktor für die Steigerung der Datenrate ist die Verwendung mehrerer Sende- und Empfangsantennen, wodurch insbesondere die Mehrwegeausbreitung, die sonst unter Umständen zur Auslöschung gegeneinander verschobener Signale führt, zur Verstärkung des Empfangssignals und zur gleichzeitigen Übertragung mehrerer Kanäle verwendet werden kann. Erste, auf Entwurfsversionen des Standards basierende Geräte sind bereits von vielen Herstellern verfügbar, insofern ist eine weite Verbreitung des kommenden Verfahrens zu erwarten.

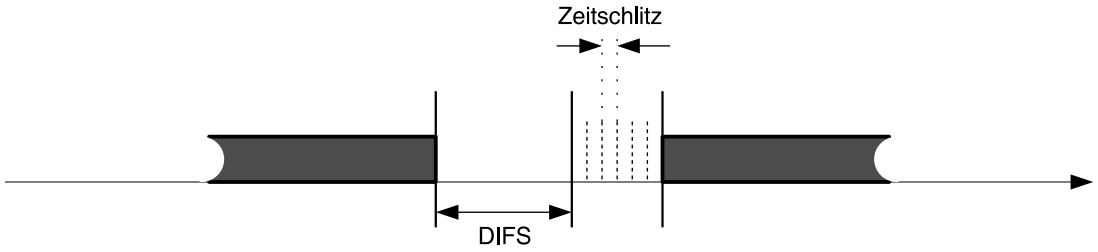


Abbildung 2.7: Das CSMA/CA-Verfahren

2.2.4.3 Medienzugriff

Zu den beiden Betriebsmöglichkeiten Infrastruktur- bzw. Ad-hoc-Modus gibt es bei IEEE 802.11 auch zwei unterschiedliche Verfahren zur Regelung des Medienzugriffs. Im Ad-hoc-Modus kann nur der verteilte Ansatz *Distributed Coordination Function* (DCF) verwendet werden, im Infrastrukturmodus liegt die Verwendung des zentral gesteuerten Ansatzes *Point Coordination Function* (PCF) nahe. Letzterer kann Kollisionsfreiheit garantieren und Dienste mit nach oben beschränkter Verzögerung anbieten, indem die Basisstation alle anderen Stationen reihum abfragt. Dieses Verfahren soll hier aber nicht näher betrachtet werden, da es für Ad-hoc-Netze ungeeignet ist. DCF und PCF können gleichzeitig verwendet werden, indem die Basisstation das Medium regelmäßig für eine DCF-Phase freigibt.

Das verteilte Medienzugriffsverfahren von IEEE 802.11 (DCF) basiert auf dem CSMA/CA-Verfahren (*Carrier Sense Multiple Access with Collision Avoidance*), welches wie folgt funktioniert: Ein sendewilliges Endgerät überprüft (anhand des CCA-Signals der Bitübertragungsschicht), ob das Medium frei ist. Wenn das Medium für eine bestimmte Zeitspanne – bezeichnet als *DCF Inter-Frame Spacing* (DIFS) – frei ist, darf das Endgerät sofort beginnen zu senden. Ist das Medium dagegen nicht frei, so wird zunächst gewartet, bis dies zumindest für die Zeitspanne DIFS der Fall ist. Anschließend hört das Gerät noch weiter für eine zufällig gewählte Anzahl kurzer vorgegebener Zeitschlitze das Medium ab. Bleibt es in dieser Zeit frei, so darf das Gerät senden (siehe Abbildung 2.7). Wird das Medium dagegen während dieser letzten Wartephase durch ein anderes Gerät belegt, das eine kürzere Verzögerung gewählt hatte, so beginnt der Wartevorgang von vorn.

Um bei diesem Verfahren zusätzlich zu berücksichtigen, dass ein sendewilliges Gerät möglicherweise schon seit Längerem wartet, weil das Medium immer wieder durch andere Geräte belegt wurde, wird bei IEEE 802.11 die Anzahl abzuwartender Zeitschlitze nicht bei jedem Versuch neu zufällig bestimmt, sondern es wird der beim ersten Versuch gewählte Wert abzüglich inzwischen bereits abgewarteter Zeitschlitze verwendet. Damit werden Geräte, die bereits einige Zeit gewartet haben, statistisch gesehen bevorzugt.

Bei Punkt-zu-Punkt-Übertragungen sendet der Empfänger direkt nach Empfang eines anhand seiner Prüfsumme als korrekt übertragen erkannten Pakets eine Bestätigung an den Sender. Er wartet dazu nicht die DIFS-Zeitspanne ab, sondern nur eine kürzere, mit *Short Inter-Frame Spacing* (SIFS) bezeichnete Zeit. Dadurch ist garantiert, dass das Medium immer frei ist und bei Übertragungsbestätigungen keine Kollisionen auftreten können. Das Ausbleiben einer Übertragungsbestätigung zeigt an, dass das Paket durch eine Kollision verfälscht wurde und wiederholt werden muss. Eine wiederholte Übertragung wird genau wie die erste Übertragung behandelt, also nicht bevorzugt. Bei Broadcast-Übertragung erfolgt keine Übertragungsbestätigung.

Der Wertebereich, aus dem die Anzahl zu wartender Zeitschlitze beim ersten Sendeversuch eines Pakets zufällig gewählt wird, ist zunächst klein, wird aber bei jeder erkannten Kollision verdoppelt. Da-

mit wird die Kollisionswahrscheinlichkeit verringert, und das Kollisionsvermeidungsverfahren passt sich so an ein gesteigertes Verkehrsaufkommen an.

Beim bisher beschriebenen Verfahren kann ein Problem durch so genannte „versteckte Endgeräte“ entstehen: Zwei Endgeräte, die sich gegenseitig nicht hören können, können nicht feststellen, ob ihre Übertragungen Kollisionen bei Dritten verursachen, die sich in der Mitte zwischen den beiden ersten befinden und beide hören können. IEEE 802.11 bietet zur Lösung dieses Problem den folgenden optional anzuwendenden Mechanismus an: Statt eines Datenpakets wird zunächst ein RTS-Steuerpaket (*Request-to-Send*) gesendet. Dieses gibt die gesamte, vom Sender im Voraus berechnete Übertragungsdauer für das zu sendende Paket und die zugehörige Empfangsbestätigung an. Jede Station, die das Paket hört, ermittelt anhand dieser Übertragungsdauer den nächstmöglichen Zeitpunkt, zu dem das Medium wieder frei sein kann. Der Empfänger beantwortet das RTS-Steuerpaket sofort mit einem CTS-Steuerpaket (*Clear-to-Send*), welches ebenfalls die Übertragungsdauer enthält und von allen Mithörenden ebenso wie das RTS-Paket behandelt wird. Alle Stationen innerhalb der Sendereichweiten des Senders und des Empfängers sind damit nun über die Dauer der Belegung des Mediums informiert, und das Medium ist praktisch reserviert. Die Übertragung beginnt nach Empfang des CTS-Pakets durch den Sender und wird vom Empfänger normal bestätigt. Kollisionen können nur bei der Übertragung des RTS-Pakets auftreten. Das RTS/CTS-Verfahren verursacht zusätzlichen Verkehr und kann insbesondere für größere Rahmen eingesetzt werden, bei denen Kollisionen und die nötigen Übertragungswiederholungen mehr Aufwand verursachen würden.

Das RTS/CTS-Verfahren kann zusätzlich mit einem Fragmentierungsmechanismus kombiniert werden, der dazu dient, lange Rahmen in mehrere kürzere Fragmente aufzuteilen, die einzeln bestätigt werden, so dass bei Verlust durch Übertragungsfehler nur wenig wiederholt werden muss. Die RTS/CTS-Pakete werden in diesem Fall nur für das erste Fragment versendet und enthalten auch nur die Dauer für das erste Fragment und dessen Bestätigung. Die Belegungsdauer für weitere Fragmente ist im jeweils vorherigen Fragment (sowie in der zugehörigen Bestätigung) angegeben. Der Sender darf nach Empfang einer Übertragungsbestätigung und einer zusätzlichen SIFS-Zeitspanne jeweils sofort das nächste Fragment senden, solange, bis der gesamte Rahmen übertragen ist.

2.2.4.4 Verwaltungsfunktionen

Neben den direkt mit der Datenübertragung und Verwaltung des Mediums befassten Funktionen sind von jeder Station noch einige zusätzliche Aufgaben zu erledigen:

- *Energiesparmechanismen:* Um in mobilen Endgeräten Energie zu sparen, ist es wünschenswert, die Sende- und Empfangselektronik bei Nichtgebrauch abzuschalten. Für das Senden ist dies offensichtlich unproblematisch, aber auch die Empfangsbereitschaft kostet Energie [WESW98], und um jederzeit erreichbar zu bleiben, muss eine Station im Voraus wissen, wann ein Paket für sie gesendet wird. Bei Infrastrukturnetzen wird dies ermöglicht, indem die Basisstation Pakete zwischenspeichert und in regelmäßigen Zeitintervallen bekannt gibt, für welche Stationen Pakete vorliegen. Es muss also jeweils diese Nachricht empfangen werden, und wenn keine Pakete für eine Station vorliegen, kann diese ihren Empfänger für ein Zeitinterval abschalten. Im Ad-hoc-Modus muss jede Station selbst puffern, wenn sie an eine nicht empfangsbereite Station senden will, und es wird versucht, eine Liste gepufferter Rahmen in einer gemeinsamen Wachperiode zu verteilen, die ebenfalls in regelmäßigen Zeitabständen stattfindet. Jede Station versucht dort mitzuteilen, für welche anderen Stationen sie Daten gepuffert hat (wenn dies der Fall ist). Je mehr Stationen in einem Ad-hoc-Netz im Energiesparmodus sind, desto wahrscheinlicher werden Kollisionen im Wachfenster, und die Zugriffsverzögerung wird größer und schwer vorhersagbar.

- *Synchronisation:* Alle Stationen enthalten Uhren, die wegen der Energiesparmechanismen, bei Verwendung von PCF sowie zur Sprungfolgensynchronisation bei FHSS synchronisiert werden müssen. Dazu wird in festgelegten Intervallen eine als *Beacon* bezeichnete Nachricht mit Zeitstempel ausgesendet. In Infrastrukturnetzen tut dies die Basisstation, in Ad-hoc-Netzen versucht jede Station, Beacons zu senden, bricht ihren Versuch aber für das laufende Zeitinterval ab, sobald sie von einer anderen Station ein Beacon empfängt.

- *Roaming:* In Infrastrukturnetzen sind noch Mechanismen für den Wechsel von Stationen zwischen verschiedenen Basisstationen vorgesehen. Da diese aber für Ad-hoc-Netze nicht relevant sind, werden sie hier nicht näher betrachtet.

2.3 Netzwerksicherheit

2.3.1 Begriffe

Ziel einer Sicherung von Netzwerkdiensten ist es, die für ihre regulären Nutzer wertvollen Eigenschaften der Dienste und der durch sie verarbeiteten Daten so zu schützen, dass sie nicht durch Angriffe Dritter in Frage gestellt werden können. Solche schützenswerten Eigenschaften lassen sich grob in die folgenden Kategorien unterteilen:

- Die *Authentizität* der an einem Kommunikationsvorgang beteiligten Parteien, also die nachweisliche Korrektheit der von ihnen angegebenen Identität, ist häufig eine grundlegende Voraussetzung für die sinnvolle Nutzung von kommunikationsbasierten Diensten. Insbesondere in offenen Netzen, wo die Menge der möglichen Teilnehmer nicht feststeht, stellt der Nachweis der Authentizität einer Partei gegenüber einer anderen, die so genannte *Authentisierung*, häufig eine Herausforderung dar. Angriffe gegen die Authentizität von Teilnehmern durch Vorspiegeln einer falschen Identität werden auch als *Maskerade* bezeichnet. Wenn ein Angreifer sich (in Bezug auch eine Kommunikationsbeziehung) zwischen zwei Teilnehmern positioniert und sich gegenüber jedem der beiden als der jeweils andere maskiert, spricht man von einem *Man-in-the-middle*-Angriff.

  Neben der Authentizität von Teilnehmern ist oft auch die Authentizität von Nachrichten von Bedeutung, also die nachweisliche Korrektheit der Zuordnung zwischen Nachricht und Absender.

- Die Authentizität von Nachrichten beinhaltet auch ihre *Integrität*, also die Eigenschaft, während der Übertragung oder Verarbeitung nicht unvorhergesehen von Dritten verändert worden zu sein. Insofern wird Nachrichtenauthentizität häufig mit Integrität gleichgesetzt. Grundsätzlich kann die Integrität von Nachrichten aber auch unabhängig davon betrachtet werden, ob deren Absender überhaupt bekannt ist. Die Sicherstellung der Integrität von Nachrichten wird damit als eine Leistung des verwendeten Kommunikationsdienstes aufgefasst und weniger als eine den einzelnen Nachrichten innewohnende Eigenschaft.

  Eine weitere Ausprägung der Eigenschaft Integrität besitzen Dienste und Komponenten, die sich entsprechend ihrer Bestimmung verhalten und nicht von Unbefugten manipuliert wurden.

- *Vertraulichkeit* kann für kommunikationsbasierte Dienste in vielerlei Hinsicht von Bedeutung sein. Eine häufig benötigte Dienstleistung ist der Schutz übertragener vertraulicher Information vor unbefugter Einsichtnahme. Der Versuch, vertrauliche Kommunikationsinhalte abzuhören,

ist bereits durch so genannte *passive Angriffe* möglich, bei denen der Angreifer im Gegensatz zu den im Allgemeinen schwieriger durchzuführenden *aktiven Angriffen* (etwa auf die Integrität von Inhalten) nicht in der Lage sein muss, übertragene Daten zu verändern oder zusätzlich zu generieren.

Für manche Anwendungen kann eine vertrauliche Behandlung anderer Informationen als des Inhalts übertragener Nachrichten von Bedeutung sein: die Identitäten beteiligter Parteien (gegenseitig oder gegenüber Dritten), das bloße Bestehen von Kommunikationsbeziehungen oder die Lokation mobiler Teilnehmer.

- *Verbindlichkeit* (auch *Nichtabstreitbarkeit*) wird benötigt, wenn nachweisbar sein soll, dass eine bestimmte Aktion oder Nachricht von einer bestimmten Partei durchgeführt bzw. abgesandt oder empfangen wurde. Dies kann entweder direkt für die realisierte Anwendung relevant sein, oder zum Zweck der Abrechnung oder der Verfolgung von Missbrauch auch innerhalb eines Kommunikationsdienstes.

- Die *Verfügbarkeit* von Komponenten und Diensten ist Voraussetzung für ihre Nutzbarkeit. Angriffe gegen die Verfügbarkeit, die typischerweise entweder zur Unterstützung anderer Angriffe oder zum Zweck des rein destruktiven Störens von abhängigen Prozessen oder der Schädigung des Ansehens der für die Aufrechterhaltung der Verfügbarkeit zuständigen Parteien durchgeführt werden, werden als *Denial-of-Service*-Angriffe bezeichnet.

Zum Schutz der genannten Eigenschaften dienen *Sicherheitsmechanismen*; wenn sie in Form von Diensten angeboten werden, werden diese als *Sicherheitsdienste* bezeichnet. Die zum Schutz von Authentizität, Integrität und Vertraulichkeit bestimmten Sicherheitsmechanismen basieren meist auf den im nachfolgenden Abschnitt 2.3.2 kurz vorgestellten kryptographischen Algorithmen. Die Authentisierung wird in Abschnitt 2.3.3 noch genauer behandelt. Zur Gewährleistung von Verbindlichkeit ist zudem eine unabhängige dritte Partei erforderlich, die das Vertrauen der anderen genießt. Verfügbarkeit kann im Allgemeinen nicht mit einzelnen Maßnahmen umfassend geschützt werden, da sie auf sehr unterschiedliche Weise angreifbar ist. Ein Sicherheitsdienst, der zum Schutz von Verfügbarkeit beitragen kann, ist die *Zugangskontrolle*, auf welche in Abschnitt 2.3.6 näher eingegangen wird.

2.3.2 Kryptographische Algorithmen

Sicherheitsmechanismen zum Schutz von Authentizität, Integrität und Vertraulichkeit beruhen in der Regel auf der Anwendung kryptographischer Algorithmen.

2.3.2.1 Verschlüsselung

Die Verschlüsselung dient primär der Sicherung der Vertraulichkeit übertragener Daten. Man unterscheidet bezüglich der zugrunde liegenden Algorithmen zwei Ausprägungen: *symmetrische* und *asymmetrische* Verschlüsselung.

Bei symmetrischen Verfahren wird für Chiffrierung und Dechiffrierung ein und derselbe Schlüssel benutzt, der Sender und Empfänger bekannt sein muss, ansonsten aber geheimzuhalten ist. Abbildung 2.8 veranschaulicht das Verfahren: Die Verschlüsselung zum Schutz der Vertraulichkeit übertragener Daten erfolgt bei symmetrischen Verfahren durch den Sender mit dem beiden Parteien bekannten geheimen Schlüssel; die Entschlüsselung durch den Empfänger unter Verwendung desselben Schlüssels liefert wieder den Klartext.

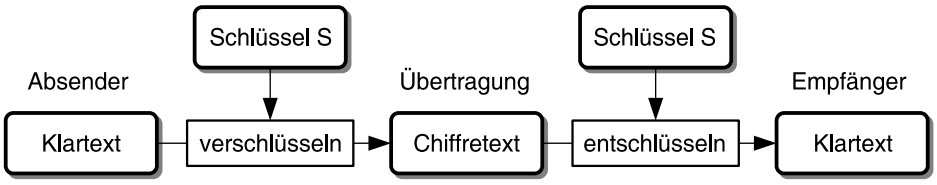


Abbildung 2.8: Vertraulichkeitssicherung mit symmetrischen kryptographischen Verfahren

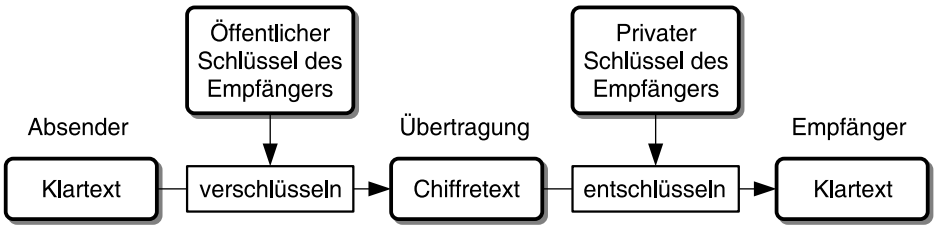


Abbildung 2.9: Vertraulichkeitssicherung mit asymmetrischen kryptographischen Verfahren

Bekannte und bewährte Verfahren zur symmetrischen Verschlüsselung sind beispielsweise der *Data Encryption Standard* (DES) [NIS99] und der *Advanced Encryption Standard* (AES) [NIS01]. Ersterer ist allerdings mittlerweile aufgrund seiner Schlüssellänge von 56 bit (die außerdem aufgrund von Exportbestimmungen der USA häufig künstlich auf 40 bit verringert wurde) angreifbar geworden, da mit modernen Rechensystemen eine erschöpfende Suche nach einem Schlüssel zu einem gegebenen Chiffretext in für manche Angriffe akzeptabler Zeit durchführbar ist [KPPP$^+$06]. Durch dreimaliges Anwenden in Folge mit zwei Schlüsseln (Chiffrierung mit dem ersten – Dechiffrierung mit dem zweiten – Chiffrierung mit dem ersten; der Gesamtalgorithmus wird dann auch als Triple-DES bezeichnet) verlängert sich die Schlüssellänge auf 112 bit; dies ist bezüglich der Sicherheit akzeptabel, dafür aber relativ rechenaufwändig. Als Nachfolger AES von DES wurde deshalb im Jahr 2001 ein Algorithmus namens Rijndael (nach seinen Erfindern Rijmen und Daemen) mit einer Schlüssellänge von mindestens 128 bit standardisiert. Weitere symmetrische Chiffrierverfahren sind z. B. IDEA, RC4 und RC5.

Bei asymmetrischen Verfahren existieren zwei Schlüssel, wobei eine Chiffrierung mit dem einen der beiden Schlüssel unter Verwendung des anderen wieder rückgängig gemacht werden kann. Zur Anwendung des Verfahrens besitzt jede Partei ein solches Schlüsselpaar. Einer der Schlüssel – der *private Schlüssel* – wird dabei geheimgehalten, der andere – der *öffentliche Schlüssel* – kann öffentlich bekannt gemacht werden. Für die beidseitige Sicherung von Kommunikationsvorgängen benötigt jede Partei außer ihrem eigenen Schlüsselpaar auch den öffentlichen Schlüssel des Kommunikationspartners. Asymmetrische Verfahren werden auch als *Public-Key*-Verfahren bezeichnet. Abbildung 2.9 veranschaulicht das Vorgehen bei asymmetrischer Vertraulichkeitssicherung: Der Sender verwendet zum Verschlüsseln den öffentlichen Schlüssel des Empfängers. Nur dieser kann aus dem Ergebnis unter Verwendung seines privaten Schlüssels den Klartext wieder ermitteln. Außerdem können asymmetrische Verfahren (durch umgekehrtes Vorgehen) zur Erstellung digitaler Signaturen verwendet werden (siehe dazu Abschnitt 2.3.2.5).

Einer der meistverwendeten asymmetrischen Algorithmen ist der nach seinen Erfindern Rivest, Shamir und Adleman benannte RSA-Algorithmus [RiSA78], der darauf aufbaut, dass es rechnerisch äußerst aufwändig ist, sehr große Primzahlen zu faktorisieren. Eine derzeit als „sicher“ eingeschätzte Schlüssellänge (vergleichbar mit AES mit 128 bit Schlüssellänge) beträgt bei RSA 2048 bit [Schn96, Kap. 7].

Die Verschlüsselung von Nachrichten kann neben dem Vertraulichkeitsschutz gleichzeitig auch eine Integritätssicherung darstellen, da nur die Inhaber der jeweiligen Schlüssel den Klartext in die chif-

frierte Form bringen können. Allerdings müssen die zu verschlüsselnden Daten genügend Redundanz aufweisen, anhand derer nach der Entschlüsselung erkannt werden kann, dass es sich tatsächlich um eine Nachricht handelt, und nicht um das sinnlose Ergebnis der „Entschlüsselung“ eines von einem Angreifer blind erzeugten Chiffretexts.

2.3.2.2 Schlüsselvereinbarung nach Diffie und Hellman

Mit dem Diffie-Hellman-Verfahren [DiHe76] kann ein gemeinsames Geheimnis zwischen zwei entfernten Parteien vereinbart werden, ohne dass Mithörende dieses erfahren. Das Verfahren beruht darauf, dass die Berechnung diskreter Logarithmen in einem endlichen Körper, die von einem Angreifer zur Ermittlung des Geheimnisses durchgeführt werden müsste, wesentlich schwieriger ist als die Potenzierung, die von den kommunizierenden Parteien zur Bestimmung des gemeinsamen Geheimnisses durchgeführt wird.

Durch die Möglichkeit zur Vereinbarung eines gemeinsamen symmetrischen Schlüssels ist die Grundlage für eine effiziente verschlüsselte und integritätsgesicherte Kommunikation zwischen zwei Parteien gelegt. Allerdings müssen diese noch die Authentizität ihres jeweiligen Kommunikationspartners sicherstellen, da das Diffie-Hellman-Verfahren auch von jedem anderen als dem eigentlich gewünschten Kommunikationspartner durchgeführt werden kann. Insbesondere sind ohne Authentisierung auch Man-in-the-middle-Angriffe möglich, bei denen der zwischen den Kommunikationspartnern platzierte Angreifer mit jeder Partei unter Verwendung des Diffie-Hellman-Verfahrens einen Schlüssel aushandelt und sich gegenüber beiden als der jeweils andere ausgibt. Er kann dann die gesicherten Nachrichten beider Parteien lesen und sie, um unbemerkt zu bleiben, nach Gutdünken an die jeweils andere Partei weiterleiten, nachdem er sie mit dem entsprechenden Schlüssel gesichert hat.

2.3.2.3 Message Authentication Codes

So genannte *Message Authentication Codes* (MAC) werden zur Sicherung der Integrität übertragener Daten eingesetzt. Dazu wird vom Absender nach einem bestimmten Algorithmus aus den zu schützenden Daten unter Einbeziehung eines symmetrischen Schlüssels ein MAC (eine Art „kryptographische Prüfsumme“) fester Länge errechnet und zusammen mit der Nachricht verschickt. Der Empfänger, der den Schlüssel ebenfalls besitzt, errechnet genauso den MAC aus der Nachricht und vergleicht ihn mit dem übertragenen. Da vorausgesetzt wird, dass kein Dritter den Schlüssel besitzt, kann bei Übereinstimmung davon ausgegangen werden, dass die Nachricht unverändert ist, denn eine Änderung hätte eine Neuberechnung des übertragenen MACs erfordert, die nur mit Schlüssel möglich ist.

Die zur Berechnung von MACs verwendeten Algorithmen müssen die folgenden Eigenschaften besitzen:

1. Die Bestimmung einer passenden Nachricht zu einem gegebenen MAC ist nicht durchführbar (da rechnerisch zu aufwändig). Ohne diese Eigenschaft könnten Dritte unter Umständen die übertragene Nachricht unbemerkt durch eine andere mit demselben MAC ersetzen.

2. Bei Betrachtung aller möglichen Nachrichten sind alle MAC gleich wahrscheinlich. Dies erschwert die erschöpfende Suche nach einer zu einem bestimmten MAC passenden Nachricht, die in manchen Fällen versucht werden kann, wenn eine Möglichkeit besteht, beliebige Nachrichten in großer Zahl vorzulegen und mit einem MAC versehen zu lassen.

3. Der MAC einer auf irgendeine bestimmte Weise transformierten Nachricht sollte ebenfalls gleichverteilt über alle möglichen MACs sein (und nicht etwa mit erhöhter Wahrscheinlichkeit mit dem MAC der ursprünglichen Nachricht übereinstimmen). Unregelmäßigkeiten an dieser Stelle könnte ein Angreifer ebenfalls ausnutzen, um schneller eine veränderte Nachricht zu einem gegebenen MAC zu finden.

Die MAC-Berechnung ähnelt der Verschlüsselung, ist aber im Allgemeinen nicht umkehrbar. Verschlüsselungsalgorithmen können häufig mit geringen Änderungen auch zur MAC-Berechnung eingesetzt werden. Gängige MAC-Verfahren sind HMAC-MD5 und HMAC-SHA1, die auf den Hash-Funktionen MD5 und SHA1 und dem HMAC-Verfahren (siehe folgender Abschnitt) basieren.

2.3.2.4 Hash-Funktionen

Auch Hash-Funktionen, also Funktionen, die Nachrichten auf Hash-Werte fester Länge abbilden, ohne dabei geheime Information einzubeziehen, können zur Integritätssicherung verwendet werden, wenn sie geeignet mit anderen Verfahren kombiniert werden.

Beispielsweise kann ein Hash-Wert nach seiner Ermittlung aus einer Nachricht verschlüsselt und dann mit der Nachricht (die selbst entweder ebenfalls verschlüsselt oder aber im Klartext übertragen wird) übermittelt werden. Der Empfänger der Nachricht kann durch Vergleich des selbst ermittelten Hash-Werts mit dem entschlüsselten erhaltenen Wert die Integrität der Nachricht prüfen, denn Angreifer können wegen der öffentlich bekannten Hash-Funktion zwar den richtigen Hash-Wert zu einer von ihnen veränderten Nachricht ermitteln, diesen aber mangels Schlüssel nicht verschlüsseln.

Diese Kombination von Hash-Ermittlung und Verschlüsselung entspricht bei Verwendung eines symmetrischen Algorithmus genau einem MAC (wenn der Empfänger nicht den erhaltenen verschlüsselten Hash-Wert entschlüsselt, sondern den selbst ermittelten Hash-Wert verschlüsselt und dann mit dem erhaltenen vergleicht). Bei Verwendung eines asymmetrischen Algorithmus entspricht das Verfahren einer digitalen Signatur (siehe Abschnitt 2.3.2.5).

Eine andere Möglichkeit zur Erzeugung eines MAC-Verfahrens aus einer Hash-Funktion besteht darin, die zu schützende Nachricht zunächst mit einem Schlüssel zu verknüpfen (z. B. durch Anhängen des Schlüssels) und dann erst die Hash-Funktion anzuwenden. Dieses Verfahren ist unter der Bezeichnung HMAC standardisiert [KrBC97].

Anforderungen an zur Integritätssicherung geeignete Hash-Funktionen, auch als *kryptographische Hash-Funktionen* bezeichnet, sind die Folgenden:

1. Einweg-Eigenschaft: Es ist praktisch unmöglich, Nachrichten zu finden, die einen bestimmten vorgegebenen Hash-Wert besitzen. Dies ist insbesondere dann wichtig, wenn ein geheimer Schlüssel wie oben angedeutet zur Ermittlung eines MAC eingesetzt wird; dieser könnte von einem Angreifer ermittelt werden, wenn die Hash-Funktion umkehrbar wäre.

2. Kollisionssicherheit: Es ist praktisch unmöglich, zu einer gegebenen Nachricht eine weitere zu finden, die denselben Hash-Wert besitzt. Dies verhindert den Austausch einer unverschlüsselten Nachricht durch Angreifer unter Beibehaltung eines verschlüsselten Hash-Werts.

3. Starke Kollisionssicherheit: Es ist praktisch unmöglich, zwei beliebige Nachrichten zu finden, die denselben Hash-Wert besitzen. Dies ist erforderlich für bestmögliche Sicherheit gegen Angriffe, bei denen der Angreifer versucht, zwei Nachrichten mit gleichem Hash-Wert zu produzieren, von denen eine für das Opfer akzeptabel ist, während die andere dem Interesse des

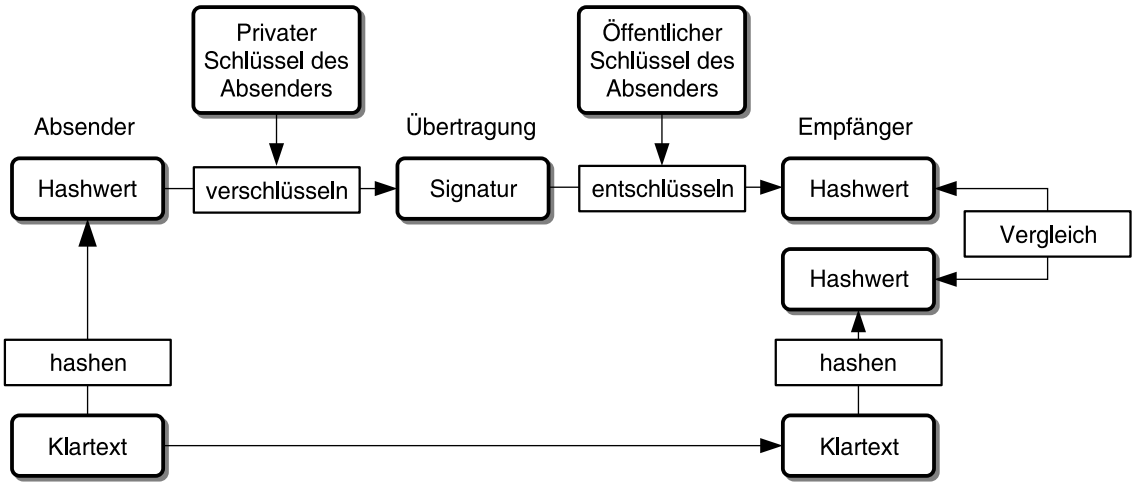


Abbildung 2.10: Digitale Signatur

Angreifers entspricht. Wenn der Angreifer dann die akzeptable Nachricht vom Opfer mit einem verschlüsselten Hash-Wert signieren lassen kann, kann er anschließend die Nachrichten unter Beibehaltung des Hash-Werts austauschen.

2.3.2.5 Digitale Signatur

Digitale Signaturen dienen dazu, die Authentizität von Dokumenten und Nachrichten zu sichern. Ein signiertes Dokument kann nur vom Inhaber des für die Signatur verwendeten Schlüssels signiert worden sein (was dieser auch nur unter der Behauptung, sein Schlüssel sei kompromittiert worden, abstreiten kann). Das Dokument kann nachträglich nicht verändert oder ausgetauscht werden.

Digitale Signaturen lassen sich beispielsweise durch Verschlüsselung eines kryptographischen Hash-Werts über das zu signierende Dokument mit dem geheimen Schlüssel eines asymmetrischen Schlüsselpaars realisieren. Da der geheime Schlüssel nur dessen Inhaber bekannt ist, kann nur er die Signatur erstellt haben. Da aber der zugehörige öffentliche Schlüssel öffentlich verfügbar ist, kann jeder den Hash-Wert entschlüsseln und überprüfen, ob er mit einem selbst aus dem Dokument berechneten übereinstimmt (siehe Abbildung 2.10).

2.3.3 Authentisierung

Die Authentisierung dient – noch etwas allgemeiner als im einführenden Abschnitt 2.3.1 formuliert – dazu, eine vorliegende Information einer bestimmten Identität zuzuordnen, welche die Information erzeugt hat oder „besitzt“ oder auf andere Weise mit derselben in Verbindung steht.

Handelt es sich bei der Information um eine einzelne erhaltene Nachricht, welche ihrem Absender zugeordnet werden soll, so spricht man von Nachrichtenauthentisierung. Sie kann durch die Überprüfung eines vom Absender angebrachten Message Authentication Codes (Abschnitt 2.3.2.3) oder einer digitalen Signatur (Abschnitt 2.3.2.5) durchgeführt werden.

Bei der anderen Form der Authentisierung, der Instanzauthentisierung, soll nicht nur eine einzelne Nachricht ihrem Absender zugeordnet werden, sondern vielmehr soll ein ganzer Kommunikationskanal, über den aus vielen Nachrichten bestehende Kommunikationsvorgänge abgewickelt werden können, bzw. insbesondere die am anderen Ende des Kanals befindliche Instanz identifiziert werden. Dies ist in seiner Wirkung grundsätzlich äquivalent mit der Authentisierung jeder einzelnen Nachricht des

Kommunikationspartners, kann aber häufig effizienter umgesetzt werden: Wenn beispielsweise zur Nachrichtenauthentisierung asymmetrische Kryptographie verwendet werden soll, so kann sehr viel Rechenaufwand gespart werden, indem ein mittels schneller symmetrischer Kryptographie integritätsgeschützter Kanal aufgebaut und die aufwändige asymmetrische Kryptographie nur noch für eine einmalige anfängliche Authentisierung des Kommunikationspartners verwendet wird.

Während bei der Nachrichtenauthentisierung alle für die Authentisierung notwendigen Informationen zusammen mit der einzelnen Nachricht übertragen werden, kann bei der Instanzauthentisierung zwischen den Kommunikationspartnern ein interaktives Protokoll mit mehreren Schritten ablaufen, bei dem die Teilnehmer jeweils auf die Vorgaben bzw. Anfragen der anderen Seite angemessen reagieren. Dieser zusätzliche Gestaltungsspielraum wird von Verfahren zur Instanzauthentisierung in der Regel auch genutzt, so dass sie nicht zur Nachrichtenauthentisierung verwendbar sind.

Die Authentisierung stellt einen grundlegenden Sicherheitsdienst dar, da es normalerweise nicht sinnvoll ist, einen Kommunikationsvorgang anderweitig zu schützen, wenn nicht wenigstens eine Partei über die Identität der anderen Gewissheit besitzt. Denn ansonsten kann ein Angreifer die Stelle des vermeintlichen Kommunikationspartners einnehmen und sich unbemerkt für diesen ausgeben, was jegliche weitere Sicherung des Kommunikationsvorgangs ad absurdum führt.

Nicht immer ist allerdings eine gegenseitige Authentisierung erforderlich. Bei einem öffentlich und gebührenfrei angebotenen Dienst spielt es beispielsweise für den Diensterbringer nicht unbedingt eine Rolle, mit wem er kommuniziert, während andererseits der Dienstnehmer unter Umständen sowohl sicher sein möchte, dass er den richtigen Diensterbringer (und nicht etwa einen falsche Ergebnisse liefernden oder zur Verfügung gestellte Daten veruntreuenden Betrüger) kontaktiert, als auch die Vertraulichkeit übertragener Daten gegenüber Dritten absichern möchte.

2.3.3.1 Ablauf

Die Authentisierung bei Kommunikationsvorgängen in Rechnernetzen muss zwangsläufig ausschließlich anhand ausgetauschter Nachrichten erfolgen. Eine Partei versichert der anderen ihre Identität dabei im Allgemeinen dadurch, dass sie den Besitz von Information – im Folgenden auch als *Geheimnis* bezeichnet – nachweist, die nur dem Inhaber dieser Identität zugänglich ist.

Es gibt unterschiedliche Verfahren, um den Besitz des Geheimnisses nachzuweisen:

- *Einfache Authentisierung durch Übertragung des Geheimnisses*

  Die einfachste und direkteste Methode besteht darin, das Geheimnis an den Kommunikationspartner zu übertragen. Dieser muss es ebenfalls kennen und kann dann einfach erkennen, ob es richtig angegeben wurde. Eine häufig anzutreffende Form dieser Authentisierungsmethode ist das Abfragen eines Passworts.

  Das Verfahren hat allerdings wesentliche Nachteile:

  – Jeder, der durch Mithören der Übertragung in den Besitz des Geheimnisses gelangt, wird in die Lage versetzt, sich für dessen Besitzer auszugeben und erfolgreich zu authentisieren. Deshalb muss entweder der Kanal, über den das Geheimnis übertragen wird, gegen Mithören gesichert sein, oder jedes Geheimnis darf nur einmal verwendet werden. Bei der zweiten Lösung muss außerdem sichergestellt sein, dass ein Mithörender das Geheimnis nicht schneller gegenüber der authentisierenden Stelle anwenden kann als der tatsächliche Eigentümer, beispielsweise indem er die ursprüngliche Übertragung auf späteren Teilstrecken stört oder schnellere Wege benutzt.

Bei Passwort-Authentisierung auf Web-Seiten (z. B. von Banken) wird dieses Problem häufig durch Verschlüsselung der Übertragung unter Verwendung von SSL (siehe Abschnitt 2.3.7.1) gelöst.

– Auch der Kommunikationspartner muss das Geheimnis kennen, und kann sich deshalb im Prinzip gegenüber Dritten als dessen Besitzer ausgeben. Deshalb muss für jede Paarung von Netzteilnehmern ein anderes Geheimnis verwendet werden, und derjenige Partner, der beispielsweise eine Nachricht als authentisch erkannt hat, kann gegenüber Dritten nicht nachweisen, dass er sie nicht selbst erzeugt hat.

  Zur Authentisierung bei Web-Zugängen verschiedener Dienstleister wird in der Praxis oft aus Bequemlichkeit dasselbe Passort verwendet, was eine erhebliche Schwachstelle darstellt, die von den Dienstleistern allerdings bisher glücklicherweise selten missbraucht wird. Teilweise schützen auch zusätzliche Mechanismen (etwa nur einmalig verwendbare Transaktionsnummern bei Bankzugängen) vor ernsthaftem Missbrauch.

- *Indirekte Authentisierung durch Anwendung des Geheimnisses*

  Statt das Geheimnis selbst zu übertragen, wird es auf zusätzliches Datenmaterial nach vorher vereinbarten Regeln angewendet. Dieses zusätzliche Datenmaterial, das beispielsweise Zufallszahlen, Zeitstempel oder Verbindungskennungen umfassen kann, ist nicht geheim und beiden Kommunikationspartnern bekannt, und es muss für jeden Authentisierungsvorgang verschiedenes Material verwendet werden. Soweit das Zusatzmaterial für den Empfänger nicht aus den Umständen ableitbar ist, kann es (gegebenenfalls ohne weitere Sicherung) zusammen mit dem Ergebnis der Operation an den Kommunikationspartner übertragen werden.

  Falls diesem das verwendete Geheimnis bekannt ist, führt er dieselbe Operation auf demselben Datenmaterial durch und vergleicht das Ergebnis mit dem übermittelten. Beispielsweise kann als Operation ein Message Authentication Code über das Zusatzmaterial berechnet oder dieses symmetrisch verschlüsselt werden, wobei jeweils das Geheimnis als Schlüssel eingesetzt wird.

  Dass der Kommunikationspartner das Geheimnis kennen muss, führt wieder zu demselben Nachteil wie schon bei der einfachen Authentisierung: Es ermöglicht Missbrauch durch den Authentisierenden und erfordert ein eigenes Geheimnisses pro Teilnehmerpaarung.

  Im Unterschied zur direkten Authentisierung kann dieser Nachteil hier aber vermieden werden, indem als Geheimnis der private Schlüssel eines asymmetrischen Schlüsselpaars und als Operation die Verschlüsselung mit demselben verwendet wird. Der Authentisierende kann die Operation dann mit Hilfe des zugehörigen öffentlichen Schlüssels umkehren, kennt aber das Geheimnis selbst nicht und kann deshalb auch keine Authentisierungsnachrichten fälschen. Ein- und dieselbe Nachricht kann von jedem authentisiert werden, der den öffentlichen Schlüssel kennt.

  Damit keine Wiedereinspielung von Authentisierungsnachrichten möglich ist (wofür ein Angreifer das Geheimnis ja nicht zu kennen braucht), darf es nicht erlaubt sein, zweimal dasselbe Zusatzmaterial zu verwenden, was aber für den Authentisierenden im allgemeinen Fall nur zu überprüfen ist, indem er sämtliche vom Kommunikationspartner bereits verwendeten Varianten aufbewahrt, was in der Regel nicht praktikabel ist. Häufig wird deshalb ein aktueller Zeitstempel als Zusatzmaterial verwendet und verlangt, dass die Authentisierungsnachricht nur innerhalb eines schmalen Zeitschlitzes um den Erzeugungszeitpunkt herum akzeptiert werden darf. Dieses Vorgehen erfordert allerdings synchronisierte Uhren bei den Kommunikationspartnern.

- *Challenge-Response-Verfahren*

  Challenge-Response-Verfahren arbeiten ähnlich wie die eben beschriebenen indirekten, wobei aber das Zusatzmaterial vom Kommunikationspartner geliefert (und als *Challenge* bezeichnet)

wird. Es darf zwar theoretisch ebenfalls nie ein zweites Mal verwendet werden, allerdings muss diesmal der Angreifer alle verwendeten Varianten speichern, um im Wiederholungsfall die bereits gehörte richtige Antwort (die *Response*) wiedereinspielen zu können. Da die Speicherung bei einem genügend großen Raum von Möglichkeiten auch für den Angreifer nicht praktikabel ist, reicht es aus, die Challenge zufällig zu wählen.

Durch die Übertragung der Challenge benötigen Challenge-Response-basierte Verfahren häufig einen Übertragungsvorgang mehr als die vorgenannten indirekten. Dafür werden keine synchronisierten Uhren gebraucht.

- *Zero-Knowledge-Verfahren*

  Verwendet man bei den bisher genannten Verfahren, bei denen das Geheimnis nicht übertragen wird, ein asymmetrisches kryptographisches Verfahren, so kennt der Kommunikationspartner das Geheimnis nicht. Allerdings erfährt er trotzdem durch jeden Authentisierungsvorgang ein wenig mehr über das Geheimnis. Durch das Vorgeben bestimmter Challenge-Werte kann der Kommunikationspartner sich in gewissen Grenzen frei wählbare Klartexte verschlüsseln lassen, was ihm bei einer kryptologischen Analyse mit dem Ziel, das Geheimnis herauszufinden, gewisse Vorteile einräumt.

  Bei den so genannten Zero-Knowledge-Verfahren wird sowohl das Mithören als auch die Analyse des Geheimnisses durch Teilnehmer des Verfahrens verhindert, indem keinerlei mit dem Geheimnis in rechnerischem Zusammenhang stehende Information übertragen wird.

2.3.3.2 Initiale Authentisierung

Für die Authentisierung über ein Netzwerk hinweg muss entweder ein gemeinsames Geheimnis mit dem Kommunikationspartner vereinbart worden sein oder ein öffentlicher Schlüssel des anderen vorliegen.

Diese Voraussetzungen können nicht über dieselben Kommunikationskanäle erreicht werden wie die Authentisierung selbst, da sowohl für die Übertragung oder Vereinbarung eines gemeinsamen Geheimnisses als auch für die Übertragung eines öffentlichen Schlüssels bereits die Authentizität der anderen Seite gewährleistet sein muss, so dass eine korrekte Zuordnung zwischen Geheimnis und Kommunikationspartner entsteht.

Letztlich kann eine wirkliche Identifizierung nur auf einer Ebene stattfinden, auf der die Teilnehmer unfälschbare Merkmale besitzen, anhand derer sie erkannt und unterschieden werden können. Merkmale elektronischer Geräte sind hierfür in der Regel nicht geeignet, da sie meist weder eindeutig noch unfälschbar sind. In letzter Instanz werden deshalb meist reale Personen authentisiert, die zumindest mit hoher Wahrscheinlichkeit wiedererkannt und unterschieden werden können.

Wenn einmal eine *initiale Authentisierung* mittels persönlichem Kontakt zwischen realen Personen erfolgt ist, können unter deren Schutz Geheimnisse vereinbart werden, anhand derer später eine Authentisierung von Kommunikationsbeziehungen auch auf Entfernung möglich ist.

Da die Geheimnisse je nach Umfang (siehe dazu auch den folgenden Abschnitt) für Menschen schwer handhabbar sind, werden sie bereits auf elektronischem Weg ausgetauscht. Hierbei muss gewährleistet sein, dass keine Verfälschungen durch Dritte erfolgen können oder sich ein Dritter gar unbemerkt zwischen die kommunizierenden Parteien schalten kann (Man-in-the-middle-Angriff). Dies lässt sich auf zwei Wegen erreichen:

- Der Kanal wird integritätsgesichert, indem z. B. mit dem Diffie-Hellman-Verfahren (siehe Abschnitt 2.3.2.2) ein temporärer Schlüssel vereinbart wird und anschließend alle Nachrichten mit einem Message Authentication Code unter Verwendung dieses Schlüssels versehen werden. Außerdem müssen noch Man-in-the-Middle-Angriffe ausgeschlossen werden, indem z. B. die Authentizität des Kommunikationspartners durch Kontrolle des vereinbarten Schlüssels über einen von einem solchen Angreifer sicher nicht kontrollierbaren Weg sichergestellt wird. Hierzu reicht es aus, wenn die Benutzer etwa per Sprachkommunikation einen kurzen und damit gut handhabbaren Hashwert auszutauschen, der aus dem Schlüssel berechnet wird.
- Es wird ein schon physikalisch gegen Angriffe immuner Kanal verwendet, etwa eine (kurze, visuell kontrollierte) Kabelverbindung, eine durch ihre kurze Reichweite und ihren gerichteten Charakter ebenfalls recht gut kontrollierbare Infrarot-Verbindung, oder ein auszutauschender Datenträger (siehe z. B. [BSSW02]).

Es ist sehr aufwändig und bei großen Netzen praktisch unmöglich, zwischen jeweils zwei potentiellen Kommunikationspartnern eine direkte Authentisierung nach obigem Schema durchzuführen. Deshalb vermittelt in der Praxis zwischen zwei kommunikationswilligen Parteien häufig eine dritte Partei, der die beiden ersten vertrauen und mit der sie eine initiale Authentisierung durchgeführt haben. Die dritte Partei stellt den anderen dann geeignete Geheimnisse (bzw. öffentliche Schlüssel) zur Verfügung, so dass sie sich entfernt authentisieren können.

2.3.3.3 Umfang des Geheimnisses

Der Raum, aus dem zur Authentisierung verwendete Geheimnisse gewählt werden, sollte so groß sein, dass ein Angreifer keine realistische Chance hat, durch erschöpfende Suche in diesem Raum mit Ausprobieren aller Möglichkeiten eine erfolgreiche Authentisierung durchzuführen.

Da auch bei guten Verschlüsselungsalgorithmen die erschöpfende Suche nach dem Schlüssel die erfolgversprechendste Möglichkeit zur Ermittlung des Klartexts sein sollte, können die schon in Abschnitt 2.3.2.1 angegebenen „sicheren“ Schlüssellängen auch als Richtwert für die Längen von zur Authentisierung verwendeten Geheimnissen gelten.

Wenn menschliche Benutzer beteiligt sind, werden häufig Passworte zur Authentisierung verwendet, weil diese sich besser handhaben und memorieren lassen. Dabei ist zu beachten, dass die in einem Passwort enthaltene geheime Information (Entropie) relativ kurz ist, insbesondere wenn natürlichsprachliche Wörter gewählt werden, die eine hohe Redundanz enthalten. Um mit einem Passwort dieselbe Sicherheit zu erreichen wie mit einem 112 bit langen Triple-DES-Schlüssel, müsste es aus mindestens 19 zufällig gewählten alphanumerischen Zeichen (6 bit Entropie pro Zeichen) bestehen, wodurch es wieder sehr schlecht handhabbar wäre. Deshalb ist es meist sinnvoller, die erschöpfende Suche dadurch zu verhindern, dass die Anzahl erlaubter Authentisierungsversuche beschränkt wird. Dadurch ergibt sich allerdings für Angreifer die Möglichkeit, gezielt die Verfügbarkeit des gesicherten Dienstes anzugreifen, indem der Zugang durch absichtliche Fehlversuche für den regulären Benutzer gesperrt wird.

2.3.4 Schlüsselverwaltung

Voraussetzung für entfernte Authentisierung ist das Vorliegen eines gemeinsamen Geheimnisses oder eines öffentlichen Schlüssels der jeweils anderen Partei. In Abschnitt 2.3.3.2 wurde bereits angesprochen, dass in realen Netzen Geheimnisse bzw. öffentliche Schlüssel initial meist nicht direkt zwi-

schen den Kommunikationspartnern ausgetauscht werden – die Zahl erforderlicher initialer Authentisierungen wüchse quadratisch mit der Zahl der Netzteilnehmer, wenn Kommunikationsbeziehungen zwischen beliebigen Teilnehmerpaaren möglich sein sollten. Stattdessen werden Schlüsselverwaltungsverfahren verwendet, bei denen die Kommunikationspartner die benötigten Geheimnisse bzw. öffentlichen Schlüssel bei Bedarf anfordern können.

Bei der Verwendung gemeinsamer Geheimnisse zur Authentisierung muss (wie in Abschnitt 2.3.3.1 erläutert) jeweils ein eigenes Geheimnis pro Teilnehmerpaarung existieren. Eine dritte Partei, welche solche Geheimnisse an authentisierungsbedürftige Teilnehmer ausliefert, kennt es notwendigerweise ebenfalls (bzw. muss es sogar selbst erst bei Bedarf generieren, wenn man vermeiden will, dass sie eine in der Anzahl der Netzteilnehmer quadratische Anzahl von Geheimnissen vorhalten muss). Die Kommunikationspartner müssen also darauf vertrauen können, dass die dritte Partei sowohl die richtigen Geheimnisse ausliefert als auch diese nicht selbst missbraucht. Bei der Kommunikation zwischen jedem Kommunikationspartner und der dritten Partei während der Geheimnisbeschaffung muss jeweils eine Authentisierung erfolgen, die auf einem Geheimnis oder öffentlichen Schlüssel basiert, welches bzw. welcher bei einer initialen Authentisierung ausgetauscht wurde. In der Praxis werden geheimnisbasierte Schlüsselverwaltungsverfahren nur in geschlossenen Umgebungen begrenzter Größe verwendet. Sie sollen deshalb an dieser Stelle nicht näher behandelt werden. Ein Beispiel für ein solches Verfahren ist das in Abschnitt 2.3.7.3 beschriebene Kerberos-System.

2.3.4.1 Zertifikate und Zertifizierungsinstanzen

Werden statt gemeinsamer Geheimnisse asymmetrische Schlüsselpaare zur Authentisierung verwendet, so gestaltet sich die Verwaltung etwas einfacher, da zur Authentisierung eines bestimmten Teilnehmers immer derselbe öffentliche Schlüssel benötigt wird, der nicht geheim gehalten werden muss. Allerdings sind trotzdem Vorkehrungen zur Gewährleistung der Authentizität der verwalteten öffentlichen Schlüssel zu treffen, damit Angreifer diese nicht durch andere ersetzen können, zu denen sie selbst die zugehörigen privaten Schlüssel besitzen.

Um die Authentizität öffentlicher Schlüssels abzusichern, also um zu gewährleisten, dass ein erhaltener Schlüssel tatsächlich der des Kommunikationspartners ist, kann man öffentliche Schlüssel in Form von *Zertifikaten* speichern und übertragen. Ein Zertifikat wird von einer so genannten *Zertifizierungsinstanz* ausgestellt und enthält im Wesentlichen die Identität eines Teilnehmers, dessen öffentlichen Schlüssel und eine Signatur der Zertifizierungsinstanz, mit der diese die Integrität und Authentizität des Schlüssels bestätigt. Im Folgenden wird die ausstellende Zertifizierungsinstanz auch als *Aussteller* eines Zertifikats und der Teilnehmer, dessen Schlüssel zertifiziert wurde, als *Inhaber* des Zertifikats bezeichnet.

Zertifikate brauchen für jeden öffentlichen Schlüssel nur einmal erzeugt zu werden und können dann, da sie keine geheime Information enthalten und ihre Integrität durch die Signatur gesichert ist, beispielsweise in öffentlichen Verzeichnissen abgelegt werden, aus denen sie bei Bedarf von jedem Teilnehmer abgerufen werden können.

Um im Rahmen einer Authentisierung die Signatur eines Zertifikats und damit Authentizität und Integrität des enthaltenen Schlüssels überprüfen zu können, muss nun wiederum der authentische öffentliche Schlüssel der Zertifizierungsinstanz vorliegen, was eine initiale Authentisierung der Zertifizierungsinstanz gegenüber jedem Teilnehmer voraussetzt. Gibt es im betrachteten Netz mehr als eine Zertifizierungsinstanz, so wäre die Forderung, eine initiale Authentisierung zwischen jeder Zertifizierungsinstanz und jedem Teilnehmer durchzuführen, wenig effizient und unter Umständen gar nicht praktikabel. Sie lässt sich aber umgehen, indem sich Zertifizierungsinstanzen auch gegenseitig Zertifikate ausstellen. Zur Authentisierung eines Kommunikationspartners benötigt man dann eine

Kette von Zertifikaten, deren jedes den Schlüssel derjenigen Zertifizierungsinstanz zertifiziert, die das nächste Zertifikat ausgestellt hat. Das letzte Zertifikat der Kette muss den Schlüssel des zu authentisierenden Kommunikationspartners enthalten, das erste muss von einer Zertifizierungsinstanz ausgestellt sein, deren authentischer öffentlicher Schlüssel dem prüfenden System bereits bekannt ist. Eine solche Kette von Zertifikaten wird auch als *Zertifizierungspfad* bezeichnet.

Neben dem öffentlichen Schlüssel der Zertifizierungsinstanz muss bei der Überprüfung eines Zertifikats beim Prüfenden außerdem Vertrauen in die Fähigkeit und Willigkeit der Zertifizierungsinstanz zur Ausstellung korrekter Zertifikate vorhanden sein. Bezüglich der ersten Zertifizierungsinstanz der Kette kann der Prüfende dieses Vertrauen eventuell bei der initialen Authentisierung aufbauen; gegebenenfalls kann es durch eine vertragliche Regelung zwischen beiden untermauert werden. Mit weiteren Zertifizierungsinstanzen in der Kette hat der Prüfende aber nie direkten Kontakt. Das Vertrauen in sie stammt daher, dass die jeweils vorangehende Zertifizierungsinstanz mit der Erstellung des Zertifikats auch eine Aussage zur Vertrauenswürdigkeit des Zertifikatinhabers abgibt. Dabei kann entweder absolutes Vertrauen implizit durch die Ausstellung bestätigt werden, oder es wird ein Maß für die Vertrauenswürdigkeit explizit im Zertifikat angegeben.

Die Regeln, nach denen eine Zertifizierungsinstanz die Authentizität der von ihr zertifizierten Schlüssel und gegebenenfalls die Vertrauenswürdigkeit der zugehörigen Zertifizierungsinstanz bezüglich weiterer Zertifizierung prüft, werden als *Zertifizierungsrichtlinien* bezeichnet. Je strikter die Zertifizierungsrichtlinien sind, d. h. je umfassender die von ihnen vorgeschriebenen Prüfung ist, desto sicherer ist die Authentizität der zertifizierten Schlüssel. Wenn die Zertifizierungsrichtlinien mehrerer Zertifizierungsinstanzen auf einem Zertifizierungspfad sich unterscheiden, so ist für die Authentizität des zertifizierten Schlüssels nur so sicher, wie durch die am wenigsten strikten Zertifizierungsrichtlinien gewährleistet.

Bezüglich der Organisation der Zertifizierungsinstanzen lassen sich im Wesentlichen zwei Varianten unterscheiden:

- Die Zertifizierungsinstanzen stellen einen *zentralen* Dienst zur Verfügung. Sie befinden sich unter der Kontrolle von in bestimmter Hinsicht besonders vertrauenswürdigen Institutionen, beispielsweise indem sie von staatlichen Stellen oder durch gesetzliche oder vertragliche Regelungen gebundene Firmen betrieben werden. Zertifikate für andere Zertifizierungsinstanzen werden in der Regel nur ausgestellt, wenn deren Zertifizierungsrichtlinien ebenso strikt sind wie die eigenen, so dass allen Zertifizierungsinstanzen auf einem Zertifizierungpfad automatisch in gleicher Weise vertraut werden kann. Es ist in diesem Fall sinnvoll, eine Hierarchie auf den Zertifizierungsinstanzen festzulegen, wie im folgenden Abschnitt 2.3.4.2 beschrieben.

- Die Zertifizierung wird *verteilt* durchgeführt: Jeder Teilnehmer ist gleichzeitig Zertifizierungsinstanz und kann Zertifikate für andere Teilnehmer ausstellen. Die Teilnehmer sind dadurch nicht gezwungen, einer zentralen Instanz zu vertrauen, sondern können selbst bestimmen, wem sie bezüglich der Zertifizierung vertrauen wollen. In Abschnitt 2.3.4.3 wird noch näher auf diesen Ansatz eingegangen.

2.3.4.2 Pfadsuche bei hierarchisch organisierten Zertifizierungsinstanzen

Die Suche nach Zertifizierungspfaden kann erleichtert werden, indem eine Hierarchie auf Zertifizierungsinstanzen erklärt wird, diese also beispielsweise in einer Baumstruktur angeordnet werden, wobei jede Zertifizierungsinstanz nur Zertifikate für ihren Vorgänger (Rückwärtszertifikat) und ihre Nachfolger (Vorwärtszertifikate) im Baum erstellt; die Wurzel des Baums bildet eine oberste Zertifizierungsinstanz.

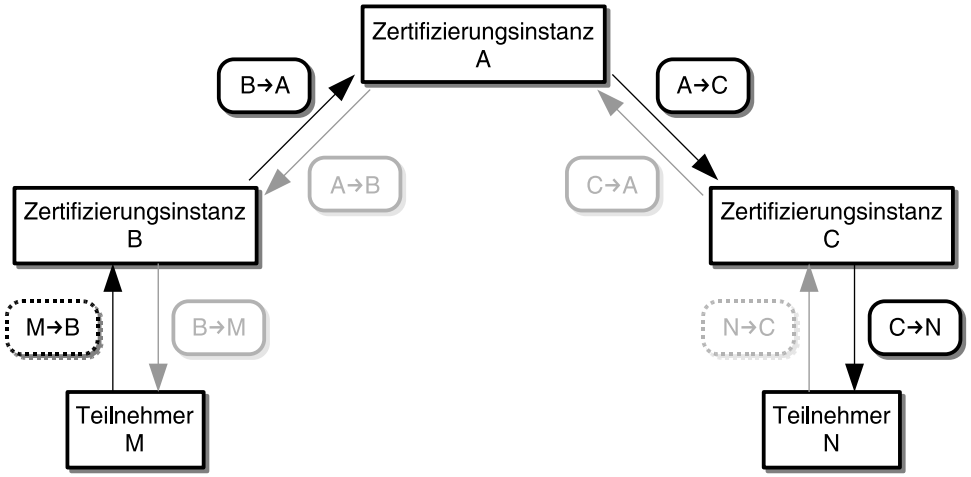


Abbildung 2.11: Hierarchie von Zertifizierungsinstanzen

Eine Kette von Zertifikaten zur Authentisierung des öffentlichen Schlüssels eines Kommunikationspartners kann ausgehend von einem Vorwärtszertifikat dieses Kommunikationspartners und dem vorhandenen Rückwärtszertifikat einer lokalen Zertifizierungsinstanz gefunden werden, indem solange jeweils Zertifikate höherer Zertifizierungsinstanzen angefordert werden, bis die dadurch erhaltenen Teilketten sich in einer gemeinsamen Zertifizierungsinstanz treffen.

Ein Beispiel anhand von Abbildung 2.11: Teilnehmer N hat Teilnehmer M seinen öffentlichen Schlüssel in Form des Zertifikats C→N mitgeteilt, welches von der Zertifizierungsinstanz C erstellt wurde. Um die Authentizität des Schlüssels zu überprüfen, fordert M ein Vorwärtszertifikat für C an und erhält A→C. Der Schlüssel der lokalen Zertifizierungsinstanz B ist M bereits bekannt (gespeichert in Form eines „Zertifikats“ M→B), M fordert nun ein Rückwärtszertifikat an und erhält B→A. In diesem einfachen Fall ergibt sich bereits eine Kette: B→A liefert den Schlüssel von A, welcher zur Prüfung von A→C verwendet werden kann. Aus dem Zertifikat A→C geht der Schlüssel von C hervor, mit welchem C→N und damit der Schlüssel von N geprüft werden kann.

2.3.4.3 Pfadsuche bei verteilter Zertifizierung

Wenn jeder Teilnehmer die Rolle einer Zertifizierungsinstanz übernehmen kann, hat der gerichtete Graph, dessen Knoten die Teilnehmer und dessen Kanten die Zertifikate repräsentieren, nicht wie im hierarchischen Fall Baumform. Die Suche nach Zertifizierungspfaden zwischen beliebigen Teilnehmern gestaltet sich entsprechend aufwändiger.

Wenn alle Zertifikate an einer Stelle vorliegen, dort also der ganze Graph bekannt ist, so kann in ihm z. B. mit Hilfe des Dijkstra-Algorithmus ein kürzester Pfad gesucht werden. Die Verwendung eines Standard-Verzeichnisdienstes, die im hierarchischen Fall noch in recht effizienter Weise möglich ist, wird dann allerdings ineffizient, da Verzeichnisdienste keine Pfadsuche anbieten und somit zur Erstellung des Graphen erst sämtliche Zertifikate abgerufen werden müssten.

Werden die Zertifikate dezentral aufbewahrt, so benötigt man zumindest ein Verfahren, um Zertifikate anhand ihres Ausstellers oder ihres Inhabers zu finden. Können sie z. B. anhand ihres Inhabers gefunden werden, so kann man einen Pfad zwischen den Teilnehmern T_{Start} und T_{Ziel} immerhin noch auf folgende Weise finden: Man beginnt bei T_{Ziel} und sucht alle für diesen Teilnehmer ausgestellten Zertifikate. Dieser Schritt ist dann rekursiv mit den Ausstellern der gefundenen Zertifikate zu wiederholen, bis man dabei entweder irgendwann ein von T_{Start} ausgestelltes Zertifikat findet, oder keine weiteren Teilnehmer mehr verfügbar sind. Letzteres Ergebnis bedeutet dabei, dass kein geeigneter

Zertifizierungspfad existiert. Das Verfahren entspricht einer Breitensuche im Graphen. Bei insgesamt n Teilnehmern sind $O(n)$ Zertifikatanfragen erforderlich, was sowohl bei großem n als auch bei teuren Anfragen sehr ungünstig ist. Ein besseres Verfahren ist ohne weitere Voraussetzungen aber nicht realisierbar.

2.3.5 Vertrauen und Authentizität

In den vorangegangenen Abschnitten wurde beschrieben, wie Authentizität mit Hilfe von Zertifikaten überprüft werden kann. Voraussetzung ist dabei allerdings, dass der Prüfende jeder Zertifizierungsinstanz auf dem gefundenen Zertifizierungspfad vertraut, denn wenn auch nur eine Zertifizierungsinstanz auf dem Pfad ein falsches Zertifikat ausstellt, ist die Authentizität aller folgenden und des gesuchten Schlüssels selbst nicht mehr erwiesen. Insbesondere bei verteilter Zertifizierung ist Vertrauen in alle potentiellen Zertifizierungsinstanzen nicht automatisch gegeben.

2.3.5.1 Vertrauensbegriff

Vertrauen ist eine menschliche Wahrnehmung, welche die Erwartung ausdrückt, dass eine andere Partei sich korrekt verhält. Welches Verhalten korrekt ist, wird dabei durch konkrete Verträge oder durch ein Wertesystem festgelegt. Im Fall der Zertifizierung wird beispielsweise zum großen Teil durch Zertifizierungsrichtlinien bestimmt, wie sich Zertifizierungsinstanzen richtig zu verhalten haben. Eine allgemeinere Möglichkeit, richtiges Verhalten zu definieren, ist etwa der Kantsche kategorische Imperativ [Kant85], der Verhaltensregeln dann als richtig erkennt, wenn sie bei gleichzeitiger Anwendung durch alle Individuen einer Gruppe nicht zu Widersprüchen führen.

Die menschliche Vertrauenswahrnehmung setzt zunächst voraus, dass die betrachteten Subjekte einen freien Willen besitzen und entscheiden können, wie sie sich verhalten wollen. Vertrauen spielt aber auch bei interagierenden Maschinen eine Rolle, weil diese als Agenten menschlicher Benutzer aufgefasst werden können und als solche immer im Interesse (mindestens) einer menschlichen Partei agieren. Vertrauensbeziehungen zwischen Menschen sind dann gegebenenfalls Voraussetzung dafür, dass die durch Interaktion der von ihnen kontrollierten Maschinen erzielten Ergebnisse für die Benutzer von Wert sind.

Es ist nicht immer automatisch garantiert, dass Maschinen ausschließlich im Interesse ihrer derzeitigen Benutzer handeln. Durch unabsichtliche Fehler oder beabsichtigte Manipulationen Dritter bei der Herstellung der Maschinen (Rechnern oder Programmen) oder in Form nachträglicher Angriffe, kann bewirkt werden, dass Maschinen sich unerwartet verhalten, wobei das unerwartete Verhalten entweder eher indeterministisch oder aber zielgerichtet im Interesse einer dritten Partei erfolgt. Man kann deshalb neben dem Vertrauen zwischen Menschen auch ein Vertrauen zwischen Menschen und Maschine definieren, als die Erwartung deterministischen Verhaltens und die Sicherheit gegen Manipulationen Dritter [Jøsa96].

Vertrauen entsteht ursprünglich entweder aus eigenen positiven Erfahrungen mit derselben Partei in vergleichbaren Situationen oder aus dem (vermeintlichen) Wissen, dass richtiges Verhalten den Interessen der Partei nur förderlich ist. Letzteres kann z. B. auch durch vertragliche Verpflichtungen erreicht werden, welche zur Folge haben, dass falsches Verhalten dem Vertragsbrüchigen Nachteile (in Form rechtlicher Verfolgung) verursacht – damit wird die Erwartung begründet, dass der Vertragspartner den Vertrag einhalten wird, um diese Nachteile zu vermeiden.

Schließlich kann Vertrauen auch durch Empfehlungen Dritter begründet sein, es handelt sich dann um so genanntes *indirektes Vertrauen*. In diesem Fall muss allerdings auch (direktes oder indirektes)

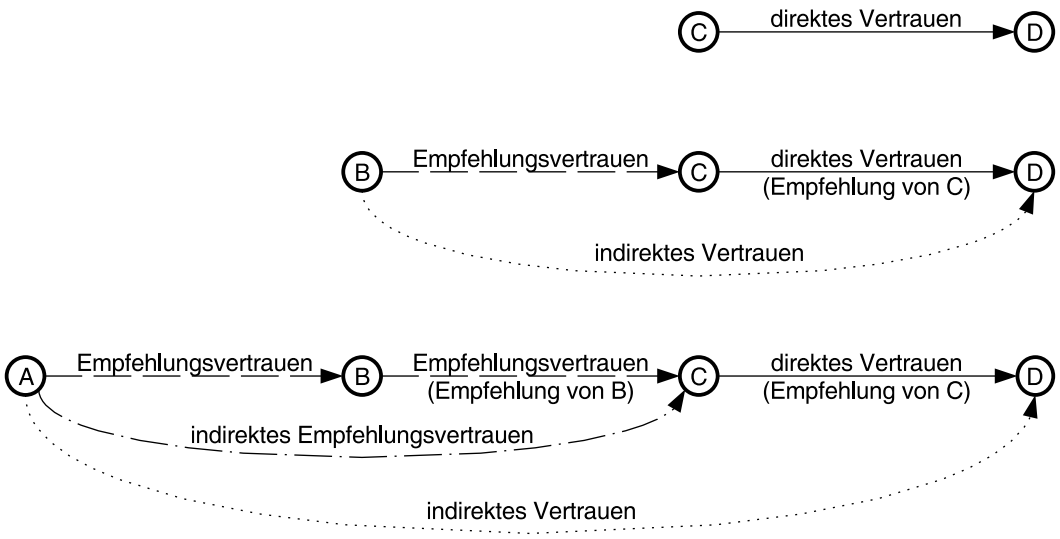


Abbildung 2.12: Beispiele für verschiedene Arten von Vertrauensbeziehungen

Vertrauen in solche Empfehlungen der empfehlenden dritten Partei vorhanden sein. Diese besondere Art von Vertrauen in Empfehlungen wird im Folgenden als *Empfehlungsvertrauen* bezeichnet.

Abbildung 2.12 zeigt Beispiele für verschiedene Vertrauensbeziehungen:

- Direktes Vertrauen einer Partei C in eine Partei D (oben; durchgezogener Pfeil) beschreibt die auf Erfahrung oder Wissen über Verpflichtungen begründete Erwartung von C, dass D sich in bestimmter Weise verhalten wird.

- Indirektes Vertrauen von B in D (Mitte; punktierter Pfeil) entsteht, wenn

  1. C sein Vertrauen in D in Form einer Empfehlung B gegenüber ausdrückt und
  2. B solchen Empfehlungen von C vertraut, also Empfehlungsvertrauen (gestrichelter Pfeil) in C besitzt. In anderen Worten erwartet B aufgrund von Erfahrung oder Wissen über Verpflichtungen, von C ehrlich und kompetent beraten zu werden.

- Eine Kette von Empfehlungen kann zu indirektem Vertrauen von A in D führen (unten), wenn

  1. A die Empfehlung von C bezüglich D bekannt ist,
  2. B sein Empfehlungsvertrauen in C in Form einer Empfehlung gegenüber A ausdrückt, und
  3. A noch Empfehlungsvertrauen in B besitzt.

  Aus den beiden letztgenannten Vertrauensbeziehungen ergibt sich bei A indirektes Empfehlungsvertrauen in C, welches die Verwendung der von C geäußerten Empfehlung ermöglicht.

Ebenso wie hier Empfehlungsvertrauen als eine besondere Art von Vertrauen aufgefasst wird, die vom nicht Empfehlungen betreffenden Vertrauen in eine Partei unterschieden und prinzipiell unabhängig davon behandelt wird, so kann man auch die beiden Vorkommnisse von (direktem) Empfehlungsvertrauen im unteren Teil der Abbildung als unterschiedliche Arten von Empfehlungsvertrauen auffassen, da sie sich auf Empfehlungen bezüglich verschiedener Arten von Vertrauen beziehen. Das führt allerdings dazu, dass für jede Länge von Ketten von Vertrauensbeziehungen eine eigene Art von Vertrauen eingeführt werden muss. Geht man andererseits vereinfachend davon aus, dass eine Partei bei allen

Arten von Empfehlungen in gleichem Maße ehrlich und kompetent ist, so kann man sich auf eine Art von Empfehlungsvertrauen beschränken und diese als transitiv betrachten.

Kryptographische Verfahren können kein Vertrauen erzeugen, sondern dieses nur dorthin übertragen, wo es benötigt wird. Im Fall der Zertifizierung wird das Vertrauen in die Zertifizierungsinstanzen unter Verwendung von Zertifikaten als kryptographischem Hilfsmittel auf die letztendlich interessierende Bindung zwischen einer Teilnehmeridentität und einem Schlüssel übertragen.

2.3.5.2 Zusammenhang von Vertrauen und Authentizität

Vertrauen und Authentizität sind eng miteinander verknüpft, weil immer dann, wenn keine direkte Authentisierung möglich ist, die erreichte Authentizität vom Vertrauen in Dritte abhängt, die an der Authentisierung beteiligt waren. Und auch bei direkter Authentisierung kann man das Vertrauen in das verwendete Verfahren berücksichtigen, welches nicht immer absolut sein muss, so dass auch dabei evtl. keine absolute Authentizität erreicht werden kann.

Abhängig davon, wie hoch das Vertrauen in den Aussteller eines Zertifikats ist, wird man den zertifizierten Schlüssel als authentisch betrachten oder nicht. Wenn die Aussteller einzelner Zertifikate für einen bestimmten Schlüssel nicht vertrauenswürdig genug sind, möchte man evtl. bei Vorliegen mehrerer solcher Zertifikate von verschiedenen Ausstellern den Schlüssel trotzdem als authentisch betrachten.

2.3.5.3 Vertrauensmetriken

Um auf rechnerischem Weg modellieren zu können, in welchem Maß ein Benutzer dem Ergebnis eines kryptographischen Verfahrens in Abhängigkeit von seinem (sozusagen als Eingabe dienenden) Vertrauen in die beteiligten Instanzen vertrauen kann, müssen ein Modell und eine Metrik für Vertrauen gefunden werden.

In der Literatur und in existierenden Authentisierungssystemen werden unterschiedliche Metriken für Vertrauen und Authentizität vorgeschlagen bzw. verwendet. Reiter und Stubblebine analysieren [ReSt97] eine Reihe von ihnen und stellen einige Kriterien auf, denen sinnvolle Metriken genügen sollten:

1. Die Bildung des Modells, auf welches die Metrik angewendet wird, sollte nicht erfordern, dass der Benutzer die Verbindung zwischen Schlüsseln und deren Besitzern herstellen muss. Insbesondere sollte für die Repräsentation von Zertifikaten im Modell berücksichtigt sein, dass Zertifikate letztlich nicht von Personen, sondern von Schlüsseln signiert sind.

   Weil Zertifikate mit Hilfe von Schlüsseln signiert sind, die sich im Besitz der Zertifikatinhaber befinden, ist die Angabe der Ausstelleridentität im Zertifikat im Prinzip redundant und darf nur als Hinweis zur Akquisition des Schlüssels verstanden werden. Erst wenn die Verifikation des Zertifikats anhand des Schlüssels erfolgreich abgeschlossen ist, darf sein Inhalt – inklusive einer Ausstelleridentität – als erwiesen gelten. Wenn im Modell zur Metrikberechnung, welche zur Auswahl geeigneter Pfade *vor* der Verifikation durchgeführt werden können soll, die Aussteller von Zertifikaten als Personen benötigt werden, so steht der Benutzer vor dem schwierigen Problem, die Besitzer der Schlüssel finden zu müssen.

2. Die Bedeutung der Parameter des Modells sollte klar definiert sein. Dies gilt insbesondere für die Bedeutung von Wahrscheinlichkeiten und Vertrauenswerten, wenn solche verwendet werden.

Wenn die Bedeutung von Parametern zweideutig ist, hängt das Ergebnis der Berechnung von der Interpretation des Ausführenden ab. Besonders problematisch ist es, wenn von Anderen bestimmte Parameter in die Berechnung eingehen.

3. Alle verfügbare Information sollte in die Berechnung eingehen.

   Dies ist erforderlich, um die auf dem Ergebnis beruhende Entscheidung darüber, ob die Authentizität „gut genug“ ist, möglichst einfach treffen zu können.

4. Die erzeugten Ausgabewerte sollten intuitiv interpretierbar sein; ihre Bedeutung sollte sich in einem einfachen Satz beschreiben lassen.

5. Die Metrik sollte unempfindlich gegenüber gezielten Manipulationen einzelner Instanzen am Modell sein, und die Empfindlichkeit gegenüber verschiedenen Formen von Fehlverhalten sollte explizit ablesbar sein.

6. Die praktische Anwendbarkeit einer Metrik sollte nicht davon abhängen, dass Entscheidungen, die das Ergebnis beeinflussen, vor dem Benutzer verborgen werden, obwohl sie nicht sinnvoll automatisiert werden können. Verborgen werden dürfen nur Entscheidungen, die anhand eindeutiger, wohldokumentierter und intuitiv verständlicher Regeln getroffen werden können.

   Ein Beispiel für eine Entscheidung, die oft kaum automatisiert werden kann, ist die Frage, ob zwei Schlüssel unabhängig voneinander sind, so dass zwei mit ihnen signierte Zertifikate als unabhängig gelten können. Besitzt beispielsweise eine Person zwei Schlüssel und erzeugt damit zwei Zertifikate für dasselbe Ziel, so sollte diesem dadurch normalerweise nicht doppelt so stark vertraut werden.

7. Eine Metrik sollte effizient berechenbar sein.

8. Die Ausgabe sollte auch bei unvollständiger Eingabe sinnvoll sein.

   In verteilten Systemen ist es beispielsweise oft schwierig oder unmöglich, alle existierenden Zertifikate zu kennen. Ein aufgrund solcher unvollständiger Information berechneter Ausgabewert sollte sich bei Ergänzung der fehlenden Information nicht willkürlich verändern.

2.3.5.4 Subjektive Vertrauensmetrik nach Jøsang

Jøsang [Jøsa98, JøIs02] bewertet Authentizität und Vertrauen anhand einer zweidimensionalen Metrik, indem er so genannte unsichere Wahrscheinlichkeiten, also Wahrscheinlichkeiten zweiter Ordnung, als Maßzahlen verwendet. Er verwendet zwei verschiedene äquivalente Modelle für unsichere Wahrscheinlichkeiten, die er als *Beweisraum (Evidence Space)* und *Meinungsraum (Opinion Space)* bezeichnet.

2.3.5.4.1 Beweisraum. Im Beweisraum werden anhand eines Modells für das Verhalten von Instanzen sowie der beiden durch Beobachtung des Verhaltens in der Vergangenheit gewonnenen Parameter r und s, welche für die Anzahlen positiver bzw. negativer Erfahrungen stehen, Aussagen über die Vertrauenswürdigkeit der Instanzen gewonnen.

Das Verhalten einer Instanz wird, indem davon ausgegangen wird, dass sie sich in einem Anteil θ aller Fälle richtig und ansonsten falsch verhält, als Urne mit Anteilen von θ roten und $(1-\theta)$ schwarzen

Kugeln modelliert. Die Anzahl Y gezogener roter Kugeln in einer Stichprobe der Größe n (mit Zurücklegen) ist damit eine binomialverteilte Zufallsvariable, und für jede bestimmte Anzahl $y \leq n$ gilt

$$P(Y = y) = \binom{n}{y} \theta^y (1-\theta)^{n-y}.$$

Tatsächlich ist dem Beobachter der Anteil roter Kugeln (also die Vertrauenswürdigkeit der betrachteten Instanz) allerdings nicht bekannt. Man kann den unbekannten Anteil roter Kugeln ebenfalls als Zufallsvariable Θ mit Wertebereich $[0,1]$ auffassen und die Wahrscheinlichkeit, eine bestimmte Anzahl roter Kugeln zu ziehen, als bedingte Wahrscheinlichkeit $P(Y = y \,|\, \Theta = \theta)$ schreiben. Das Zufallsexperiment des Ziehens aus einer Urne mit unbekanntem Anteil roter Kugeln wird dabei gedanklich ersetzt durch eines, bei dem vor dem Ziehen der Kugeln zunächst aus einer großen Zahl von Urnen mit verschiedenen, jeweils bekannten Anteilen roter Kugeln eine zufällig ausgewählt wird. $P(Y = y \,|\, \Theta = \theta)$ ist dann die Wahrscheinlichkeit, y rote Kugeln zu ziehen, wenn bereits eine Urne mit einem Anteil θ roter Kugeln ausgewählt wurde.

Der Beobachter möchte aber eigentlich eine andere Aufgabe lösen. Ihm liegt nämlich mit den Anzahlen r und s positiver bzw. negativer Erfahrungen bereits das Ergebnis einer Ziehung der Größe $r+s$ vor. Er möchte nun etwas über die Vertrauenswürdigkeit der Instanz, also den Anteil der roten Kugeln in der tatsächlich verwendeten Urne erfahren. Dazu wäre die „umgekehrte“ bedingte Wahrscheinlichkeit $P(\Theta = \theta \,|\, Y = r)$ hilfreich, welche angibt, wie wahrscheinlich es ist, dass eine Urne mit einem Anteil θ roter Kugeln ausgewählt worden ist, wenn bei der anschließenden Ziehung r rote Kugeln gezogen werden. Genauer gesagt möchte man nicht die Wahrscheinlichkeit für ein bestimmtes θ erfahren (diese ist wegen des kontinuierlichen Merkmalsraums sowieso gleich Null), sondern die Verteilung der Zufallsvariablen Θ unter der Bedingung $Y = r$, die auch als A-posteriori-Verteilung zu der A-priori-Verteilung von Θ bezeichnet wird.

Für die A-priori-Verteilung von Θ soll Gleichverteilung angenommen werden, so dass jeder Anteil roter Kugeln gleich wahrscheinlich ist. Die A-posteriori-Verteilung lässt sich nun mit Hilfe des Theorem von Bayes bestimmen, und man findet dabei, dass Θ betaverteilt mit den Parametern $a = r+1$ und $b = s+1$ ist. Die Dichte der Betaverteilung ist

$$\beta_{a,b}(\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1-\theta)^{b-1} \quad \text{mit} \quad \Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\, dt \;\text{ für }\; x \in (0,\infty).$$

Die Betaverteilung hat den Erwartungswert $E(\theta) = a/(a+b)$ und stellt wegen

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!} \quad \text{und} \quad \Gamma(x) = (x-1)! \;\text{ für }\; x \in \mathbb{N}$$

gewissermaßen eine kontinuierliche Verallgemeinerung der Binomialverteilung dar. Insgesamt erhält man im gewählten Modell als Verteilungsdichte $\varphi_{r,s}(\theta)$ der Vertrauenswürdigkeit einer Instanz, die r mal bei richtigem und s mal bei falschem Verhalten beobachtet wurde

$$\varphi_{r,s}(\theta) = \frac{\Gamma(r+s+2)}{\Gamma(r+1)\Gamma(s+1)} \theta^r (1-\theta)^s, \quad \theta \in (0,1).$$

Der Erwartungswert der zu $\varphi_{r,s}(\theta)$ gehörigen Verteilung ist $E(\theta) = (r+1)/(r+s+2)$. Der Raum aller Verteilungsdichten $\varphi_{r,s}(\theta)$ für $r,s \geq 0$ wird mit Φ bezeichnet.

In Abbildung 2.13, welche $\varphi_{r,s}(\theta)$ für verschiedene Werte von r und s zeigt, wobei $n = r+s$ immer gleich 10 ist, ist zu erkennen, wie sich die Lage des Maximums mit dem Verhältnis von r und s verschiebt. Abbildung 2.14 veranschaulicht dagegen, wie die Verteilung mit steigender Gesamtzahl n immer „schmaler“ wird.

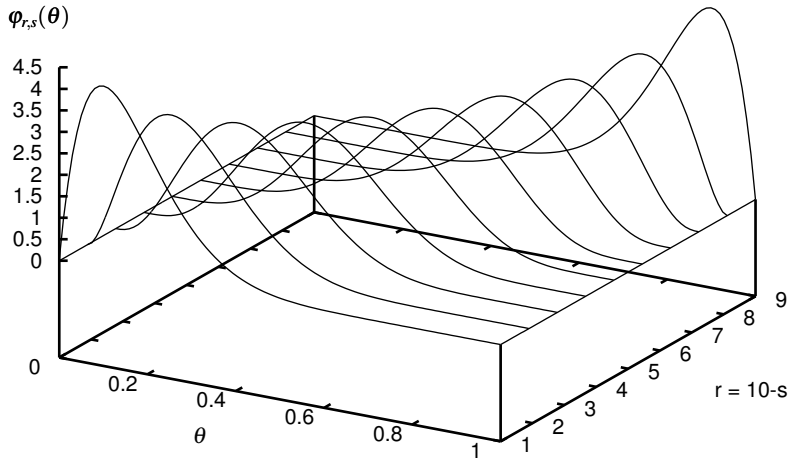


Abbildung 2.13: Abhängigkeit von $\varphi_{r,s}(\theta)$ vom Verhältnis von r und s

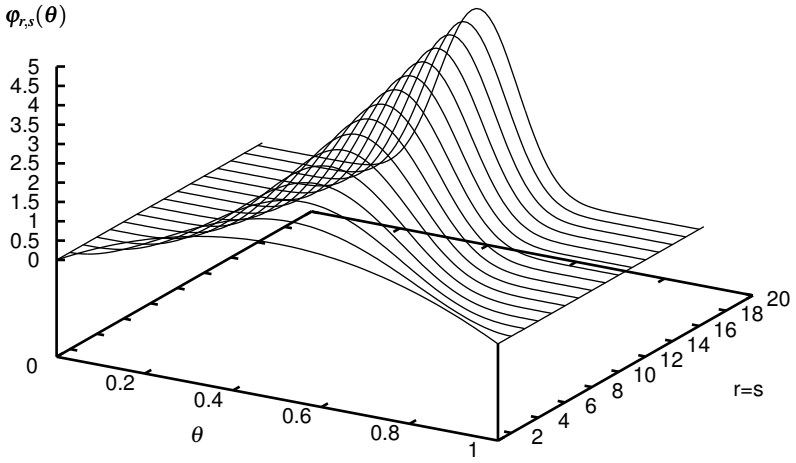


Abbildung 2.14: Abhängigkeit von $\varphi_{r,s}(\theta)$ von der Gesamtzahl $n = r + s$ der Beobachtungen

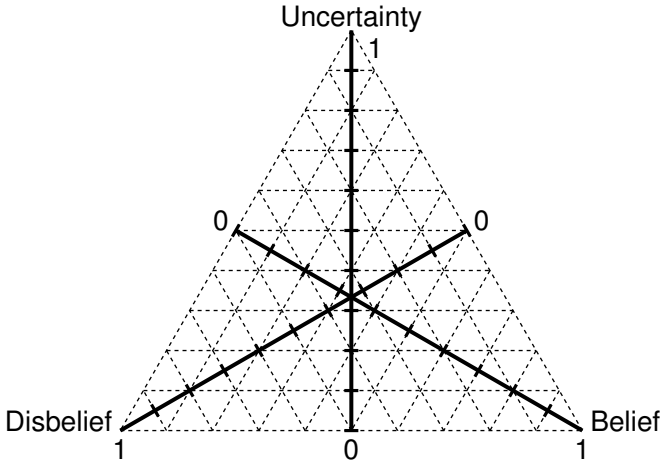


Abbildung 2.15: Meinungsdreieck (vgl. [Jøsa98])

2.3.5.4.2 Meinungsraum. Im Meinungsraum werden Aussagen bewertet, die objektiv nur entweder wahr oder falsch sein können, von denen der Bewertende aber nicht genau wissen kann, welches der beiden Extreme zutrifft. Er bildet sich deshalb eine *Meinung (Opinion)* $\omega = (b,d,u)$, die aus drei sich ergänzenden Komponenten besteht: dem Glauben b an die Richtigkeit der Aussage (Belief), dem Glauben d an ihre Falschheit (Disbelief) und schließlich der Unsicherheit u in Ermangelung ausreichender Information, die für Richtigkeit oder Falschheit sprechen würde. Dabei gilt

$$b,d,u \in [0,1] \quad \text{und} \quad b+d+u = 1.$$

Meinungen können als Punkte in dem in Abbildung 2.15 dargestellten Dreieck aufgefasst werden. Die rechte untere Ecke korrespondiert beispielsweise mit der Meinung $(1,0,0)$, die für absoluten Glauben an die Richtigkeit der Aussage steht und damit dem aussagenlogischen Wahrheitswert „wahr“ entspricht. In der linken unteren Ecke liegt der entgegengesetzte Wahrheitswert „falsch“, entsprechend der Meinung $(0,1,0)$. Die obere Spitze des Dreiecks schließlich gehört zur Meinung $(0,0,1)$, welche für völlige Unsicherheit steht. Alle Meinungen ohne Unsicherheit liegen auf der horizontalen Grundlinie zwischen $(1,0,0)$ und $(0,1,0)$ und entsprechen den eindimensionalen Bewertungszahlen anderer Metriken.

Der Unterraum von $[0,1]^3$, der alle möglichen Meinungen enthält, wird mit Ω bezeichnet. Eine Meinung einer bestimmten Instanz A zu einer Aussage x wird als ω_x^A notiert.

2.3.5.4.3 Äquivalenz zwischen Beweis- und Meinungsraum. Zwischen dem Beweisraum und dem Meinungsraum (bezogen auf die Bewertung einer bestimmten Aussage durch eine bestimmte Instanz) wird eine bijektive Abbildung $f : \Phi \to \Omega$ erklärt durch

$$f(\varphi_{r,s}) = \omega \quad \text{mit} \quad \omega = (b,d,u) \quad \text{und} \quad \begin{cases} b = \frac{r}{r+s+1} \\ d = \frac{s}{r+s+1} \\ u = \frac{1}{r+s+1} \end{cases}.$$

Für r und s muss ∞ als Wert zugelassen werden, damit f bijektiv ist. Es entspricht dann z. B. $\varphi_{\infty,0}$ – zustandegekommen bei unendlich vielen positiven Beobachtungen – der Meinung $(1,0,0)$, die absoluten Glauben an die Richtigkeit der Aussage ausdrückt. $\varphi_{0,\infty}$, das Resultat unendlich vieler negativer Beobachtungen, korrespondiert mit der absolut negativen Meinung $(0,1,0)$. Das aus dem Fehlen von Beobachtungen entstandene $\varphi_{0,0}$ wird auf die völlige Unsicherheit $(0,0,1)$ abgebildet. Die Interpretationen von $\varphi_{r,s}$ und $\omega = f(\varphi_{r,s})$ stimmen also überein.

2.3.5.4.4 Operatoren der subjektiven Logik. Auf dem Beweis- und dem Meinungsraum sind besondere logische Operatoren definiert, so dass das Meinungsmodell und die Operatoren zusammen eine Logik bilden, bezeichnet als *subjektive Logik*.

- Die *Konjunktion* $\omega_{x\wedge y} = \omega_x \wedge \omega_y$ zweier unabhängiger Meinungen $\omega_x = (b_x, d_x, u_x)$ und $\omega_y = (b_y, d_y, u_y)$ über zwei verschiedene Aussagen x und y bewertet das gleichzeitige Zutreffen beider Aussagen und ist erklärt durch

$$\omega_{x\wedge y} = (b_{x\wedge y}, d_{x\wedge y}, u_{x\wedge y}) \quad \text{mit} \quad \begin{cases} b_{x\wedge y} = b_x b_y \\ d_{x\wedge y} = d_x + d_y - d_x d_y \\ u_{x\wedge y} = b_x u_y + u_x b_y + u_x u_y \end{cases}.$$

  Die Operation $\wedge$ ist kommutativ und assoziativ und entspricht in ihrer Wirkung der Konjunktion in der binären Aussagenlogik, wenn der Meinungsraum auf die Aussagen „wahr“ $(1,0,0)$ und „falsch“ $(0,1,0)$ eingeschränkt wird. Schränkt man den Meinungsraum auf Meinungen ohne Unsicherheit ein und interpretiert diese als Wahrscheinlichkeiten, so entspricht die Konjunktion der Multiplikation der Einzelwahrscheinlichkeiten bei der Bestimmung der Wahrscheinlichkeit des gleichzeitigen Eintretens zweier unabhängiger Ereignisse.

- Der *Konsens* $\omega^{A,B} = \omega^A \oplus \omega^B$ zweier voneinander unabhängiger Meinungen $\omega^A = (b^A, d^A, u^A)$ und $\omega^B = (b^B, d^B, u^B)$ verschiedener Instanzen A und B über dieselbe Aussage beschreibt die Bewertung, die sich durch Vereinigung aller durch die beiden Instanzen getätigten Beobachtungen ergibt. Er ist zunächst im Beweisraum erklärt durch

$$\varphi_{r^{A,B},s^{A,B}} = \varphi_{r^A,s^A} \oplus \varphi_{r^B,s^B} \quad \text{mit} \quad \begin{cases} r^{A,B} = r^A + r^B \\ s^{A,B} = s^A + s^B \end{cases}$$

  und lässt sich anhand der in Abschnitt 2.3.5.4.3 eingeführten Abbildung zwischen Beweis- und Meinungsraum in letzteren übertragen als

$$\omega^{A,B} = (b^{A,B}, d^{A,B}, u^{A,B}) \quad \text{mit} \quad \begin{cases} b^{A,B} = (b^A u^B + b^B u^A)/\kappa \\ d^{A,B} = (d^A u^B + d^B u^A)/\kappa \\ u^{A,B} = u^A u^B/\kappa \end{cases}, \text{wobei } \kappa = u^A + u^B - u^A u^B.$$

  Die Operation $\oplus$ ist kommutativ und assoziativ. Durch die Vereinigung der Beobachtungen mehrerer Instanzen verringert sie die Unsicherheit. Enthält eine der Meinungen keine Unsicherheit, so dominiert sie das Ergebnis, das dann mit ihr übereinstimmt. Das Ergebnis der Verknüpfung zweier Meinungen ohne Unsicherheit (entsprechend $r + s \to \infty$) ist undefiniert, was so interpretiert werden kann, dass jede der beiden Meinungen durch das völlige Fehlen von Unsicherheit nicht mehr durch andere Meinungen beeinflussbar ist.

- Die *Empfehlungsoperation* nutzt eine Instanz A, wenn ihr eine Meinung $\omega_x^B = (b_x^B, d_x^B, u_x^B)$ der Instanz B zu einer Aussage x vorliegt, um unter Berücksichtigung der eigenen Meinung $\omega_B^A = (b_B^A, d_B^A, u_B^A)$ zu Bs Empfehlungen eine Meinung $\omega^{AB} = \omega_B^A \otimes \omega_x^B$ über die Aussage x abzuleiten. Sie ist definiert durch

$$\omega^{AB} = (b^{AB}, d^{AB}, u^{AB}) \quad \text{mit} \quad \begin{cases} b^{AB} = b_B^A b_x^B \\ d^{AB} = b_B^A d_x^B \\ u^{AB} = d_B^A + u_B^A + b_B^A u_x^B \end{cases}.$$

Die Operation $\otimes$ ist assoziativ, aber (einleuchtenderweise) nicht kommutativ. Wenn A an der Qualität von Bs Empfehlungen zweifelt ($d_B^A > 0$), so bewirkt dies, dass As empfehlungsbasierte Meinung zur Aussage x mehr Unsicherheit enthält als Bs Meinung; im Extremfall völligen Misstrauens mit $\omega_B^A = (0,1,0)$ ist $\omega_x^{AB} = (0,0,1)$, es bleibt bei A also völlige Ungewissheit. Vertraut A dagegen Bs Empfehlungen absolut ($\omega_B^A = (1,0,0)$), so ist auch $\omega_x^{AB} = \omega_x^B$.

2.3.5.4.5 Anwendung bei der Zertifizierung. Die subjektive Vertrauensmetrik kann verwendet werden, um die Authentizität eines Schlüssels anhand eines zu ihm führenden Zertifizierungspfades zu bewerten. Für den Pfad

$$A_1 \rightarrow A_2 \rightarrow \ldots \rightarrow A_{n-1} \rightarrow A_n$$

berechnet Jøsang beispielsweise das Vertrauen von A_1 in die Authentizität des Schlüssels K_n von A_n als

$$\omega_{a(K_n)}^{A_1} = (\omega_{r(A_2)}^{A_1} \wedge \omega_{a(K_2)}^{A_1}) \otimes \cdots \otimes (\omega_{r(A_{n-1})}^{A_{n-2}} \wedge \omega_{a(K_{n-1})}^{A_{n-2}}) \otimes \omega_{a(K_n)}^{A_{n-1}}$$

wobei $a(K_i)$ die Authentizität des Schlüssels K_i bedeutet und $r(A_i)$ die Eignung von A_i als Zertifizierungsinstanz.

2.3.5.4.6 Bewertung. Die subjektive Authentizitätsmetrik von Jøsang genügt den Kriterien, die in Abschnitt 2.3.5.3 für Vertrauensmetriken genannt wurden. Durch die Verwendung zweidimensionaler Vertrauenswerte ist es möglich, auch Unwissenheit bzw. Unsicherheit zu repräsentieren, was in anderen Metriken, die häufig nur eindimensionale Werte im Bereich $[0,1]$ zulassen, nicht geht. Die Wirkung der definierten Operatoren ist intuitiv verständlich.

Die für die Operatoren verlangte Unabhängigkeit der verknüpften Werte ist allerdings in der Praxis schwierig zu erreichen. Wenn Vertrauenswerte aus Beobachtungen ermittelt werden, muss dazu gewährleistet sein, das nicht mehrere Instanzen dieselben Ereignisse beobachtet haben. Jede einzelne Beobachtung zu protokollieren, um später bei der Verknüpfung doppelt auftretende Beobachtungen ausfiltern zu können, wäre doch sehr aufwändig.

2.3.6 Zugangskontrolle

Die Aufgabe von Mechanismen zur *Zugangskontrolle* ist es, sicherzustellen, dass nur bestimmte berechtigte Benutzer oder Instanzen einen bestimmten Dienst nutzen können. Der Begriff Zugangskontrolle kann als Spezialisierung des Begriffs *Zugriffskontrolle* aufgefasst werden: Zugriffskontrolle regelt allgemein den Zugriff auf Objekte aller Arten, und Zugang zu einem Dienst ist ein mögliches Objekt. Auch wenn der Gegenstand dieser Arbeit die Zu*gangs*kontrolle in Ad-hoc-Netzen ist, wird in diesem Abschnitt überall dort, wo die Ausführungen auch für den allgemeineren Fall gelten, das Wort Zu*griffs*kontrolle verwendet.

Eine wichtige Voraussetzung für die Durchführung von Zugriffskontrolle ist eine Authentisierung, denn wenn dem Zugreifenden keine Identität zugeordnet werden kann, kann auch keine Zugriffsentscheidung getroffen werden. Die Identität muss den Zugreifenden allerdings nicht unbedingt individuell auszeichnen. In manchen Fällen reicht es, ihn als Mitglied einer bestimmten Gruppe identifizieren zu können.

Wenn die Entscheidung, ob der Zugriff gewährt werden soll, im Wesentlichen nur von der Identität des Zugreifenden und einer verhältnismäßig statischen Rechtezuordnung abhängt, wird sie meist anhand einer matrixartigen Datenstruktur getroffen, in welcher zu jeder Identität und jedem Objekt vermerkt

ist, ob der Zugriff zu erlauben ist. Diese Datenstruktur wird als *Zugriffskontrollliste* (*Access Control List*, ACL) bezeichnet.

Eine andere Möglichkeit, die den Zugriff kontrollierende Instanz in die Lage zu versetzen, eine Entscheidung zu treffen, stellt die Verwendung von so genannten Autorisierungszertifikaten dar. Ein Autorisierungszertifikat enthält beispielsweise den öffentlichen Schlüssel einer Instanz und eine Liste von Objekten, auf die der Inhaber des Schlüssels zugreifen darf, und ist von derjenigen Instanz signiert, welche über die Vergabe von Zugriffsrechten entscheidet. Wenn der Inhaber eines Autorisierungszertifikats auf ein Objekt zugreifen möchte, präsentiert er der Zugriffskontrollinstanz sein Zertifikat und beweist im Rahmen einer Authentisierung, dass er tatsächlich der Besitzer des im Zertifikat genannten Schlüssels ist. Die Zugriffskontrollinstanz kontrolliert die Signatur des Zertifikats mit Hilfe des ihr bekannten Schlüssels der Rechtevergabestelle und erlaubt den Zugriff entsprechend der Angaben in Zertifikat.

2.3.7 Gebräuchliche Sicherungsverfahren

Im Folgenden sollen exemplarisch drei Sicherungsverfahren vorgestellt werden, die im Internet oder in größeren internen Netzen eingesetzt werden. Es wird dabei besonders auf die Methoden für Authentisierung und Zugriffskontrolle eingegangen.

2.3.7.1 Secure Socket Layer und Transport Layer Security

Secure Socket Layer (SSL) ist ein ursprünglich von der Firma Netscape entwickeltes Protokoll zur Verschlüsselung und Integritätssicherung von Transportschichtverbindungen. Von der IETF wurde SSL inzwischen zu *Transport Layer Security* (TLS) weiterentwickelt [DiAl99, BWNHM+03], welches zu SSL rückwärtskompatibel ist. SSL wird sehr häufig zur Sicherung von Verbindungen zu Web-Servern und Mail-Servern verwendet.

SSL liegt konzeptionell zwischen Transport- und Anwendungsschicht und kann damit grundsätzlich für Anwendungen transparent eingefügt werden. Es wird jeweils eine einzelne TCP-Verbindung oder UDP-Assoziation gesichert. Dazu muss nach dem Aufbau der Verbindung das so genannte *SSL-Handshake* angestoßen werden, das wie folgt abläuft [GuGu01]:

1. Der Client teilt dem Server eine Auswahl von ihm unterstützter asymmetrischer, symmetrischer und Hash-Algorithmen mit. Der Server wählt daraus die zu verwendenden Algorithmen aus und meldet seine Auswahl zurück.

2. Der Server sendet seinen öffentlichen Schlüssel samt einem Zertifikat dafür an den Client.

3. Der Client generiert zufällig einen Sitzungsschlüssel und sendet diesen – verschlüsselt mit dem öffentlichen Schlüssel des Servers – an den Server.

4. Der Client sendet eine mit dem Sitzungsschlüssel verschlüsselte Testnachricht an den Server, die dieser entschlüsselt und bestätigt.

5. Wenn die Anwendung es erfordert und der Client ein Zertifikat für seinen Schlüssel besitzt, wird auch der Client authentisiert.

Wenn keine Fehler auftraten, werden im Anschluss alle Daten mit dem vereinbarten Sitzungsschlüssel verschlüsselt.

Bei SSL-gesicherten Verbindungen authentisiert sich häufig nur der Server gegenüber dem Client mit Hilfe seines öffentlichen Schlüssels, während eine Authentisierung für die umgekehrte Richtung falls erforderlich mit Hilfe von Passworten bzw. PINs über die verschlüsselte Verbindung durchgeführt wird. Dies beruht nicht auf einer Einschränkung des SSL-Verfahrens, sondern ist eher durch die Handhabungsvorteile von Passworten gegenüber langen kryptographischen Schlüsseln begründet (siehe dazu auch Abschnitt 2.3.3.3).

Der Schlüssel des Ausstellers des vom Server gelieferten Zertifikats muss dem Client bekannt sein. In der Praxis sind die Schlüssel einiger Zertifizierungsstellen (hier Firmen, die den Zertifizierungsdienst gegen Entgelt anbieten) beispielsweise in Web-Browser bereits mitgeliefert.

2.3.7.2 IP Security

IP Security, kurz IPSec, ist die von der IETF standardisierte Architektur für die kryptographische Sicherung des Internet-Protokolls [KeSe05].

Zur Sicherung der Nutzdaten werden in der IP-Schicht zwei zusätzliche Protokolle eingeführt: *Encapsulating Security Payload* (ESP) [Kent05b] für Verschlüsselung und Integritätssicherung und *Authentication Header* (AH) [Kent05a] nur für Integritätssicherung. Beide definieren jeweils einen speziellen Paketkopf und können entweder im *Transportmodus* oder im *Tunnelmodus* betrieben werden. Im Transportmodus wird die Sicherung nur auf den Nutzinhalt des Pakets angewendet und der zusätzliche Paketkopf wird nach demjenigen von IP eingefügt. Im Tunnelmodus wird dagegen ein neues IP-Paket generiert, welches den zusätzlichen Paketkopf sowie das gesamte gesicherte IP-Paket enthält. Der Tunnelmodus bietet z. B. die Möglichkeit, den gesamten Datenverkehr zwischen zwei durch das Internet verbundenen internen Netzen zu schützen, indem alle Pakete beim Verlassen der internen Netze von einem so genannten *Security Gateway* mit einem IPSec- und einem neuen IP-Paketkopf versehen werden. Der neue IP-Paketkopf enthält als Zieladresse das Security Gateway des anderen internen Netzes. Dort werden die hinzugefügten Paketköpfe wieder entfernt und die Pakete ungeschützt ins interne Netz entlassen.

Für Authentisierung, Schlüsselaustausch und Verwaltung von IPSec-Verbindungen waren zunächst die Protokolle ISAKMP/IKE (Internet Security Association and Key Management Protocol / Internet Key Exchange) definiert worden [Pipe98, MSST98, HaCa98], die allerdings aufgrund ihrer hohen Komplexität sehr umstritten waren, da sie bei unvorsichtiger Verwendung auch unsichere Konfigurationen zuließen [FeSc00]. Vor Kurzem wurde eine neue, stark vereinfachte und mit dem Vorgänger nicht interoperable Version standardisiert, die mit IKEv2 bezeichnet wird [Kauf05].

Eine Zugriffskontrolle erfolgt bei IPSec durch systemweit vorgegebene Richtlinien, die vorschreiben, mit welchen Zielen kommuniziert werden kann und welche Sicherheitsmechanismen dabei eingesetzt werden müssen.

2.3.7.3 Kerberos

Kerberos [MNSS87, NYHR05] ist ein ursprünglich am Massachusetts Institute of Technology entwickeltes Zugangskontrollsystem, welches ausschließlich auf symmetrischer Kryptographie basiert und zentrale Server verwendet: einen Authentisierungs-Server (AS), der alle zur Authentisierung von Benutzern nötigen Informationen besitzt, sowie einen oder mehrere Ticket-Granting-Server (TGS) mit Zugriffskontrollinformation für alle zu schützenden Dienste des Netzwerks.

Zu Beginn jeder Sitzung muss sich ein Benutzer beim AS anmelden. Dort erfolgt die Authentisierung und der Benutzer erhält als Nachweis darüber ein sitzungsbezogenes so genanntes *Ticket*.

1. Der Benutzer meldet sich mit seinem Namen beim AS an.
2. Der AS prüft, ob der Name bekannt ist, und erzeugt einen Schlüssel K_p aus dem geheimen Passwort des Benutzers, einen zufälligen Sitzungsschlüssel K_s und ein Ticket T_s, das Namen und IP-Adresse des Benutzers, den Namen eines TGS, einen Zeitstempel und eine Lebensdauer sowie den zufälligen Sitzungsschlüssel enthält und mit einem geheimen, nur dem AS und den TGS bekannten Schlüssel K_{TGS} verschlüsselt wird.

   Der Sitzungsschlüssel K_s und das Ticket werden nochmal mit K_p verschlüsselt und an den Benutzer zurückgesandt.
3. Der Benutzer erzeugt ebenfalls K_p aus seinem Kennwort und entnimmt damit der Nachricht des AS den Sitzungsschlüssel K_s und das Ticket T_s. Die Anmeldung ist damit abgeschlossen.

Bevor ein Benutzer Dienste nutzen kann, muss er sich unter Verwendung seines sitzungsbezogenen Tickets bei einem TGS anmelden, wo die Zugriffskontrolle durchgeführt wird. Bei Erfolg erhält der Benutzer ein anwendungsbezogenes Ticket.

1. Der Benutzer erzeugt eine Nachricht X mit seinem Namen, seiner IP-Adresse und einem Zeitstempel, verschlüsselt diese mit dem Sitzungsschlüssel, und schickt sie zusammen mit dem Namen der gewünschten Anwendung sowie seinem Ticket T_s an den TGS.
2. Der TGS überprüft die Übereinstimmung der Benutzerinformation im Ticket und der vom Benutzer erzeugten Nachricht X sowie die Lebensdauer des Tickets. Falls der Benutzer zur Nutzung der Anwendung berechtigt ist, erzeugt der TGS ein weiteres Ticket T_a mit Benutzername und -adresse, Anwendungsname, Zeitstempel, Lebensdauer und einem neu erzeugten Schlüssel K_a, das mit dem nur dem TGS und der Anwendung bekannten Schlüssel K_{App} verschlüsselt wird.

   Der Schlüssel K_a und das Ticket T_a werden nochmal mit dem Sitzungsschlüssel K_s verschlüsselt und an den Benutzer zurückgesandt.
3. Der Benutzer entnimmt der Nachricht des TGS den Schlüssel K_a und das neue Ticket T_a.

Zur Dienstnutzung wendet sich der Benutzer an die entsprechende Anwendung und präsentiert ihr das anwendungsbezogene Ticket:

1. Der Benutzer sendet seinen Wunsch zur Dienstnutzung zusammen mit der Nachricht X und dem Ticket T_a an die Anwendung.
2. Die Anwendung überprüft die Übereinstimmung der Benutzerinformation im Ticket T_a und der vom Benutzer erzeugten Nachricht X sowie die Lebensdauer des Tickets. Bei Übereinstimmung wird die Dienstnutzung ausgeführt.

Für Dienstnutzung über Domänengrenzen hinweg müssen sämtliche TGS paarweise geheime Schlüssel vereinbaren und sich gegenseitig vertrauen. Der Benutzer erhält dann von seinem eigenen TGS ein spezielles Ticket, mit dem er sich bei einem fremden TGS um ein Ticket bewerben kann.

Die zentralen Server stellen kritische Punkte bezüglich Sicherheit, Zuverlässigkeit und Leistungsfähigkeit dar. Bei Ausfall können keine Dienste mehr genutzt werden, bei Kompromittierung ist die Zugriffskontrolle für den Angreifer außer Kraft gesetzt.

2.4 Ad-hoc-Netze

Ad-hoc-Netze wurden in Abschnitt 2.2.1 eingeführt als Netze, die keine ortsfest installierten Komponenten voraussetzen, weil die teilnehmenden Knoten mit Hilfe drahtloser Übertragungstechnik direkt miteinander kommunizieren, und bei denen grundsätzlich jeder Knoten nicht nur als End-, sondern auch als Zwischensystem fungiert, um Pakete zwischen anderen, außerhalb ihrer gegenseitigen Sendereichweite befindlichen Knoten weiterzuleiten.

In der Literatur werden die so beschriebenen Netze häufig auch als „mobile Ad-hoc-Netze“ bezeichnet; auch die Arbeitsgruppe der IETF, die sich mit solchen Netzen beschäftigt, trägt den Namen MANET – Mobile Ad-hoc Networking. Die Bezeichnung erscheint allerdings nicht unbedingt besonders treffend, da die Mobilität von Netzen als Ganzem, die durch die Bezeichnung streng genommen ausgedrückt wird, zwar auch gegeben ist, in ihrer Bedeutung jedoch deutlich hinter der Mobilität der einzelnen Knoten zurücksteht. Bezieht man das Adjektiv „mobil“ im Namen „mobile Ad-hoc Netze“ in sprachlicher Großzügigkeit stattdessen auf die Knoten, so stellt man fest, dass man in den in Abschnitt 2.2.1 eingeführten Begriffen dann eigentlich von „mobilen und drahtlosen Ad-hoc-Netzen“ sprechen müsste, denn auch die Verwendung drahtloser Übertragungstechnik spielt eine wesentliche Rolle für die Eigenschaften der betrachteten Netze. In der vorliegenden Arbeit wird meist einfach von „Ad-hoc-Netzen“ gesprochen, und immer, wenn nichts anderes erwähnt ist, wird dabei davon ausgegangen, dass die Knoten mobil sind und mittels drahtloser Übertragungstechnik kommunizieren.

2.4.1 Szenarien für den Einsatz von Ad-hoc-Netzen

Ein Hauptvorteil von Ad-hoc-Netzen, der auch wesentlich die Szenarien mitbestimmt, in denen ihr Einsatz besonders günstig ist, ist ihre Infrastruktur-Unabhängigkeit. Häufig wäre es zwar durchaus möglich, an der betreffenden Stelle Infrastruktur aufzubauen, aber es sind viele Situationen denkbar, in denen trotzdem aus verschiedenen Gründen keine Infrastruktur verfügbar ist:

- Infrastruktur ist nicht vorhanden, da der Bedarf so spontan entstand, dass keine Zeit für deren Planung und Aufbau war.
- Der Aufbau von Infrastruktur an der betreffenden Stelle wäre unökonomisch, weil der Bedarf zu selten auftritt oder die Anzahl der Nutzer zu gering wäre.
- Infrastruktur ist nicht vorhanden, oder die vorhandene Infrastruktur ist für eine bestimmte Benutzergruppe nicht zugänglich, weil der (potentielle) Betreiber keine Interesse daran hat oder dies ausdrücklich nicht wünscht.
- Die vorhandene Infrastruktur wurde zerstört, beispielsweise durch Katastrophen, und die Möglichkeit zur Kommunikation wird von Rettungskräften benötigt.
- Es ist Infrastruktur vorhanden, aber die Bedingungen des Betreibers sind für eine bestimmte Benutzergruppe nicht akzeptabel.

Abgesehen von solchen Situationen, in denen es keine Alternativen zu Ad-hoc-Netzen gibt, kann man Ad-hoc-Netze auch als eine Art Verallgemeinerung von Infrastrukturnetzen auffassen. Insofern wäre es evtl. auch vorstellbar, in der Zukunft vorhandene Infrastrukturnetze durch Ad-hoc-Netze zu ersetzen, wenn diese den Leistungsanforderungen ebenso genügen – einfach deshalb, weil der Aufwand für Pflege der Infrastruktur entfällt. Manche Autoren [BBČG+01] sehen sogar weltumspannende Ad-hoc-Netze im Bereich des Möglichen.

Bei größeren Ad-hoc-Netzen, die nur aus mobilen Knoten mit drahtloser Übertragungstechnik bestehen, ist zu bedenken, dass die gesamte Energie für alle Übertragungen auch zwischen weit voneinander entfernten Teilen des Netzes von den mobilen Geräten geliefert werden muss. Auch wenn durch die Aufteilung weiter Übertragungswege in kurze Strecken der minimale physikalische Energiebedarf zwar nicht quadratisch mit der Entfernung der kommunizierenden Knoten steigt (sondern bei der Gesamtentfernung r und einer Aufteilung in Teilstrecken der Länge r_0 proportional zu $r \cdot r_0$ ist) entsteht mit steigender Netzgröße eine zunehmende Belastung der Knoten durch weiterzuleitende Pakete (insbesondere in der Nähe des „Zentrums“ des Netzes, wo sich viele kürzeste Wege kreuzen). Bei ständigem Senden können heutige Geräte nur relativ kurze Zeit betrieben werden, bevor ihre Energiespeicher ersetzt oder aufgeladen werden müssen. Zwar steht zu erwarten, dass durch technische Fortschritte bei der Energiespeicherung (etwa mittels der derzeit in Entwicklung befindlichen Brennstoffzellen) noch Verbesserungen erreicht werden können, aber voraussichtlich wird der Energievorrat bei mobilen Geräte weiterhin eine knappe Ressource bleiben.

Abgesehen von der Übertragungstechnik unterscheiden sich Ad-hoc-Netze durch ihre Organisationsform, die sich durch die Infrastrukturfreiheit natürlich ergibt, von bestehenden Netzen: Alle Funktionalität liegt innerhalb der Knoten, die sich sich dynamisch zu einem Netz zusammenschließen. Diese Organisationsform, bei der die Abhängigkeit von vorgegebenen Strukturen und das erzwungene Vertrauen in diese teilweise aufgehoben werden, findet beispielsweise auch in Form so genannter Peer-to-Peer-Netze Anklang – selbstorganisierenden Netzen, die meist das Internet als Bereitsteller von Verbindungen zwischen ihren Knoten nutzen, um darauf eine eigene Netzstruktur aufzubauen. Verfahren zum Aufbau und zur Organisation von Ad-hoc-Netzen könnten in ähnlicher Weise, indem zusätzlich zu mobilen Knoten und drahtlosen Verbindungen auch ortsfeste Knoten und Kabelverbindungen zugelassen werden, auch für größere Netze verwendet werden und in manchen Bereichen existierende Strukturen ersetzen.

Ad-hoc-Netze mit mobilen Knoten und drahtloser Übertragung benötigen eine gewisse Dichte potentieller Teilnehmer, um funktionieren zu können. Da bisher keine Standards für Ad-hoc-Netze existieren – von funktionsfähigen Produkten ganz zu schweigen –, sind derzeit eher überschaubare Szenarien von Interesse, bei denen eine Menge von potentiellen Teilnehmern in einer Situation zusammentrifft, in der Kommunikationsbedarf besteht; dort wäre die erforderliche Knotendichte relativ leicht zu erreichen. Beispielsweise kommen die folgenden Szenarien in Betracht:

- *Kooperatives Arbeiten*: Datenaustausch und kooperatives Arbeiten in kurzfristig entstehenden und kurzzeitig bestehenden Gruppen, beispielsweise zwischen Teilnehmern von Konferenzen oder in Teams wie bei Katastropheneinsätzen; gemeinsame Nutzung von Ressourcen wie z. B. eines Internet-Zugangs.

- *Zentral gesteuertes Arbeiten*: Übertragung präsentierter Inhalte oder anderer Begleitmaterialien vom Rechner eines Dozenten auf die von Teilnehmern; überwachtes Arbeiten mit beschränkter Eigenkontrolle, z. B. im Klassenzimmer.

- *Fahrzeug-Fahrzeug-Kommunikation*: Kommunikation zwischen sich bewegenden Fahrzeugen (z. B. auf der Autobahn, wo die Fahrzeugdichte gleichmäßig relativ hoch ist) zum Zweck des Informationsaustauschs (Verkehrsinformation, Fahrzeuggeschwindigkeit und Ähnliches, etwa zur Unterstützung bei der Vermeidung von Unfällen).

- *Zugriff auf ortsfeste oder mobile Infrastruktur*: Bequeme Verwendung öffentlich angebotener oder von anderen mobilen Teilnehmern zur Verfügung gestellter Infrastrukturkomponenten, etwa Druckern oder Internet-Zugängen an Flughäfen.

- *Messdatenübermittlung*: Übermittlung und verteilte Vorverarbeitung von Messdaten innerhalb von Netzwerken kleiner automatischer Sensoren (ohne Beteiligung menschlicher Benutzer).

Wenn die Menge der teilnehmenden Knoten von vornherein feststeht (etwa die Mitglieder eines bestimmten Teams), so kann man von einem *geschlossenen Ad-hoc-Netz* sprechen. Bei den *offenen Ad-hoc-Netzen* können dagegen neue, den anderen Teilnehmern vorher unbekannte Teilnehmer jederzeit hinzukommen.

2.4.2 Unterschiede zu Infrastrukturnetzen

Ergänzend zu den sich aus der Verwendung drahtloser Übertragungstechnik ergebenden Unterschieden, die bereits in Abschnitt 2.2.2 behandelt wurden, werden im Folgenden noch einige Punkte genannt, in denen sich Ad-hoc-Netze von Infrastrukturnetzen unterscheiden:

- *Hohe Dynamik der Netztopologie:* Während bei Infrastrukturnetzen wenigstens die Zugangspunkte ortsfest sind und Konnektivitätsverluste meist durch Bewegung desjenigen Knotens selbst verursacht sind, den sie auch als einzigen betreffen, kann bei Ad-hoc-Netzen das Abreißen den Funkkontakts zwischen zwei Knoten sowohl durch beide verursacht werden als auch viele andere Knoten beeinflussen; ungünstigstenfalls zerfällt das Netz in zwei einzelne. Ständige Änderungen der günstigsten Wege zwischen zwei Knoten stellen hohe Anforderungen an die Netzwerkschicht der Knoten, die normalerweise gleichzeitig End- und Zwischensysteme sind. Es sind spezielle Wegfindungsverfahren erforderlich, die sich von den in Infrastrukturnetzen verwendeten unterscheiden.

- *Beschränkte Ressourcen:* Die bereits früher genannten, durch die drahtlose Übertragungstechnik bzw. die Mobilität der Endgeräte verursachten Beschränkungen bezüglich Energievorrat, Rechenleistung und Speicherkapazität der Knoten sowie die beschränkte Übertragungsbandbreite und erhöhte Fehleranfälligkeit wirken sich bei Ad-hoc-Netzen evtl. noch gravierender aus als bei Infrastrukturnetzen, weil die Verfahren zur Unterhaltung des Netzes und der Dienste komplexer sind und höhere Ressourcenanforderungen haben und weil die Knoten voneinander und nicht nur von einer Verbindung zur Basisstation abhängen.

- *Einfache Konfiguration:* Für Netze, die spontan bei Entstehen des Bedarfs aufgebaut werden, sollte keine aufwändige Konfiguration erforderlich sein.

- *Rolle der Teilnehmer:* Alle Komponenten von Ad-hoc-Netzen sind in der Regel Eigentum der Teilnehmer. Die Teilnehmer sind damit nicht nur Nutzer, sondern auch Dienstleister, und haben ein gesteigertes Interesse am Schutz des Netzes vor Missbrauch.

- *Kooperation:* Da alle Dienste innerhalb von Ad-hoc-Netzen (inklusive der Weiterleitung von Nachrichten) allein von den Netzteilnehmern erbracht werden, können Ad-hoc-Netze nur dann funktionieren, wenn zumindest ein großer Teil der Teilnehmerschaft sich auch aktiv beteiligt und die eigenen Ressourcen zum Nutzen der Allgemeinheit einbringt.

- *Verzicht auf zentrale Komponenten:* Weil weder die ständige Erreichbarkeit bestimmter Knoten noch überhaupt deren Verfügbarkeit garantiert werden kann, dürfen zur Realisierung von Diensten keine zentralen Komponenten eingesetzt werden. Vielmehr müssen diese auf möglichst viele Knoten verteilt werden.

2.4.3 Netzwerkfunktionalität in Ad-hoc-Netzen

Ad-hoc-Netze unterscheiden sich in einigen wichtigen Eigenschaften von Infrastrukturnetzen, so dass dort etablierte Verfahren und Protokolle in Ad-hoc-Netzen nur eingeschränkt funktionieren. Anpassungen sind vor allem in der Netzwerkschicht erforderlich, wo durch die hohe Dynamik der Netztopologie besondere Herausforderungen entstehen. Ab der Transportschicht aufwärts müssen evtl. Anpassungen vorgenommen werden, um mit der größeren Paketverlustrate und Latenz bzw. häufigeren Verbindungsabbrüchen fertigzuwerden.

Im Unterschied zu Infrastrukturnetzen, wo Adressen immer von einer zentralen Instanz vergeben werden, müssen in Ad-hoc-Netzen spezielle verteilte Adresszuweisungsverfahren verwendet werden. Während IP-Adressen im Internet eine topologische Strukturierung aufweisen, die auch für die Skalierbarkeit unverzichtbar ist, kann eine solche Strukturierung in Ad-hoc-Netzen nur sehr schwierig umgesetzt werden. Es gibt zwar Adressvergabeverfahren, welche die zugewiesenen Adressen bei Bewegung des Knoten an die neue Umgebung anpassen [Perk01, Kap. 4], aber dies ist recht aufwändig und schafft außerdem neue Probleme durch ständig wechselnde Adressen.

Wesentliche Änderungen sind bei den Wegfindungsverfahren erforderlich. Die in Abschnitt 2.1.4 genannten Verfahren sind aus mehreren Gründen für Ad-hoc-Netze schlecht geeignet:

- Jeder Knoten ist Router; dadurch wächst die zu verwaltende Informationsmenge.
- Topologieänderungen erfordern die Versendung von Update-Nachrichten, d. h. je höher die Mobilität, desto mehr Bandbreite geht dafür verloren. Versucht man, Update-Nachrichten zu vermeiden, indem man differenziertere Zustandinformation in den Knoten speichert, so besteht die Gefahr, dass diese schnell veraltet.
- Schleifen in den gefundenen Wegen, die bei hoher Mobilität leichter auftreten würden, wirken sich in Ad-hoc-Netzen wegen der geringeren verfügbaren Bandbreite gravierender aus.
- Asymmetrische Verbindungen werden z. B. von Distanz-Vektor-Verfahren, die ihre Weginformation „rückwärts“ aufbauen, nicht verkraftet.

Es gibt eine beträchtliche Menge von Entwürfen für Wegfindungsprotokolle für Ad-hoc-Netze. Sie lassen sich grob in zwei Kategorien unterteilen: *Proaktive* oder *tabellengesteuerte* Verfahren tauschen ständig Informationen aus, um stets ein aktuelles Bild von der Topologie des Netzes zu haben. *Reaktive* oder *anforderungsgesteuerte* Verfahren versuchen erst dann, einen Weg zu einem bestimmten Ziel zu finden, wenn tatsächlich Bedarf besteht. Proaktive Verfahren haben den Vorteil, Wege wesentlich schneller liefern zu können als reaktive. Dafür verschwenden sie einen unter Umständen beträchtlichen Anteil der verfügbaren Bandbreite, um Wege zu ermitteln, die möglicherweise gar nicht gebraucht werden. Es gibt neben den beiden Extremen auch hybride Verfahren, die versuchen, Vorteile beider Seiten zu nutzen.

Exemplarisch soll im Folgenden ein einfaches reaktives Wegfindungsverfahren für Ad-hoc-Netze kurz vorgestellt werden: das *Dynamic Source Routing* (DSR). DSR ist, wie der Name auch besagt, ein so genannten Source-Routing-Verfahren, d. h. jedes Paket erhält bereits von der Quelle den gesamten Weg einbeschrieben, auf dem es zum Ziel gelangen soll.

Zur Wegfindung senden Knoten, die mit einem bisher unbekannten Ziel kommunizieren möchten, eine Wegfindungsanfrage (Route Request), die einen Weg in Form einer (anfänglich leeren) Liste passierter Knoten enthält, per Broadcast aus. Alle Knoten, die diese Nachricht empfangen, fügen sich selbst zum Weg hinzu und senden das Paket wieder per Broadcast aus, wenn sie dieselbe Nachricht

nicht kurz zuvor schon erhalten hatten. Wenn das Ziel die Nachricht erhält, sendet es sie ebenfalls nicht weiter, sondern antwortet stattdessen mit einer Nachricht (Route Reply), welche den aus der Anfrage übernommenen Weg enthält. Diese Antwortnachricht kann entweder auf dem umgekehrten ermittelten Weg zurückgesendet werden, oder durch eine eigene Wegfindungsanfrage. Die Quelle erhält möglicherweise mehrere Antworten und kann sich dann den kürzesten oder schnellsten Weg aussuchen und in Zukunft benutzen.

Falls auf einem bestehenden Weg die Weiterleitung fehlschlägt, weil ein nächster Knoten nicht mehr erreichbar ist, erfolgt eine Fehlermeldung an die Quelle, die dann versuchen kann, die fehlerhafte Stelle zu umgehen oder eine neue Wegfindungsanfrage absendet.

2.4.4 Netzwerksicherheit in Ad-hoc-Netzen

Die besonderen Eigenschaften von Ad-hoc-Netzen im Unterschied zu Infrastrukturnetzen, die in Abschnitt 2.4.2 allgemein angesprochen wurden, haben teilweise auch veränderte Anforderungen und Randbedingungen für Sicherheitsmechanismen zur Folge:

- *Physikalischer Zugang:* Der physikalische Zugang zum Netz ist bei drahtloser Kommunikationstechnik für jeden in einer gewissen räumlichen Umgebung möglich. Sowohl passives Abhören als auch das Einspielen zusätzlicher Nachrichten sind damit für potentielle Angreifer sehr einfach machbar. In gewisser Hinsicht kann die Möglichkeit der Beobachtung des Verhaltens anderer Teilnehmer in Ad-hoc-Netzen allerdings andererseits auch als Vorteil gewertet werden, da so die Kooperativität ansonsten völlig unbekannter Teilnehmer eingeschätzt und missbräuchliche Nutzung leichter erkannt werden kann.

- *Verzicht auf zentrale Komponenten:* Verfahren, die auf ständig verfügbare Infrastrukturkomponenten angewiesen sind, können nicht eingesetzt werden. Dies betrifft insbesondere die Schlüsselverwaltung, bei der in Infrastrukturnetzen häufig zentrale Komponenten zum Einsatz kommen, etwa in Form von Zertifizierungsstellen und Zertifikatverzeichnissen (bei Public-Key-Infrastrukturen) oder Authentisierungs-Servern (bei Kerberos). Da Sicherheitsdienste ständig verfügbar sein müssen, müssen sie in allen Teilen vollständig verteilt realisiert werden.

- *Angreifbarkeit der Infrastruktur:* Da die „Infrastruktur“ von Ad-hoc-Netzen auf alle Knoten verteilt ist und nicht wie in Infrastrukturnetzen durch dedizierte und besonders geschützte Geräte gebildet wird, ist sie wesentlich leichter angreifbar. Es muss deshalb mehr Aufwand darauf verwendet werden, grundlegende Mechanismen wie das Routing vor Angriffen zu schützen (siehe z. B. [HuPJ02]), um die Verfügbarkeit des Netzwerks gewährleisten zu können.

- *Keine gesicherten Teilnetze:* Es existiert keine natürliche „Verteidigungslinie“, wie sie beispielsweise in Infrastrukturnetzen am häufig durch Firewalls geschützten Übergang von öffentlichen in private Netzteile mit erhöhter Sicherheit zu finden ist. Jeder Knoten ist „allein“ und von einem entsprechend platzierten Angreifer direkt erreichbar. Angreifern auf die Systemsicherheit einzelner Netzknoten stellen sich damit keine Hindernisse außer den durch jeden Knoten selbst getroffenen Maßnahmen in den Weg. Auf einem System abgelegte sensitive Informationen (wie z. B. Schlüssel) sind deshalb im Allgemeinen weniger sicher als in Infrastrukturnetzen. Sofern Sicherungsmaßnahmen von der Unversehrtheit der Systemsicherheit anderer Knoten abhängen, muss beachtet werden, dass deren Kompromittierung wahrscheinlicher ist, als wenn sie durch Maßnahmen im Netz zusätzlich gesichert wären.

- *Einfluss der Kooperation:* Jeder einzelne Knoten hat Aufgaben für die Gemeinschaft zu erfüllen. Eine faire Verteilung von Rechten (Ressourcennutzung) und Pflichten (Diensterbringung) ist erforderlich; dabei ist Missbrauch zu erwarten und muss möglichst verhindert werden. Durch die Ressourcenknappheit entsteht bei den einzelnen Knoten die für das Netz insgesamt ungünstige Motivation, möglichst wenig Leistungen für andere Knoten zu erbringen, um die vorhandene Energie für eigene Zwecke aufzusparen.

- *Angriffe auf begrenzte Ressourcen:* Beschränkte Ressourcen wie Bandbreite, Rechenleistung und Energievorrat ermöglichen Angriffe auf die Verfügbarkeit von Knoten oder des Netzes, indem die Ressourcen gezielt belegt bzw. verbraucht werden.

- *Ressourcenanforderungen kryptographischer Verfahren:* Die besonders bei Kleinstgeräten beschränkte Rechenleistung kann den Einsatz kryptographischer Verfahren einschränken, so dass z. B. keine asymmetrischen kryptographischen Verfahren verwendet werden können. Geringe verfügbare Bandbreite und beschränkter Energievorrat stehen im Konflikt damit, dass die Sicherung vorhandener Protokolle und die Einführung von Diensten z. B. zur Schlüsselverwaltung eine höhere Netzbelastung verursachen.

- *Offenheit erschwert Authentisierung:* In offenen Ad-hoc-Netzen liegen den Knoten im Allgemeinen zunächst keinerlei Informationen über andere Knoten vor, insbesondere weder über deren Identität noch über deren Vertrauenswürdigkeit. Es gibt keine zentrale Stelle, die z. B. öffentliche Schlüssel zur Verfügung stellen könnte. Damit ist eine Authentisierung auf technischer Ebene zunächst nicht möglich.

- *Autonomie der Knoten:* In offenen Ad-hoc-Netzen stellt in der Regel jeder Knoten eine eigene administrative Domäne dar. Schon deshalb wäre es schwierig, für alle Teilnehmer akzeptable vertrauenswürdige zentrale Instanzen zu etablieren, auch wenn sich dies nicht schon aus anderen Gründen verbieten würde. Außerdem bedeutet es, dass Sicherheitsrichtlinien zunächst nur für jeden Knoten selbst bestimmt werden können; verbindliche gemeinsame Regelungen zu vereinbaren ist sehr aufwändig.

Die grundlegendste Schwierigkeit bei der Anwendung von Sicherheitsmechanismen in offenen Ad-hoc-Netzen ist sicherlich das Fehlen der Möglichkeit zur Authentisierung aufgrund nicht vorhandener und auch nicht mit herkömmlichen Schlüsselverwaltungsverfahren beschaffbarer Geheimnisse bzw. öffentlicher Schlüssel. In Abschnitt 2.4.5 werden einige existierende Ansätze zur Lösung dieses Problems beschrieben. Auf das im Rahmen der vorliegenden Arbeit entworfene Verfahren wird in Kapitel 3 eingegangen.

Für die Zusammenarbeit wird in der Regel außerdem eine Sicherung zweiseitiger oder gruppenartiger Kommunikationsbeziehungen bezüglich der Integrität und/oder Vertraulichkeit der übertragenen Daten gefordert. Für eine solche Sicherung von Nutzkanälen können bekannte und in Infrastrukturnetzen bereits bewährte Verfahren (z. B. IPSec, SSL) eingesetzt werden, sobald sichere Identifizierung und Authentisierung möglich sind. Die Verfahren sind meist nicht speziell auf die Eigenschaften drahtloser Übertragung abgestimmt; hier sind Optimierungen denkbar, aber nicht zwingend erforderlich [GuGu01].

2.4.5 Vertrauen und Authentizität in Ad-hoc-Netzen

Da die Schwierigkeiten bei Authentisierung und Schlüsselverwaltung durch die mangelnde Verfügbarkeit zentraler Dienste wie Zertifizierungsinstanzen und Verzeichnisdiensten verursacht werden, liegt es nahe zu versuchen, solche Dienste verteilt zu realisieren.

Die bisher entwickelten Lösungsansätze lassen sich in zwei Kategorien unterteilen. Diejenigen der ersten Kategorie basieren auf einer logisch gesehen zentralen, aber mit Hilfe von Schwellwertkryptographie verteilt realisierten Zertifizierungsinstanz, an der entweder eine ausgewählte Teilmenge aller Knoten (Abschnitt 2.4.5.1) oder sogar die ganze Menge aller Knoten beteiligt ist (Abschnitt 2.4.5.2). Zertifikate können dann jeweils von einem Teil (z. B. einer Mehrheit) aller an der Zertifizierungsinstanz beteiligten Knoten gemeinsam erstellt werden, wodurch sowohl gegen Ausfall als auch gegen Kompromittierung einzelner Knoten Vorsorge getroffen ist. Der öffentliche Schlüssel der Zertifizierungsinstanz wird im ganzen Netz bekannt gemacht.

Die Ansätze der zweiten Kategorie verzichten zugunsten der in Abschnitt 2.3.4.3 beschriebenen verteilten Variante auf zentrale Zertifizierung: Jeder Knoten wird als eigenständige Zertifizierungsinstanz gesehen, welche für andere Knoten Zertifikate ausstellen kann. Solche Zertifikate werden von Knoten akzeptiert, die dem Aussteller (evtl. indirekt) vertrauen und seinen Schlüssel authentisieren können (Abschnitt 2.4.5.3).

Die Aufbewahrung und bedarfsgerechte Verteilung von Zertifikaten wird häufig nicht behandelt – auch hierfür wird eine verteilte Lösung benötigt. Bei dem in Abschnitt 2.4.5.3 beschriebenen Ansatz mit verteilter Zertifizierung speichert jeder Knoten eine bestimmte Auswahl an Zertifikaten.

2.4.5.1 Verteilung der Zertifizierungsinstanz über ausgewählte Netzknoten

Zhou und Haas schlagen ein Schlüsselverwaltungsverfahren vor [ZhHa99], bei dem eine logisch gesehen zentrale Zertifizierungsinstanz über eine bestimmte Menge von Knoten verteilt wird. Dazu werden Verfahren der so genannten Schwellwertkryptographie eingesetzt, die es erlauben, den geheimen Schlüssel eines asymmetrischen Schlüsselpaars in n Teile so aufzuspalten, dass mindestens k Teile ($k \leq n$) benötigt werden, um den Schlüssel einsetzen zu können.

Die Teile des geheimen Schlüssels werden also auf n verschiedene Knoten verteilt, während der öffentliche Schlüssel im ganzen Netz bekannt gemacht wird. Zur Ausstellung von Zertifikaten arbeiten jeweils k der Zertifizierungsinstanz-Knoten zusammen. Die Verfügbarkeit der Zertifizierungsinstanz ist damit erheblich besser als bei einer zentral realisierten Lösung. Außerdem kann das System auch die Kompromittierung von bis zu $k-1$ Knoten der Zertifizierungsinstanz verkraften, ohne unsicher zu werden.

Um auch gegen mobile Angreifer schützen zu können, die versuchen, nach und nach verschiedene Knoten zu kompromittieren und deren Schlüsselteile zu erhalten, wird die Aufteilung des geheimen Schlüssels regelmäßig erneuert, d. h. die verteilten Schlüsselteile werden durch neue ersetzt, ohne dass der geheime Schlüssel dabei geändert oder aufgedeckt wird. Alte Schlüsselteile werden damit wertlos.

Ein Nachteil des Verfahrens ist, dass die Träger der Schlüsselteile ausgewählt und initialisiert werden müssen, was nicht in vollständig verteilter Weise erfolgen kann.

2.4.5.2 Verteilung der Zertifizierungsinstanz über alle Netzknoten: URSA

Das URSA-System [LZKL+02, LuLu00, LKZL+04] stellt ein vollständiges Zugangskontrollsystem für Ad-hoc-Netze dar, welches als notwendige Grundlage auch die Authentisierung von Knoten ermöglicht, indem es einen logisch gesehen zentralen, aber auf alle Knoten des Netzes verteilt realisierten Schlüsselverwaltungsdienst anbietet. Authentisierung und Zugangskontrolle sind eng miteinander verknüpft: Nur authentisierte Knoten erhalten Zugang zum Netz. Der Ansatz wird in Abschnitt 2.4.6.4 zusammen mit anderen Zugangskontrollverfahren noch näher beschrieben.

2.4.5.3 Verteilte Zertifizierung nach Hubaux et al.

Im Rahmen eines längerfristigen Projekts, bei dem es um die Entwicklung von Mechanismen für große (bis hin zu weltumspannenden) offene Ad-hoc-Netze geht [BBČG+01], wurde von Hubaux et al. unter Anderem ein Schlüsselverwaltungsverfahren nach dem Prinzip der verteilten Zertifizierung entwickelt [HuBČ01].

In der Beschreibung des Verfahrens wird nicht zwischen den die Netzknoten bildenden Geräten und ihren menschlichen Benutzern unterschieden, sondern es wird nur von „Benutzern“ gesprochen.

Jeder Benutzer besitzt ein asymmetrisches Schlüsselpaar und kann Zertifikate für andere Benutzer ausstellen, wenn er von der Korrektheit der Zuordnung zwischen Schlüssels und dessen Benutzer überzeugt ist. Dabei wird zunächst davon ausgegangen, dass alle Benutzer diese Zertifizierungsfunktion gewissenhaft ausführen, also in dieser Hinsicht absolut vertrauenswürdig sind. Der (virtuelle) gerichtete Graph, dessen Knoten alle Benutzer und dessen Kanten alle erzeugten Zertifikate repräsentieren, wird als Vertrauensgraph des Netzes bezeichnet.

Die Aufbewahrung erzeugter Zertifikate erfolgt ebenfalls dezentral, indem jeder Benutzer eine begrenzte Anzahl von Zertifikaten lokal speichert. Einerseits sind dies alle Zertifikate, die er selbst ausgestellt hat; damit ist gewährleistet, dass Zertifikate nicht verloren gehen können. Außerdem werden nach einem für alle Benutzer gleichen Algorithmus weitere Zertifikate zur Aufbewahrung ausgewählt.

Wenn ein Benutzer einen anderen authentisieren möchte, sucht er in der Vereinigungsmenge der von beiden Benutzern gespeicherten Zertifikate einen Zertifizierungspfad, also einen gerichteten Weg in dem durch die Vereinigungsmenge lokal gespeicherter Zertifikate induzierten Untergraphen des gesamten Vertrauensgraphen. Ist die Suche erfolgreich, so kann die Authentisierung durchgeführt werden. Ansonsten ist die Authentisierung fehlgeschlagen, da keine Möglichkeit vorgesehen ist, andere Zertifikate zu finden, auch wenn diese möglicherweise existieren. Die These der Autoren ist, dass auch bei einer im Vergleich zur Gesamtzahl der Benutzer des Netzes geringen Zahl von jedem Benutzer gespeicherter Zertifikate mit hoher Wahrscheinlichkeit Pfade zwischen beliebigen Benutzern existieren. Auf diese Weise soll der Ansatz für große Ad-hoc-Netze skalieren.

Von großem Einfluss auf die genannte Wahrscheinlichkeit ist der Algorithmus, anhand dessen die lokal zu speichernden Zertifikate ausgewählt werden. Um verschiedene Algorithmen vergleichen zu können, wird als Maß für deren Leistungsfähigkeit berechnet, für welchen Anteil der möglichen Benutzerpaarungen, für die im gesamten Vertrauensgraphen ein Zertifizierungspfad existiert, dies auch schon in der Vereinigungsmenge der lokalen, durch den jeweiligen Algorithmus erzeugten Zertifikatspeicher der beiden Benutzer der Fall ist. Weitere Charakteristika von Algorithmen sind die Größe der erzeugten Menge und die Menge und Art des zur Ausführung erforderlichen Wissens. Zu vermeiden ist, dass etwa alle Teilmengen einen bestimmten Knoten enthalten, was zwar die Leistungsfähigkeit steigert, dafür aber die Sicherheit schmälert.

Zwei konkrete Algorithmen werden vorgestellt, und ihre Leistungsfähigkeit wird anhand dreier verschieden großer Untergraphen eines Vertrauensgraphen gemessen, der sich aus den in öffentlichen Verzeichnissen gespeicherten PGP-Zertifikaten ergibt. Für diese Graphen (der Größen 2124, 3211 und 8695 Knoten) zeigt sich, dass der zweite Algorithmus, der eine Verbesserung des ersten darstellt, eine Leistungsfähigkeit von etwa 95% erreicht, wenn die Anzahl der selektierten Zertifikate etwa zweimal der Quadratwurzel aus der Gesamtknotenzahl entspricht.

Für den Fall, dass der Zertifizierung nicht unbedingt vertraut werden kann, wird eine modifizierte Leistungsfähigkeitsmetrik vorgeschlagen, die ein Maß für die über die betrachteten Pfade erreichbare Authentizität in die Bewertung einbezieht. Die Wirksamkeit dieser Maßnahme wird allerdings nicht evaluiert.

2.4.6 Zugangskontrolle in Ad-hoc-Netzen

Alle Komponenten offener Ad-hoc-Netze sind in der Regel Eigentum der Teilnehmer. Bei diesen besteht deshalb berechtigtes Interesse daran, das Netz vor unerwünschter Benutzung durch Individuen zu schützen, die zwar Netzdienste nutzen möchten, aber zwecks Schonung der eigenen Ressourcen nicht zu deren Erbringung beitragen wollen. Analog zu Infrastrukturnetzen, wo der Betreiber der Infrastruktur in der Regel die Benutzung und unter Umständen Beeinträchtigung durch nicht zahlende Nutznießer verhindern möchte, ist deshalb auch in Ad-hoc-Netzen die Durchführung einer Zugangskontrolle sinnvoll.

Man kann den erwünschten Effekt übrigens auch von anderen Standpunkten aus betrachten, deshalb sind die folgenden Bezeichnungen nahezu äquivalent, auch wenn die jeweilige Motivation etwas unterschiedlich klingt:

- *Zugangskontrolle:* Ähnlich der Vergabe von Zugriffsrechten bei herkömmlichen Zugangskontrollverfahren (Abschnitt 2.3.6) versucht man, Kriterien zu finden, anhand derer für die Knoten von Ad-hoc-Netzen entschieden werden kann, ob bzw. in welchem Umfang sie Zugang erhalten sollen. Es erscheint sinnvoll, diese Entscheidung vom Wohlverhalten (kein Missbrauch, faire Beteiligung an gemeinsam erbrachten Diensten) der Knotens abhängig zu machen.

- *Förderung der Kooperation:* Da alle Dienste innerhalb von Ad-hoc-Netzen allein von den Netzteilnehmern erbracht werden, können Ad-hoc-Netze nur dann funktionieren, wenn zumindest ein großer Teil der Teilnehmerschaft sich auch aktiv beteiligt und die eigenen Ressourcen zum Nutzen der Allgemeinheit einbringt. Es müssen Verfahren gefunden werden, um die Knoten zur Zusammenarbeit zu motivieren.

- *Schutz der Verfügbarkeit:* Die Verfügbarkeit der von den Netzknoten gemeinsam erbrachten Dienste hängt vom korrekten Verhalten jedes Knotens ab. Das Verhalten der Knoten muss überwacht werden, um Missbrauch feststellen zu können und die betreffenden Knoten gegebenenfalls aus dem Netz auszuschließen

Bisherige Arbeiten auf diesem Gebiet verfolgen im Wesentlichen die beiden folgenden verschiedenen Ansätze, um Knoten zur Zusammenarbeit zu motivieren:

- *Detektionsbasierter Ansatz*

  Der erste Ansatz basiert auf der gegenseitigen Überwachung benachbarter Knoten. Wird falsches Verhalten erkannt, so verweigern die Nachbarn die weitere Zusammenarbeit und benachrichtigen bei manchen Ansätzen außerdem andere Knoten.

  Nachteilig an solchen Verfahren ist, dass der Ausschluss aus dem Netz eine recht drastische Maßnahme ist in Anbetracht dessen, dass Fehlverhalten nicht immer zuverlässig erkannt werden kann: Die Beobachtung wird durch die Mobilität der Knoten erschwert, und unbeabsichtigte Fehler kommen aufgrund der Eigenschaften drahtloser Übertragungstechnik relativ häufig vor. Unter Umständen kann der Anschein von Fehlverhalten absichtlich durch Dritte erzeugt werden. Unklar ist außerdem, ob z. B. Dienstverweigerung wegen Energiemangel Fehlverhalten ist. Die Prüfung der Glaubwürdigkeit von Berichten über Fehlverhalten anderer Knoten ist schwierig. Außerdem kann durch einen Wechsel der eigenen Identität eine Geschichte eigenen Fehlverhaltens vergessen gemacht werden, wenn dies nicht durch geeignete Maßnahmen verhindert wird.

Drei detektionsbasierte Ansätze werden in den folgenden Abschnitten etwas näher beschrieben:

- Missbrauchserkennung nach Paul und Westhoff (Abschnitt 2.4.6.1: Manipulationen während der Wegfindungsphase des DSR-Protokolls werden anhand von in den Nachrichten angebrachter Hash-Werte erkannt.
- Das CONFIDANT-System (Abschnitt 2.4.6.3): Fehlverhalten wird registriert und führt einerseits zur Dienstverweigerung gegenüber dem Angreifer und andererseits zu Alarmmeldungen, die an vorkonfigurierte „Freunde“ verschickt werden und dort ebenfalls zur Dienstverweigerung führen.
- Das URSA-System (Abschnitt 2.4.6.4): Zugangskontrolle erfolgt über signierte Tickets mit relativ kurzer Gültigkeitsdauer, die immer von einer Nachbarschaft eines Zugang suchenden Knotens ausgestellt werden. Wenn Fehlverhalten festgestellt wird, werden für den betreffenden Knoten keine Tickets mehr erstellt.

- *Motivationsbasierter Ansatz*

  Beim zweiten Ansatz werden Knoten zur Zusammenarbeit motiviert, indem sie für Leistungen, die sie zum Gemeinwohl erbracht haben, bezahlt werden, beispielsweise in einer virtuellen Währung [LaPW03]; in dieser Währung werden in Anspruch genommene Leistungen dann wiederum bezahlt. Die Beträge werden entweder in den weitergeleiteten Paketen (sozusagen in bar) mitgeführt, oder zwischen virtuellen Bankkonten transferiert, nachdem der Bank die erfolgten Dienstleistungen und -nutzungen mitgeteilt wurden. Eine solche Bank muss das Vertrauen aller Knoten genießen und fast immer verfügbar sein; bei einem Ansatz liegt sie deshalb in einem Infrastrukturnetz, das zumindest intermittierend erreichbar sein muss.

  Ein Problem motivationsbasierter Ansätze ist, dass Knoten am Rand des Netzes nicht in der Lage sind, Dienste für die Allgemeinheit zu erbringen (zumindest, wenn – wie bisher üblich – nur die Paketweiterleitung betrachtet wird); hier wird auf Ausgleich durch Mobilität gehofft. Außerdem problematisch ist die korrekte Abrechnung, die hochverfügbare Banken oder manipulationsgesicherte Einheiten in den Knoten erfordert.

  Unter den nachfolgend beschriebenen Ansätzen ist einer motivationsbasiert, nämlich das „Nuglet“-Verfahren nach Hubaux et al. (Abschnitt 2.4.6.2): Als „Bezahlung“ für das Weiterleiten von Paketen werden „Nuglets“ eingesetzt; diese werden entweder jedem Paket vom Absender mitgegeben und von den Weiterleitenden entnommen, oder sie dienen als Währung für den jeweils gewinnbringenden „Weiterverkauf“ von Paketen an den nächsten Knoten auf dem Übertragungsweg.

2.4.6.1 Missbrauchserkennung nach Paul und Westhoff

Der detektionsbasierte Ansatz von Paul und Westhoff [PaWe02] ist stark auf den Schutz des Wegfindungsverfahrens „Dynamic Source Routing“ (DSR) spezialisiert. Bei dem Ansatz wird vorausgesetzt, dass jeder Knoten mit jedem Kommunikationspartner a priori eine „geheime Zufallszahl“ (also einen Schlüssel, auch wenn er nicht als solcher bezeichnet wird) vereinbart hat.

Die Weganfragenachricht des DSR-Protokolls, in der während ihres Transports zum Zielknoten der durchlaufene Weg gespeichert wird, wird erweitert: Der erste Knoten speichert einen Hash-Wert[2] über die Identitäten seiner selbst und des Zielknotens sowie über die geheime Zufallszahl im Paket. Jeder

[2]Es wird ausdrücklich von einem un-keyed Hash gesprochen, aber durch die Einbeziehung der geheimen Zufallszahl entspricht das Ergebnis im ersten Schritt einem schlüsselbasierten Hash-Wert, also einem Message Authentication Code.

Knoten bildet bei der Weitergabe (per Broadcast) des Pakets wieder einen Hash über seine Identität und den vorhanden Hashwert und ersetzt letzteren durch das Ergebnis. Der Empfänger kann anhand der Liste der durchlaufenen Knoten die Hash-Kette nachvollziehen; daraus kann er erkennen, dass kein Knoten die Einträge seines Vorgängers manipuliert hat.

Weitere Angriffe sind allerdings denkbar: Ein Knoten kann es unterlassen, seine eigene Identität ins Paket zu schreiben, oder er kann beliebige andere Knoten direkt vor oder hinter seiner eigenen Identität einfügen. Solche Angriffe können von Nachbarn des Angreifers entdeckt werden, die dazu teilweise schon gesehene Anfragen aufbewahren müssen. Wird ein Angreifer entdeckt, so wird eine spezielle Nachricht an den Absender der Anfrage gesendet (entlang des im bereits früher gesehenen Paket gespeicherten Wegs oder – falls nicht vorhanden – per Broadcast oder mittels einer eigenen Weganfragenachricht). Es werden Regeln angegeben, nach denen bestimmt wird, ob die über einen Angreifer informierte Quelle einer solchen Anschuldigung glaubt, oder ob die Beschuldigung selbst als Angriff bewertet wird. Problematisch an diesem Vorgehen erscheint, dass Beobachter davon abgehalten werden könnten, Anschuldigungen auszusprechen, aus Angst, selbst beschuldigt zu werden.

Für die Antwort auf die Weganfragenachricht sowie für die Fehlernachricht im Fall des Verschwindens eines Wegs werden ebenfalls Erkennungsmöglichkeiten angegeben.

2.4.6.2 Das „Nuglet“-Verfahren nach Hubaux et al.

Im Rahmen desselben Projekts, aus dem der oben (Abschnitt 2.4.5.3) beschriebene Ansatz zur verteilten Zertifizierung stammte, wurde auch ein motivationsbasiertes Zugangskontrollverfahren [HuBu00, BBČG$^+$01] entwickelt. Bei diesem Verfahren werden so genannte „Nuglets“ eingesetzt, eine virtuelle Währung, in der für die Dienstleistung der Weiterleitung von Paketen bezahlt wird. Es werden zwei verschiedene Modelle vorgeschlagen: Das „Geldbörsen-Modell“ (Packet Purse Model) und das „Pakethandelsmodell“ (Packet Trade Model).

Beim Geldbörsen-Modell stattet der Erzeuger eines Pakets dieses mit einer Anzahl von Nuglets aus seinem Vorrat aus. Jeder weiterleitende Knoten entnimmt eine feste oder eine aufwandsabhängige Anzahl von Nuglets aus dem Paket und fügt sie seinem Vorrat hinzu. Falls nicht mehr genügend Nuglets im Paket enthalten sind, wird es verworfen. Überzählige Nuglets verfallen bei Ankunft des Pakets oder können zur Bezahlung anderer Dienstleistungen im Zielknoten verwendet werden. Dass die notwendige Anzahl an Nuglets also vom Absender möglichst genau eingeschätzt werden muss, stellt einen Nachteil dieses Modells dar, insbesondere weil kein geeignetes Verfahren angegeben wird. Ein Vorteil ist, dass das Senden von Paketen bezahlt werden muss, was zu sparsamem Umgang mit der Ressource Bandbreite motiviert.

Beim Pakethandelsmodell tragen Pakete keine Nuglets mit sich, sondern werden selbst bei der Weiterleitung weiterverkauft: Jeder weiterleitende Knoten kauft ein Paket von seinem Vorgänger und verkauft es etwas teurer an seinen Nachfolger. Auf diese Weise braucht der Absender eines Pakets den Nuglet-Bedarf nicht im Voraus zu berechnen. Andererseits ist aber das Senden für ihn damit kostenlos oder sogar gewinnträchtig, weshalb er im Gegensatz zum Packet-Purse-Modell nicht direkt davon abgehalten wird, das Netz unnötig zu belasten. Allerdings kann eine Überlastung von seinen Nachbarn beschränkt werden, indem diese ihm Pakete nicht in beliebiger Rate abkaufen. Diese erweiterte Möglichkeit, den Kauf von Paketen abzulehnen, wurde von den Autoren noch nicht genauer untersucht.

Beide Modelle motivieren dazu, möglichst lange verfügbar zu sein und Pakete weiterzuleiten, da dies die einzige Möglichkeit ist, den eigenen Nuglet-Vorrat aufzufüllen.

Damit die Knoten ihren Nuglet-Vorrat nicht manipulieren oder bei der Weiterleitung mehr Nuglets entnehmen können als ihnen zustehen, werden die entsprechenden Funktionen in einem manipulationsgeschützten Sicherheitsmodul, beispielsweise auf einer Chipkarte, die nur von vertrauenswürdigen Hersteller produziert wird, realisiert. Diese enthalten bereits bei der Auslieferung ein gewisses „Startkapital“.

Die Sicherheitsmodule werden als richtlinientreue Instanzen in einer potentiell feindlichen Umgebung aufgefasst. So vereinbaren sie z. B. mit ihren jeweiligen Nachbarn unter Rückgriff auf eine vorausgesetzte Public-Key-Infrastruktur Geheimnisse und Nachrichtenzähler, unter deren Verwendung sie ihre Kommunikation vor dem Zugriff durch andere, möglicherweise in betrügerischer Absicht manipulierte Teile der Knotenfunktionalität schützen. Die Weiterleitung eines jeden Pakets wird dem Vorgänger-Sicherheitsmodul bestätigt, und eine ausbleibende Bestätigung führt zur Erhöhung eines Missbrauchszählers, der durch das Bezahlen einer Strafgebühr wieder gelöscht werden kann. Bei Überschreiten einer gewissen Grenze werden keine Pakete mehr an den entsprechenden Knoten weitergeleitet.

Das Protokoll zur Implementierung des Verfahrens ist zwischen MAC- und Netzwerkschicht angesiedelt. Hier wird vom Sicherheitsmodul ein besonderer, durch einen Message Authentication Code geschützter Paketkopf (Länge ca. 80 byte) zwischen MAC- und Netzwerk-Header eingefügt, der unter Anderem die Paket-Geldbörse bzw. den Preis-Eintrag enthält und bei der Weiterleitung in jedem Knoten gelesen und ersetzt wird. Er beinhaltet auch einen Hashwert über den Inhalt des Pakets, so dass er nicht missbräuchlich mit anderen Paketen verwendet werden kann. Simulationen ergaben, dass durch die Verwendung der Nuglets kaum Auswirkungen auf den Durchsatz eines Netzes zu spüren waren.

2.4.6.3 Das CONFIDANT-System nach Buchegger und Le Boudec

Das CONFIDANT-System (Cooperation of Nodes: Fairness in Dynamic Ad-hoc Networks)[BuLB02] stellt einen detektionsbasierten Ansatz zum Schutz reaktiver Source-Routing-Protokolle dar, wobei konkret auf das DSR-Protokoll Bezug genommen wird.

Das System registriert Fehlverhalten anderer Knoten, welches einerseits zur Dienstverweigerung gegenüber dem Angreifer und andererseits zu Alarmmeldungen führt, die an vorkonfigurierte „Freunde“ verschickt werden und dort ebenfalls Dienstverweigerung veranlassen. Das System besteht aus den folgenden Komponenten, die auf jedem Knoten vorhanden sind:

- Die *Monitor*-Komponente dient dazu, das Verhalten des jeweils nächsten Knoten auf den durch den eigenen Knoten führenden Wegen zu überwachen und Anomalien bei der Weiterleitung oder der Bearbeitung von Wegfindungsnachrichten zu erkennen; zunächst wurde dabei nur das Verwerfen von Nachrichten betrachtet. Wenn Fehlverhalten beobachtet wird, wird das Reputationssystem (siehe unten) informiert.

- Der *Vertrauensverwalter* ist für die Erzeugung und Verarbeitung von ALARM-Nachrichten zuständig. Solche Nachrichten werden verschickt, wenn anhand eigener Erfahrung, durch Beobachtung oder durch Nachricht von Dritten festgestellt wurde, dass sich ein Knoten böswillig verhält. Die Empfänger der ALARM-Nachrichten sind so genannte *Freunde*, die fest vorkonfiguriert werden können. Bei Empfang von ALARM-Nachrichten wird anhand einer Tabelle[3] überprüft, ob die Quelle der jeweiligen Nachricht als vertrauenswürdig gilt. Ist dies der Fall, wird das Reputationssystem (siehe unten) informiert; ansonsten wird die Nachricht ignoriert.

[3] Ob diese Tabelle mit der des Reputationssystems übereinstimmt oder vom Benutzer definiert werden muss, geht aus [BuLB02] nicht klar hervor.

- Das *Reputationssystem* pflegt eine Tabelle, die alle bekannten Knoten und deren Einstufungen enthält. Einstufungen werden nur geändert, wenn ausreichende Beweise über böswilliges Verhalten vorliegen und die Anzahl entsprechender Vorkommnisse einen gewissen Schwellwert übersteigt. Die sich ergebenden Änderungen von Einstufungen sind am größten für eigene Erfahrungen mit den betreffenden Knoten, etwas geringer für eigene Beobachtungen und noch geringer für Mitteilungen anderer. Falls die Einstufung eines Knotens durch eine Änderung aus dem tolerierbaren Rahmen fällt, werden der Wegeverwalter (siehe unten) und der Vertrauensverwalter informiert.

- Der *Wegeverwalter* bewertet alle dem Knoten bekannten Wege anhand der Einstufungen der auf den Wegen liegenden Knoten. Wege, die als böswillig erkannte Knoten enthalten, werden gelöscht. Wegeanfragen von böswilligen Knoten werden ignoriert, und Wegeanfragen, in deren akkumulierten Wegen sich bereits böswillige Knoten befinden, werden nicht weitergeleitet, sondern stattdessen werden die Quellen der Anfragen informiert.

Bei Simulationen konnte unter Einsatz des CONFIDANT-Systems ein Anteil von bis zu 60% böswilliger (d. h. weiterzuleitende Pakete verwerfender) Knoten innerhalb eines Ad-hoc-Netzes verkraftet werden, ohne dass merkliche Leistungseinbußen auftraten.

2.4.6.4 Zugangskontrolle im URSA-System

Das detektionsbasierte URSA-System wurde bereits in Abschnitt 2.4.5.2 angesprochen, da es auch Authentisierung und Schlüsselverwaltung beinhaltet. Im Folgenden soll es etwas näher beschrieben werden.

URSA verwendet ein spezielles Vertrauensmodell, das auf lokalen Gruppen basiert: Ein Knoten gilt dann netzwerkweit als vertrauenswürdig, wenn ihm k Knoten aus seiner lokalen Umgebung vertrauen.

Jeder Knoten benötigt ein gültiges Ticket, um am Netzbetrieb teilnehmen zu dürfen. Gültig ist ein Ticket, wenn es zertifiziert und nicht abgelaufen ist. Tickets binden eine Identität des Knotens – als Beispiele dafür werden MAC- oder IP-Adressen genannt – an den öffentlichen Schlüssel des Knotens. Jeder Knoten tauscht mit allen seinen Nachbarn beim ersten Kontakt die Tickets aus, um eine gegenseitige Vertrauensbeziehung aufzubauen. Knoten, die kein Ticket vorweisen können, bleiben isoliert.

Tickets sind nur begrenzte Zeit gültig. Wenn diese abläuft, bittet der betreffende Knoten seine Nachbarn, es gemeinsam zu erneuern. Die Nachbarn tun dies, wenn sie während der letzten Überwachungsperiode (die in etwa der typischen Verweildauer eines Knotens in der Nachbarschaft eines anderen entspricht) kein Fehlverhalten des Knotens beobachtet haben.

Der Ticketerneuerungsdienst kann nur von Knoten in Anspruch genommen werden, die bereits ein noch gültiges Ticket besitzen. Für die Ausstellung ganz neuer Tickets gibt es verschiedene Möglichkeiten, etwa die Ausstellung durch eine bestimmte Zahl von Nachbarn unter Verwendung eines Authentizitätsnachweises auf anderer Ebene (z. B. zwischen den menschlichen Benutzern oder die vorläufige Zulassung mit einem Ticket, dass zwar die Diensterbringung, nicht aber die Dienstnutzung erlaubt, und das nach einer gewissen Probezeit durch ein vollwertiges Ticket ersetzt wird. Auch die Möglichkeit einer Aufnahme gegen Bezahlung wird angesprochen.

Wird anhand eines (als austauschbar angesehenen und deshalb nicht näher spezifizierten) Detektionsverfahrens ein Fehlverhalten eines Knotens beobachtet, so wird der beobachtete Knoten in einer

Ticketwiderrufsliste des Beobachters als Angreifer markiert und damit von der Erneuerung ausgeschlossen. Der Beobachter sendet außerdem eine Warnung an alle Knoten in einer gewissen Umgebung. Knoten, die k solche Warnungen von anderen empfangen, markieren den betreffenden Knoten ebenfalls als Angreifer. Die Umgebung, in der die Warnung verbreitet wird, wird so groß gewählt, dass der Angreifer sie möglichst innerhalb der Gültigkeitsperiode seines Tickets nicht verlassen kann. Sein Ticket verfällt damit, und er hat netzweit keinen Zugang mehr. Der Eintrag in den Ticketwiderrufslisten kann dann ebenfalls gelöscht werden.

Es gibt einen gemeinsamen geheimen Systemschlüssel, der aber auf alle Knoten verteilt ist und den kein Knoten allein kennt. Alle Knoten erhalten ihren Anteil am Systemschlüssel, den sie zur Erbringung des Ticketerneuerungsdienstes brauchen, bei der Initialisierung des Netzwerks. Später hinzukommende Knoten können auch nachträglich noch durch eine bestimmte Anzahl von k Nachbarn gemeinsam initialisiert werden. Der öffentliche Teil des Systemschlüssels ist allen Knoten bekannt; er wird zur Prüfung von Signaturen verwendet. Für die Initialisierung des Netzes wird eine zentrale Instanz verwendet, welche den ersten k Knoten Anteile am Systemschlüssel mitteilt. Alle übrigen Knoten werden dann jeweils durch k ihrer Nachbarn initialisiert.

Fälschliche Anschuldigungen von bis zu $k-1$ zusammenarbeitenden Angreifern führen nicht zum Ausschluss des Beschuldigten. Dadurch, dass immer nur mit Nachbarn kommuniziert wird, wird der durch den Ansatz verursachte Kommunikationsaufwand recht gering gehalten.

2.4.7 Schwachstellen existierender Ansätze

Bei der Untersuchung anderer Ansätze, die sich mit Zugangskontrolle und Authentisierung in Ad-hoc-Netzen befassen, zeigte sich, dass häufig Annahmen gemacht bzw. Voraussetzungen gefordert werden, die in offenen Ad-hoc-Netzen eigentlich nicht haltbar sind. Einige Beispiele dafür sind die folgenden:

- *Eindeutige Teilnehmerkennungen:* Teilnehmerkennungen auf maschineller Ebene – darunter fallen z. B. auch Adressen – sind beliebig synthetisierbar. Die Möglichkeit, nach Belieben die Teilnehmerkennung zu wechseln, kann direkt nicht verbaut werden, und je nach Beschaffenheit der Kennung besteht häufig auch die Möglichkeit, Kennungen anderer Teilnehmer zu benutzen.

  Auch der Versuch, Teilnehmerkennungen durch von anderen Teilnehmern erzeugte Zertifikate eindeutig an menschliche Benutzer zu binden, wird im Allgemeinen nicht erfolgreich sein, da die als Zertifizierungsstellen arbeitenden Teilnehmer sich dazu streng an einheitliche Zertifizierungsrichtlinien halten müssten, was nicht garantiert werden kann; in der Regel wird es einem Angreifer durchaus gelingen, sich mehrere Zertifikate für unterschiedliche Teilnehmerkennungen zu beschaffen, indem er sich an unterschiedliche Zertifizierungsstellen wendet.

- *Schwellwertkryptographie:* Der Einsatz von Schwellwertkryptographie, durch den kryptographische Operationen nur durch eine bestimmte Mindestzahl gemeinsam agierender Teilnehmer durchgeführt werden können, wirkt zunächst verlockend, da (zumindest scheinbar) verhindert werden kann, dass einzelne Angreifer und sogar kleine Gruppen gemeinsam agierender Angreifer in sicherheitskritische Mechanismen eingreifen können. Dafür ist allerdings eine Initialisierung durch eine für alle Teilnehmer vertrauenswürdige Instanz erforderlich, und deren Vorhandensein darf nicht angenommen werden.

- *Gruppenentscheidungen:* Bei Verfahren, bei denen Gruppen gemeinsam Entscheidungen treffen bzw. Leistungen erbringen – etwa auch unter Verwendung von Schwellwertkryptographie –

wird davon ausgegangen, dass damit einzelne Angreifer nichts bewirken könnten. Man verlässt sich dabei allerdings implizit darauf, dass Angreifer nicht etwa mehrfach, d. h. unter mehreren verschiedenen Identitäten, an einer Gruppe teilnehmen, wodurch sie Entscheidungen leicht dominieren könnten. Wegen der oben beschriebenen Problematik der eindeutigen Teilnehmerkennungen ist diese Annahme aber nicht haltbar.

Wenn ein Angreifer bei Einsatz von Schwellwertkryptographie unter Verwendung verschiedener Identitäten genügend Schlüsselteile sammeln kann, kann er alle vermeintlich geschützten Operationen allein durchführen; dies kann einen recht hohen Anreiz darstellen.

- *Richtlinientreue Komponenten:* Da wie oben bereits erläutert nicht davon ausgegangen werden kann, dass Teilnehmer sich an gemeinsame Richtlinien halten, wenn diese nicht ihren (momentanen) Interessen entsprechen, werden teilweise richtlinientreue gesicherte Komponenten in den Teilnehmerknoten vorausgesetzt, die bestimmte Operationen immer korrekt durchführen. Auch diese Voraussetzung ist wohl kaum haltbar in Anbetracht der Tatsache, dass jeder Teilnehmer die alleinige Hoheit über seine eigenen Geräte besitzt und diese beliebig modifizieren kann. Es dürfte illusorisch sein, etwa anzunehmen, dass eine völlige Kontrolle über den ganzen Markt an verfügbaren Geräten möglich wäre. Die Verwendung linientreuer Geräte beispielsweise durch zentrale Zertifizierung aller Geräte und Ausschluss nicht zertifizierter zu erzwingen, ist in Anbetracht des technischen und organisatorischen Aufwands unrealistisch, abgesehen davon, dass es dem Ad-hoc-Paradigma diametral widerspricht.

2.5 Zusammenfassung

In diesem Kapitel wurde zunächst eine Einführung in die Architektur von Telekommunikationssystemen gegeben. Diese sind, um die hohe Komplexität durch modularen Aufbau beherrschbar zu machen, in horizontale Schichten gegliedert. Bezüglich der Aufteilung der Aufgaben auf Schichten wird heute de facto ein fünfschichtiges Modell verwendet, welches sich aus dem siebenschichtigen ISO/OSI-Basisreferenzmodell und dem ursprünglichen Internet-Modell entwickelt hat.

Zur Datenkommunikation werden heute meist paketvermittelte Netze eingesetzt, bei denen Pakete über eine Reihe von speichernden Zwischenstationen bis zum Ziel weitergeleitet werden. Der Weiterleitung liegt Topologieinformation zugrunde, die von Wegfindungsverfahren ermittelt wird. In großen Netzen wie dem Internet sind solche Verfahren, die in die beiden Klassen der Link-State- und der Distanz-Vektor-Verfahren aufgeteilt werden können, auf eine topologische Strukturierung des Adressraums angewiesen.

In der Mobilkommunikation muss zwischen infrastrukturbasierten Netzen, die durch zentrale, ortsfest installierte Stationen gesteuert und an das Internet angebunden werden, und infrastrukturfreien Ad-hoc-Netzen unterschieden werden. In Ad-hoc-Netzen wird alle Funktionalität ausschließlich von den Teilnehmern erbracht. Die zwei gebräuchlichsten drahtlosen Übertragungstechniken sind IEEE 802.11, das zu anderen Standards für lokale Netze kompatibel ist und sowohl infrastrukturbasierte als auch Ad-hoc-Netze unterstützt, und Bluetooth, welches zur infrastrukturlosen Nahbereichsvernetzung von Geräten gedacht ist.

Zur Sicherstellung der Authentizität von Kommunikationspartnern, der Integrität und Vertraulichkeit übertragener Daten und der Verfügbarkeit der Kommunikationsdienste werden Sicherheitsmechanismen eingesetzt, die meist auf kryptographischen Algorithmen und Protokollen beruhen. Ein sehr wichtiger und grundlegender Mechanismus ist die Authentisierung. Um sie über ein Netzwerk

durchzuführen, muss die sich authentisierende Partei den Besitz von nur ihr bekannter Information nachweisen. Zur Kontrolle benötigt die andere Partei beispielsweise einen passenden öffentlichen Schlüssel. Solche Schlüssel können geschützt durch Zertifikate verteilt werden, wobei dann allerdings den Zertifizierungsinstanzen vertraut werden muss, die sie ausgestellt haben. Der Grad des Vertrauens und damit auch der resultierenden Authentizität kann mit Hilfe von Vertrauensmetriken modelliert und verrechnet werden. Die Vertrauensmetrik von Jøsang modelliert auch die Unsicherheit, die entsteht, wenn weder vertrauensbildende noch vertrauenszerstörende Fakten bekannt sind.

Ad-hoc-Netze können immer dann sinnvoll eingesetzt werden, wenn keine Infrastruktur vorhanden ist oder diese nicht genutzt werden kann oder soll. Ihre Topologie ist entsprechend der Mobilität der Teilnehmer sehr dynamisch, was unter Anderem hohe Anforderungen an Wegfindungsverfahren stellt und herkömmliche Verfahren ungeeignet macht. Interessant sind in Ad-hoc-Netzen auch reaktive Verfahren, die Wege erst bei Bedarf suchen, sowie Verschmelzungen solcher mit den herkömmlichen proaktiven Verfahren.

Bezüglich der Netzwerksicherheit ergeben sich in Ad-hoc-Netzen besondere Probleme dadurch, dass aufgrund mangelnder Verfügbarkeit einzelner Knoten keine zentralen Komponenten wie Zertifizierungsinstanzen und Zertifikatverzeichnisse eingesetzt werden können, sowie dadurch, dass in offenen Ad-hoc-Netzen a priori keinerlei Informationen über die Teilnehmer vorliegen. Ein Lösungsansatz für diese Schwierigkeiten ist die Verwendung verteilter Zertifizierungsinstanzen, bei denen mit speziellen kryptographischen Verfahren erreicht wird, dass jeweils mehrere Knoten zusammenarbeiten müssen, um Zertifikate zu erstellen. Ein Nachteil solcher Verfahren ist, dass sie besonders initialisiert werden müssen. Ein anderer Ansatz sieht vor, alle Teilnehmer Zertifikate erstellen zu lassen. Hier fällt es nicht leicht, die zur Authentisierung benötigten Zertifikate bei Bedarf zu beschaffen.

Da die Teilnehmer von Ad-hoc-Netzen ihre Ressourcen für die Netzgemeinschaft zur Verfügung stellen, möchten sie missbräuchliche Nutzung durch Knoten, die nichts beitragen wollen, verhindern. Dazu werden Zugangskontrollmechanismen benötigt. Existierende Ansätze basieren entweder darauf, Fehlverhalten anderer Knoten zu erkennen und diese auszuschließen, oder darauf, alle Teilnehmer durch Belohnung richtigen Verhaltens zur Mitwirkung zu motivieren. Detektionsbasierte Verfahren leiden dabei unter Schwierigkeiten bei der Erkennung von Fehlverhalten, während der vorgestellte motivationsbasierte Ansatz, bei dem erbrachte Leistungen bezahlt werden, zur Abrechnung auf manipulationsgeschützte Sicherheitsmodule vertraut, die von Angreifern emuliert werden könnten.

Kapitel 3

Verteilte Zugangskontrolle in offenen Ad-hoc-Netzen

Im Rahmen dieser Arbeit wurde ein Konzept entwickelt, welches die verteilte Durchführung von Zugangskontrolle in offenen Ad-hoc-Netzen ermöglicht. Dieses Konzept wird im Folgenden vorgestellt. In Abschnitt 3.1 werden dazu zunächst – Bezug nehmend auf die in Kapitel 2 eingeführten Grundlagen – nochmals die Ziele formuliert, konkret betrachtete Einsatzszenarien beschrieben, Anforderungen gesammelt, die sich aus den in Ad-hoc-Netzen gegenüber Infrastrukturnetzen gegebenen Einschränkungen ergeben, sowie sonstige angenommene Randbedingungen genannt. Die Lösungsidee wird in Abschnitt 3.2 vorgestellt, bevor in Abschnitt 3.3 ein Überblick über alle Bestandteile des Konzepts und deren Zusammenwirken gegeben wird. Im Anschluss daran wird auf einzelne Bestandteile im Detail eingegangen. In Abschnitt 3.11 wird die Architektur des entworfenen Zugangskontrollsystems nochmals in ihrer Gesamtheit dargestellt, wobei auch auf die Schnittstellen zum regulären Kommunikationssubsystem eingegangen wird.

3.1 Ziele und Voraussetzungen

Das Hauptziel des in der vorliegenden Arbeit entwickelten Konzepts ist die Ermöglichung von Zugangskontrolle in offenen Ad-hoc-Netzen. Die Zugangskontrolle dient dem Zweck, die Funktionsfähigkeit eines Netzes (bzw. die Verfügbarkeit der Dienste des Netzes) zu schützen, die in Ad-hoc-Netzen, in denen alle Dienste von den Netzteilnehmern gemeinsam erbracht werden müssen, gefährdet wird, wenn Netzteilnehmer zwar Dienste nutzen, aber selbst nicht zu deren Erbringung beitragen, um ihre eigenen Ressourcen (etwa Rechenleistung, Übertragungsbandbreite, Energievorat) zu schonen. Solche unkooperativen Teilnehmer werden durch eine Zugangskontrolle von der Nutzung der Dienste ausgeschlossen.

Als Gegenstand der Zugangskontrolle wird in dieser Arbeit in erster Linie der Zugang zum ganzen Netz, also dem Dienst der Netzwerkschicht betrachtet: Den per Zugangskontrolle ausgeschlossenen Knoten wird die Weiterleitung von Paketen verweigert – alle Dienste höherer Schichten sind damit ebenfalls unerreichbar. Die Weiterleitung ist wichtigster Gegenstand der Zugangskontrolle, weil die durch das Weiterleiten von Paketen im ganzen Netz belegte Übertragungsbandbreite und die für das Senden benötigte Energie in mobilen Ad-hoc-Netzen knappe und wertvolle Ressourcen darstellen.

Als Grundlage für den Zugangskontrolldienst wird außerdem ein Verfahren benötigt, mit dessen Hilfe es (nach entsprechender Vorbereitung) möglich ist, Nachrichten entfernter Knoten auf ihre Authentizität hin zu überprüfen.

3.1.1 Einsatzszenario

Aus der Vielzahl verschiedener Einsatzszenarien mobiler Ad-hoc-Netze (siehe Abschnitt 2.4.1) müssen beim Entwurf von Sicherungsverfahren in der Regel bestimmte Szenarien ausgewählt werden, da sich sowohl Anforderungen als auch Randbedingungen je nach Einsatzzweck unterscheiden. In der vorliegenden Arbeit wurde angenommen, dass die mobilen Rechner, welche die Knoten des Ad-hoc-Netzes bilden, persönliche Geräte unterschiedlicher menschlicher Benutzer sind (wobei allerdings das tatsächliche Vorhandensein von Benutzern für die Funktionsfähigkeit des entworfenen Konzepts nur am Rande von Bedeutung ist).

Jeder Knoten gilt administrativ und politisch als abgeschlossene Einheit. Sicherheitsrichtlinien sind nicht a priori über mehrere Knoten hinweg vorgegeben, sondern werden durch den Benutzer jedes einzelnen Knotens nur für diesen Knoten bestimmt. Jeder fremde Knoten muss im Prinzip als potentieller Angreifer betrachtet werden.

Dieses Szenario ist von recht allgemeiner Natur in Hinsicht auf die Anforderungen, die an ein Zugangskontrollsystem gestellt werden. Ein diesbezüglich einfacherer und sozusagen speziellerer Fall ergibt sich beispielsweise, wenn eine gemeinsame administrative oder politische Kontrolle über eine gewisse Anzahl von Knoten oder über alle Knoten gegeben ist, beispielsweise innerhalb einer durch Bindungen außerhalb des Kommunikationssystems zur Zusammenarbeit verpflichteten Gruppe, etwa im militärischen Umfeld. Innerhalb solcher Gruppen kann jeder Teilnehmer jedem anderen in einer Weise vertrauen, die die gemeinsame Diensterbringung wesentlich vereinfacht. Als potentielle Angreifer müssen nur noch Außenseiter betrachtet werden. Da derartige Gruppen in aller Regel auch nicht ad hoc entstehen, sondern sich zumindest aus einer a priori bekannten, zahlenmäßig beschränkten Menge möglicher Teilnehmer zusammensetzen (also *keiner* offenen Teilnehmermenge), können beispielsweise im Voraus geheime Schlüssel vereinbart oder authentische öffentliche Schlüssel sicher verteilt werden – die schwierigsten Fragestellungen der Sicherung von Netzwerkdiensten fallen damit weg.

Ein ähnlicher Fall entsteht, wenn mehrere Geräte pro Benutzer zugelassen werden – beispielsweise können dies mehrere Kleinstgeräte für verschiedene Aufgaben sein, die teilweise in „intelligente“ Gebrauchsgegenstände oder auch Kleidungsstücke des Benutzers eingebettet sind und mittels eines *Personal Area Network* miteinander kommunizieren. Da diese Geräte unter gemeinsamer Kontrolle stehen, können sie mittels vorgegebener Schlüssel sicher miteinander kommunizieren und sich sozusagen nach außen hin „abschotten“. Gegenüber Anderen können sie als einzelner Netzknoten eines offenen Ad-hoc-Netzes gelten.

Wenn die Netzknoten jeweils von menschlichen Benutzern bedient werden, erlaubt dies die Annahme, dass die Benutzung von Netzdiensten in der Regel im Auftrag des Benutzers erfolgt, und dass dieser bei bestimmten Vorgängen aktiv eingreifen und Entscheidungen treffen kann. Ist keine solche Benutzerinteraktion möglich oder erwünscht – etwa bei Geräten ohne Benutzer (z. B. zur allgemeinen Verfügung ins Netz integrierte Drucker) oder weil der Benutzer zeitweise nicht abgelenkt werden darf (z. B. im Fahrzeug) – so sollte das entworfene System trotzdem funktionieren, wenn auch gegebenenfalls mit kleinen Einschränkungen. Es wird davon ausgegangen, dass in diesem Fall durch im Voraus festgelegte Richtlinien das Verhalten in bestimmten Situationen vorgegeben werden kann.

Ein Beispielszenario, für das die genannten Bedingungen zutreffen, ist ein Ad-hoc-Netz zwischen Teilnehmern einer Konferenz zum Zweck des Datenaustauschs und der rechnergestützten Zusammenarbeit. Aus diesem Szenario, das beim Entwurf des Konzepts eine wesentliche Rolle spielte, leitet sich auch die weiterhin getroffene Annahme ab, dass die in dieser Arbeit betrachteten Netze eine beschränkte Größe von maximal einigen hundert Teilnehmern haben.

Bezüglich der Wahl der verwendeten Endgeräte wird davon ausgegangen, dass der Einsatz der immer relativ rechenaufwändigen asymmetrischen kryptographischen Verfahren, etwa zum digitalen Signieren von Nachrichten, möglich ist, ohne dass für die normale Benutzung untragbare Verzögerungen entstehen. Diese Anforderung wird von vollwertigen mobilen Rechnern zweifellos erfüllt, und in Anbetracht der fortschreitenden Leistungssteigerungen bei Kleinstgeräten wie „persönlichen digitalen Assistenten“ (PDAs) und Mobiltelefonen stellt sie wahrscheinlich auch dort bald keine Einschränkung mehr dar.

3.1.2 Anforderungen

Grundlegende Anforderung an das zu entwerfende Konzept zur Ermöglichung von Authentisierung und Zugangskontrolle in offenen Ad-hoc-Netzen ist es, die eingangs des Abschnitts 3.1 genannten Ziele zu erreichen. Ein wichtiger Punkt ist dabei die geforderte Eignung des Verfahrens für offene Netze, die zu zwei tendenziell zueinander in Konflikt stehenden Anforderungen führt, die dennoch beide sinnvoll sind und bestmöglich umgesetzt werden sollten:

- *Offenheit:* Einerseits soll gewährleistet sein, dass neue Teilnehmer, über die im bestehenden Netz noch nichts bekannt ist, jederzeit die Möglichkeit zum Beitritt und zur Nutzung des Netzes haben.
- *Robustheit gegenüber Umgehung der Zugangskontrolle durch Identitätswechsel:* Andererseits sollen Teilnehmer, die aufgrund ihres unkooperativen Verhaltens aus dem Netz ausgeschlossen wurden, sich nicht einfach erneut unbeschränkten Zugang verschaffen können, indem sie ihre Identität wechseln und sich als neue Teilnehmer ausgeben.

Aus den Eigenschaften von Ad-hoc-Netzen (vgl. Abschnitt 2.4.2) ergeben sich weiterhin die folgenden Anforderungen:

- *Vollständig verteilte Erbringung aller Dienste:* In Infrastrukturnetzen kommen sowohl bei der Zugangskontrolle als auch bei der Authentisierung von Teilnehmern regelmäßig ständig verfügbare und mit speziellen Privilegien ausgestattete Infrastrukturkomponenten zum Einsatz (etwa bei Kerberos, beschrieben in Abschnitt 2.3.7.3). Da in Ad-hoc-Netzen im Unterschied zu ortsfesten Netzen keine zentralen Komponenten eingesetzt werden können (vgl. Abschnitte 2.4.2 und 2.4.4), müssen Verfahren entwickelt werden, durch welche die bei Authentisierung und Zugangskontrolle anfallenden Aufgaben verteilt von den gleichberechtigten Teilnehmern des Netzes erbracht werden können. Einzelne Teilnehmer sollen möglichst keine hervorgehobenen Rollen übernehmen, da die ständige Verfügbarkeit einzelner Knoten in Anbetracht der Mobilität, der Anfälligkeit drahtloser Verbindungen und der meist knappen Energiereserven nicht gewährleistet werden kann.
- *Einfache Konfiguration:* Ein Vorteil von Ad-hoc-Netzen ist ihre spontane Verfügbarkeit bei Bedarf. Die Notwendigkeit aufwändiger Konfigurationsmaßnahmen würde die Nützlichkeit von Ad-hoc-Netzen stark einschränken und den Benutzer von seinem eigentlichen Ziel ablenken. Auch sollte kein Fachwissen beim Benutzer vorausgesetzt werden, so dass Ad-hoc-Netze von beliebigen Verbrauchern ohne spezielle Vorkenntnisse betrieben werden können [KrSc04]. Entscheidungen, die während des Betriebs vom Benutzer getroffen werden sollen, müssen diesem verständlich darstellbar sein; dies gilt insbesondere bei Fragen, die Auswirkungen auf die Sicherheit des eigenen Systems haben.

Die Verwendung drahtloser Übertragungstechnik für den Aufbau von Ad-hoc-Netzen und die Mobilität von Teilnehmern bzw. Endgeräten führt zu den folgenden Anforderungen (vgl. Abschnitt 2.2.2):

- *Sparsamer Ressourcenverbrauch:* Die durch Verwendung eines geteilten Medium stets relativ knappe Übertragungsbandbreite soll durch das entwickelte Verfahren in möglichst geringem Maße zusätzlich beansprucht werden, d. h. es sollen möglichst wenige und kurze zusätzliche Nachrichten ausgetauscht werden bzw. soll möglichst wenig zusätzliche Information bei existierenden Protokollnachrichten ergänzt werden. Dies schont auch die zweite besonders wichtige Ressource, nämlich den Energievorrat der mobilen Endgeräte – Sendevorgänge sind relativ energieaufwändig.

- *Robustheit gegenüber Verzögerungen und Verbindungsausfällen:* Die Unzuverlässigkeit der drahtlosen Übertragung, insbesondere auch in Verbindung mit der Mobilität der Teilnehmer, soll das entwickelte Verfahren nicht übermäßig beeinträchtigen. Insbesondere sollen der Verlust von Paketen, die Nichterreichbarkeit bestimmter Knoten und die Partitionierung des Netzes toleriert werden können.

- *Robustheit gegenüber passiven und aktiven Angriffen auf das Übertragungsmedium:* Da es bei drahtloser Übertragung für physikalisch benachbarte Knoten relativ einfach ist, übertragene Nachrichten an der Luftschnittstelle mitzuhören, zusätzliche Nachrichten einzuspielen oder gar übertragene Nachrichten durch Überlagerung zu verändern, muss das entworfene Verfahren gegen solche Veränderungen weitmöglichst unempfindlich sein.

Schließlich sollen an dieser Stelle noch einige Anforderungen formuliert werden, die dazu dienen, Schwächen anderer Ansätze zur Authentisierung und/oder Zugangskontrolle in Ad-hoc-Netzen (vgl. Abschnitt 2.4.5) zu vermeiden:

- *Gleichstellung aller Teilnehmer:* Alle Teilnehmer sollen dieselbe Funktionalität bezüglich des entworfenen Verfahrens ausüben; es soll keine ausgezeichneten Knoten mit besonderer Funktion geben. Dies ist einerseits durch die bereits erwähnten Verfügbarkeitsprobleme begründet, andererseits aber auch dadurch, dass mit der (dauerhaften und mehrseitigen) Übernahme ausgezeichneter Rollen häufig auch besonderes Vertrauen verbunden ist, das dem ausübenden Knoten von Anderen entgegengebracht werden muss. Bei gleichberechtigten und einander im Allgemeinen unbekannten Teilnehmern, die zunächst keine Grundlage für die gegenseitige Beurteilung ihrer Vertrauenswürdigkeit haben, kann eine solche Hervorhebung Einzelner kaum legitimiert werden.

- *Alleinstellung aller Teilnehmer:* Verwandt mit dem vorigen Punkt ist die Forderung, Gruppenstrukturen und Hierarchien zur Organisation des Verfahrens oder auch von Vertrauensbeziehungen zu vermeiden. Solchen Gruppen bzw. Hierarchieebenen (und gegebenenfalls einem aus den Mitgliedern ausgewähltem Gruppenleiter oder -vertreter) muss ebenfalls von den Mitgliederknoten ein gewisses (durch die jeweilige Funktion der Gruppe näher bestimmtes) Vertrauen entgegengebracht werden, welches wie oben zwischen einander unbekannten Teilnehmern nicht a priori vorausgesetzt werden darf. Hinzu kommen schwierige, kaum zufriedenstellend lösbare Fragestellungen bei gruppenbezogenen Aktivitäten wie der Auswahl eines Vertreters oder der dynamischen Aufteilung (unter Umständen erzwungen durch Partitionierung des Netzes) oder Vereinigung bestehender Strukturen, wo grundsätzlich jeweils die Zustimmung aller Mitglieder eingeholt werden müsste.

Ein Beispiel für unerwünschte Gruppenbildung ergibt sich z. B. bei der Verteilung von Zertifizierungsinstanzen auf mehrere Knoten mittels Schwellwertkryptographie (siehe dazu Abschnitt 2.4.5.1), wo zumindest eine Gruppe aller an einer Zertifizierungsinstanz beteiligten Knoten entsteht, deren privilegierte Mitglieder auf irgendeine Weise ausgewählt und legitimiert werden müssen.

- *Möglichst weitgehende Unabhängigkeit vom Wegfindungsverfahren:* Bei existierenden Ansätzen ist teilweise eine enge Beziehung zum verwendeten Wegfindungsverfahren festzustellen (vgl. z. B. Abschnitt 2.4.6.3). Das entworfene Verfahren soll möglichst mit Wegfindungsverfahren aller Art funktionieren.

- *Vermeidung der jeweiligen Nachteile der verschiedenen Verfahren zur Erkennung von Fehlverhalten bzw. Motivation zur Zusammenarbeit:* Die beiden unterschiedlichen Ansätze, entweder falsches Verhalten von Teilnehmern zu erkennen und zu sanktionieren oder die Teilnehmer durch Belohnung richtigen Verhaltens zu solchem zu motivieren (siehe Abschnitt 2.4.6), weisen jeweils spezifische Nachteile auf. Diese sollen nach Möglichkeit vermieden werden.

3.1.3 Angreifermodell

Angreifer sind Teilnehmer, die *Angriffe* durchführen, bei denen sie versuchen, Sicherungsziele zu beschädigen. In Bezug auf den Zugangskontrolldienst bedeutet das, dass Angreifer sich unberechtigt Zugang zu Diensten verschaffen, also etwa das Netz nutzen, ohne selbst Pakete für andere Knoten weiterzuleiten, oder einen unverhältnismäßig großen Anteil der zur Verfügung stehenden Ressourcen belegen. Mit Angriffen auf das Authentisierungsverfahren zielt der Angreifer darauf, die Identität eines anderen Teilnehmers anzunehmen (als Absender einer Nachricht oder als Kommunikationspartner).

Angriffe können direkt Schwächen in den verwendeten Verfahren ausnutzen oder diese Verfahren in nicht vorgesehener Weise missbrauchen, um ihre Ziele zu erreichen – beides sollte selbstverständlich verhindert werden.

In der vorliegenden Arbeit wird davon ausgegangen, dass Angriffe nur durchgeführt werden, wenn sie dem Angreifer auch einen eigenen Nutzen innerhalb des betrachteten Kommunikationssystems bringen, indem er beispielsweise Sendeenergie spart. Es wird deshalb darauf verzichtet, gegen „Angriffe“ zu schützen, die den Angreifer genauso viel Energie kosten wie korrektes Verhalten, ihm keinen erkennbaren Nutzen bringen und ansonsten keine nennenswerten Schäden anrichten.

Es sei angemerkt, dass es Klassen von Angriffen gegen die Verfügbarkeit von Ad-hoc-Netzen gibt, gegen die ein Zugangskontrollsystem nicht schützen kann oder gegen die gar kein Schutz möglich ist. Beispiele dafür sind:

- *Störung des Funkverkehrs:* Wenn die Signalübertragung gezielt durch geeignete Störsignale überlagert wird, kann keine Kommunikation mehr stattfinden. Die Empfindlichkeit gegenüber solchen Angriffen hängt vom Übertragungsverfahren ab, und dort besteht auch die einzige Möglichkeit zur Vorbeugung.

- *Verwerfen von Paketen anstelle von Weiterleitung:* Dass ein Netzknoten die ihm zur Weiterleitung übergebenen Nachrichten verwirft, kann natürlich nicht verhindert werden. Das in dieser Arbeit vorgestellte Konzept stellt Möglichkeiten zur Verfügung, solche Knoten zu erkennen und ihr Verhalten auch anderen Teilnehmern bekannt zu machen. Dafür zu sorgen, dass keine weiteren Nachrichten über derart verlustbehaftete Wege geleitet werden, ist aber eine Aufgabe der Wegfindung und nicht der Zugangskontrolle.

- *Verwendung fremder Adressen:* Die Verwendung der Adressen anderer Teilnehmer und Erzeugung von Nachrichten in deren Namen kann ein nicht dagegen gesichertes Wegfindungsprotokoll und damit unter Umständen die Verfügbarkeit des Netzes zumindest für einzelne Knoten stören, bis hin zur Funktionsunfähigkeit. Ein Zugangskontrollsystem kann davor nicht schützen; es wird lediglich dafür gesorgt, dass das Zugangskontrollsystem selbst durch falsche Adressangaben nicht getäuscht werden kann.

- *Täuschung der Topologieerfassung der Wegfindung:* Angreifer können durch gezielte Täuschung des Wegfindungsverfahren unter Umständen erreichen, dass sie selbst keine Pakete weiterleiten müssen, weil sie angeblich niemals auf dem Weg zu irgendeinem anderen Teilnehmer liegen. Solche Täuschungen können und sollten durch eine Sicherung des Wegfindungsverfahrens verhindert oder zumindest erschwert werden. Vom Zugangskontrollsystem werden weder die Angaben der Wegfindung überprüft, noch wird kontrolliert, ob andere Knoten sich an die von außen nicht einsehbaren Vorgaben ihrer Wegfindungsinstanz halten.

3.2 Idee

Die wesentlichsten Merkmale, die das in dieser Arbeit entworfene Zugangskontrollverfahren von anderen solchen Verfahren unterscheidet, ergeben sich aus den Anforderungen, die durch die besonderen Eigenschaften offener Ad-hoc-Netze entstehen: Alle Dienste müssen verteilt realisiert werden, und alle Teilnehmer sind gleich gestellt. Die wichtigsten Teilprobleme, für die jeweils neuartige Lösungen gefunden werden mussten, sind die Vertrauensbildung, die authentische Identifikation von Netzknoten sowie die eigentliche Zugangskontrolle.

- *Vertrauensbildung*

  Vertrauen in andere Teilnehmer ist eine grundlegende Voraussetzung für die Funktionsfähigkeit jedes verteilten Systems, denn es müssen immer Dienstleistungen erbracht und in Anspruch genommen werden. Ohne Vertrauen in die Willigkeit und Fähigkeit anderer Teilnehmer, ihre Aufgaben im Gesamtsystem korrekt zu erledigen, bliebe als Alternative nur, sämtliche Arbeiten selbst durchzuführen und auf Funktionalität zu verzichten, wo dies nicht möglich ist. Bei existierenden verteilten Systemen wird allerdings häufig einfach generell in korrektes Verhalten aller Teilnehmer vertraut. Dies funktioniert zwar in der Regel, wenn alle Teilnehmer daran interessiert sind, die korrekte Funktion herzustellen und zu erhalten – dies ist häufig auch bei Prototypen wie z. B. dem ursprünglichen Internet der Fall. In der späteren Nutzungsphase eines offenen Systems, wo die Nutzer im System nur noch ein Werkzeug für die Verfolgung ihrer anderweitigen Interessen sehen, überwiegen häufig diese eigenen Interessen. Wenn die verwendeten Verfahren Einzelnen erlauben, sich Vorteile zu verschaffen, wird diese Möglichkeit auch früher oder später wahrgenommen – insbesondere dann, wenn eine gewisse Anonymität und die Abwesenheit verbindlicher, durchsetzbarer und tatsächlich auch durchgesetzter Verhaltensregeln vor Verfolgung schützt. Dies ist an vielen für das ursprüngliche Internet entworfenen Verfahren feststellbar, etwa dem E-Mail-System, das mittlerweile sehr umfassend missbraucht wird, um z. B. unerwünschte Werbung zu verbreiten.

  In offenen Ad-hoc-Netzen, wo die Teilnehmer einander häufig nicht kennen und sich gegenseitig auch keinen Organisationen, Netzzugangsanbietern oder sonstigen rechtlich fassbaren Institutionen zuordnen können, die gegebenenfalls bereit und in der Lage wären, korrektes Verhalten durchzusetzen, sollte Vertrauen in andere Teilnehmer nicht einfach vorausgesetzt werden

– zumindest keines, was über die Erwartung hinausgeht, dass jeder versuchen wird, seine persönlichen Interessen zu verfolgen, ohne dabei unbedingt Rücksicht auf die Interessen Anderer zu nehmen.

Es gibt im Wesentlichen zwei Möglichkeiten, das benötigte Vertrauen zu anderen Teilnehmern zu beschaffen, die beide im hier entworfenen Ansatz genutzt werden können: Vertrauen muss entweder aus der realen Welt der Benutzer in die virtuelle des Netzwerks übertragen oder mit der Zeit innerhalb des Systems aufgebaut werden. Voraussetzung für die Nutzung der ersten Möglichkeit ist, dass Benutzeridentitäten mit Identitäten von im Netz agierenden Rechnern sicher verknüpft werden können. Die Idee für die Umsetzung der zweiten Möglichkeit zur Vertrauensbeschaffung, also für den Aufbau von Vertrauen innerhalb des Systems, ist es zu versuchen, menschliches Verhalten beim Aufbau sozialer Beziehungsnetzwerke auf der Basis zwischenmenschlicher Kontakte in den gegebenen Grenzen nachzuahmen, indem die menschlichen – häufig intuitiven – Vorgehensweisen dabei auf die Maschinenebene übertragen werden. Wesentlicher Bestandteil der Vertrauensbildung ist es, aufgrund der Beobachtung des Verhaltens anderer Knoten mit der Zeit Einschätzungen über deren Vertrauenswürdigkeit zu gewinnen. Sobald gewisse Vertrauensbeziehungen zu anderen Teilnehmern vorhanden sind, können diese außerdem genutzt werden, um Einschätzungen bezüglich der Vertrauenswürdigkeit Dritter auszutauschen.

- *Identifikation, Authentisierung und Schlüsselverwaltung*

  Die Teilnehmer identifizieren sich auf maschineller Ebene im Netz durch kryptographische Schlüssel. Diese Methode erlaubt einen unmittelbaren Identitätsnachweis und die Nachrichtenauthentisierung anhand von digitalen Signaturen; eine Zertifizierungsinfrastruktur ist dabei nicht erforderlich. Die zur Signaturprüfung benötigten Schlüssel müssen im Netz verteilt werden, aber es wird gezeigt, dass für die Zwecke der Zugangskontrolle keine Sicherung durch Zertifikate und kein aufwändiges Schlüsselverzeichnis benötigt werden.

  Durch die lokale Zuordnung von Benutzeridentitäten zu den Schlüsseln – etwa als Folge einer auf zwischenmenschlicher Ebene erfolgten Identitätsprüfung – können bei Bedarf auch Vertrauensbeziehungen aus der realen Welt in die virtuelle übertragen werden.

- *Zugangskontrolle*

  Die Zugangskontrolle zum Netz wird durch die Kontrolle der Weiterleitung von Paketen realisiert, und diese Art von Zugangskontrolle führt jeder Knoten auf dem Weg eines im Netz übertragenen Pakets durch: Nur wenn der Nutzer des Weiterleitungsdienstes, also der Absender eines Pakets, von anderen Knoten als kooperativer Teilnehmer eingeschätzt wird, der auch selbst Pakete anderer weiterleitet, darf er auch den Weiterleitungsdienst nutzen.

  Grunsätzlich kann auch die Zugangskontrolle zu anderen Diensten des Netzwerks aufgrund der Einschätzungen realisiert werden, die mit Hilfe der oben beschriebenen Verfahren zur Vertrauensbildung ermittelt werden. Auf diese Möglichkeit wird in dieser Arbeit allerdings nicht detailliert eingegangen.

3.3 Überblick

Im Folgenden wird zunächst ein Überblick über die Bestandteile des entworfenen Konzepts gegeben. Entwurfsentscheidungen werden anhand der zugrundeliegenden Anforderungen motiviert.

3.3.1 Identifikation und Authentisierung von Netzknoten

Als Basis für die gegenseitige Authentisierung der Knoten im Netz besitzt jeder Knoten ein asymmetrisches Schlüsselpaar, dessen öffentliche Hälfte bei Bedarf im Netz bekannt gemacht wird (mit Hilfe der weiter unten beschriebenen Verfahren zur Schlüsselverwaltung). Dieser öffentliche Schlüssel dient direkt als Identifikator des Schlüsselinhabers im Netz. Außerdem kann mit Hilfe des Schlüsselpaars die Authentizität von Nachrichten gesichert werden, indem ihr Erzeuger sie digital signiert; diese Signatur ist durch jeden anderen Knoten, der eine Kopie des öffentlichen Schlüssels des Absenders besitzt, prüfbar.

Auf die verwendeten Verfahren zur Identifikation und Authentisierung von Netzknoten und deren Nachrichten wird in Abschnitt 3.4 näher eingegangen; dort werden auch potentielle Manipulationsmöglichkeiten untersucht.

3.3.2 Vertrauensmodell

Vertrauen wurde in Abschnitt 2.3.5.1 beschrieben als die Erwartung, dass eine andere Partei sich in bestimmter, „korrekter“ Weise verhält. Als Quellen für Vertrauen wurden dort eigene positive Erfahrungen mit dieser Partei, gleich gerichtete Interessen – gegebenenfalls begründet durch vertragliche Verpflichtungen – sowie Empfehlungen Dritter (deren Empfehlungen bereits vertraut wird) genannt.

Vertragliche Bindungen zwischen Teilnehmern offener Ad-hoc-Netze sind nicht praktikabel, da einerseits schon die Aushandlung gemeinsamer Regeln sehr schwierig wäre und andererseits eine rechtlich bindende Verpflichtung praktisch nicht machbar ist. Als Quelle für Vertrauen in offenen Ad-hoc-Netzen kommen damit nur eigene Erfahrungen und Empfehlungen in Betracht. Beide kommen jeweils auf zwei unterschiedlichen Ebenen vor: einerseits zwischen menschlichen Benutzern und andererseits zwischen den Rechnern, die die eigentlichen Netzknoten darstellen. Von der menschlichen Ebene kann Vertrauen auf die Rechnerebene übertragen werden, indem der Benutzer seinem Rechner gegenüber explizit Angaben zur Vertrauenswürdigkeit anderer Benutzer (und damit derer Rechner[1]) macht. Auf der Rechnerebene können zusätzliche Einschätzungen der Vertrauenswürdigkeit Anderer durch automatisierte Verhaltensbeobachtung gewonnen werden. Solche Einschätzungen können als ebenfalls automatisiert ausgetauschte Empfehlungen an andere Rechner weitergegeben werden. Gegebenenfalls können auf der Rechnerebene gewonnene Einschätzungen auch das Vertrauen auf der menschlichen Ebene beeinflussen, wenn der Benutzer die automatisiert ermittelten Einschätzungen einsieht.

3.3.3 Vertrauensprofile

Entscheidungen darüber, ob anderen Knoten Zugang zum Netz gewährt wird, ob also Dienste für sie erbracht werden, werden aufgrund von Einschätzungen der Kooperativität der Kandidaten getroffen. Einschätzungen werden wiederum aus der Beobachtung richtigen und falschen Verhaltens abgeleitet.

Um Vertrauen und Einschätzungen in quantitativer Form auszudrücken und so Berechnungen und Vergleiche zu ermöglichen, wird im vorgeschlagenen Konzept die in Abschnitt 2.3.5.4 vorgestellte Metrik von Jøsang [Jøsa98] verwendet, die subjektive Einschätzungen zu objektiv entweder wahren oder falschen Aussagen durch Wertetripel $(b,d,u) \in [0,1]^3$ mit $b+d+u=1$ ausdrückt (im Folgenden als *Meinungen*, *Einschätzungen* oder *Vertrauensmaßzahlen* bezeichnet), wobei b für den Glauben an

[1] Hier wird angenommen, dass die Rechner nicht fehlerhaft oder manipuliert sind und tatsächlich vollständig im Interesse ihrer Benutzer handeln.

die Richtigkeit der Aussage (Belief), *d* für den Glauben an die Falschheit (Disbelief) und *u* für die in Ermangelung ausreichender Information bestehende Unsicherheit (Uncertainty) steht.

Zu dieser Metrik existieren sowohl Methoden, um anhand positiver und negativer Beobachtungen Einschätzungen zu gewinnen, als auch, um Einschätzungen miteinander zu verknüpfen, wenn beispielsweise mehrere Aussagen gleichzeitig zutreffen müssen oder wenn eine Einschätzung für eine Aussage benötigt wird, über die nur indirekte Informationen von anderen vorliegen, deren Vertrauenswürdigkeit in die Einschätzung einbezogen werden muss. Ein in dieser Arbeit neu entwickeltes Verfahren zur Verknüpfung (siehe Abschnitt 3.5) erlaubt zuverlässig die Auswertung ganzer Graphen von Vertrauensbeziehungen zwischen verschiedenen Knoten, was mit den von Jøsang eingeführten Operationen nur sehr eingeschränkt möglich ist.

Die Einschätzung der Kooperativität anderer Knoten ist relevant für die Zugangskontrolle zum Netzwerk. Außerdem können aber Einschätzungen bezüglich weiterer Charakteristika des Verhaltens anderer Knoten ermittelt und genutzt werden, beispielsweise bezüglich der Willigkeit und Fähigkeit zur Ausstellung korrekter Zertifikate oder bezüglich der Übereinstimmung zwischen von anderen Knoten geäußerten Einschätzungen mit den eigenen. Weitere Kategorien können für alle zusätzlich betrachteten Dienste hinzugefügt werden. Insgesamt entstehen so genannte *Vertrauensprofile* über andere Knoten, die eine ganze Reihe von Einschätzungen zu unterschiedlichen Verhaltensmerkmalen enthalten. Vertrauensprofile werden in Abschnitt 3.7 behandelt.

3.3.4 Verhaltensbeobachtung

Die Beobachtung des Verhaltens benachbarter Knoten zum Zweck der Beurteilung in Bezug auf Korrektheit und Fairness kann auf maschineller Ebene durch Mithören des Netzverkehrs erfolgen – aufgrund der drahtlosen Übertragungstechnik ist dies in mobilen Ad-hoc-Netzen ohne besondere technische Vorkehrungen möglich. Beobachtet und beurteilt werden kann dabei grundsätzlich die Korrektheit des Verhaltens bezüglich verschiedener Protokolle, angefangen beim Netzwerkprotokoll zur Weiterleitung von Paketen im Netz über Wegfindungsprotokolle und die Protokolle des Zugangskontrollsystems selbst bis hin zu Anwendungsprotokollen. Im Allgemeinen werden positive Beobachtungen registriert, wenn ein Nachbar auf eine beobachtete Anfrage korrekt antwortet; reagiert er dagegen in offensichtlich falscher Weise oder gar nicht, so wird er negativ bewertet. Abgesehen von der prozeduralen Korrektheit kann auch der Umfang der Nutzung bzw. Erbringung oder Bereitstellung von Diensten und Ressourcen in die Bewertung mit einbezogen werden.

Das wichtigste Beispiel für die Beobachtung anderer Knoten betrifft die Weiterleitung von Paketen: Immer, wenn ein Knoten ein Paket weiterleitet, erhält er von seinen Nachbarn – soweit diese den Vorgang beobachtet haben – eine positive Bewertung. Leitet er dagegen ein an ihn zur Weiterleitung gesendetes Paket nicht innerhalb einer gewissen Zeitspanne weiter, so erhält er eine negative Bewertung. Die Verhaltensbeobachtung enthält immer gewisse Unsicherheiten – beispielsweise kann sich der zur Weiterleitung verpflichtete Knoten zwischen Empfang und Weiterleitung aus der Reichweite des Beobachters bewegen oder auf andere Weise von ihm abgeschirmt werden, so dass der Beobachter die Weiterleitung nicht beobachten kann und fälschlich annimmt, dass sie unterlassen wurde. Da die Bedingungen für eine richtige Bewertung aber in einem Großteil der Fälle erfüllt sind und einzelne Beobachtungen immer nur zu graduellen Anpassungen von Einschätzungen führen, können wenige falsch bewertete Beobachtungen in der Regel ohne wesentliche Verfälschungen der Gesamteinschätzungen verkraftet werden.

Allgemeine Verfahrensweisen bei der Verhaltensbewertung sowie insbesondere und im Detail das Verfahren zur Beobachtung und Bewertung des Weiterleitungsverhaltens von benachbarten Knoten sind Gegenstand von Abschnitt 3.6.

3.3.5 Zugangskontrolle anhand von Einschätzungen

Die Zugangskontrolle zum Netzwerkdienst wird implementiert, indem Pakete, die nicht aus als kooperativ eingeschätzten Quellen[2] stammen, nicht weitergeleitet werden. Jeder Knoten prüft also vor dem Weiterleiten seine Einschätzung bezüglich der Kooperativität der Quelle des jeweils weiterzuleitenden Pakets und leitet nur weiter, wenn diese Einschätzung positiv ist. Für die sichere Zuordnung des Pakets zu seiner Quelle muss die Authentizität der Quellangabe im Paket überprüft werden. Zu diesem Zweck wird jedes Paket von seiner Quelle digital signiert. Zur Überprüfung dieser Signatur benötigen die weiterleitenden Knoten den authentischen Schlüssel des Quellknotens. Das in Abschnitt 3.3.8 beschriebene Schlüsselverwaltungsverfahren wurde im Hinblick darauf entworfen, d. h. es ist in der Lage, benötigte Schlüssel schnell zu liefern, ohne das Netzwerk dabei übermäßig zu belasten.

Wenn die Zugangskontrolle fehlschlägt, wird die Quelle durch eine Fehlermeldung darüber informiert (sofern dasselbe Problem nicht unmittelbar zuvor schon auftrat – die Rate der Fehlermeldungen wird stark begrenzt), so dass sie gegebenenfalls mit Hilfe der in den folgenden Abschnitten beschriebenen Verfahren für eine Verbesserung ihrer Einschätzung sorgen kann.

Die Abläufe bei der Zugangskontrolle zum Netzwerkdienst werden in Abschnitt 3.9 detailliert beschrieben und auf potentielle Angriffspunkte hin untersucht. Dort wird auch der so genannte Bootstrapping-Mechanismus beschrieben, der dazu dient, in neu entstandenen Ad-hoc-Netzen die Zugangskontrolle anfangs für kurze Zeit auszusetzen, um eine schnelle Gewinnung gegenseitiger Einschätzungen zu fördern.

Einschätzungen können grundsätzlich auch verwendet werden, um den Zugang zu anderen Diensten im Netzwerk zu kontrollieren. Im Allgemeinen muss dazu immer dann, wenn eine Dienstleistung für einen anderen Knoten erbracht werden soll, anhand der vorliegenden Einschätzungen zunächst überprüft werden, ob dieser Knoten in der Vergangenheit selbst bereits Dienste für die Allgemeinheit (evtl. speziell bezüglich des betreffenden Dienstes) erbracht hat. Nur wenn dies der Fall ist, darf dem Dienstleistungswunsch entsprochen werden. Diese Form der Zugangskontrolle – auf die allerdings in dieser Arbeit nicht detailliert eingegangen wird – ist in der Regel von denjenigen Knoten durchzuführen, welche die entsprechenden Dienste anbieten oder verwalten.

Da Vertrauensmaßzahlen Einschätzungen repräsentieren, werden sie durch Inanspruchnahme von Diensten nicht verringert, das entgegengebrachte Vertrauen wird also sozusagen durch seine Nutzung nicht verbraucht (anders als bei währungsbasierten Ansätzen; siehe Abschnitt 2.4.6). Allerdings kann man die Rate erlaubter Dienstnutzungen von der Höhe des Vertrauens abhängig machen. Versuchte Überschreitung dieser Rate führt dann zunächst zu einer Warnung und wird später negativ als Missbrauch gewertet.

3.3.6 Weitergabe von Einschätzungen

Wenn für eine Zugangsentscheidung eine Einschätzung über einen bestimmten Knoten benötigt wird, die lokal nicht vorhanden ist, weil die Knoten z. B. in der Vergangenheit nie benachbart waren und sich deshalb auch nicht beobachten konnten, so können andere Knoten, etwa die eigenen Nachbarn, nach ihrer Einschätzung gefragt werden. Selbstverständlich können fremde Einschätzungen nur in dem Maße berücksichtigt werden, wie der anfragende Knoten den gefragten in Bezug auf die Gewinnung und ehrliche Weitergabe von Einschätzungen vertraut.

[2] Als „Quelle“ wird auch im Folgenden immer der ursprüngliche Erzeuger eines Schicht-3-Pakets bezeichnet, also der Absender auf Schicht 3.

Die Antwort auf eine solche „Einschätzungsanfrage“ wird von dem Knoten, dessen Einschätzung sie repräsentiert, digital signiert. Sie ähnelt damit einem Zertifikat insofern, als sie beliebig weitergegeben und – beispielsweise vom bewerteten Knoten – auch zur späteren Verwendung gespeichert werden kann. Allerdings darf die Gültigkeitsdauer eines solchen Einschätzungzertifikats nur relativ kurz sein, damit es nicht dazu missbraucht werden kann, positive Einschätzungen vorzuspiegeln, nachdem das eigene Verhalten zum Schlechteren geändert wurde.

Die Durchführung von Einschätzunganfragen wird im Zusammenhang mit der Zugangskontrolle in den Abschnitten 3.9.3.3 und 3.9.3.4 beschrieben. Einschätzungszertifikate werden schon vorher im Zusammenhang mit Vertrauensprofilen in Abschnitt 3.7.3.1 vorgestellt. Außerdem benötigt wird ein Verfahren zur Bewertung der Vertrauenswürdigkeit fremder Einschätzungen bzw. Teilnehmer; dieses ist Gegenstand von Abschnitt 3.8.

3.3.7 Bürgschaften

Da über neu zum Netz hinzugekommene Knoten noch keinerlei Beobachtungen vorliegen, könnten diese bei alleiniger Anwendung obiger Mechanismen nur mühsam Zugang zu den Diensten des Netzwerks erlangen[3]. Aus diesem Grund wurde zusätzlich die Möglichkeit geschaffen, für andere Knoten zu bürgen: Ein bereits etablierter Knoten gibt dabei auf eine Bürgschaftsanfrage hin mittels eines von ihm signierten Einschätzungszertifikats anderen Knoten einen zeitlich beschränkten Vertrauensvorschuss. Neue Knoten haben die (stark ratenbeschränkte) Möglichkeit, Bürgschaftsanfragen an Kommunikationspartner innerhalb des bestehenden Netzes zu senden (siehe Abschnitt 3.7.3.2). Dabei wird davon ausgegangen, dass ein solcher Kommunikationspartner, wenn er an der Kommunikation mit dem neuen Knoten interessiert ist, häufig bereit sein wird, eine solche Bürgschaft zu übernehmen (beispielsweise aufgrund persönlicher Bekanntschaft oder einer vorangegangenen Absprache zwischen den Benutzern).

Der Antragsteller kann ein per Bürgschaft erhaltenes Einschätzungszertifikat in der Folge während dessen Gültigkeitsdauer allen Knoten zur Verfügung stellen, deren Dienste er in Anspruch nehmen möchte; es wird dann jeweils in die Einschätzungsauswertung einbezogen und bewirkt eine Verbesserung der Einschätzung des Dienstnutzers, wenn der Bürge beim Dienstgeber Vertrauen genießt. Der neue Knoten übermittelt Diensterbringern seine Einschätzungszertifikate immer dann, wenn er von diesen eine Meldung über fehlendes Vertrauen erhält.

Der bürgende Knoten trägt das Risiko, dass unkooperatives Verhalten des neuen Knotens seine eigene Einschätzung verschlechtert, und wird deshalb Bürgschaften nicht allzu leichtsinnig vergeben. Letzteres würde die Wirksamkeit des Zugangskontrollmechanismus aushöhlen.

Der relativ aufwändige Bürgschaftsmechanismus könnte vermieden werden, indem neuen Knoten von Anfang an ein gewisser Vertrauensbonus eingeräumt würde. Dies würde aber dazu führen, dass durch Fehlverhalten in Misskredit geratene Knoten einfach ihre Identität wechseln könnten, um ihr Fehlverhalten vergessen zu machen und umgehend in den Genuss des Neueinsteigerbonus zu gelangen. Die grundsätzliche Möglichkeit des Identitätswechsels kann wegen der gewünschten Offenheit nicht verbaut werden, aber dadurch, dass für neue Knoten zunächst eine gewisse Schwelle zu überwinden ist, ist sie relativ unattraktiv. Teilnehmer, die sich bereits einige Zeit unkooperativ verhalten haben, haben nur geringe Chancen, ihre neue anfängliche Einschätzung nach Identitätswechsel durch Bürgschaften zu verbessern.

[3] Neue Knoten können Zugang erhalten, indem sie sich zunächst durch Teilnahme am Wegfindungsprotokoll bekannt machen und dann so lange bedingungslos Pakete anderer Knoten weiterleiten (ohne selbst aber eigene Nachrichten absenden zu können), bis ihre Einschätzung gut genug ist.

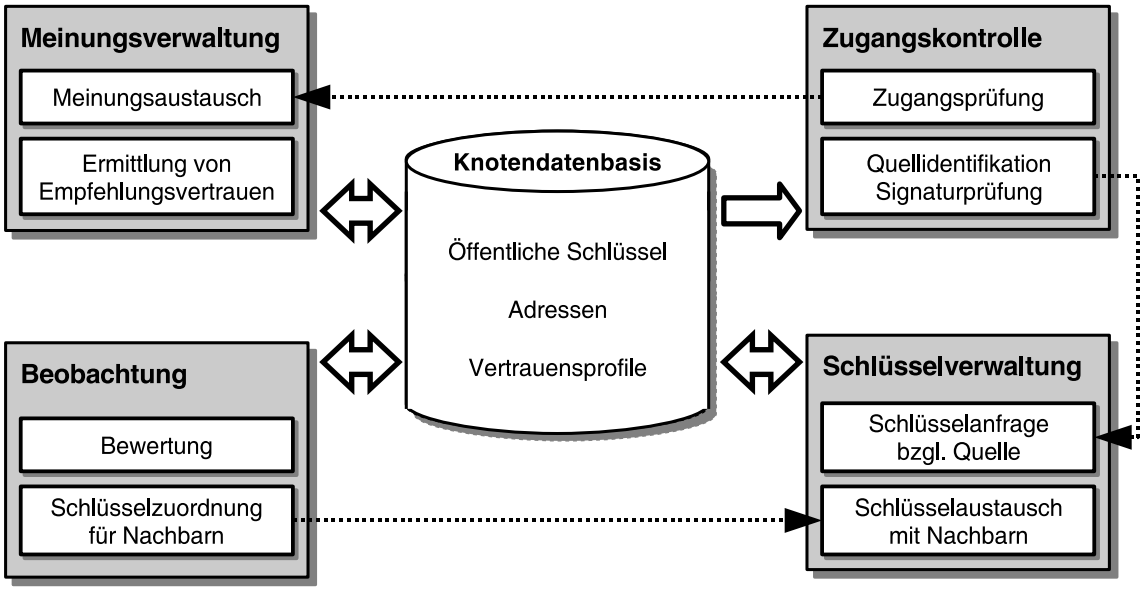


Abbildung 3.1: Komponenten des Zugangskontrollsystems

3.3.8 Schlüsselverwaltung

Die öffentlichen Schlüssel der Teilnehmer, die gleichzeitig zu ihrer Identifikation dienen, müssen anderen Teilnehmern für zwei unterschiedliche Zwecke verfügbar gemacht werden: Erstens müssen Beobachtungen den jeweils aktiven Teilnehmern zugeordnet und die zugehörigen Vertrauensprofile gefunden werden, in denen Bewertungen registriert werden können. Zweitens muss bei der Zugangskontrolle jeweils ein Vertrauensprofil identifiziert werden, anhand dessen die Zugangsentscheidung getroffen wird.

Für Zwecke der Beobachtung muss der Schlüssel jedes Teilnehmers genau seinen Nachbarn bekannt sein, denn nur diese können ihn beobachten. Um das zu gewährleisten, wird immer beim Kontakt mit neuen Nachbarn ein Schlüsselaustauschprotokoll durchgeführt, in dessen Verlauf jeder der beiden Teilnehmer dem anderen seine Identität nachweist. Anschließend erfolgt die Zuordnung zwischen beobachtbaren Nachrichten und jeweils sendendem Nachbarn anhand der Absenderadresse der Schicht 2 (siehe Abschnitt 3.6.4). Bei der detaillierten Bedrohungsanalyse (Abschnitte 3.6.3 und 3.10.4) wird nachgewiesen, dass eine weitere Sicherung nicht erforderlich ist.

Zum Zweck der Zugangskontrolle wird der Schlüssel der Quelle bei jedem Weiterleitungsschritt benötigt. Die Beschaffung kann deshalb sehr einfach durch eine Anfrage beim vorigen Weiterleiter erfolgen. Die Sicherheitsanalyse (Abschnitt 3.10.5) zeigt auch hier, dass die Sicherheit der Zugangskontrolle damit gewährleistet werden kann.

Die Schlüsselverwaltung wird in Abschnitt 3.10 näher behandelt. Insbesondere werden dort die Protokolle zum Schlüsselaustausch zwischen Nachbarn (Abschnitt 3.10.7) und zur Beschaffung des Quellschlüssels vom Weiterleitungs-Vorgänger (Abschnitt 3.10.8) beschrieben.

3.3.9 Zusammenspiel der Komponenten

Abbildung 3.1 soll einen groben Überblick über das Zusammenspiel der Komponenten des entworfenen Konzepts geben. Datenfluss wird dort durch breite umrandete Pfeile dargestellt, Kontrollfluss durch punktierte Pfeile. Ein zentraler Bestandteil ist eine Datenbasis, in der Informationen über andere Knoten gespeichert werden, insbesondere ihre Schlüssel und Adressen sowie Einschätzungen zu

ihrem Verhalten. Beeinflusst wird die dort vorgehaltene Information aus drei Richtungen: Schlüssel und Adressen werden von der Schlüsselverwaltungskomponente eingetragen, Einschätzungen entstehen entweder durch die Beobachtung des Verhaltens von Nachbarn oder aufgrund von Empfehlungen anderer Knoten (Meinungsaustausch). Abgesehen davon können auch explizite Vorgaben des Benutzers in die Datenbasis eingehen.

Empfehlungen müssen selbst aus vertrauenswürdigen Quellen stammen, d. h. sie können nur in dem Maße berücksichtigt werden, wie bereits Empfehlungsvertrauen in ihre jeweilige Quelle vorhanden ist. Eine Aufgabe der Meinungsverwaltungkomponente ist es neben der Durchführung des Meinungsaustauschs mit anderen Knoten deshalb auch, aufgrund der vorhandenen eigenen und fremden Einschätzungen Maße für Empfehlungsvertrauen zu ermitteln und wiederum in der Datenbasis abzulegen.

Für die Verhaltensbeobachtung ist es erforderlich, den beobachteten Nachrichten bzw. den sie absendenden und empfangenden Netzknoten die zugehörigen Vertrauensprofile zuzuordnen, so dass dort Änderungen vorgenommen werden können. Diese Zuordnung wird anhand von Adressinformation aus der Knotendatenbasis hergestellt.

Sobald die Beobachtungskomponente Nachrichten von einem bisher nicht in der Datenbasis enthaltenen Knoten wahrnimmt, wird durch die Schlüsselverwaltung ein Schlüsselautauschprotokoll mit Nachbarn durchgeführt, um die benötigte Zuordnung herzustellen. Außerdem werden von der Schlüsselverwaltung die Schlüssel der Quellen weitergeleiteter Pakete beim jeweiligen vorigen Weiterleiter angefordert werden, wenn bei der Zugangskontrolle der erforderliche Quellschlüssel fehlt.

Eigentlicher Hauptzweck der Vertrauensprofile innerhalb der Knotendatenbasis ist es, eine Grundlage für Zugangskontrollentscheidungen zu liefern. Hierzu müssen den Dienstnutzungswünschen die Vertrauensprofile der zugehörigen Knoten zugeordnet werden, wobei die Identität der Quelle anhand der von ihr erstellten Signatur überprüft wird. Falls die in der Knotendatenbasis enthaltene Information über die Quelle für die sich anschließende Zugangsprüfung nicht ausreicht, wird die Durchführung von Einschätzungsanfragen angestoßen.

Im Anschluss an die nun folgende detaillierte Behandlung der einzelnen Komponenten wird in Abschnitt 3.11 nochmals eine etwas verfeinerte Darstellung der Architektur aus Abbildung 3.1 gegeben.

3.4 Identifikation und Authentisierung

Als Basis für die Zugangskontrolle soll bei dem in dieser Arbeit entwickelten Ansatz das bisherige Verhalten desjenigen Teilnehmers herangezogen werden, der eine Dienstleistung in Anspruch nehmen will. Damit Information über das bisherige Verhalten anderer Teilnehmer tatsächlich zur Verfügung steht, wenn sie für die Zugangskontrolle benötigt wird, beobachtet jeder Teilnehmer automatisch und ständig das Verhalten seiner Nachbarn und speichert seine Beobachtungen lokal in einem dem jeweiligen Nachbarn zugeordneten Vertrauensprofil, welches bei der Zugangskontrolle wiederum konsultiert wird.

Auf der Ebene maschineller Verfahren handelt es sich bei den Teilnehmern um Netzknoten in Form von Rechnern, deren Möglichkeiten der gegenseitigen Wahrnehmung auf den Empfang der von anderen Knoten ausgesandten Nachrichten beschränkt sind. Allein auf empfangene Nachrichten gründet sich also alles Wissen um Existenz, Anwesenheit und Verhalten anderer Teilnehmer. Allein anhand einer empfangenen Nachricht müssen die Teilnehmer identifiziert werden, zwischen denen die Nachricht ausgetauscht wurde. Diese *Identifikation* der an einem Kommunikationsvorgang beteiligten Teilnehmer bedeutet in Bezug auf Zugangskontrolle und automatisierte Beobachtung letztendlich, dass

die diesen Teilnehmern zugeordneten Vertrauensprofile in der lokalen Datenbasis gefunden werden können, so dass sie aufgrund von Beobachtungen modifiziert oder zum Zweck der Zugangskontrolle konsultiert werden können.

Verfahren zur Identifikation liefern *Teilnehmerkennungen*, anhand derer gespeicherte Vertrauensprofile gefunden werden können. Solche Teilnehmerkennungen müssen netzweit eindeutig sein, so dass gewährleistet ist, dass nicht versehentlich mehrere Teilnehmer demselben Vertrauensprofil zugeordnet werden.

Neben der Identifikation von Netzknoten (Maschinen) kann außerdem die Identifikation von Benutzern (Menschen) sinnvoll sein. Durch eine Abbildung zwischen Vertrauensprofilen und Benutzeridentitäten kann die Möglichkeit geschaffen werden, bestehende Vertrauensbeziehungen zwischen Benutzern auf die maschinelle Ebene zu übertragen (siehe das Vertrauensmodell in Abschnitt 3.3.2). Im Unterschied zur Abbildung zwischen Netzknoten und Vertrauensprofilen – welche für die Funktion des Zugangskontrollverfahren zwingend erforderlich ist – ist die Abbildung zwischen Vertrauensprofilen und Benutzeridentitäten aber optional, und sie sollte auch nur auf Wunsch des jeweiligen Benutzers durchführbar sein: Der Benutzer kann, indem er die automatisierte Zuordnung seiner Identität zum Vertrauensprofil seines Netzknotens nicht zulässt, eine gewisse Anonymität wahren. In dieser Arbeit wird kein besonderes Verfahren zur Benutzeridentifikation beschrieben, grundsätzlich können aber existierende Verfahren zusammen mit dem entworfenen System verwendet werden.

3.4.1 Bedrohungen im Zusammenhang mit Identifikationsverfahren

Die folgenden generellen Angriffsschemata stehen im Zusammenhang mit der Identifikation von Teilnehmern. Sie sind zunächst unabhängig davon, wie die Identifikation konkret realisiert wird. Die konkrete Ausprägung der Angriffe sowie die Maßnahmen zur Verhinderung solcher Angriffe können sich aber je nach Identifikationsverfahren unterscheiden. Bei der weiter unten folgenden Beschreibung konkreter Identifikationsverfahren wird jeweils noch näher auf Angriffe und Gegenmaßnahmen eingegangen.

- *Fälschung der Identität*

  Der naheliegendste Angriff ist das Vortäuschen einer fremden Identität. Kann ein Angreifer gezielt die Identität eines anderen Teilnehmers annehmen, so kann er dadurch einerseits das Vertrauensprofil des anderen Teilnehmers modifizieren, indem er sich unter falscher Identität in bestimmter Weise verhält, oder er kann andererseits bewirken, dass die Zugangskontrolle bei eigener Dienstnutzung aufgrund des Vertrauensprofils des anderen Knotens erfolgt.

  Hier zeigt sich ein entscheidender Unterschied zwischen zwischenmenschlichen Kontakten und Kontakten zwischen Rechnern: Menschen können sich in der Regel gegenseitig anhand äußerlicher Merkmale eindeutig identifizieren, wenn sie miteinander kommunizieren (zumindest in einem ausreichend großen Anteil der Fälle), und zwar auch dann, wenn eine Seite eigentlich gerne verhindern will, dass ihr Kommunikationspartner sie wiedererkennt. Bei Rechnerkommunikation ist das anders: Zwar können die Rechner sich (bzw. ihre Nachrichten) z. B. anhand kryptographischer Signaturen sicher erkennen, wenn das erwünscht ist (vorausgesetzt, die Zuordnung zwischen Schlüssel und Identität ist sicher möglich, was aber in beiderseitigem Einverständnis unter kurzzeitiger Verwendung eines manipulationssicheren Kanals so eingerichtet werden kann). Wenn ein Wiedererkennen aber nicht gewünscht ist, kann sich ein Rechner jederzeit eine neue Identität zulegen.

Eine Sicherung gegen Fälschung der Identität ist nicht immer notwendig, da in manchen Fällen keine Motivation für eine Fälschung vorhanden ist, beispielsweise wenn durch die Fälschung höchstens „erreicht“ werden kann, dass der Fälscher selbst *keine* positive Bewertung mehr erhält. Bei der Beschreibung des Verfahrens zur Verhaltensbeobachtung und -bewertung wird darauf noch näher eingegangen (Abschnitt 3.6.3.1).

- *Verwendung mehrerer Identitäten*

  Dass ein Netzknoten mehrere verschiedene Identitäten besitzt und diese innerhalb des gleichen Zeitraums abwechselnd verwendet, ist grundsätzlich nicht verboten; dies kann sogar sehr sinnvoll sein, um Anonymität zu erreichen.

  Eine Angriffsmöglichkeit ergibt sich aber durch das verwendete Verfahren zur Verhaltensbewertung bezüglich der Weiterleitung: Ein Knoten kann unter einer seiner Identitäten Pakete erzeugen, um sie dann selbst unter einer anderen Identität „weiterzuleiten“ – ohne dass sie bis dahin den physikalischen Netzknoten überhaupt verlassen haben. Tatsächlich handelt es sich hier einfach um das Absenden eigener Pakete, was keine Dienstleistung an der Allgemeinheit ist und deshalb auch nicht bewertet werden soll. Dadurch, dass der Angreifer den ausgesendeten Paketen zwei verschiedene Identitäten einschreibt – eine für die Quelle des Pakets und eine andere für einen weiterleitenden Knoten –, kann aber für Beobachter der Anschein entstehen, dass tatsächlich eine Weiterleitung durchgeführt wurde, die eine positive Bewertung rechtfertigt. Damit würde beim Absenden eigener Pakete jeweils eine Identität des Knotens positiv bewertet werden; durch Rollenwechsel unter den eigenen Identitäten könnten auf die Dauer sämtliche Identitäten positive Bewertungen sammeln, ohne dass eine einzige Dienstleistung für andere Knoten erbracht werden musste.

  Das genannte Problem rührt auch daher, dass einzelne, eine Weiterleitung beobachtende Knoten beurteilen müssen, ob der beobachtete Vorgang der Allgemeinheit nützt. Das ist anhand einzelner Pakete eigentlich nur möglich, wenn der Beobachter genau weiß, aus welchen physikalischen Knoten die Allgemeinheit sich zusammensetzt, und die Pakete diesen Knoten sicher zuordnen kann. Knoten (und nicht Identitäten) sind hier die maßgeblichen Instanzen, weil die erbrachte Leistung im Wesentlichen an der aufgebrachten Sendeenergie gemessen wird, und Energie nur bei Übertragungen zwischen verschiedenen physikalischen Knoten verbraucht wird. Ein Lösungsansatz wäre, nur die Quelle eines jeden Pakets die Weiterleitung ihrer Pakete positiv bewerten zu lassen, denn von der Weiterleitung profitiert letztlich nicht „die Allgemeinheit“, sondern konkret zunächst die Quelle. Wenn allerdings nur die Quelle positiv bewertet, entsteht die Gefahr, dass die Bewertungen insgesamt zu negativ werden, denn negative Bewertungen für Nicht-Weiterleitung müssen von allen Knoten auf einem Übertragungsweg vergeben werden (sonst bleibt die Nicht-Weiterleitung fast immer ungestraft) und können mobilitätsbedingt gelegentlich ungerechtfertigt erfolgen.

  Bei der genauen Beschreibung des Verfahrens zur Verhaltensbeobachtung und -bewertung in Abschnitt 3.6 wird noch näher auf die Problematik eingegangen, und dort werden auch die vorgesehenen Maßnahmen zur Verhinderung des Angriffs beschrieben.

- *Verwendung exzessiv vieler verschiedener Identitäten*

  Eine weitere Angriffsmöglichkeit ergibt sich daraus, dass jeder Knoten Informationen über die Identitäten anderer Teilnehmer sowie diesen zugeordnete Vertrauensprofile speichern muss. Durch ständiges Wechseln der eigenen Identität kann ein Angreifer versuchen, die Speicherkapazität anderer Teilnehmer zu strapazieren, indem diese zum Anlegen immer neuer Einträge angeregt werden.

- *Identitätswechsel*

  Unter den Anforderungen an das Zugangskontrollverfahren wurde in Abschnitt 3.1.2 auch Robustheit gegenüber Umgehung der Zugangskontrolle durch Identitätswechsel gefordert. Deshalb soll an dieser Stelle erwähnt werden, dass eine Umgehung durch Identitätswechsel automatisch ausgeschlossen ist, wenn das Verfahren zur Identifikation gegen Fälschung der Identität gesichert ist, da die Zugangskontrolle immer auf dem Vertrauensprofil basiert, das dem Dienstnutzer mit der Identifikation zugeordnet wird.

3.4.2 Identifikation von Netzknoten

Im Folgenden wird auf die Identifikation der beteiligten Parteien bei der wichtigsten Dienstleistung innerhalb eines Netzwerks eingegangen, nämlich bei der Weiterleitung von Paketen. Die an einem einzelnen Weiterleitungsschritt beteiligten und für Beobachtung und Zugangskontrolle relevanten Parteien sind dabei die Quelle des betrachteten Pakets sowie der jeweilige vorige und nächste Weiterleiter. Diese Parteien müssen von Beobachtern und Weiterleitern identifiziert werden können.

3.4.2.1 Identifikation anhand von Netzwerkadressen

Grundsätzlich kann die Quelle eines beobachteten oder empfangenen Pakets an der von der Vermittlungsschicht (Schicht 3) an jedem Paket angebrachten Quelladresse abgelesen werden. Der vorige und der nächste Weiterleiter werden durch die Absender- bzw. Empfängeradresse auf Schicht 2 angegeben (sofern eine Weiterleitung stattgefunden bzw. stattzufinden hat).

Dabei treten allerdings einige Schwierigkeiten auf: Erstens entstammen die Adressen der Schichten 2 und 3 zwei verschiedenen Adressräumen, zwischen denen zunächst eine Abbildung erfolgen muss, um in beiden Fällen eine einheitliche Knotenidentität ablesen zu können. Zweitens sind die Adressangaben recht einfach zu fälschen. Möchte man das Angeben falscher Absenderadressen mit kryptographischen Methoden verhindern, so muss man jedem Knoten außerdem einen Schlüssel zuordnen, so dass eine weitere Abbildung zwischen Adresse und Schlüssel erforderlich wird. Jede Abbildung muss so durchgeführt werden, dass sie nicht durch Angreifer entsprechend deren Absichten beeinflusst werden kann. Erschwerend kommt schließlich noch hinzu, dass sich die Vermittlungsschicht-Adressen der Knoten bei manchen Adressvergabeverfahren in Abhängigkeit von der Netztopologie verändern.

3.4.2.1.1 Technische Machbarkeit der Fälschung von Netzwerkadressen. Schicht-3-Adressen können in der Regel vom Systemadministrator eines Rechners beliebig gewählt werden, insofern ist eine Fälschung hier leicht durchzuführen. Wie schwierig es ist, Schicht-2-Adressen zu fälschen, hängt von der verwendeten Übertragungstechnologie und der speziellen Ausprägung der Übertragungsadapter ab. Beim verbreiteten Standard IEEE 802.11 für drahtlose lokale Netze kann beispielsweise im Ad-hoc-Modus aus Sicht der Protokolle der unteren Schichten jede beliebige Adresse angegeben werden, ohne dass dies Probleme verursachen oder auffallen würde, und viele Produkte erlauben es auch, die zunächst vom Hersteller vorgegebene, weltweit eindeutige Adresse des Adapters zu verändern. Für das Umfeld wirkt eine Adressänderung so, als wäre ein zusätzlicher Knoten im Empfangsbereich aufgetaucht.

3.4.2.1.2 Verwendung physikalischer Merkmale zur Unterstützung der Identifikation. Theoretisch wäre es denkbar, anhand physikalischer Merkmale der Übertragung – etwa aus Richtung und Feldstärke der empfangenen Signale – festzustellen, ob ein einzelner Knoten unterschiedliche Adressen verwendet. Abgesehen davon, dass aber nicht alle erforderlichen Größen von gängigen Übertragungsadaptern ermittelt werden können, wäre die Positionsabschätzung immer mit Fehlern behaftet, die durch gezielte Manipulation von Angreifern weiter vergrößert werden könnten. Eine zuverlässige Zuordnung empfangener Signale zu sendenden Knoten ist auf diese Weise nicht durchführbar, so dass sie nur aufgrund der empfangenen Nachrichten erfolgen kann.

3.4.2.2 Identifikation per kryptographischem Schlüssel

Zur Sicherung von Identitätsangaben können kryptographische Verfahren verwendet werden, etwa digitale Signaturen über Teile versandter Nachrichten. Damit solche Signaturen von allen Beobachtern zwar prüfbar, aber nicht fälschbar sind, muss ein asymmetrisches kryptographisches Verfahren verwendet werden.

Es bietet sich nun an, als Teilnehmerkennung den öffentlichen Schlüssel heranzuziehen, der zu dem privaten Schlüssel gehört, mit dem die Signaturen erstellt werden. Dies hat den Vorteil, dass bei Vorliegen der Teilnehmerkennung (also des Schlüssels) mit sehr hoher Sicherheit festgestellt werden kann, ob eine Nachricht von diesem Teilnehmer erstellt wurde. Die Identifikation ist also durch eine solche Prüfung praktisch fälschungssicher. Da Schlüssel und Teilnehmerkennung identisch sind, sind keine zusätzlichen Maßnahmen zur Authentisierung des Signaturschlüssels – also zur Prüfung der Zuordnung zwischen Schlüssel und Teilnehmerkennung – (etwa über Zertifikatketten) erforderlich.

Es sei hier ausdrücklich darauf hingewiesen, dass eine Nachricht durchaus nach ihrer Erstellung von einem Angreifer modifiziert und mit einer neuen Signatur versehen werden kann. Dabei kann der Angreifer aber nur einen eigenen Schlüssel verwenden, so dass sich also auch die am Paket ablesbare Teilnehmerkennung ändert, die den Ersteller des Pakets identifiziert: Ersteller ist nach einer solchen Modifikation immer der Angreifer.

Zu beachten ist außerdem, dass die beschriebene Sicherung nur für die Identität des Erstellers einer Signatur gilt. Solche Signaturen können bei der Weiterleitung von Paketen von der Quelle und den Weiterleitern erstellt werden, sobald diese das Paket erhalten haben. Die Identität des Empfängers einer Übertragung ist für Beobachter der Übertragung aber nicht in derselben Weise gesichert; eine solche Sicherung ist logischerweise ausgeschlossen, denn erst nachdem der Empfänger die Nachricht vollständig erhalten hat, die Übertragung also abgeschlossen ist, kann ein Empfänger selbst wissen und damit auch bestätigen, dass er tatsächlich Empfänger ist.

3.4.2.2.1 Schlüsselkennung. Da ein öffentlicher Schlüssel wegen seines Umfangs (z. B. 1024 bit entsprechend 128 byte) nicht mit jeder Nachricht übertragen werden sollte, muss auf andere Weise festgestellt werden, welcher Schlüssel zur Prüfung einer Signatur erforderlich ist.

Eine Möglichkeit zur Reduzierung der zu übertragenden Datenmenge ist die Übertragung einer kürzeren Schlüsselkennung an Stelle des ganzen Schlüssels. Anhand der Schlüsselkennung sollte der Schlüssel entweder innerhalb einer beim Empfänger bereits vorliegenden Menge von Schlüssel identifiziert oder aber mit Hilfe eines Schlüsselverwaltungsverfahrens beschafft werden können.

Eine Schlüsselkennung kann beispielsweise durch eine Hashfunktion aus dem öffentlichen Schlüssel abgeleitet werden. Die Verwendung des Ausgabewerts einer kryptographischen Hashfunktion (Umfang z. B. 160 bit entsprechend 20 byte bei SHA-1) als Schlüsselkennung hat den Vorteil, dass es für Angreifer wegen der für kryptographische Hashfunktionen angenommenen Kollisionssicherheit

Verfahren	*Vor- und Nachteile*
Abbildung Adresse → Schlüssel	+ kein Platzbedarf im Paket - Abbildungsfunktion muss gesichert werden
Schlüsselkennung im Paket	+ geringe Abhängigkeit von Zusatzinformation - besonderes Paketformat erforderlich
lange Kennung (z. B. 160 bit)	+ Schlüsselkennung sicher aus Schlüssel ableitbar - hoher Bandbreitenverbrauch
kurze Kennung (z. B. 16–32 bit)	- zusätzliche Maßnahmen für Eindeutigkeit erforderlich

Tabelle 3.1: Möglichkeiten zur Identifikation von Sendern bzw. Empfängern anhand von Paketen

praktisch unmöglich ist, einen anderen Schlüssel mit derselben Schlüsselkennung zu generieren. Eine solche Schlüsselkennung ist also eindeutig.

Möchte man den Umfang der zu übertragenden Information weiter verringern, so kann man beispielsweise nur einen Teil eines kryptographischen Hashwerts verwenden. Deren Eindeutigkeit ist aber mit geringer werdendem Umfang immer weniger sicher, so dass je nach Verwendungszweck der Schlüsselkennung durch zusätzliche Maßnahmen sichergestellt werden muss, dass anhand einer übertragenen Schlüsselkennung der richtige Schlüssel identifiziert wird. Eine kurze Schlüsselkennung darf eigentlich nur als Hinweis[4] bei der Suche nach einem passenden Schlüssel verstanden werden, wobei die Korrektheit der Wahl anschließend z. B. durch Verifikation einer digitalen Signatur über die Nachricht geprüft werden muss.

3.4.2.3 Auswahl geeigneter Verfahren

Wegen der Vorteile der Verwendung öffentlicher Schlüssel als Teilnehmerkennung wurde dieser Ansatz in dem in der vorliegenden Arbeit vorgestellten Konzept umgesetzt. Jede Nachricht, die mit einer digitalen Signatur versehen ist, kann damit zweifelsfrei einer eindeutigen Teilnehmerkennung zugeordnet werden, nämlich dem öffentlichen Schlüssel, mit dem die Signatur verifiziert werden kann. Durch diesen Schlüssel wird gleichzeitig das dem Teilnehmer zugeordnete Vertrauensprofil identifiziert.

Voraussetzung für diese fälschungssichere Zuordnung ist, dass die Teilnehmerkennung (bzw. der Schlüssel) dem beobachtenden bzw. empfangenden Knoten bekannt ist. Ist dies nicht der Fall, so wird gegebenenfalls das in Abschnitt 3.10 beschriebene Schlüsselverwaltungsverfahren zu ihrer Beschaffung eingesetzt.

Schließlich musste noch eine Methode ausgewählt werden, mit welcher der richtige Schlüssel zur Verifikation einer Nachrichtensignatur aus der Menge der dem verifizierenden Knoten bekannten Schlüssel herausgefunden werden kann. Tabelle 3.1 fasst die Möglichkeiten dafür zusammen, die in den vorangegangenen Abschnitten angesprochen wurden.

Die Verfahren, bei denen eine Schlüsselkennung im Paket übertragen wird, haben zwar den Vorteil, dass keine zusätzliche, von außen beeinflussbare Information benötigt wird, um den passenden Schlüssel anhand der Schlüsselkennung zu finden, da die Schlüsselkennung ja direkt aus den lokal vorliegenden Schlüsseln abgeleitet werden kann. Ihr Nachteil ist aber, dass die Schlüsselkennung die übertragenen Pakete vergrößert, was Übertragungsbandbreite und Energie kostet.

Deshalb wird im vorgeschlagenen Konzept eine Abbildung von Netzwerkadressen auf Schlüssel realisiert, so dass keine zusätzliche Information übertragen werden muss. Sofern die betrachtete Nachricht

[4] Die Bezeichung „Schlüssel*kennung*“ ist deshalb auch nicht mehr ganz treffend.

eine Signatur des zu identifizierenden Teilnehmers trägt, kann die Korrektheit der Abbildung unmittelbar überprüft werden. Bei nicht signierten Nachrichten und bei der Identifikation des Empfängers, die – wie oben schon erwähnt – nicht durch diesen selbst gesichert werden kann, ist es entscheidend, dass die Abbildungsfunktion nicht durch Angreifer manipuliert werden kann, denn damit könnte ohne Manipulation der übertragenen Nachrichten eine falsche Zuordnung zu Teilnehmern erreicht werden. Die Realisierung der Abbildung und die Vorkehrungen zu ihrer Sicherung werden in Abschnitt 3.9 beschrieben.

3.4.3 Identifikation von Benutzern

Die Identität von Benutzern kann nur von anderen Benutzern festgestellt werden, entweder aufgrund persönlicher Bekanntschaft oder etwa anhand eines Lichtbildausweises. Im Anschluss an die Identifikation wird die Identität z. B. in Form eines Namens durch Erstellung eines Zertifikats an einen kryptographischen Schlüssel gebunden, der auf sicherem Weg zum Ersteller des Zertifikats gelangt sein muss. In der Folge kann das Zertifikat als Nachweis dafür verwendet werden, dass diese Identifikation erfolgreich stattgefunden hat. Für eine beliebige andere Partei kann ein Zertifikat genau dann als Identitätsnachweis gelten, wenn sie im Besitz eines authentischen Schlüssels des Erstellers ist und diesem hinsichtlich der korrekten Zertifizierung vertraut. Näheres zum Vorgehen dabei wurde in den Ausführungen zu Authentisierung und Schlüsselverwaltung in Kapitel 2 gesagt (insbesondere Abschnitte 2.3.3.2 und 2.3.4.1).

Der durch ein Zertifikat an die Benutzeridentität gebundene Schlüssel kann nun entweder direkt als Teilnehmerkennung verwendet werden, oder er wird verwendet, um ein Zertifikat für den als Teilnehmerkennung verwendeten Schlüssel auszustellen. Letzteres Vorgehen hat den Vorteil, dass der als Teilnehmerkennung verwendete Schlüssel ausgetauscht werden kann – etwa weil er kompromittiert wurde, oder um die Wahrscheinlichkeit der Kompromittierung durch allzu häufigen Gebrauch zu verringern – ohne dass die sehr aufwändigen Zertifizierungen wiederholt werden müssen.

Durch die Bindung der Teilnehmerkennung an die Identität eines Benutzers wird Missbrauch, der auf der Verwendung unterschiedlicher Teilnehmerkennungen durch einen einzelnen Knoten beruht, zumindest erschwert, weil für jede Teilnehmerkennung mindestens ein unter Mitwirkung von Benutzern erstelltes Zertifikat gebraucht wird. Bei der Beobachtung der Weiterleitung könnte z. B. anhand von Zertifikaten überprüft werden, ob die zu Quelle und vorigem Weiterleiter gehörigen Schlüssel unterschiedlichen Benutzern gehören.

Völlig ausgeschlossen wird Missbrauch so allerdings nicht, denn es ist wahrscheinlich trotzdem möglich, Zertifikate mit verschiedenen Identitäten (für ein und denselben oder auch für verschiedene Schlüssel) zu erhalten, insbesondere wenn sie von verschiedenen „Zertifizierungsinstanzen“, also von verschiedenen anderen Benutzern ausgestellt werden. Ein aufwändiges Vorspielen einer anderen Identität auf Benutzerebene während der Identifikation ist dabei nicht einmal unbedingt erforderlich, denn unter Umständen reicht eine unterschiedliche Schreibung des Namens, das Weglassen eines Vornamens oder eine andere unauffällige Modifikation bei der Angabe des Namens aus, um einen automatischen Test auf bereits existierende Zertifikate zu unterlaufen – falls und insoweit ein solcher bei der Erstellung eines Zertifikats überhaupt durchgeführt wird und werden kann.

Da die Bindung von Teilnehmerkennungen an Benutzeridentitäten also nicht den gewünschten Vorteil bringt, wird ein Verfahren zur Herstellung einer solchen Bindung in der vorliegenden Arbeit nicht vorgesehen. Eine spätere Ergänzung eines solchen Verfahrens wäre voraussichtlich unproblematisch möglich und würde es Benutzern erlauben, Vertrauensbeziehungen zu anderen menschlichen

Abbildung 3.2: Die grau unterlegten Teile jeder Nachricht werden von der Quellsignatur abgedeckt.

Benutzern auch ohne unmittelbaren Kontakt zu diesen auf die Maschinenebene zu übertragen. Allerdings besitzt die Erstellung, Bekanntmachung, Verteilung und Verwendung der zugehörigen Zertifikate auch eine recht hohe Komplexität und verursacht erheblichen Aufwand, der gegen den erwarteten Nutzen abgewogen werden muss.

3.4.4 Nachrichtenauthentisierung

Für die Zugangskontrolle und eine eventuelle Nutzungskontrolle ist es erforderlich, die Quelle eines jeden Pakets authentisch identifizieren zu können. Dies wird ermöglicht, indem alle Pakete bei ihrer Erzeugung (also von ihrer Quelle) digital signiert werden. Jeder andere Knoten, dem ein Paket und ein Schlüssel der Quelle vorliegen, kann die Signatur überprüfen und damit zweifelsfrei feststellen, ob das Paket vom Inhaber des Schlüssels erzeugt und unterwegs nicht manipuliert wurde.

Die Quellsignatur deckt alle Teile einer Nachricht ab, die unterwegs nicht von Vermittlungsschichtinstanzen verändert werden müssen. Insbesondere sind die Quell- und Zieladresse der Vermittlungsschicht sowie die Nutzdaten abgedeckt (siehe Abbildung 3.2).

3.4.4.1 Verbleibende Manipulationsmöglichkeiten

Zu beachten ist zunächst, dass der Quellknoten selbst grundsätzlich jede beliebige Identität als Quelle angeben kann. Dadurch, dass diese Identität durch den Schlüssel repräsentiert wird, mit der die Quellsignatur erzeugt wird, muss der Quellknoten allerdings immer auch den zur Identität gehörigen privaten Schlüssel kennen. Ein einzelner Knoten *kann* sich also mehrere Identitäten zulegen, indem er mehrere Schlüsselpaare erzeugt, und er kann diese in beliebigem Wechsel einsetzen; jeder Identität wird dann ein eigenes Vertrauensprofil zugeordnet. Er kann aber *keine* Identitäten anderer Knoten verwenden, da er deren private Schlüssel nicht kennt.

Ein weiterleitender Knoten kann jederzeit unter Verwendung eines eigenen Schlüssels eine neue Signatur für ein weitergeleitetes Paket generieren, welche die alte ersetzt. Sorgt er außerdem dafür, dass bei der Verifikation der Signatur der von ihm beim Signieren eingesetzte Schlüssel verwendet wird[5], so können andere Knoten die Manipulation (bei der natürlich auch andere Inhalte des Pakets verändert werden können) nicht erkennen – das Paket scheint nun vom Angreifer zu stammen.

Tatsächlich stellt dies aber kein Problem dar, sondern ist sogar so erwünscht: Der Angreifer konnte nicht etwa unbemerkt eine Nachricht eines anderen Knotens verändern, sondern er hat sozusagen dessen Nachricht verworfen (woran ihn selbstverständlich sowieso niemand hindern kann) und selbst eine neue Nachricht erzeugt, die möglicherweise einen ähnlichen Inhalt hat, wie die verworfene. Letzteres spielt aber keine Rolle, da der Empfänger den Inhalt immer in Bezug zur Quelle der Nachricht setzen wird, und diese ist bei der neuen Nachricht ja nachweislich der Angreifer (bzw. eine unter möglicherweise mehreren Identitäten des Angreifers).

[5] Dazu muss er entweder die Quelladressangabe auf eine Adresse ändern, die auf den richtigen Schlüssel abgebildet wird, oder die Abbildungsfunktion manipulieren, so dass die ursprüngliche Adresse auf seinen eigenen Schlüssel abgebildet wird. Dies wird bei der Analyse von Bedrohungen gegen die Schlüsselverwaltung näher untersucht (Abschnitt 3.10.5).

Im Kontext der Beobachtung der Weiterleitung führt das Verwerfen der ursprünglichen Nachricht korrekterweise zu einer negativen Bewertung, und die Aussendung der veränderten Nachricht liefert keine positive Bewertung, da es sich nicht um Weiterleitung sondern Neuerzeugung handelt. In Abschnitt 3.10.4.1 wird dies nochmals genauer beschrieben.

Bezüglich der Zugangskontrolle wird für ein verändertes Paket auch korrekterweise ein Vertrauensprofil des Angreifers herangezogen, das Vertrauensprofil der ursprünglichen Quelle spielt keine Rolle mehr. In Abschnitt 3.9.2.1 wird darauf nochmals eingegangen.

3.4.4.2 Signaturfolgenummer

Zum Schutz vor Wiedereinspielung einmal gesendeter Nachrichten deckt jede erstellte Signatur neben der eigentlichen Nachricht auch eine Signaturfolgenummer ab. Jede Folgenummer darf nur ein einziges mal verwendet werden, deshalb muss der Nummernraum so groß sein, dass er für die gesamte erwünschte Lebensdauer des Schlüssels ausreicht. Damit der Empfänger die Signatur prüfen kann, muss er auch die Signaturfolgenummer kennen. Dies kann sichergestellt werden, indem die Folgenummer jeweils in der signierten Nachricht angegeben wird. Bei einem großen Nummernraum nimmt die Folgenummer dann allerdings relativ viel Raum in der Nachricht ein, so dass bei der Übertragung viel Bandbreite dafür verbraucht wird.

Bei vollkommen zuverlässiger Übertragung könnte in manchen Situation alternativ davon ausgegangen werden, dass der Empfänger die Folgenummer bereits kennt, beispielsweise zwischen Nachbarn, die jede Nachricht des anderen Knotens beobachten und so Sequenznummern selbst mitzählen können. In diesem Fall könnte auf die Übertragung von Sequenznummern ganz verzichtet werden. Gerade in Ad-hoc-Netzen ist eine solche zuverlässige Übertragung allerdings nicht gegeben; auch ein Nachbar kann hier leicht einzelne Pakete oder mobilitätsbedingt auch längere Folgen von Paketen überhören. Außerdem haben weiter entfernte Knoten keine Möglichkeit, gesendete Pakete mitzuzählen, die nicht zufällig auf ihrem Übertragungsweg den Beobachter passieren.

In dem vorgeschlagenen Konzept wird deshalb eine Hybridlösung verwendet, bei der die Länge des übertragenen Anteils einer langen Folgenummer dem Bedarf angepasst werden kann. Dazu führt jeder Knoten intern mit seinem Schlüssel sowie auch mit gespeicherten Schlüsseln anderer Knoten einen Folgenummernzähler, dessen Nummernraum groß genug für lange Schlüssellebensdauern ist. Diese lange Folgenummer des eigenen Schlüssels wird in Signaturen einbezogen und auch jeweils zusammen mit dem Schlüssel verbreitet, während die signierten Pakete nur noch einen kurzen Teil enthalten. Vor der Überprüfung einer Signatur entnimmt der prüfende Knoten den niederwertigen Teil der Folgenummer dem Paket und fügt den fehlenden Teil aus der mit dem Schlüssel erhaltenen Information hinzu. Führt dies nicht zum Erfolg, so erhöht er den höherwertigen Teil um eins und versucht es erneut. Verläuft die Verifikation dann erfolgreich, so kann davon ausgegangen werden, dass ein Überlauf in dem kurzen, im Paket übertragenen Teil stattgefunden hat. Es wird dann der neue Wert mit dem Schlüssel gespeichert. Führt das Erhöhen des nicht übertragenen Teils um wenige Schritte nicht zum Erfolg, so wird der aktuelle Stand des Folgenummernzählers per Schlüsselverwaltung angefordert.

Konkret werden in den Nachrichten der entworfenen Protokolle drei unterschiedliche Längen des übertragenen niederwertigen Teils der Signaturfolgenummer verwendet:

- Bei Nachrichten, die an Nachbarn gerichtet sind (welche die eigene zuletzt verwendete Signaturfolgenummer in der Regel kennen), genügt eine sehr kurze Angabe (ca. 1 byte).

- Bei Nachrichten an weiter entfernte Knoten, mit denen aber regelmäßig kommuniziert wird oder kürzlich kommuniziert wurde, wird ein etwas längerer Teil der Folgenummer angegeben (2-4 byte).

- Wenn davon ausgegangen werden muss, dass die eigene Folgenummer dem Kommunikationspartner noch gar nicht bekannt ist, wird sie in voller Länge angegeben (8-10 byte).

3.5 Repräsentation von Einschätzungen und Vertrauen

Vertrauen in andere Teilnehmer bezüglich bestimmter Vorgänge – beispielsweise bezüglich der korrekten Ausstellung von Zertifikaten – sowie Einschätzungen anderer Teilnehmer bezüglich bestimmter Verhaltensaspekte – beispielsweise bezüglich ihrer Kooperativität – werden bei dem hier vorgestellten Konzept als *Meinungen* der Form $\omega = (b,d,u)$ mit $b,d,u \in [0,1]$ und $b+d+u = 1$ in der Vertrauensmetrik nach Jøsang repräsentiert (siehe Abschnitt 2.3.5.4.2). Insofern Einschätzungen aus der Beobachtung von Verhalten hervorgehen, werden sie zunächst in Form von je zwei Zählern r und s für positive bzw. negative Beobachtungen geführt; die im Beweisraum (siehe Abschnitt 2.3.5.4.1) resultierende Verteilungsdichte $\varphi_{r,s}(\theta)$ wird dann bei Bedarf über die (in Abschnitt 2.3.5.4.3 genannte) Äquivalenzabbildung in den Meinungsraum übertragen.

Durch den Austausch von Meinungen mit anderen Knoten erhält jeder Knoten auch Kenntnis von Meinungen Anderer über Dritte. Jeder Knoten baut damit lokal einen Untergraphen des gesamten Vertrauensgraphen auf, der sich ergeben würde, wenn sämtliche im Netz existierenden Vertrauensbeziehungen zusammengenommen würden.

In diesem Abschnitt wird auf das im Rahmen der vorliegenden Arbeit entwickelte neuartige Verfahren zur Auswertung von Vertrauensgraphen eingegangen, welches es erstmals erlaubt, jederzeit ein aus allen relevanten Vertrauensbeziehungen des gesamten bekannten Untergraphen resultierendes Gesamtvertrauen in bestimmte Knoten zu bestimmen.

3.5.1 Verknüpfung von Einschätzungen

Immer, wenn – etwa als Grundlage für eine Zugangsentscheidung – eine Einschätzung benötigt wird, sollen möglichst alle bekannten und relevanten Vertrauensbeziehungen berücksichtigt und zu einer resultierenden Meinung zusammengefasst werden. Ein Beispiel anhand des in Abbildung 3.3 links dargestellten Vertrauensgraphen: Angenommen, Knoten A benötigt eine Einschätzung zur Kooperativität von Knoten E. Aus eigener Erfahrung hat A sich bisher lediglich eine Meinungen über die Vertrauenswürdigkeit von Einschätzungen des Knotens B gebildet. A kennt aber außerdem Bs Meinung über die Vertrauenswürdigkeit von Einschätzungen Cs und Ds sowie die Meinungen von C und D über die Kooperativität von E. Durch die Verknüpfung der genannten Meinungen kann A eine Einschätzung zur Kooperativität von E erhalten. Keine Rolle spielen die Meinungen von G über D (weil A keine Einschätzung zu Vertrauenswürdigkeit von Einschätzungen Gs erhalten kann) und von C über F (weil A keine von F geäußerten Meinungen bekannt sind, die in irgendeiner Weise mehr Information über E liefern).

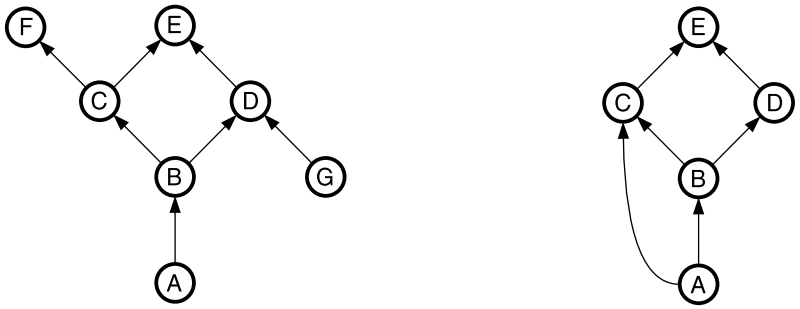


Abbildung 3.3: Beispiele für einfache Vertrauensgraphen

3.5.1.1 Verknüpfung mit Hilfe der subjektiven Logik nach Jøsang

Mit Hilfe der (in Abschnitt 2.3.5.4.4 eingeführten) Operatoren der subjektiven Logik kann die Verknüpfung in der beschriebenen Weise wie folgt dargestellt werden:

$$\omega^A_{E,\text{res}} = \omega^A_B \otimes ((\omega^B_C \otimes \omega^C_E) \oplus (\omega^B_D \otimes \omega^D_E))$$

In der Notation wurde nicht gekennzeichnet, dass ω^A_B, ω^B_C und ω^B_D Meinungen über die Vertrauenswürdigkeit geäußerter Einschätzungen (also Empfehlungsvertrauen) repräsentieren, während ω^C_E und ω^D_E Einschätzungen der Kooperativität darstellen. Wie jedoch leicht nachzuvollziehen ist, müssen immer genau die Meinungen „des letzten Schrittes“, die sich also direkt auf das Ziel der gesuchten resultierenden Meinung beziehen, auch die gesuchte Eigenschaft des Ziels zum Gegenstand haben, während alle übrigen Meinungen jeweils die Vertrauenswürdigkeit von Einschätzungen ihres Ziels betreffen. Wegen dieser Regelmäßigkeit soll zugunsten größerer Übersichtlichkeit auf eine kompliziertere Notation verzichtet werden.

Man beachte, dass die Verknüpfung nur in der angegebenen Reihenfolge entsprechend der Klammerung durchgeführt werden darf. Für die Empfehlungsoperation $\otimes$ und die Konsensoperation $\oplus$ gilt kein Distributivgesetz, d. h. der rechts stehende Term ist nicht äquivalent zu $(\omega^A_B \otimes \omega^B_C \otimes \omega^C_E) \oplus (\omega^A_B \otimes \omega^B_D \otimes \omega^D_E)$. Dies ist so, weil die Konsensoperation nur auf unabhängige Meinungen angewendet werden darf, welche nicht vorlägen, wenn in beide zu verknüpfenden Terme die Meinung ω^A_B einginge. ω^A_B würde dann nämlich sozusagen „doppelt ins Ergebnis eingehen“.

Wenn Meinungen, die auf mehreren möglichen Pfaden durch einen Vertrauensgraphen liegen, mehrfach berücksichtigt würden, könnten Angreifer unter Umständen zusätzliche Pfade gezielt erzeugen, um das Ergebnis zu beeinflussen. Ein Beispiel hierzu zeigt Abbildung 3.4: Wenn der Knoten Y in der links gezeigten Situation zusätzliche Knoten $\mathrm{W}_1 \ldots \mathrm{W}_n$ und zugehörige Meinungen erfindet, so dass die rechts gezeigte Situation entsteht, so hat er damit n neue Pfade durch sich selbst konstruiert. Würden diese unabhängig voneinander berücksichtigt, obwohl sie ω^A_Y gemeinsam haben, so hätte Y damit auch seinen Einfluss auf die resultierende Gesamteinschätzung vervielfacht.

3.5.1.2 Grenzen der Verknüpfbarkeit nach Jøsang

Statt des linken soll nun der rechts in Abbildung 3.3 dargestellte Vertrauensgraph betrachtet werden, der sich aus dem linken Graphen durch Weglassen der beiden irrelevanten Vertrauensbeziehungen und Ergänzen einer zusätzlichen Beziehung zwischen den Knoten A und C ergibt. Für den menschlichen Betrachter ist die neue Situation durchaus nachvollziehbar: A hat selbst eine Meinung zur Vertrauenswürdigkeit von Einschätzungen Cs und berücksichtigt diese neben Bs Meinung zum selben Gegenstand. Bei dem Versuch, dies mit Hilfe der Operatoren der subjektiven Logik auszudrücken, stellt

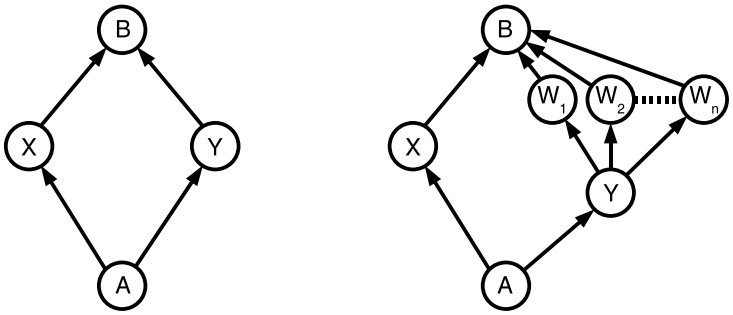


Abbildung 3.4: Knoten Y erzeugt neue Pfade durch erfundene Knoten W_i und sich selbst.

man nun aber fest, dass das nicht möglich ist, ohne dass entweder ω_E^C oder ω_B^A doppelt vorkommt, was jedoch beides nicht erlaubt ist.

Jøsang schreibt für Fälle wie den beschriebenen vor, dass Vertrauensbeziehungen unberücksichtigt bleiben müssen – im Graphen aus Abbildung 3.3 rechts führt das Weglassen einer beliebigen Vertrauensbeziehung auf einen Graphen, der mit Hilfe der subjektiven Logik erfasst werden kann. Es stellt sich dabei allerdings die Frage, wie entschieden werden soll, welche Vertrauensbeziehung wegzulassen ist. In manchen Fällen ist diese Frage zwar relativ leicht zu beantworten, weil das Weglassen einer Vertrauensbeziehung zum Erfolg führt, die aufgrund hoher enthaltener Unsicherheit sowieso nur sehr wenig zum Gesamtergebnis beitragen könnte. In anderen Fällen kann aber jede Entscheidung einen gleich großen Informationsverlust fordern.

Für die automatische Auswertung von Vertrauensgraphen wird ein Verfahren benötigt, das immer ein eindeutiges Ergebnis liefert und dabei alle bekannten Vertrauensbeziehungen berücksichtigt. Da die Auswertung mit Hilfe der subjektiven Logik diese Bedingungen nicht erfüllt, wurde im Rahmen der vorliegenden Arbeit ein neues Verfahren entwickelt, welches im Folgenden vorgestellt wird.

3.5.2 Auswertung von Vertrauensgraphen mit Hilfe von Widerstandsnetzwerken

Da Vertrauen eine menschliche Wahrnehmung ist, welche durch Vertrauensmetriken modelliert werden soll, ist maßgebend beim Entwurf eines Verfahrens zur Verknüpfung von Vertrauensmaßzahlen, dass das Ergebnis der Verknüpfung dem menschlichen Empfinden entspricht (welches allerdings gewisse Spielräume bietet). Dies ist auch bei der subjektiven Logik nach Jøsang der Fall, weshalb ein neues Verfahren sich auf jeden Fall ähnlich verhalten muss. Dass allerdings manche Vertrauensgraphen nicht ausgewertet werden können, stellt einen offensichtlichen Mangel dar, dessen Behebung durchaus graduelle Abweichungen vom Jøsang-Verfahren rechtfertigt.

Um die Vorgänge bei der Verknüpfung vieler Vertrauensbeziehung transparenter zu machen, soll zunächst betrachtet werden, was das Hinzunehmen einzelner Vertrauensbeziehungen bewirkt, und zwar sowohl nach dem menschlichen Empfinden als auch bei Anwendung der entsprechenden Operatoren der subjektiven Logik:

- Fügt man zu einer Kette von aufeinander aufbauenden Meinungen ein weiteres Glied hinzu, so verringert sich in der Regel die Sicherheit der resultierenden Meinung; sie bleibt nur dann gleich, wenn das hinzukommende Glied die Meinung $(1,0,0)$ repräsentiert, also absolutes Vertrauen ausdrückt.

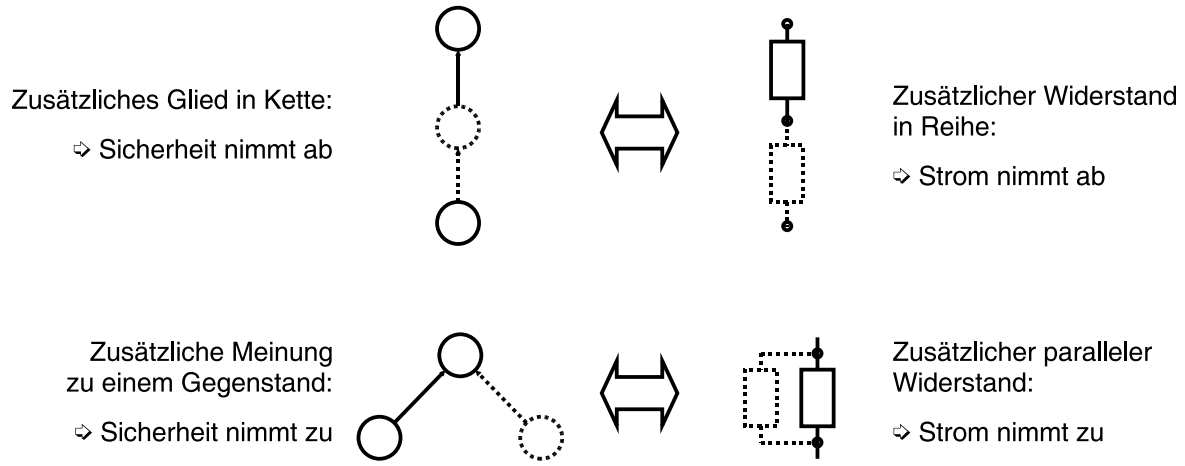


Abbildung 3.5: Analogie zwischen Vertrauen und elektrischem Widerstand

Genauer gesagt reduzieren sich die b- und die d-Komponenten des Resultats in gleichem Maße und antiproportional zur Größe der b-Komponente des zusätzlichen Gliedes. Die u-Komponente des Resultats erhöht sich entsprechend.

- Zusätzliche unabhängige Meinungen zum selben Gegenstand erhöhen in der Regel die Sicherheit der resultierenden Meinung; sie bleibt nur dann gleich, wenn die zusätzliche Meinung absolut unsicher ist, entsprechend dem Wert $(0, 0, 1)$.

  Genauer gesagt erhöhen b- und d-Komponenten der zusätzlichen Meinung auch die b- bzw. d-Komponente des Ergebnisses (jeweils in geringerem Maß). Die u-Komponente des Ergebnisses wird entsprechend reduziert.

In ähnlicher Weise verhalten sich elektrische Ströme durch Widerstandsnetzwerke, wenn zusätzliche Widerstände hinzugefügt werden (siehe Abbildung 3.5):

- Das Hinzufügen eines zusätzlich in Reihe geschalteten Widerstand zu einer Reihenschaltung verringert den resultierenden Strom.
- Das Hinzufügen eines zusätzlich parallel geschalteten Widerstands zu einer Parallelschaltung erhöht den resultierenden Strom.

Das Anlegen einer Spannung an zwei beliebige Punkte eines Widerstandsnetzwerks führt immer zu einem Strom, der von der Größe aller Widerstände abhängt, die auf zyklenfreien Pfaden zwischen den beiden Punkten liegen.

Die Ähnlichkeit legt nahe, dass die Modellierung von Vertrauensgraphen durch Widerstandsnetzwerke zu einem Verfahren zur Berechnung eines eindeutigen resultierenden Wertes für das Vertrauen zwischen beliebigen Knoten eines Vertrauensgraphen führen könnte.

3.5.2.1 Abbildung von Vertrauensgraphen auf Widerstandsnetzwerke

Um einen Vertrauensgraphen auf ein Widerstandsnetzwerk abzubilden, wird wie folgt vorgegangen:

- Jedem Knoten des Vertrauensgraphen, also jeder Instanz, die Vertrauen in andere hat oder der vertraut wird, entspricht ein Knotenpunkt im Widerstandnetzwerk.

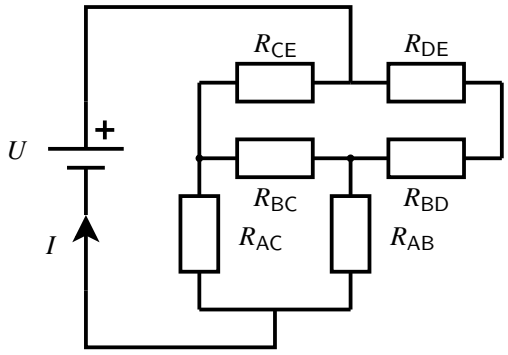


Abbildung 3.6: Widerstandsnetzwerk zum Vertrauensgraphen aus Abbildung 3.3 rechts

- Jede Kante zwischen zwei Knoten X und Y im Vertrauensgraphen wird repräsentiert durch einen Widerstand R_{XY} zwischen den X und Y entsprechenden Knotenpunkten im Widerstandsnetzwerk. Der Widerstandswert ist dabei – grob gesagt – reziprok zur Größe des Vertrauens, das durch die Beziehung ausgedrückt wird; genauere Angaben hierzu folgen weiter unten.

- Eine Spannungsquelle mit einem festen Spannungswert wird an die beiden Knotenpunkte des Widerstandsnetzwerks angeschlossen, die den auswertenden Knoten und den Zielknoten der Auswertung repräsentieren.

Abbildung 3.6 zeigt als Beispiel das Schaltbild des Widerstandsnetzwerks, das sich bei Anwendung dieser Regeln auf den in Abbildung 3.3 rechts dargestellten Vertrauensgraphen ergibt.

Mit Hilfe bekannter, auf den Kirchhoffschen Gesetzen [Kirc45] basierenden Verfahren wird der Strom durch das so konstruierte Widerstandsnetzwerk berechnet. Daraus lässt sich auf einen resultierenden Widerstandswert R_{eff} für das gesamte Widerstandsnetzwerk schließen. Dieser Widerstandswert wird nun wieder in eine Vertrauensmaßzahl zurücktransformiert, wobei die Inverse zu derjenigen Abbildung verwendet wird, mit der vorher die Widerstandswerte bestimmt wurden. Die so erhaltene Vertrauensmaßzahl beschreibt das resultierende Vertrauen im Vertrauensgraphen.

Widerstandswerte sind skalare Größen, deshalb sind sowohl die in ein konstruiertes Widerstandsnetzwerk eingehenden als auch die als Resultat erhaltenen Werte zunächst eindimensional. Vertrauen kann jedoch durch eindimensionale Werte nur unzureichend modelliert werden, weshalb die Meinungen nach Jøsang auch zweidimensionale Werte (bzw. äquivalent dazu) sind. Es stellt sich deshalb nun die Frage nach einem Vorgehen, durch welches die volle Informationsmenge zweidimensionaler Meinungen mit Hilfe eines (oder evtl. mehrerer) eindimensionale Werte verknüpfenden Widerstandsnetzwerks verarbeitet werden kann.

3.5.2.2 Der Meinungsraum Π

Bei der Suche nach einer geeigneten Abbildung zwischen Vertrauensmaßzahlen und Widerstandswerten stellt man fest, dass die drei skalaren Komponenten b, d und u der Meinungen aus dem Meinungsraum Ω wegen ihrer gegenseitigen Abhängigkeit über die Beziehung $b+d+u=1$ schwer handzuhaben sind. Aus diesem Grund soll zunächst ein weiterer Raum eingeführt werden – bezeichnet als Meinungsraum Π –, der zum Meinungsraum Ω äquivalent ist, dessen Elemente aber voneinander unabhängige Komponenten haben.

Meinungen in Π haben die Form $\pi = (p,c)$ mit $p,c \in [0,1]$. Zwischen dem Meinungsraum Ω und dem Meinungsraum Π wird eine bijektive Abbildung $g : \Omega\backslash(0,0,1) \rightarrow \Pi\backslash\{(p,c) \in \Pi | c = 0\}$ erklärt durch

$$g(\omega) = \pi \quad \text{mit} \quad \omega = (b,d,u), \pi = (p,c) \quad \text{und} \quad \begin{cases} p = \frac{b}{b+d} \\ c = 1-u = b+d \end{cases}$$

Dass der Wert $(0,0,1) \in \Omega$ (völlige Unsicherheit) keine eindeutige Entsprechung in Π hat, sondern durch jeden Wert $(p,c) \in \Pi$ mit $c = 0$ dargestellt werden kann, stellt in der Praxis kein Problem dar.

Die Werte $(p,c) \in \Pi$ sollen folgendermaßen interpretiert werden: Die Komponente p wird als *Position* bezeichnet und gibt an, wo die beschriebene Meinung zwischen den Extremen völligen Vertrauens ($(1,0,0) \in \Omega$, entspricht $p = 1$) und völligen Misstrauens ($(0,1,0) \in \Omega$, entspricht $p = 0$) angesiedelt ist. Die Komponente c, die als *Sicherheit (Certainty)* bezeichnet werden soll, stellt ein Gegenstück zur Unsicherheitskomponente u in Ω dar: Je größer c, desto sicherer ist sich der Inhaber der Meinung.

3.5.2.2.1 Empfehlung und Konsens im Meinungsraum Π. Die Operatoren der subjektiven Logik lassen sich mit Hilfe der Abbildungsfunktion g auf den Meinungsraum Π übertragen.

Die *Empfehlungsoperation*, welche die Meinung $\pi_B^A = (p_B^A, c_B^A)$ der Instanz A bezüglich Aussagen der Instanz B verknüpft mit der Meinung $\pi_x^B = (p_x^B, c_x^B)$ der Instanz B bezüglich einer Aussage x liefert dann eine Meinung $\pi^{AB} = \pi_B^A \otimes \pi_x^B$ durch

$$\pi^{AB} = (p^{AB}, c^{AB}) \quad \text{mit} \quad \begin{cases} p^{AB} = p_x^B \\ c^{AB} = p_B^A c_B^A c_x^B \end{cases} . \tag{3.1}$$

Dass die p-Komponente des Resultats einer Empfehlungsoperation nur von der p-Komponente des zweiten Operanden abhängt, erweist sich für den Entwurf des widerstandsnetzwerksbasierten Auswertungsverfahrens als sehr vorteilhaft: Wie oben bereits angesprochen kann unmittelbar anhand der Ströme in einem Widerstandnetzwerk bzw. anhand des Ersatzwiderstands des Netzwerks nur eine skalare Größe ermittelt werden. Da die p-Komponente des Resultats einer Empfehlung offensichtlich gar nicht vom ganzen umgebenden Vertrauensgraphen abhängt, entsteht hier schon die Idee, das Widerstandsnetzwerk hauptsächlich zur Berechnung der c-Komponente einzusetzen. Die p-Komponenten der Kanten des Vertrauensgraphen müssen dann allerdings wegen des Vorkommens von p_B^A in der Bestimmungsgleichung für c^{AB} trotzdem in die Widerstandswerte des Netzwerks eingehen.

Der *Konsens* $\pi^{A,B} = \pi^A \oplus \pi^B$ zweier unabhängiger Meinungen $\pi^A = (p^A, c^A)$ und $\pi^B = (p^B, c^B)$ verschiedener Instanzen A und B über dieselbe Aussage berechnet sich zu

$$\pi^{A,B} = (p^{A,B}, c^{A,B}) \quad \text{mit} \quad \begin{cases} p^{A,B} = \frac{p^A c^A (1-c^B) + p^B c^B (1-c^A)}{c^A + c^B - 2c^A c^B} \\ c^{A,B} = \frac{c^A + c^B - 2c^A c^B}{1 - c^A c^B} \end{cases} . \tag{3.2}$$

Die c-Komponente des Resultats einer Konsensoperation hängt also nur von den c-Komponenten der Operanden ab. Auch dies ist günstig für die Idee, mit Hilfe des Widerstandsnetzwerks die c-Komponenten des Resultats zu bestimmen. Lediglich die Abhängigkeiten der beim Konsens resultierenden p-Komponenten von den c-Komponenten der Operanden erscheint noch etwas problematisch. Bevor in Abschnitt 3.5.2.4 das entworfene Verfahren zur Auswertung von Vertrauensgraphen mit Hilfe von Widerstandsnetzwerken im Detail beschrieben wird, soll die Bestimmung der p-Komponente beim Konsens noch etwas näher untersucht werden.

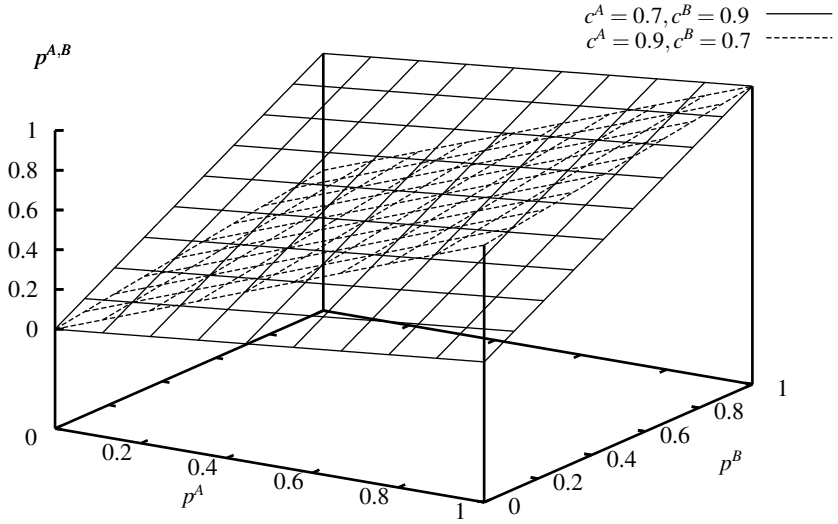


Abbildung 3.7: Abhängigkeit der p-Komponente einer Konsensoperation von den Eingangs-p-Komponenten

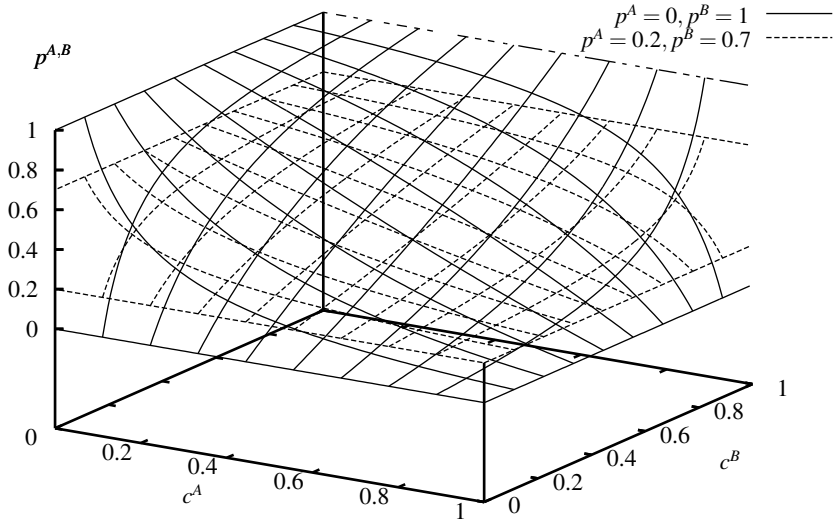


Abbildung 3.8: Abhängigkeit der p-Komponente einer Konsensoperation von den Eingangs-c-Komponenten

3.5.2.2.2 Veranschaulichung der Konsens-Operation im Meinungsraum Π. In Abbildung 3.7 wird das Verhalten der p-Komponente des Ergebnisses einer Konsensoperation in Abhängigkeit von den p-Komponenten der eingehenden Meinungen bei festen Werten für deren c-Komponenten gezeigt ($c^A = 0.7, c^B = 0.9$ durchgezogen, $c^A = 0.9, c^B = 0.7$ gestrichelt). Es ist anhand der Linearität der Graphen anschaulich zu erkennen, dass die p-Komponente des Ergebnisses einfach eine gewichtete Summe der Eingangs-p-Komponenten darstellt: Die Summe der Gewichte, die in Gleichung (3.2) als Faktoren bei p^A und p^B stehen, ist Eins. Eine eingehende p-Komponente wird umso höher gewichtet, je sicherer die zugehörige Meinung und je unsicherer die andere eingehende Meinung ist.

Abbildung 3.8 zeigt das Verhalten der p-Komponente des Ergebnisses einer Konsensoperation in Abhängigkeit von den c-Komponenten der eingehenden Meinungen bei festen Werten für deren p-Komponenten ($p^A = 0, p^B = 1$ durchgezogen, $p^A = 0.2, p^B = 0.7$ gestrichelt). Wenn A sich völlig unsicher oder B sich absolut sicher ist – also für $c^A = 0$ oder $c^B = 1$ (linke bzw. hintere Kante des Graphen) –, ist die Position des Ergebnisses identisch mit der von B: $p^{A,B} = p^B$. Ist sich umgekehrt A absolut sicher, B sich aber völlig unsicher – also für $c^A = 1$ oder $c^B = 0$ (rechte bzw. vordere Kante des Graphen) –, so stimmt die Position des Ergebnisses einleuchtenderweise mit der von A überein:

$p^{A,B} = p^A$. Dazwischen ist der Graph stetig und monoton in beiden Parametern. Lediglich wenn sich sowohl A als auch B völlig sicher oder völlig unsicher sind – also an den Punkten $c^A = c^B = 0$ und $c^A = c^B = 1$ (linke vordere bzw. rechte hintere Ecke) –, ist die Ergebnisposition nicht definiert.

3.5.2.3 Bestimmung der Widerstandswerte

Fließende Ströme durch Widerstände innerhalb des Widerstandsnetzwerks sollen „genutztes Vertrauen“ der zugehörigen Kanten des Vertrauensgraphen modellieren, also Vertrauen, das auch für das resultierende Gesamtvertrauen des Graphen relevant ist. Der Strom, der durch einen bestimmten Widerstand fließen kann und mit dem er also zum Gesamtstrom durch ein Widerstandsnetzwerk beitragen kann, ist schon dadurch beschränkt, dass ihm andere Widerstände auf dem Weg zur Quelle vorangehen bzw. auf dem Weg zur Senke nachfolgen. Dadurch wird modelliert, dass eine bestimmte Vertrauensbeziehung innerhalb eines Vertrauensgraphen nur dann für das resultierende Gesamtvertrauen relevant sein kann, wenn sowohl die vorangehenden Vertrauensbeziehungen auf dem Weg vom auswertenden Knoten her als auch die nachfolgenden bis hin zum Ziel der Auswertung ein entsprechend hohes Vertrauen rechtfertigen. Der durch die vorigen Widerstände beschränkte maximale Strom, der noch durch einen Widerstand fließen kann, entspricht also im Modell in etwa dem Vertrauen, das aufgrund dorthin führender anderer Vertrauensbeziehungen in den Inhaber einer bestimmten Meinung gesetzt wird und das resultierende Vertrauen in den Gegenstand der geäußerten Meinung beschränkt.

Während im vorigen Absatz allgemein von „Vertrauen“ als Entsprechung zu Strömen in Widerstandsnetzwerken die Rede war, liegt es nach den Überlegungen in Abschnitt 3.5.2 zur Analogie zwischen Verknüpfungen von Vertrauensbeziehungen und Widerstandsschaltungen, bei denen die *Sicherheit* der betrachteten Meinungen eine Rolle spielte, nahe, die c-Komponente von Meinungen als Entsprechung von Strömen aufzufassen.

Widerstände wirken Strömen entgegen, deshalb soll hohe Sicherheit einer Vertrauensbeziehung einem Widerstandswert nahe Null entsprechen, eine geringe Sicherheit dagegen einem sehr hohen Widerstandswert. Wir setzen deshalb vorläufig

$$R = \left(\frac{1}{c} - 1\right) \Omega,$$

was wunschgemäß zu $R = 0\,\Omega$ für $c = 1$ und $R \to \infty\,\Omega$ für $c \to 0$ führt.

Aus der angelegten Spannung U_0, die zweckmäßigerweise auf den konstanten Wert 1 V gesetzt wird, und dem resultierenden Strom I_{res} im Netzwerk lässt sich nach Auswertung des Widerstandsnetzwerks der resultierende Gesamtwiderstand R_{res} berechnen zu $R_{\text{res}} = \frac{U_0}{I_{\text{res}}} = \frac{1\,\text{V}}{I_{\text{res}}}$. Für c_{res} ergibt sich durch Umkehrung der obigen Transformation

$$c_{\text{res}} = \frac{1}{\frac{R_{\text{res}}}{1\,\Omega} + 1} = \frac{1\,\Omega}{R_{\text{res}} + 1\,\Omega} = \frac{1}{\frac{1\,\text{A}}{I_{\text{res}}} + 1} = \frac{I_{\text{res}}}{I_{\text{res}} + 1\,\text{A}}.$$

Da die Hintereinanderschaltung von Widerständen einer Empfehlungsoperation entsprechen soll, für welche die Beziehung aus Gleichung 3.1 gilt, muss nun allerdings nicht nur die c-, sondern auch die p-Komponente in den Widerstandswert eingehen, weshalb als endgültige Vorschrift zur Berechnung des Widerstandswerts die im folgenden Abschnitt vorgestellte verwendet wird.

3.5.2.4 Verfahren zur Auswertung

Das folgende Verfahren kann zur Berechnung des aus einem Vertrauensgraphen resultierenden Gesamtvertrauens zwischen dem auswertenden Knoten und dem Zielknoten verwendet werden. Für die Beschreibung wird davon ausgegangen, dass der Graph nur solche Kanten enthält, die auf einem zyklenfreien Weg zwischen auswertendem Knoten und Zielknoten liegen.

1. Stelle nach dem in Abschnitt 3.5.2.1 beschriebenen Verfahren ein Widerstandsnetzwerk zum auszuwertenden Vertrauensgraphen auf.

2. Jeder Kante $X \to Y$ des Vertrauensgraphen wird eine Variable $\bar{c}_Y^X$ zugeordnet, die mit dem Wert c_Y^X aus der Meinung $\pi_Y^X = (p_Y^X, c_Y^X)$ initialisiert wird.

3. Bestimme den Wert jedes einer Kante $X \to Y$ des Vertrauensgraphen zugeordneten Widerstands R_{XY} als
$$R_{XY} = \left(\frac{1}{\bar{p}_X c_Y^X} - 1 \right) \Omega. \tag{3.3}$$
Dabei stammt c_Y^X aus der zur Kante $X \to Y$ gehörigen Meinung $\pi_Y^X = (p_Y^X, c_Y^X)$, und $\bar{p}_X$ wird so bestimmt:

   (a) Wenn der Knoten X des Vertrauensgraphen der auswertende Knoten ist, so ist $\bar{p}_X = 1$.

   (b) Wenn der Knoten X des Vertrauensgraphen nur einen einzelnen Vorgänger V hat und zu der Kante $V \to X$ die Meinung $\pi_X^V = (p_X^V, c_X^V)$ gehört, so ist $\bar{p}_X = p_X^V$.

   (c) Wenn der Knoten X des Vertrauensgraphen mehrere Vorgänger $V_1 \ldots V_n$ hat, wobei den zugehörigen Kanten die Meinungen $\pi_X^{V_1} \ldots \pi_X^{V_n}$ mit $\pi_X^{V_i} = (p_X^{V_i}, c_X^{V_i})$ zugeordnet sind, so berechnet sich $\bar{p}_X$ als
$$\bar{p}_X = \frac{\sum_i p_X^{V_i} \bar{c}_X^{V_i} \prod_{j \neq i} (1 - \bar{c}_X^{V_j})}{\sum_i \bar{c}_X^{V_i} \prod_{j \neq i} (1 - \bar{c}_X^{V_j})}.$$
   Dies ist eine Verallgemeinerung der Gleichung (3.2) auf beliebig viele am Konsens teilnehmende Meinungen, die man findet, indem man das Ergebnis einer Konsensoperation über zwei Meinungen sukzessive mit weiteren Meinungen verknüpft.
   Bezogen auf die zu den $\bar{c}_X^{V_i}$ gehörigen Ströme I_{V_iX} (aus denen die $\bar{c}_X^{V_i}$ nach Formel 3.4 (siehe unten) im vorigen Iterationsschritt berechnet wurden), entspricht das
$$\bar{p}_X = \frac{\sum_i p_X^{V_i} I_{V_iX}}{\sum_i I_{V_iX}},$$
   also einer nach den Strömen gewichteten Summe der p-Komponenten der Meinungen der Vorgänger.

4. Berechne alle Ströme im Widerstandsnetzwerk. Bestimme für jede Kante $X \to Y$ anhand des Stroms I_{XY} durch den Widerstand R_{XY} einen neuen Wert
$$\bar{c}_Y^X := \frac{I_{XY}}{I_{XY} + 1\,\mathrm{A}}. \tag{3.4}$$
   Wenn sich eines der $\bar{c}_Y^X$ dadurch nennenswert ändert, beginne wieder bei Schritt 3.

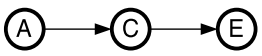


Abbildung 3.9: Vertrauensgraph zu einer einfachen Empfehlung

5. Die resultierende Meinung $\pi_{\text{res}} = (p_{\text{res}}, c_{\text{res}})$ ergibt sich aus

$$c_{\text{res}} = \frac{I_{\text{res}}}{I_{\text{res}} + 1\,\text{A}} = \frac{1\,\text{V}}{R_{\text{res}} + 1\,\text{V}} \tag{3.5}$$

$$p_{\text{res}} = \bar{p}_X \tag{3.6}$$

wobei I_{res} der Gesamtstrom durch das Netzwerk bzw. R_{res} dessen Ersatzwiderstand sowie X der Zielknoten ist. $\bar{p}_X$ wird wie in Schritt 3 berechnet.

3.5.2.5 Beispiel

Für die Auswertung des Beispielgraphen aus Abbildung 3.3 rechts werden folgende Widerstände benötigt: R_{AB}, R_{BC}, R_{BD}, R_{CE}, R_{DE}, R_{AC}. Sie können in folgender Reihenfolge bestimmt werden:

1. Berechne R_{DE} aus vorgegebenen p_D^B und c_E^D, denn hier liegt eine reine Empfehlung vor.
2. R_{CE} berechnet sich (Empfehlung) aus $\bar{p}_C$ und c_E^C. Dabei stammt $\bar{p}_C$ aus dem Konsens zwischen den Meinungen von A und B bezüglich C, in den die festen p_C^A und p_C^B sowie die Variablen für $\bar{c}_C^A$ bzw. $\bar{c}_C^B$ eingehen.
3. R_{BC} berechnet sich aus den vorgegebenen p_B^A und c_C^B (Empfehlung).
4. Ebenso berechnet sich R_{BD} aus den vorgegebenen p_B^A und c_D^B.
5. R_{AB} wird aus dem vorgegebenen c_B^A bestimmt („triviale Empfehlung“).
6. Ebenso bestimmt sich R_{AC} aus dem vorgegebenen c_C^A.

3.5.3 Vergleich der Verfahren anhand einfacher Vertrauensgraphen

Anhand der Ergebnisse bei der Auswertung einfacher Vertrauensgraphen wird im Folgenden das neue Verfahren zur Auswertung über Widerstandsnetzwerke mit der subjektiven Logik nach Jøsang verglichen.

3.5.3.1 Einfache Empfehlung

Abbildung 3.9 zeigt einen Vertrauensgraphen zur Auswertung einer Meinung des Knotens C über Knoten E (die auch als „Empfehlung“ bezeichnet werden soll, obwohl sie nicht unbedingt positiv sein muss) durch Knoten A, der dazu seine eigene Meinung über C einbezieht.

Nach Gleichung (3.1) hängt bei der Empfehlungsoperation der subjektiven Logik die p-Komponente des Ergebnisses nur von der p-Komponente der Meinung von C über E ab, es gilt also $p_{\text{res/Jøsang}} = p_E^C$ unabhängig von π_C^A und c_E^C. Dasselbe trifft für die p-Komponente bei der Auswertung über Widerstandsnetzwerke zu, da Knoten C nur einen Vorgängerknoten im Vertrauensgraphen hat und damit dieselbe Vorschrift zu dessen Berechnung angewendet wird (Schritt 3b in Abschnitt 3.5.2.4). Es ist hier also $p_{\text{res/Jøsang}} = p_{\text{res/Netz}}$.

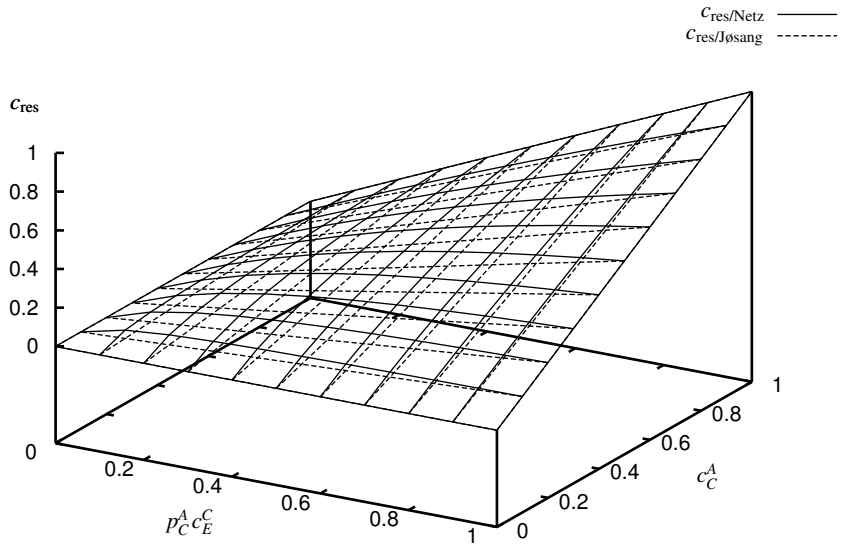


Abbildung 3.10: c-Resultate einer Empfehlung mit Widerstandsnetzwerk (durchgezogen) und subjektiver Logik (gestrichelt)

In der c-Komponente unterscheiden sich jedoch die Ergebnisse. In der subjektiven Logik errechnet sie sich nach Gleichung (3.1) zu

$$c_{\text{res/Jøsang}} = c_C^A \, p_C^A \, c_E^C,$$

während die Reihenschaltung der nach Gleichung (3.3) bestimmten Widerstände $R_{AC} = (\frac{1}{c_C^A} - 1)\,\Omega$ und $R_{CE} = (\frac{1}{p_C^A c_E^C} - 1)\,\Omega$ im Widerstandsnetzwerkmodell zu

$$c_{\text{res/Netz}} = \frac{1\,\Omega}{R + 1\,\Omega} = \frac{1\,\Omega}{R_{AC} + R_{CE} + 1\,\Omega} = \frac{c_C^A \, p_C^A \, c_E^C}{(1 - c_C^A)\, p_C^A \, c_E^C + c_C^A}$$

führt.

Abbildung 3.10 zeigt Graphen der Resultate $c_{\text{res/Jøsang}}$ (gestrichelt) und $c_{\text{res/Netz}}$ (durchgezogen) über den Parametern c_C^A und $p_C^A c_E^C$. Das Maximum der Differenz

$$c_{\text{res/Netz}} - c_{\text{res/Jøsang}} = \frac{(1 - c_C^A)\,(1 - p_C^A c_E^C)\, c_C^A \, p_C^A \, c_E^C}{(1 - c_C^A)\, p_C^A \, c_E^C + c_C^A}$$

ist $\frac{10\sqrt{5}-22}{4} \approx 0.090$ und wird angenommen, wo $c_C^A = p_C^A c_E^C = \frac{3-\sqrt{5}}{2} \approx 0.382$ gilt.

3.5.3.2 Ketten von Empfehlungen

Für Vertrauensgraphen mit einer zusätzlichen Empfehlung entsprechend Abbildung 3.11 links bleibt die Übereinstimmung zwischen subjektiver Logik und Widerstandsnetzwerkmodell bezüglich der p-Komponente des Ergebnisses erhalten. Für die c-Komponente gilt nach der subjektiven Logik

$$c_{\text{res/Jøsang}} = p_C^B \, p_B^A \, c_B^A \, c_C^B \, c_E^C.$$

Abbildung 3.11: Vertrauensgraphen zu Ketten mit 3 bzw. 5 Vertrauensbeziehungen

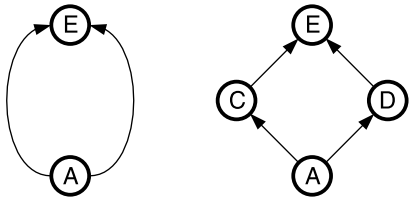


Abbildung 3.12: Vertrauensgraph zum einfachen Konsens; links mit mehrfachen Kanten, rechts Ersatzgraph bei $\pi_C^A = \pi_D^A = (1,1)$

Die Reihenschaltung der Widerstände $R_{AB} = (\frac{1}{c_B^A} - 1)\,\Omega$, $R_{BC} = (\frac{1}{p_B^A c_C^B} - 1)\,\Omega$ und $R_{CE} = (\frac{1}{p_C^B c_E^C} - 1)\,\Omega$ ergibt im Widerstandsnetzwerkmodell

$$\begin{aligned} c_{\text{res/Netz}} &= \frac{1\,\Omega}{R + 1\,\Omega} = \frac{1\,\Omega}{R_{AB} + R_{BC} + R_{CE} + 1\,\Omega} \\ &= \frac{p_B^A p_C^B c_B^A c_C^B c_E^C}{(1 - c_B^A)\, p_B^A p_C^B c_C^B c_E^C + c_B^A (1 - (1 - p_B^A c_C^B)(1 - p_C^B c_E^C))}. \end{aligned}$$

Die Differenz $c_{\text{Netz}} - c_{\text{Jøsang}}$

$$= -\frac{\left(\left(2\,{c_B^A}^2 - c_B^A\right) {c_C^B}^2 {c_E^C}^2 {p_B^A}^2 - {c_B^A}^2 c_C^B c_E^C {}^2 p_B^A\right) {p_C^B}^2 + \left(c_B^A c_C^B c_E^C p_B^A - {c_B^A}^2 {c_C^B}^2 c_E^C {p_B^A}^2\right) p_C^B}{\left(\left(2\,c_B^A - 1\right) c_C^B c_E^C p_B^A - c_B^A c_E^C\right) p_C^B - c_B^A c_C^B p_B^A}$$

nimmt in den Punkten mit $c_B^A = p_B^A c_C^B = p_C^B c_E^C = 0.5$ ihr Maximum von 0.125 an. Mit einem weiteren Kettenglied steigt das Differenzmaximum auf ca. 0.143.

Eine solche moderate Abweichung zwischen den mit der subjektiven Logik und dem neuen Widerstandsnetzwerkmodell bestimmten Werten für das Gesamtvertrauen bei Ketten von Empfehlungen erscheint durchaus tolerierbar. In der Praxis dürften längere Ketten sowieso selten auftreten, da mit jedem zusätzlichen Glied die Forderung erheblich schwerer zu erfüllen wird, dass *alle* Meinungen einer Kette ein relativ hohes Vertrauen repräsentieren müssen, damit das Endergebnis noch eine nutzbare Sicherheit aufweist.

3.5.3.3 Einfacher Konsens

Der einfachste Vertrauensgraph, bei dem ein Konsens zwischen zwei Meinungen bestimmt werden muss, ist links in Abbildung 3.12 dargestellt. Um mehrfache Kanten zwischen zwei Knoten zu vermeiden, kann man ihn durch den rechten Graphen ersetzen, wobei man die Meinungen π_C^A und π_D^A fest auf den Wert $(1,1)$ für volles Vertrauen setzt und π_E^C und π_E^D variiert.

Die Auswertung im Widerstandsnetzwerkmodell liefert exakt dasselbe Ergebnis wie diejenige mit Hilfe der subjektiven Logik. Für die Widerstandswerte setzt man nämlich

$$\begin{aligned} R_{CE} &= \frac{1\,\Omega}{p_C^A c_E^C} - 1\,\Omega = \frac{1\,\Omega}{c_E^C} - 1\,\Omega \\ R_{DE} &= \frac{1\,\Omega}{p_D^A c_E^D} - 1\,\Omega = \frac{1\,\Omega}{c_E^D} - 1\,\Omega \end{aligned}$$

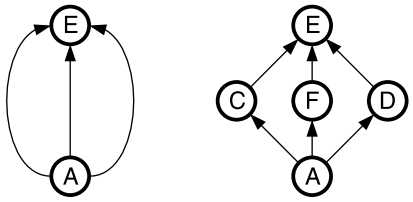


Abbildung 3.13: Vertrauensgraph zum Konsens mehrerer Meinungen

sowie gegebenenfalls $R_{AC} = R_{AD} = 0\,\Omega$. Der Ersatzwiderstand des Netzwerks berechnet sich bei der dem Vertrauensgraphen entsprechenden Parallelschaltung von R_{CE} und R_{DE} zu

$$R_{\text{res}} = \frac{R_{CE} R_{DE}}{R_{CE} + R_{DE}}.$$

Als c-Komponente des resultierenden Gesamtvertrauens erhält man

$$c_{\text{res/Netz}} = \frac{1\,\Omega}{R + 1\,\Omega} = \frac{c_E^C + c_E^D - 2c_E^C c_E^D}{1 - c_E^C c_E^D},$$

was genau der Berechnungsvorschrift (3.2) der Konsensoperation entspricht. Und auch die p-Komponente wird in Schritt 3b des Verfahrens aus Abschnitt 3.5.2.4 genau entsprechend der Konsensdefinition der subjektiven Logik berechnet. Wegen

$$\begin{aligned} \bar{c}_E^C &= \frac{I_{CE}}{I_{CE} + 1\,\text{A}} = \frac{1\,\Omega}{R_{CE} + 1\,\Omega} = c_E^C \\ \bar{c}_E^D &= \frac{I_{DE}}{I_{DE} + 1\,\text{A}} = \frac{1\,\Omega}{R_{DE} + 1\,\Omega} = c_E^D \end{aligned}$$

ändern sich in Schritt 4 des Verfahrens auch die in die Berechnung der p-Komponente eingehenden Gewichte $\bar{c}_E^C$ und $\bar{c}_E^D$ nicht gegenüber ihren Startwerten c_E^C bzw. c_E^D, weshalb keine weitere Iteration erforderlich ist.

3.5.3.4 Konsens mehrerer Meinungen

Abbildung 3.13 zeigt einen Vertrauensgraphen, in dem ein Konsens über drei Meinungen gefunden werden muss, sowie einen Ersatzgraphen zur Vermeidung mehrfacher Kanten, bei dem wieder fest $\pi_C^A = \pi_F^A = \pi_D^A = (1,1)$ gesetzt wird. Es zeigt sich, dass auch in diesem Fall und auch beim Konsens von mehr als drei Meinungen die Ergebnisse aus Widerstandsnetzwerkmodell und subjektiver Logik exakt übereinstimmen. Der Nachweis gestaltet sich sehr ähnlich zum dem im vorigen Abschnitt und soll hier nicht ausführlich dargestellt werden.

3.5.4 Auswertung problematischer Vertrauensgraphen

Ziel bei der Entwicklung des Widerstandsnetzwerksverfahrens zur Auswertung von Vertrauensgraphen war es, die Grenzen der subjektiven Logik bezüglich der Auswertbarkeit von Graphen zu überwinden. Dass das Verfahren grundsätzlich in der Lage ist, jeden beliebigen Graphen auszuwerten, folgt bereits aus seinem Entwurf, denn jeder Vertrauensgraph lässt sich in ein Widerstandsnetzwerk überführen und dort lässt sich immer ein Gesamtstrom berechnen. Am Beispiel des schon in Abschnitt 3.5.1.2 gezeigten Vertrauensgraphen aus Abbildung 3.14 (a), der mittels der subjektiven Logik

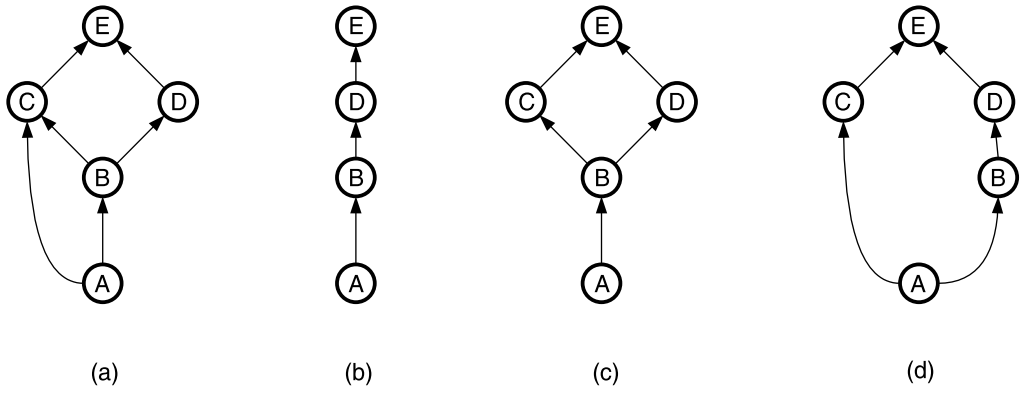


Abbildung 3.14: Problematischer Vertrauensgraph und unproblematische Vergleichsgraphen

nicht auswertbar ist, soll nun überprüft werden, wie sich das Widerstandsnetzwerksverfahren bei solchen Graphen verhält. Das Widerstandsnetzwerk zum Graphen aus Abbildung 3.14 (a) war schon in Abbildung 3.6 gezeigt worden.

Zunächst soll das Verhalten bei „extremen“ Werten für das durch die Kante $A \rightarrow C$ ausgedrückte Vertrauen betrachtet werden.

- Für $\pi_C^A = (0,1)$, also sicheres Misstrauen von A gegenüber C, ergibt sich $R_{AC} = 0\,\Omega$; durch diesen Kurzschluss im Widerstandsnetzwerk wird R_{BC} irrelevant. Außerdem ist $\bar{p}_C = p_C^A = 0$, was zu $R_{CE} = \infty\,\Omega$ führt. Der ganze linke Zweig des Vertrauensgraphen wird damit bedeutungslos. Der verbleibende Rest ist in Abbildung 3.14 (b) dargestellt. Er ist auch mit Hilfe der subjektiven Logik auswertbar, und das Ergebnis dabei stimmt mit dem mittels des Widerstandnetzwerk erhaltenen tendenziell überein (siehe Abschnitt 3.5.3.2).

- Für $c_C^A = 0$, also völlige Unsicherheit bei der Meinung von A gegenüber C, ergibt sich $R_{AC} = \infty\,\Omega$. Die Kante $A \rightarrow C$ wird damit bedeutungslos, und es verbleibt der Graph aus Abbildung 3.14 (c), der in der in Abschnitt 3.5.1.1 beschriebenen Weise mittels subjektiver Logik auswertbar ist. Die Ergebnisse stimmen auch hier tendenziell überein.

- Für $\pi_C^A = (1,1)$, also sicheres Vertrauen von A gegenüber C, ergibt sich wie im ersten Fall $R_{AC} = 0\,\Omega$ und R_{BC} wird irrelevant. Der verbleibende, in Abbildung 3.14 (d) dargestellte Graph ist wieder mittels subjektiver Logik auswertbar. Er ähnelt dem aus Abbildung 3.9 rechts; lediglich durch die Empfehlungskette im rechten Zweig ergibt sich eine leichte Abweichung zum Ergebnis der Auswertung mit dem Widerstandsnetzwerk.

3.5.5 Ähnliche Entwicklungen

Parallel zur Entstehung der vorliegenden Arbeit [KrSc04] wurde die Idee, Bayes-basierte zweidimensionale Vertrauensmaßzahlen zur Repräsentation von gegenseitigen Einschätzungen in Ad-hoc-Netzen zu verwenden, auch in anderen Ansätzen aufgegriffen.

So haben auch die Entwickler des CONFIDANT-Systems (Abschnitt 2.4.6.3) erkannt, dass die Verwendung sowohl positiver als auch negativer Beobachtungen eine wesentlich bessere Bewertung erlaubt, und sie kommen in [BuLB04] zu einer Einschätzungsrepräsentation ähnlich derjenigen im Meinungsraum Π. Weiterhin ermitteln sie auch Empfehlungsvertrauen durch Vergleich der von anderen geäußerten Meinungen mit eigenen Einschätzungen und beheben damit eine grundlegende Schwäche

des ursprünglichen CONFIDANT-Systems, das keine wirkungsvolle Beurteilung der Glaubwürdigkeit von Meldungen anderer erlaubte. Allerdings wird nur eine Stufe der Indirektion erlaubt, Ketten von Vertrauensbeziehungen oder gar ganze Vertrauensgraphen können also nicht betrachtet und der damit verbundene Informationsgewinn nicht genutzt werden.

Ein Verfahren zur Auswertung von Vertrauensgraphen wird in [ThBa04] beschrieben. Die Auswertung ist relativ einfach, aber als Voraussetzung ist erforderlich, dass der Meinungsraum mit der Konsens- und der Empfehlungsoperation einen Halbring darstellt. Dazu muss für Empfehlung und Konsens ein Distributivgesetz gelten; nach den Ausführungen in Abschnitt 3.5.1.1 ist diese Annahme nicht sinnvoll.

3.6 Beobachtung und Bewertung

Das Verhalten eines Knotens bei der Weiterleitung von Paketen liefert eine gute Grundlage für die Einschätzung von dessen Bereitschaft, Dienstleistungen im Interesse der Allgemeinheit zu erbringen; hauptsächlich aufgrund dieser Einschätzung wird im Rahmen der Zugangskontrolle die Entscheidung darüber getroffen, ob der Knoten selbst den Weiterleitungsdienst und damit das Netzwerk überhaupt nutzen darf. Deshalb behandelt der größte Teil des folgenden Abschnitts recht ausführlich das Verfahren zur Beobachtung und Bewertung des Weiterleitungsverhaltens anderer Knoten. Erst abschließend (Abschnitt 3.6.6) wird noch kurz auf Beobachtung und Bewertung einer anderen Klasse von Diensten eingegangen.

Ein erster Ansatz zur Beobachtung des Weiterleitungsverhaltens in Ad-hoc-Netzen wurde bereits früher [MGLM00] für Netze entworfen, die das Wegfindungsverfahren DSR verwenden, und in ähnlicher Form mittlerweile mehrmals aufgegriffen [BuLB02, HeWK04, BaBa03]. Dabei bewahrt eine „Watchdog“ genannte Komponente Kopien aller selbst gesendeten Pakets auf und startet mit der Aussendung jeweils einen Zeitgeber. Wird bis zu dessen Ablauf keine Weiterleitung des Pakets beobachtet, so erfolgt eine Fehlermeldung an die Quelle, die in der Folge zum Ausschluss des betroffenen Weges führt. Ein Nachteil dieses Ansatzes ist, dass nur ein Bruchteil der verfügbaren Information verwendet wird, da weder mitgehörte Weiterleitungsvorgänge noch die Wegfindungsvorgänge überwacht werden. Außerdem werden nur negative Beobachtungen berücksichtigt. Diese Nachteile besitzt das in diesem Abschnitt vorgestellte verbesserte Verfahren nicht.

Nach einigen allgemeinen Bemerkungen zu Beobachtung und Bewertung wird im Folgenden zunächst kurz die Idee des Verfahrens zur Beobachtung der Weiterleitung vorgestellt, bevor im Anschluss analysiert wird, ob und auf welche Weise das Verfahren getäuscht und wie dies gegebenenfalls verhindert werden kann. Danach folgt die detaillierte Beschreibung des Verfahrens, in der auch die getroffenen Sicherungsmaßnahmen enthalten sind.

Zur Klärung der Begriffe in den Ausführungen dieses Abschnitts sind in Abbildung 3.15 einige an der Übertragung eines Pakets beteiligte Knoten dargestellt und aus der Sicht eines Beobachters bezeichnet. An dem ebenfalls dargestellten Auszug aus dem beobachteten Paket ist zu erkennen, in welcher Weise den Adressangaben der Netzwerkprotokolle der Schichten 2 und 3 auf die bezeichneten Knoten hinweisen.

3.6.1 Allgemeines zur Beobachtung und Bewertung

Ein allgemeiner Grundsatz für die Verfahren zur Beobachtung und Bewertung ist, dass durch sie kein zusätzlicher Energieaufwand und keine zusätzliche Netzbelastung entstehen dürfen. Insbesondere dürfen nicht zum Zweck der Bewertung zusätzliche Pakete verschickt werden. Zu rechtfertigen

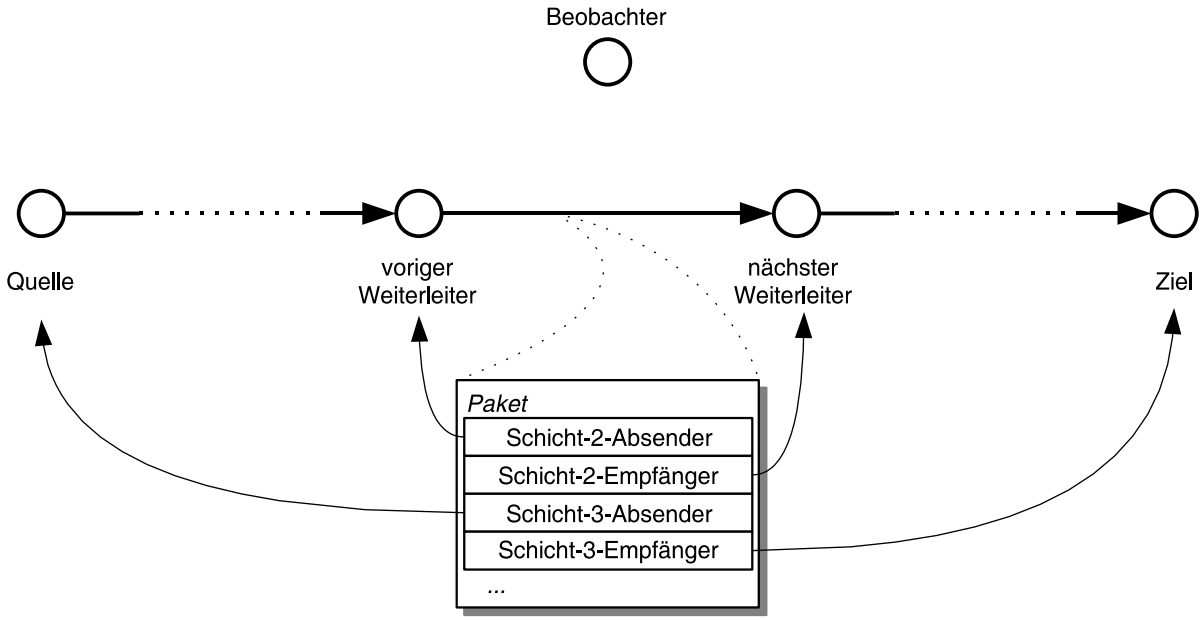


Abbildung 3.15: Bezeichnung der an der Übertragung eines Pakets beteiligten Knoten und der Adressangaben im Paket

wäre dies höchstens, wenn die Bewertung sehr wichtig wäre und extreme Folgen wie etwa den sofortigen Ausschluss nach sich zöge. Solche Fälle treten bei den hier beschriebenen Verfahren nicht auf.

Nachweislich absichtlich durchgeführte Angriffe dürfen zwar grundsätzlich zu einer starken generellen Abwertung führen (d. h. nicht nur bezogen auf eine einzelne Kategorie des Verhaltens des Angreifers), denn wenn ein Teilnehmer gezielt Angriffe durchführt, gibt es kaum einen Grund, ihm noch in irgendeiner Hinsicht zu vertrauen. Häufig enthalten Beobachtungen jedoch Unsicherheiten, die einen sicheren Nachweis der Absicht verhindern, und in diesem Fall bewirkt eine moderate negative Bewertung, dass im Fall des Irrtums kein allzu großer Schaden angerichtet wird, während sich bei wiederholten Angriffen mit der Zeit doch ein deutliches Bild abzeichnet.

3.6.2 Idee des Verfahrens

Jeder Knoten hält für jeden seiner Nachbarn ein Paar von Zählern, mit denen er Beobachtungen korrekten bzw. unkorrekten Weiterleitungsverhaltens registriert. Mittels der Transformation aus dem Beweis- in den Meinungsraum kann aus den beiden Zählerständen jederzeit eine Vertrauensmaßzahl berechnet werden, die eine Einschätzung des Weiterleitungsverhaltens des jeweiligen Nachbarn darstellt.

Jeder Knoten, der beobachtet, wie ein anderer Knoten ein Paket weiterleitet, erhöht seinen diesem weiterleitenden Knoten zugeordneten Zähler für positive Weiterleitungsbeobachtungen um eins (siehe Abbildung 3.16; Beobachter ist Knoten B).

Um auch den Fall erfassen zu können, dass ein Knoten die Weiterleitung verweigert, wird außerdem für jedes beobachtete oder selbst ausgesandte Paket, das noch von einem Nachbarn weitergeleitet werden muss, ein Eintrag in einer Liste derzeit beobachteter Pakete angelegt. Zu jedem solchen Eintrag wird ein Zeitgeber gestartet, der abläuft, wenn eine vorgegebene Zeitspanne vergangen ist, die normalerweise für die Weiterleitung eines Pakets ausreicht. Wird während der Laufzeit des Zeitgebers die erwartete Weiterleitung beobachtet, so werden der Zeitgeber und der zugehörige Eintrag gelöscht.

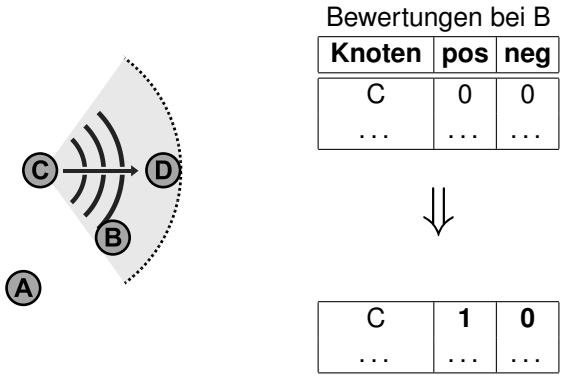


Abbildung 3.16: Positive Bewertung von Knoten C durch Knoten B bei beobachteter korrekter Weiterleitung

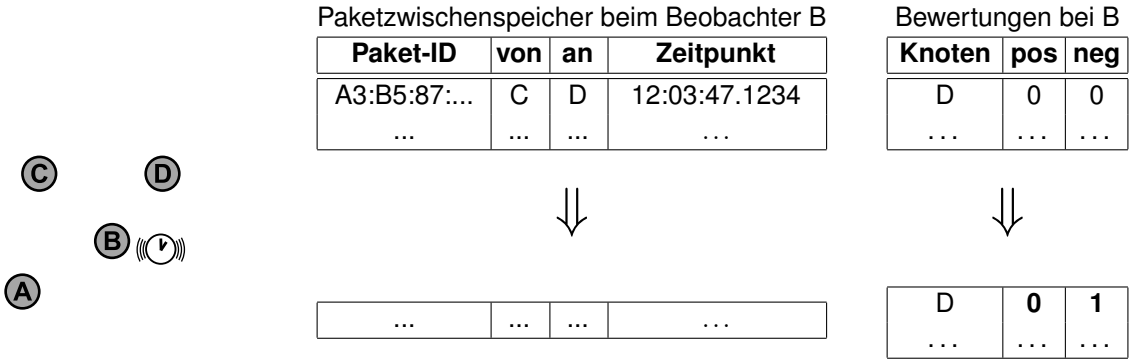


Abbildung 3.17: Negative Bewertung von D durch B bei beobachteter Nichtweiterleitung

Läuft der Zeitgeber aber ab, ohne dass die erwartete Weiterleitung beobachtet werden konnte, so wird dem säumigen Weiterleiter eine negative Beobachtung angerechnet (siehe Abbildung 3.17).

Außerdem Einfluss auf die Weiterleitung hat die Zugangskontrolle: Da die Zugangsentscheidung von jedem Knoten aufgrund seiner lokalen Vertrauensdatenbasis getroffen wird, ist für Beobachter zunächst nicht vorhersehbar oder erkennbar, wie diese Entscheidung ausfällt. Beobachter können so also angesichts von Fällen unterlassener Weiterleitung nie entscheiden, ob die Weiterleitung berechtigterweise aufgrund einer Zugangsentscheidung oder unberechtigterweise aufgrund unkooperativer Verhaltensweise unterlassen wurde. Schon aus diesem Grund muss jeder Knoten, der eine negative Zugangsentscheidung trifft, diese durch Aussenden einer entsprechenden Fehlermeldung bekanntgeben. Wenn eine solche Fehlermeldung beim Beobachter eintrifft, wird das zugehörige Paket aus der Liste beobachteter Pakete gelöscht (siehe Abbildung 3.18).

3.6.2.1 Rolle der Adressabbildung und der Weiterleitungsinformation

Bei dem soeben beschriebenen Beobachtungs- und Bewertungsverfahren werden an zwei Stellen Informationen über Adressen benötigt, die nicht aus dem beobachteten Paket entnommen werden können. Hier muss die in Abschnitt 3.10 beschriebene Abbildungsfunktion zwischen Schicht-3- und Schicht-2-Adressen zu Rate gezogen werden; in manchen Fällen kann stattdessen auch die von der Wegfindung ermittelte Weiterleitungsinformation (siehe Abschnitt 2.1.4) herangezogen werden. Gemeint sind die beiden folgenden Entscheidungen:

- Wenn ein Sendevorgang beobachtet wird, muss zunächst entschieden werden, ob es sich um einen Weiterleitungsvorgang handelt, oder ob die beobachtete Nachricht von ihrem ursprünglichen Erzeuger (also ihrer Quelle) ausgesandt wurde.

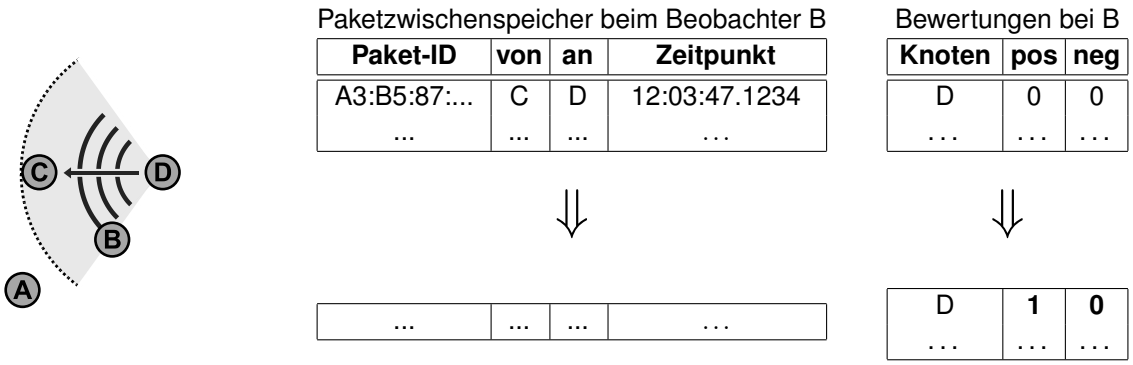


Abbildung 3.18: Fehlermeldung wegen nicht erfolgter Weiterleitung durch D

Dies kann erkannt werden, indem die Schicht-3-Absenderadresse aus dem Paket mit Hilfe der Adressabbildungsfunktion auf eine Schicht-2-Adresse abgebildet wird. Wenn diese mit der im Paket angegebenen Schicht-2-Absenderadresse übereinstimmt, wurde das Paket von seiner Quelle versandt, es hat also keine Weiterleitung stattgefunden.

- Außerdem muss ein Beobachter erkennen, ob ein beobachtetes Paket nochmals weitergeleitet werden muss, oder ob es am Ziel angekommen ist.

  Dazu kann die Empfängeradresse der Schicht 3 aus dem Paket entnommen und mit Hilfe der Adressabbildungsfunktion auf eine Schicht-2-Adresse abgebildet werden. Wenn diese mit der im Paket angegebenen Schicht-2-Empfängeradresse übereinstimmt, ist das Paket am Ziel angekommen und braucht also nicht weiter beobachtet zu werden.

  Wenn der Beobachter selbst Weiterleiter ist, das Paket also auch auf Schicht 2 an ihn adressiert war, benötigt er keine Adressabbildung zur Beantwortung dieser Frage. Denn bei der Weiterleitung ermittelt er sowieso anhand der ihm vorliegenden Weiterleitungsinformation (die vom Wegfindungsverfahren ermittelt wurde) die Schicht-3-Adresse des nächsten Knotens, so dass er die Übereinstimmung mit der im Paket angegebenen Zieladresse direkt überprüfen kann. Um das Paket dann allerdings tatsächlich weiterzuleiten, muss er die Adressabbildung anschließend noch durchführen.

3.6.3 Bedrohungsanalyse

Im Folgenden wird analysiert, welche grundsätzlichen Möglichkeiten zur Beeinflussung bei der Beobachtung und Bewertung nach dem oben grob umrissenen Verfahren bestehen, und inwieweit Angreifer diese nutzen können, um sich Vorteile zu verschaffen.

3.6.3.1 Angriffsziele und -motivation

Bei dem Versuch, die Resultate des Verfahrens zur Verhaltensbeobachtung und -bewertung zu verfälschen, kommen für Angreifer folgende Ziele in Frage:

1. Beschaffung positiver Bewertungen für den Angreifer oder einen „Komplizen“ (in der Regel auf Gegenseitigkeit), wobei

   (a) der für reguläre positive Bewertung zu erbringende Aufwand (Dienstleistung an der Allgemeinheit) entfällt oder

(b) die sich durch mangelnde Nachfrage nach zu erbringenden Dienstleistungen ergebende Einschränkung (etwa dadurch, dass keine Pakete zur Weiterleitung zur Verfügung stehen) wegfällt.

Eine Art natürlicher Beschränkung für die Beschaffung ungerechtfertigter positiver Bewertungen der erstgenannten Art besteht dadurch, dass immer nur bei Beobachtung eines tatsächlich gesendeten Pakets eine positive Bewertung erfolgt. Dass positive Bewertungen völlig ohne Aufwand (d. h. ohne Energieverbrauch durch Senden) massenhaft erzeugt werden können, ist damit schon ausgeschlossen. Es kann also höchstens erreicht werden, dass ein normalerweise nicht positiv bewerteter Sendevorgang (wie die Erzeugung eines eigenen Pakets) als Dienstleistung an der Allgemeinheit (z. B. Weiterleitung) erscheint.

2. Verhindern negativer Bewertung des Angreifers oder eines „Komplizen“. Dabei darf allerdings nicht mehr Aufwand entstehen, als durch die negativ zu bewertende Unterlassung eingespart wird.

3. Erzeugung negativer Bewertungen oder Verhindern positiver Bewertung für andere Knoten, entweder

   (a) um die Verfügbarkeit des Netzes (für bestimmte oder für alle Knoten) zu beeinträchtigen oder

   (b) damit sich eine eigene schlechte Bewertung weniger von denen anderer Knoten abhebt.

   Angriffe gegen die Verfügbarkeit sind in Ad-hoc-Netzen schwierig zu bekämpfen, weil bei genügend hoher Motivation und mit genügend hohem Aufwand sowieso einfach jede Kommunikation zumindest in der Nachbarschaft des Angreifers durch Störsignale unterbunden werden kann. Ob der Angriffsversuch (b) Erfolg haben kann, hängt vom Zugangskontrollverfahren ab.

Die unter 1. und 2. genannten Angriffsmotivationen dürften vermutlich die größte Gruppe potentiell interessierter Angreifer anziehen, denn Möglichkeiten, sich selbst Vorteile zu verschaffen, ohne dabei durch allzu offensichtliche Benachteiligung anderer auffällig zu werden, werden erfahrungsgemäß immer gerne genutzt.

3.6.3.2 Möglichkeiten der Einflussnahme durch Angreifer

Die Möglichkeiten eines Angreifers umfassen ganz allgemein die Unterschlagung, Zerstörung oder Verfälschung bewertungsrelevanter Information in Nachrichten anderer, die Wiedereinspielung fremder Nachrichten und die Angabe falscher bewertungsrelevanter Information in eigenen Nachrichten. Einige dieser Möglichkeiten schließen sich durch das verwendete Beobachtungs- und Bewertungsverfahren von vornherein aus:

- Da nur Nachrichten von Nachbarn für die Beobachtung von Bedeutung sind, fällt die Unterschlagung (welche nur bei erforderlicher Weiterleitung möglich wäre) weg.

- Die gezielte Zerstörung oder Verfälschung bestimmter bewertungsrelevanter Informationen innerhalb fremder Nachrichten ist technisch nur in den seltensten Fällen realisierbar, weil aufgrund der Eigenschaften der Signalausbreitung eine gezielte Überlagerung einer fremden Übertragung nur sinnvoll möglich ist, wenn sich der Angreifer zwischen Sender und Empfänger befindet. Normalerweise gibt es aber mehrere benachbarte Knoten, die alle Beobachtungen

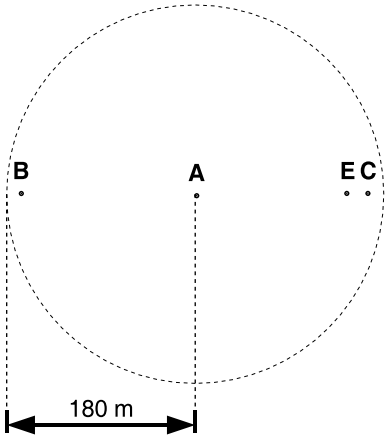


Abbildung 3.19: Beispielszenario für versuchte Manipulation übertragener Nachrichten

durchführen, und nur der geringste Anteil dieser Knoten dürfte in der Regel vom Sender aus gesehen ungefähr hinter dem Angreifer liegen.

Ein Beispiel: Bei einer Datenrate von 1 Mbit/s dauert die Übertragung eines Bits ca. $1\,\mu$s, und damit besitzt ein Bit bei seiner Übertragung eine ungefähre räumliche Ausdehnung von $10^{-6}\,\mathrm{s} \cdot 3 \cdot 10^{8}\,\mathrm{m/s} = 300\,\mathrm{m}$ in Ausbreitungsrichtung. Beispielsweise in der in Abbildung 3.19 dargestellten Situation, bei der eine für die Übertragungstechnik IEEE 802.11 realistische Sendereichweite angenommen wurde, kann ein Angreifer E damit ein von A ausgesandtes Paket nicht für die Empfänger B und C in gleicher Weise manipulieren, denn ein von E ausgesandtes Signal erreicht B mehr als $1\,\mu$s später als C, verändert also ein anderes Bit der von A ausgesandten Nachricht. Bei höheren Datenraten sinkt die Ausdehnung eines Bits entsprechend, und die für den Angreifer ungünstigen Verschiebungen treten schon bei kürzeren Entfernungen zwischen Angreifer und Sender auf.

- Die Falschangabe bewertungsrelevanter Information ist dann sinnlos, wenn der einzige Zweck der betreffenden Information darin liegt, dem Sender eine positive Bewertung zuzuordnen.

Es bleiben die Möglichkeiten der vollständigen Zerstörung fremder Nachrichten, der Wiedereinspielung fremder Nachrichten und der Falschangabe bewertungsrelevanter Information in eigenen Nachrichten (evtl. auch mit dem Ziel, die Nachricht so erscheinen zu lassen, als sei sie von einem anderen Knoten erzeugt worden). Diese Möglichkeiten werden bei der nachfolgenden Bedrohungsanalyse berücksichtigt.

3.6.3.3 Bewertungsrelevante Information für die positive Bewertung

Im Folgenden wird betrachtet, welche Informationen als Voraussetzung für eine positive Bewertung vorliegen müssen, welche Vorteile sich ein Angreifer durch ihre Manipulation beschaffen kann, und welche Gegenmaßnahmen in Frage kommen.

- *Identität des Weiterleiters*

  Die Identität des vorigen Weiterleiters wird benötigt, um das ihm zugeordnete Vertrauensprofil identifizieren zu können, so dass dort eine positive Beobachtung vermerkt werden kann.

  Angriffsmotivation: Nicht vorhanden, da eine falsche Angabe die positive Bewertung dessen, der die Angabe erstellt hat, verhindern würde.

Vorgehen: Die Identität des Weiterleiters wird dem weitergeleiteten Paket entnommen: Entweder muss dort die Schlüsselkennung des Weiterleiters enthalten sein, oder der Beobachter muss von der Schicht-2-Adresse auf die Identität schließen können. Eine Sicherung ist nicht erforderlich.

- *Existenz und Urheberschaft der Quelle und Aktualität des Pakets*

  Der Beobachter sollte sicher sein, dass der als Quelle des Pakets angenommene Knoten tatsächlich existiert und das Paket unmittelbar zuvor erzeugt und abgesendet hat. Denn positive Bewertungen dürfen nur für tatsächlich erbrachte Dienstleistungen vergeben werden, und eine solche liegt nicht vor, wenn der angebliche Weiterleiter das Paket selbst erzeugt und einen existierenden oder einen ebenfalls erfundenen Knoten als Quelle angegeben hat oder wenn er ein früher bereits übertragenes Paket nochmal wiederholt.

  Angriffsmotivation: Mittel. Der „Weiterleiter“ erschleicht sich eine positive Bewertung auf Kosten eines ansonsten sinnlosen Sendevorgangs. Dies könnte für ihn höchstens dann interessant sein, wenn nicht genügend echte Pakete zur Weiterleitung zur Verfügung stehen, so dass es sehr lange dauern würde, die eigene Bewertung zu verbessern.

  Wäre die erschlichene positive Bewertung der einzige Effekt des Angriffs, so könnte man ihn unter Umständen ignorieren, da der „Angreifer“ immerhin Energie dafür aufwenden muss. Der Gesamtaufwand zum Aufbau einer guten Einschätzung im ganzen Netz wäre damit schon recht hoch, was als Abschreckung gegen einen Identitätswechsel genügen könnte – und dies ist die primäre Anforderung, die durch das Beobachtungs- und Bewertungsverfahren erfüllt werden soll. Durch das Erzeugen sinnloser Pakete oder die Wiederholung alter Pakete, die in der Folge bis zu ihrem Ziel weitergeleitet werden, werden aber auch andere Knoten belastet, deshalb sollten solche Angriffe verhindert werden.

  Vorgehen: Es gibt zwei Möglichkeiten für Beobachter, die genannten Angriffe auszuschließen:

  - Werte nur positiv, nachdem eine von der Quelle am Paket angebrachte Signatur überprüft wurde.

    Durch die Signaturprüfung ist die Urheberschaft der Quelle sichergestellt. Die Aktualität des Pakets muss zusätzlich geprüft werden, etwa indem die Quelle alle von ihr erzeugten Pakete fortlaufend nummeriert und jeder Beobachter sich die zuletzt gesehene Sequenznummer jeder Quelle merkt; bewertet werden nur Pakete mit größeren Sequenznummern.

    Das Verfahren ist recht anspruchsvoll: Neben der Erfordernis des Vorhandenseins einer Signatur der Quelle muss auch der Schlüssel der Quelle beim Beobachter vorliegen; ist dies nicht der Fall, kann keine Bewertung vorgenommen werden. Beobachter explizit mit Hilfe des Schlüsselverwaltungsdienstes versuchen zu lassen, sich den benötigten Schlüssel zu beschaffen, erscheint erstens zu aufwändig für die Beobachter und kann außerdem eine starke Zusatzbelastung für das Netzwerk verursachen. Deshalb sollte auf diese Möglichkeit verzichtet werden.

    Die Existenz der Quelle muss durch das Schlüsselverwaltungsverfahren sichergestellt werden – kann ein Angreifer einfach einen zusätzlichen Schlüssel generieren und diesen den Beobachtern zukommen lassen, so kann er auch die Existenz einer Quelle mit diesem Schlüssel vortäuschen. Um das zu verhindern oder zumindest zu erschweren, müssen die Schlüssel an Benutzeridentitäten gebunden werden (siehe Abschnitt 3.4.3), was wiederum die Beschaffung und Prüfung einer Zertifikatkette durch den Beobachter erfordert, falls eine solche nicht schon vorgenommen wurde.

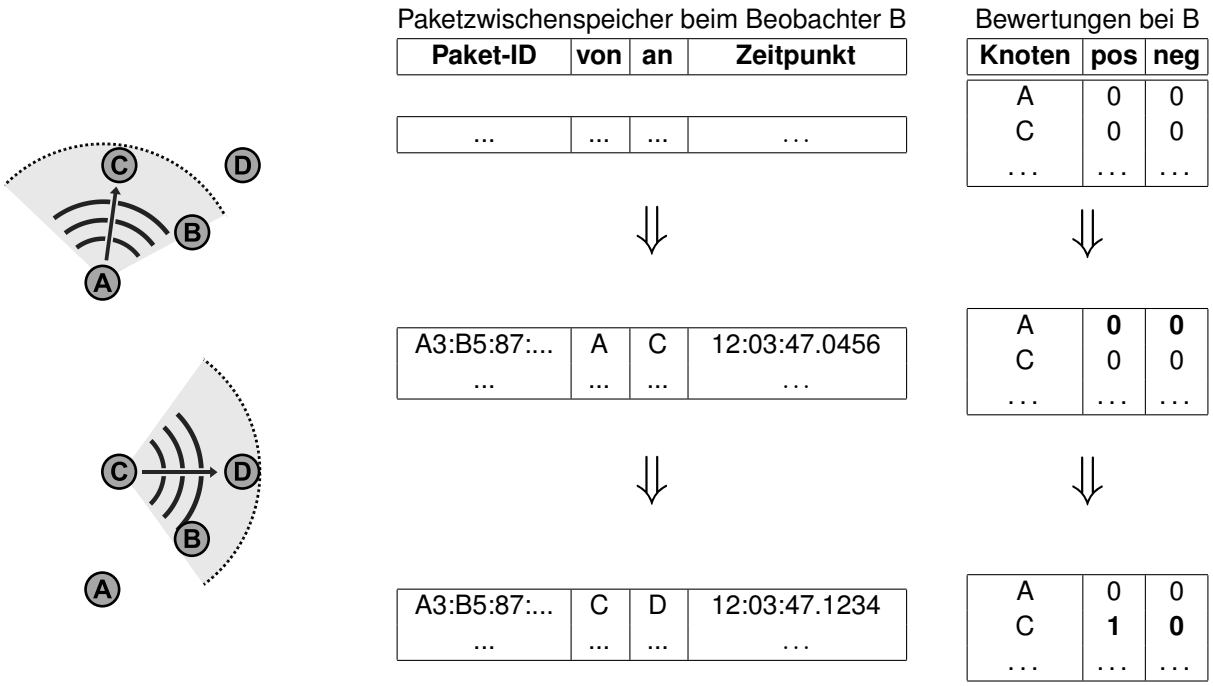


Abbildung 3.20: B bewertet beobachtete Weiterleitung nur dann positiv, wenn auch der vorige Weiterleitungsschritt beobachtet worden ist.

  - Werte nur positiv, wenn der vorige Weiterleitungsschritt beobachtet wurde.

    Anhand der Beobachtung des vorigen Weiterleitungsschrittes kann erkannt werden, dass das Paket nicht erfunden und nicht wiederholt wurde. Jeder Knoten muss dazu zu jedem beobachteten Paket für kurze Zeit einen Hash-Wert speichern, anhand dessen er das Paket wiedererkennen kann, wenn es vom nächsten Weiterleiter erneut ausgesendet wird (siehe Abbildung 3.20). Dies ist wesentlich weniger aufwändig als das obige benutzeridentitätsbasierte Verfahren. Allerdings fallen für die Bewertung alle Beobachter weg, die den vorigen Weiterleitungsschritt nicht beobachten konnten, so dass insgesamt weniger positive Bewertungen vergeben werden können.

- *Physikalische Verschiedenheit von Quelle und Weiterleiter*

  Dieser Punkt ist mit dem vorigen verwandt: Da selbst erzeugte Pakete nicht positiv bewertet werden dürfen, muss sichergestellt sein, dass es sich beim Weiterleiter um einen anderen Knoten als den Quellknoten des Pakets handelt. Indem ein Knoten mehrere Identitäten unterhält, kann er nämlich versuchen, als Weiterleiter der von ihm selbst erzeugten Pakete aufzutreten und dafür positiv bewertet zu werden – diese Möglichkeit wurde bereits bei den Ausführungen zu Bedrohungen im Zusammenhang mit Verfahren zur Teilnehmeridentifikation angesprochen (Abschnitt 3.4.1).

  Angriffsmotivation: Hoch. Der „Weiterleiter“ erschleicht sich eine positive Bewertung ohne zusätzliche Kosten (vorausgesetzt, das eigene Paket sollte sowieso gesendet werden).

  Vorgehen: Die Möglichkeiten, Betrugsversuche zu erkennen, ähneln stark den beim vorigen Punkt genannten:

  - Werte nur positiv, nachdem eine Paketsignatur der Quelle geprüft und außerdem sichergestellt wurde, dass deren Identität zu einem anderen Knoten gehört als die des vorigen Weiterleiters.

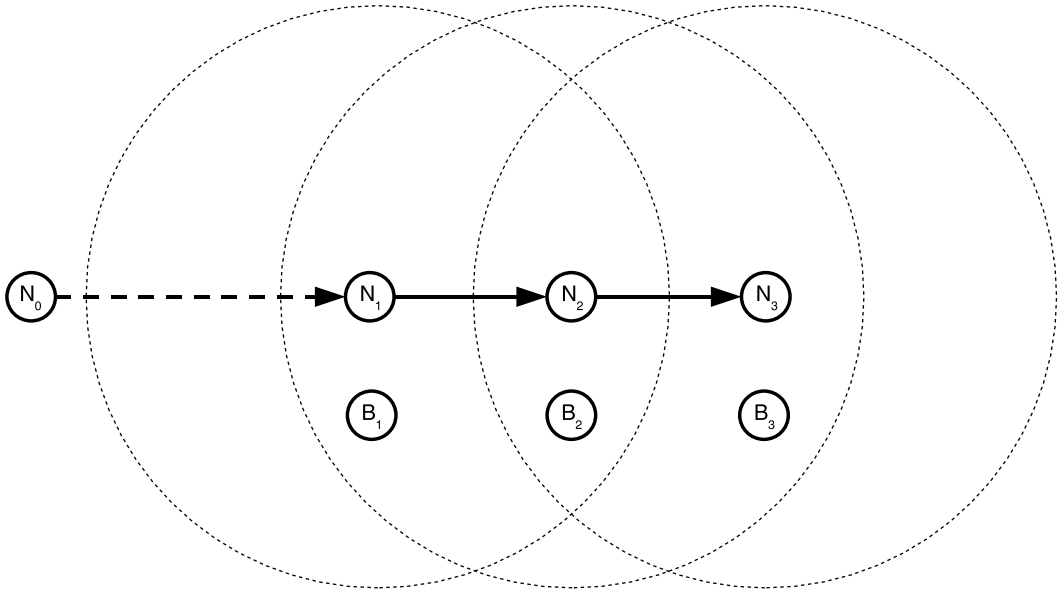


Abbildung 3.21: Weiterleitung bei N_2, beobachtet durch B_1, B_2 und B_3

Diese letztere Prüfung kann nur durchgeführt werden, indem sowohl der Quelle als auch dem vorigen Weiterleiter jeweils eine authentische Benutzeridentität zugeordnet wird; diese beiden müssen sich unterscheiden. Wie aber bei den Ausführungen zur Benutzeridentifikation begründet wurde (Abschnitt 3.4.3), schließt auch diese einzige identitätsbasierte Möglichkeit Betrug nicht völlig aus. Weil die Zuordnung von Benutzeridentitäten immer die Beschaffung von Zertifikatketten beinhaltet, ist das Verfahren extrem teuer bzw. kann es nur angewendet werden, wenn die Zuordnung bereits früher geprüft wurde.

– Werte nur positiv, wenn der vorige Weiterleitungsschritt beobachtet wurde. Dadurch ist ausgeschlossen, dass der Weiterleiter das Paket selbst erzeugt hat.

Es erscheint zunächst auch vorstellbar, dass anhand von Beobachtungen zweifelsfreier Fälle im Lauf der Zeit herausgefunden werden kann, welche Identitäten tatsächlich zu verschiedenen Knoten gehören, oder dass Einschätzungen bezüglich der Ehrlichkeit anderer Knoten ausgebildet werden können.

Zur Unterscheidung zweifelhafter und zweifelsfreier Fälle siehe Abbildung 3.21: Drei verschiedene Beobachter B_1, B_2 und B_3 beobachten die Weiterleitung durch Knoten N_2. B_1 und B_2 sind innerhalb der Sendereichweiten von N_1 und N_2 und "sehen" deshalb sowohl die Übertragung von N_1 nach N_2 als auch die Weiterleitung durch N_2. Sie gehen deshalb davon aus, dass es sich bei N_1 und N_2 tatsächlich um verschiedene Knoten handelt, da es für einen einzelnen, N_1 und N_2 vortäuschenden physikalischen Knoten wenig Sinn machen würde, das zwischen N_1 und N_2 übertragene Paket tatsächlich auszusenden. Da das Paket also zwischen N_1 und N_2 tatsächlich weitergeleitet wurde, versucht N_2, wenn er es nun nochmals (mit derselben Quellangabe) aussendet, offensichtlich nicht, eine nicht existente Quelle vorzutäuschen. B_1 und B_2 notieren deshalb eine positive Beobachtung bezüglich der Verschiedenheit von N_2 und N_0. B_3 dagegen sieht nur die Sendung durch N_2 und kann ohne weitere Information nicht sicher sein, ob N_0 tatsächlich existiert oder von N_2 nur vorgetäuscht wurde.

Was nun die Erfassung zu Informationen zu den zweifelsfreien Beobachtungen von B_1 und B_2 betrifft, so ist zu beachten, dass leider aus der Tatsache, dass N_2 in diesem Fall anscheinend nicht betrogen, sondern regulär weitergeleitet hat, nicht geschlossen werden kann, dass N_2 nicht (in

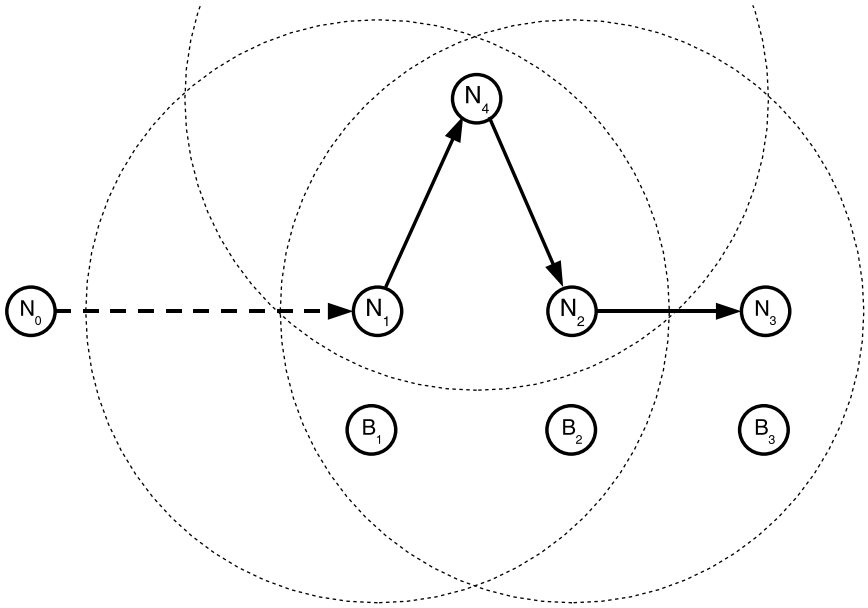


Abbildung 3.22: Unsichtbarer Weiterleiter

anderen Fällen) eine andere, von N_0 verschiedene Zweitidentität verwendet, um seine eigenen Pakete als weitergeleitet auszugeben. Lediglich die Verwendung von N_0 als Zweitidentität von N_2 ist widerlegt (vorausgesetzt, anhand eines beim Beobachter vorliegenden Schlüssels von N_0 kann die Quellsignatur des Pakets geprüft werden). Diese Information generell festzuhalten, also jeweils zu speichern, welche Identitäten zu verschiedenen Knoten gehören, erfordert Speicherplatzaufwand proportional zum Quadrat der Anzahl der Identitäten, was nicht unbedingt praktikabel ist. Denkbar wäre es höchstens, die oben genannte benutzeridentitätsbasierte Prüfung insofern zu ergänzen, als Zusatzinformation in Fällen gespeichert wird, wo Benutzeridentitäten einander so stark ähneln, dass man Betrug bei der Zertifikatserstellung vermuten könnte.

Statt tatsächlich Information darüber zu speichern, welche Identitäten zu verschiedenen Knoten gehören, könnte man auch versuchen, Einschätzungen darüber zu gewinnen, ob beobachtete Knoten sich in Hinsicht auf den betrachteten Angriff ehrlich verhalten. Um Einschätzungen gewinnen zu können, müssten allerdings auch negative Beobachtungen erfasst werden.

Negative Beobachtungen (d. h. die Erkennung von Betrugsfällen) sind aber nicht zuverlässig möglich. Siehe dazu Abbildung 3.22: Beobachtet B_2 die Übertragung von N_1 an N_4 und im Anschluss die Übertragung desselben Pakets von N_2 an N_3, ohne dass eine Übertragung von N_4 nach N_2 beobachtet wurde (weil B_2 zu weit von N_4 entfernt ist, um eine solche Übertragung empfangen zu können), so kann B_2 nicht feststellen, ob tatsächlich die Situation aus Abbildung 3.22 vorliegt oder etwa diejenige aus Abbildung 3.21, wobei N_4 die Zweitidentität N_2 verwendet. Regulärer Fall und Betrugsfall sind also nicht zu unterscheiden. Inhärent unsichtbare Übertragungswege wie z. B. Kabelverbindungen würden übrigens ebenfalls regelmäßig zu negativen Beobachtungen führen.

Da also mangels negativer Beobachtungen keine Einschätzung speziell bezüglich der Vertrauenswürdigkeit hinsichtlich des betrachteten Angriffs gewonnen werden kann und da die benutzeridentitätsbasierte Lösung außerdem zu aufwändig erscheint, bleibt als einzig praktikable Lösung übrig, nur in zweifelsfreien Fällen positive Bewertungen für die Weiterleitung zu registrieren.

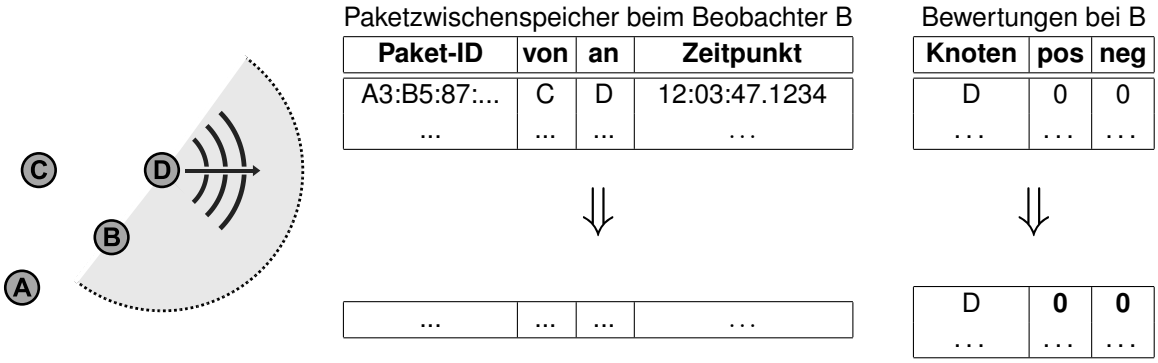


Abbildung 3.23: Wenn das Ziel der Weiterleitung kein Nachbar des Beobachters ist, erfolgt keine Bewertung.

- *Erreichbarkeit des nächsten Weiterleiters*

  Wenn der nächste Weiterleiter nicht erreichbar ist, sich also nicht in der Nachbarschaft des vorigen Weiterleiters befindet, findet genau genommen auch keine Weiterleitung statt, und der Sendevorgang ist sinnlos. In diesem Fall sollte auch auf eine positive Bewertung des „Weiterleiters“ verzichtet werden.

  Angriffsmotivation: Mittel bis gering. Der „Weiterleiter“ erhält eine positive Bewertung auf Kosten eines ansonsten sinnlosen Sendevorgangs.

  Vorgehen: Werte – unter der Annahme, dass andere Knoten für den letzten Weiterleiter ähnlich gut erreichbar sind wie für den Beobachter – nur dann positiv, wenn der Knoten, über den weitergeleitet wird, in der Nachbarschaft *des Beobachters* ist (siehe Abbildung 3.23), also bis vor Kurzem gehört wurde oder – falls technisch vorgesehen – den Empfang des Pakets sogar bestätigt hat. In letzterem Fall ist Betrug ausgeschlossen, ansonsten sind die Möglichkeiten sehr beschränkt. Es fallen einige eigentlich gerechtfertigte positive Bewertungen weg, was aber nicht sehr stark ins Gewicht fällt.

- *Korrekte Wahl des nächsten Weiterleiters*

  Wenn der Knoten, über den weitergeleitet wird, nicht auf einem sinnvollen, möglichst direkten Weg zum Ziel liegt, verursacht die folgende Weiterleitung anderen Knoten unnötigen Aufwand. Idealerweise sollte die Weiterleitung nur positiv bewertet werden, wenn sie einem sinnvollen Weg zum Ziel folgt.

  Angriffsmotivation: Mittel bis gering. Der „Weiterleiter“ erhält eine positive Bewertung (die er auch für korrekte Weiterleitung erhalten würde) und belastet gleichzeitig das Netz ungebührlich, beides zusammen auf Kosten eines Sendevorgangs. Möglicherweise kann der Angreifer etwas Aufwand bei der Durchführung des Wegfindungsprotokolls einsparen, da er keine Weiterleitungsinformation braucht, wenn er z. B. immer an einen zufällig gewählten Nachbarn weiterleitet. Oder er kann versuchen, einem bestimmten anderen Knoten gezielt zu schaden, indem er diesen bevorzugt als nächsten Weiterleiter einträgt, um seinen Energievorrat zu strapazieren.

  Vorgehen: Da die Vorteile für den Angreifer gering sind, der Angriff hauptsächlich destruktiven Charakter und dabei nur geringe Auswirkungen hat und eine Sicherung zu aufwändig wäre, weil die Güte der Wahl des Weiterleitungswegs von anderen Knoten nicht ausreichend genau beurteilt werden kann, wird auf eine Sicherung verzichtet. Die Teilnahme an einem Wegfindungsprotokoll kann unabhängig davon bewertet werden.

3.6.3.4 Bewertungsrelevante Information für die negative Bewertung

Ebenso wie vorher für die positive Bewertung bei erfolgter Weiterleitung soll nun für die negative Bewertung bei unterlassener Weiterleitung betrachtet werden, welche Informationen vorliegen müssen, welche Vorteile sich ein Angreifer durch ihre Manipulation beschaffen kann, und welche Gegenmaßnahmen in Frage kommen.

- *Identität dessen, der weiterleiten sollte*

  Die Identität desjenigen Knotens, der ein Paket weiterleiten sollte, wird benötigt, um das Vertrauensprofil zu identifizieren, in welchem eine negative Bewertung vermerkt werden kann, falls innerhalb einer vorgegebenen Zeitspanne keine Weiterleitung beobachtet wird.

  Angriffsmotivation: Mittel bis gering, da die negative Bewertung einen anderen Knoten betrifft als den, der die Identitätsangabe erstellt hat. Der Angreifer hat also keinen direkten eigenen Vorteil, sondern versucht höchstens, die Einschätzungen eines Anderen in den Augen Dritter zu verschlechtern.

  Der betroffene Knoten selbst, dessen Identität als die des nächsten Weiterleiters im Paket vermerkt ist, kann selbstverständlich nicht verhindern, dass Pakete an ihn adressiert werden. Er kann höchstens versuchen, seine Identität zu verschleiern, indem er die Übertragung eines an ihn übermittelten Pakets so überlagert, dass bei einigen Beobachtern der Eindruck entsteht, das Paket sei nicht an ihn adressiert. Den vorigen Weiterleiter kann er damit selbstverständlich nicht täuschen, und in der Regel auch höchstens einen Teil der Beobachter (siehe Abschnitt 3.6.3.1 zum gezielten Verfälschen fremder Nachrichten). Außerdem erfordert die Verfälschung der Nachricht ebenfalls einen Sendevorgang (wenn auch einen kurzen), so dass auch bei Erfolg nur ein geringer Vorteil zu erkennen ist.

  Vorgehen: Wenn die Identität des nächsten Weiterleiters mit der Schicht-2-Adresse verknüpft ist, kann sie nicht regelrecht gefälscht werden, denn das Paket muss auf Schicht 2 genau von demjenigen Knoten weitergeleitet werden, dessen Adresse im Paket angegeben ist (vorausgesetzt, es erreicht ihn überhaupt). Möglich sind damit nur noch Angriffe, bei denen der Angreifer absichtlich Schicht-2-Empfänger angibt, die die jeweiligen Pakete gar nicht empfangen und deshalb auch nicht weiterleiten können, um dafür zu sorgen, dass diese angeblichen Schicht-2-Empfänger schlechte Bewertungen erhalten – siehe dazu den folgenden Punkt.

- *Paketankunft beim säumigen Weiterleiter*

  Falls der als nächster Weiterleiter adressierte Knoten sich außerhalb der Reichweite des Senders befindet und das Paket damit gar nicht empfangen und deshalb auch nicht weiterleiten kann, entsteht bei (einigen) Beobachtern der Eindruck, der Knoten habe seine Weiterleitungspflicht vernachlässigt, so dass er negativ bewertet wird.

  Falls eine explizite Nachweismöglichkeit für den Empfang in Form einer Empfangsbestätigung existiert, kann der adressierte Weiterleiter diesen absichtlich unterschlagen, um den Eindruck zu erwecken, er habe das Paket nicht erhalten und könne es nicht weiterleiten. Hier führt also der nächste Weiterleiter – abgesehen von seinem unkooperativen Verhalten bei der Weiterleitung – einen Angriff gegen die Beobachtung durch.

  Falls kein expliziter Nachweis existiert, kann nur angenommen werden, dass das vom vorigen Weiterleiter (hoffentlich) korrekt adressierte Paket auch angekommen ist. In diesem Fall kommt der vorige Weiterleiter als Angreifer in Frage, der z. B. absichtlich einen anderen Knoten außerhalb seiner Reichweite als nächsten Weiterleiter einträgt, um diesem eine schlechte Bewertung zu verschaffen.

Angriffsmotivation: Abhängig von der Verwendung von Empfangsbestätigungen:

- *ohne Empfangsbestätigung* (potentieller Angreifer ist der vorige Weiterleiter): Mittel bis gering. Der Angreifer verschafft einem anderen Knoten eine negative Bewertung, ohne aber einen wesentlichen Vorteil zu haben (keine Energieersparnis).
- *mit Empfangsbestätigung* (potentieller Angreifer ist der adressierte Weiterleiter): Hoch. Der Angreifer vermeidet eine negative Bewertung für eigenes unkooperatives Verhalten, durch welches er selbst Energie spart.

Vorgehen: Gehe im Fall einer fehlenden oder gar nicht vorgesehenen Empfangsbestätigung immer vom attraktiveren der beiden Angriffe aus. Prüfe aber vor dem Starten des Zeitgebers, dessen Ablauf ohne beobachtete Weiterleitung zu einer negativen Bewertung des adressierten Weiterleiters führt, ob dieser sich wahrscheinlich innerhalb der eigenen Nachbarschaft befindet (als Näherung für die Prüfung, ob er sich in der Reichweite des aktuellen Weiterleiters befindet). Ist das nicht der Fall, wird der Zeitgeber nicht gestartet und eine negative Bewertung damit ausgeschlossen.

Welche Knoten sich in der eigenen Nachbarschaft befinden, wird von manchen Schicht-2-Übertragungsverfahren überwacht, und auch von proaktiven Wegfindungsverfahren. Ist beides nicht der Fall, so wird diese Funktionalität hinzugefügt, indem zu jedem Nachbarn immer der Empfangszeitpunkt des letzten beobachteten vom Nachbarn gesendeten Pakets gespeichert wird. Werden über eine gewisse Zeitspanne keine Pakete mehr von einem vermeintlichen Nachbarn empfangen, so wird er aus der Liste der Nachbarn gestrichen.

- *Nichtweiterleitung richtig beobachtet*

  Unter Umständen hat der Beobachter die Weiterleitung nicht bemerkt, obwohl sie stattgefunden hat, etwa wegen einer Störung der Funkübertragung oder weil der Weiterleiter sich zwischen Empfang und Wiederaussendung des Pakets aus der Reichweite des Beobachters bewegt hat. In diesem Fall erhält er eine ungerechtfertigte negative Bewertung.

  Angriffsmotivation: Da kein expliziter Nachweis verwendet wird, kann auch keiner manipuliert werden. Es handelt sich hier nicht um einen Angriff, sondern um eine Schwäche des Verfahrens.

  Vorgehen: Mittels der bereits genannten Nachbarschaftserkennung wird sowohl vor dem Starten als auch bei Ablauf des Zeitgebers überprüft, ob die Weiterleitung voraussichtlich beobachtet werden kann bzw. konnte. Gelegentliche Fehler verursachen geringfügige Fehlbewertungen, die aber toleriert werden können.

- *Fehlermeldung bei Nichtweiterleitung wegen Zugangskontrolle*

  Von der Zugangskontrolle abgelehnte Pakete werden nicht weitergeleitet, sondern statt des abgelehnten Pakets muss der ablehnende Knoten eine Fehlermeldung senden, um seine Entscheidung sichtbar zu machen und Beobachtern die Möglichkeit zu geben, zwischen berechtigter und unberechtigter Nichtweiterleitung zu unterscheiden. Bei Eingang der Fehlermeldung unterlässt der Beobachter die sonst fällige negative Bewertung. Angreifer, die Pakete nicht weiterleiten wollen, können eine Zugangsfehlermeldung absenden, um einer negativen Bewertung zu entgehen.

  Angriffsmotivation: Gering. Der Angreifer spart keinen Sendevorgang ein, wenn er statt des weiterzuleitenden Pakets eine Fehlermeldung aussendet. Lediglich, wenn die Fehlermeldung weniger umfangreich als das weiterzuleitende Paket ist, kann etwas Energie gespart werden.

Vorgehen: Fehlermeldungen aufgrund von Zugangsentscheidungen müssen mindestens dieselbe Größe haben wie die abgelehnten Pakete. Jede Fehlermeldung soll deshalb das abgelehnte Paket enthalten. Dadurch ist auch gewährleistet, dass jeder Beobachter nach Empfang der Fehlermeldung das abgelehnte Paket in seiner Liste beobachteter Pakete identifizieren und löschen kann.

3.6.4 Sicherungsmaßnahmen

Im vorigen Abschnitt wurden bei den Angriffen jeweils geeignete Gegenmaßnahmen angesprochen, wobei teilweise mehrere verschiedene Möglichkeiten angegeben wurden. Im Folgenden werden alle diejenigen Sicherheitsmaßnahmen nochmals zusammenfassend vorgestellt, die schließlich für das Verfahren zur Beobachtung und Bewertung der Weiterleitung in dem hier entwickelten Konzept ausgewählt wurden und in denen sich dieses endgültige Verfahren von dem in Abschnitt 3.6.2 einleitend grob umrissenen unterscheidet.

- *Ermittlung der Identität des vorigen Weiterleiters anhand Schicht-2-Absender*

  Zur Identifikation des Vertrauensprofils des vorigen Weiterleiters zum Zweck der positiven Bewertung eines durchgeführten Weiterleitungsvorgangs wird die im Paket enthaltene Schicht-2-Absenderadresse herangezogen und anhand einer dafür von jedem Knoten aufgebauten Liste der Schlüssel und Adressen benachbarter Knoten auf den zugehörigen Schlüssel abgebildet. Von einer Falschangabe der Adresse hat der Sender selbst keinen Vorteil, und andere Knoten können die Angabe nicht sinnvoll beeinflussen, deshalb ist keine Sicherung – etwa durch eine digitale Signatur über das Paket mit Hilfe des zugehörigen Schlüssels – erforderlich. Falls die Absenderadresse eines beobachteten Pakets in der Liste der Nachbarn eines Beobachters nicht enthalten ist, erfolgt keine Bewertung des Vorgangs. Es werden dann aber sofort Maßnahmen zur Ermittlung des Schlüssels des anscheinend noch unbekannten Nachbars eingeleitet.

  Es besteht für Angreifer eine relativ hohe Motivation, die Abbildungsfunktion zu manipulieren, um positive Bewertungen für von Anderen erbrachte Weiterleitungsdienste ihrer eigenen Schlüsselkennung zuzuordnen. Bei der genauen Beschreibung des Verfahrens zum Aufbau der Datenbasis aller Schlüssel und Adressen von Nachbarn (Abschnitt 3.10) wird näher auf Angriffsmöglichkeiten und Sicherungsmaßnahmen dagegen eingegangen.

- *Positive Wertung erfolgt nur, wenn auch der vorige Weiterleitungsschritt beobachtet wurde*

  Wenn die Weiterleitung eines Pakets beobachtet wird, wird zunächst anhand der Liste derzeit beobachteter Pakete überprüft, ob auch der *vorige* Weiterleitungsschritt, also die Übertragung des Pakets an den Knoten, der es nun wieder aussendet, schon beobachtet (oder selbst durchgeführt) worden war. Nur wenn dies der Fall ist, erfolgt eine positive Bewertung des Weiterleiters. Auf diese Weise wird ausgeschlossen, dass ein Angreifer die Aussendung von ihm selbst erzeugter Pakete durch geeignete Wahl der Adressangaben als Weiterleitung erscheinen lässt.

- *Ermittlung der Schlüsselkennung des nächsten Weiterleiters anhand Schicht-2-Empfänger*

  Die Schlüsselkennung des nächsten Weiterleiters (die für eine negative Bewertung desselben benötigt wird, falls dieser nicht weiterleitet) wird vom Beobachter aus der Schicht-2-Empfängeradresse im Paket ermittelt. Würde stattdessen eine Schlüsselkennung explizit im Paket angegeben, so könnte der vorige Weiterleiter eine gegebenenfalls fällige negative Bewertung vom adressierten Knoten ab- und auf einen anderen umlenken; dieser Angriff wird durch die

implizite Angabe verhindert. Außerdem wird natürlich etwas Platz im Paket gespart. Eine Abbildungsfunktion von Schicht-2-Adressen auf Schlüsselkennungen wird sowieso in jedem Fall benötigt, denn bei expliziter Angabe müsste sie vom vorigen Weiterleiter konsultiert werden.

Die nötigen Daten für die Abbildung von Schicht-2-Adressen auf Schlüsselkennungen werden mit Hilfe des in Abschnitt 3.10 beschriebenen Schlüsselaustauschprotokolls ermittelt.

- *Nachbarschaftsprüfung vor negativer Bewertung*

  Sowohl vor dem Start des Zeitgebers für die Beobachtung der Weiterleitung als auch bei dessen Ablauf wird geprüft, ob der Knoten, der die Weiterleitung durchführen soll bzw. hätte durchführen sollen, sich in der Nachbarschaft des Beobachters befindet. Ist dies nicht der Fall, wird nicht negativ bewertet bzw. wird der Zeitgeber gar nicht erst gestartet. Damit wird einerseits die negative Bewertung in Situationen verhindert, wo die Weiterleitung gar nicht beobachtet werden kann, auch wenn sie stattfindet. Andererseits werden Angriffe des vorigen Weiterleiters erschwert, bei denen er absichtlich unerreichbare Knoten als nächsten Weiterleiter angibt, um ihnen negative Bewertungen zu verschaffen.

  Ob ein Knoten sich in der eigenen Nachbarschaft befindet, wird entweder schon vom Schicht-2-Übertragungsverfahren oder vom Wegfindungsverfahren erfasst, oder es wird zusätzlich implementiert, indem zu jedem Nachbarn der Zeitpunkt des letzten empfangenen Pakets festgehalten wird. Als Nachbar zählt ein Knoten dann, wenn dieser Zeitpunkt nicht zu lange zurückliegt.

- *Fehlermeldungen der Zugangskontrolle enthalten abgelehnte Pakete*

  Eine Anforderung an die Zugangskontrolle ist, dass deren Fehlermeldungen bei Ablehnung auch die abgelehnten Pakete selbst enthalten müssen. Damit kann einerseits das Paket identifiziert werden, das aus der Liste beobachteter Pakete gelöscht werden muss, und andererseits ist gewährleistet, dass sich kein Knoten Vorteile verschaffen kann, indem er willkürlich Pakete ablehnt.

3.6.5 Detaillierte Beschreibung des verfeinerten Verfahrens

3.6.5.1 Integration in die Vermittlungsschicht

Abbildung 3.24 zeigt, wie die zusätzlichen Funktionen für Beobachtung und Bewertung der Weiterleitung in die normalen Abläufe in der Vermittlungsschicht bei der Weiterleitung integriert werden; letztere sind durch grau unterlegte Kästen gekennzeichnet.

Nach Empfang eines jeden Pakets wird zunächst – falls eine explizite Nachbarschaftserkennung implementiert werden muss – der Empfangszeitpunkt im Eintrag des Schicht-2-Absenders in der Liste der Nachbarn festgehalten. Anschließend wird – falls angezeigt – ein voriger Weiterleiter bewertet und geprüft, ob das Paket weiter zu beobachten ist; auf den genauen Ablauf wird unten noch eingegangen.

Bevor nun die eigentliche Tätigkeit der Vermittlungsschicht beginnt, muss anhand der Empfängerangabe der Schicht 2 festgestellt werden, ob das Paket überhaupt für den eigenen Knoten bestimmt ist, oder ob es sich um eine mitgehörte Übertragung handelt. Diese Adressprüfung gehört eigentlich zu den Aufgaben der Schicht 2 und wird normalerweise schon vom Netzwerkadapter durchgeführt, der dann nur die für den eigenen Knoten bestimmten Pakete nach oben weitergibt. Diese Prüfung auf Schicht 2 muss aber, um die Beobachtung fremden Verkehrs zu ermöglichen, abgeschaltet werden, indem der Netzwerkadapter in den so genannten Promiskuitätsmodus versetzt wird; er liefert dann alle gehörten Pakete an die Vermittlungsschicht. Die nachträgliche Prüfung in der Vermittlungsschicht

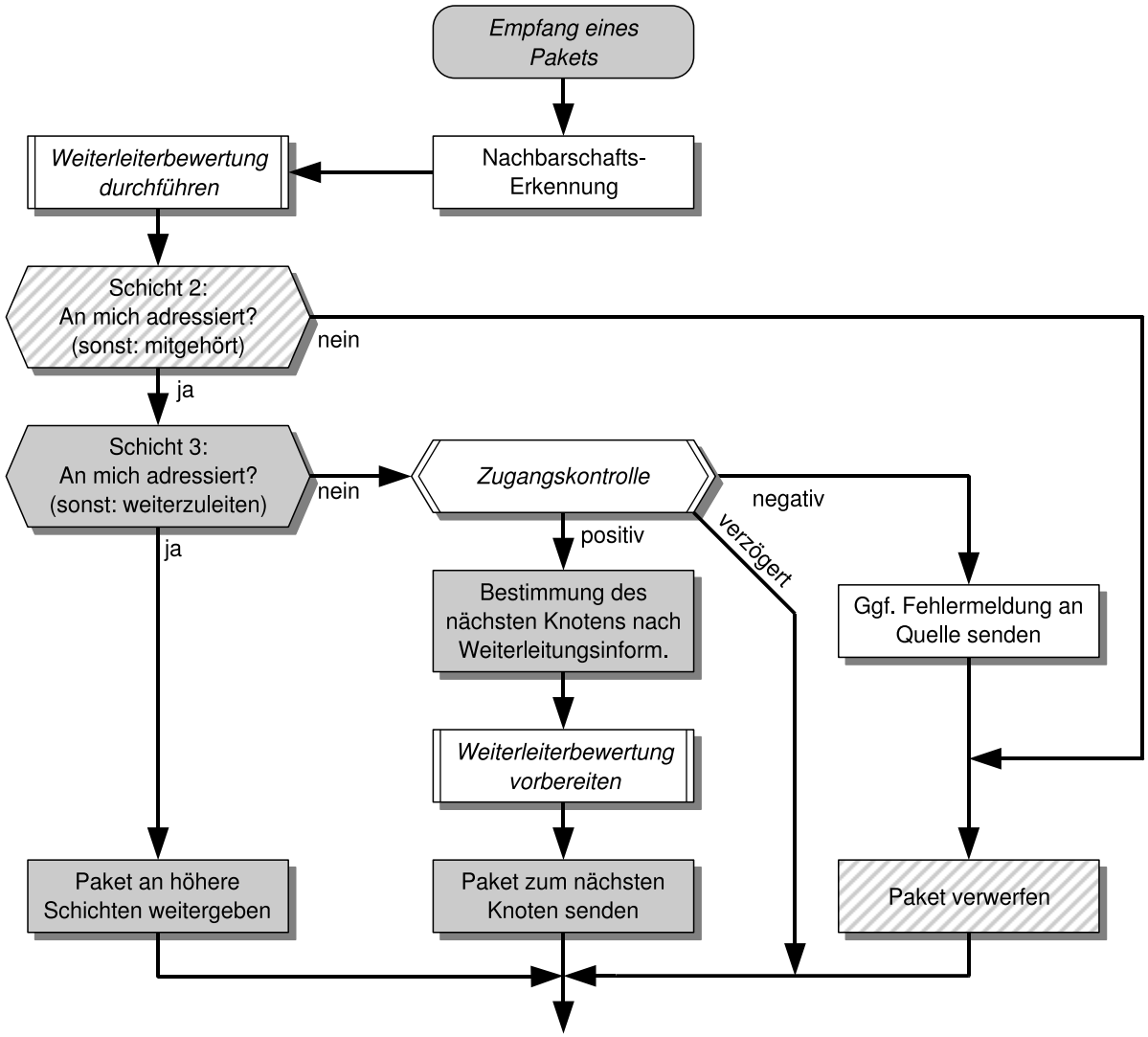


Abbildung 3.24: Integration der Weiterleitungsbeobachtung in die Abläufe bei der Bearbeitung empfangener Pakete

(die schraffiert dargestellt ist, um anzudeuten, dass sie zwar nicht neu ist, aber normalerweise nicht an dieser Stelle stattfindet) führt zum Ende der Bearbeitung an dieser Stelle, falls es sich um ein mitgehörtes Paket handelt.

Die nächste Entscheidung betrifft die Weiterleitung in Schicht 3: Wenn die Schicht-3-Empfängeradresse des Pakets die eigene Adresse ist, so ist das Paket am Ziel angelangt und wird an höhere Schichten weitergegeben. Ist dies nicht der Fall, so ist es weiterzuleiten. An dieser Stelle wird nun zunächst die Zugangskontrolle durchgeführt; sie wird in Abschnitt 3.9 noch näher beschrieben und liefert entweder ein positives oder ein negatives Ergebnis. Im positiven Fall ist das Paket zum nächsten Weiterleiter zu senden, der anhand der lokalen Weiterleitungsinformation bestimmt wird; wird ein reaktives Wegfindungsverfahren verwendet, so muss die Weiterleitungsinformation unter Umständen auch erst jetzt noch ermittelt werden. Bevor das Paket zum nächsten Weiterleiter ausgesendet wird, wird noch die weitere Beobachtung des Pakets vorbereitet, wie im folgenden Abschnitt beschrieben. Verläuft die Zugangskontrolle negativ, so wird statt des weiterzuleitenden Pakets eine Fehlermeldung an die Quelle des Pakets versendet, und damit ist die Bearbeitung abgeschlossen. Sollte die lokal verfügbare Information für die Durchführung der Zugangskontrolle nicht ausreichen, so muss die weitere Verarbeitung des Pakets eventuell verzögert werden, bis zusätzliche Information beschafft wurde; das Paket wird dann solange zwischengespeichert.

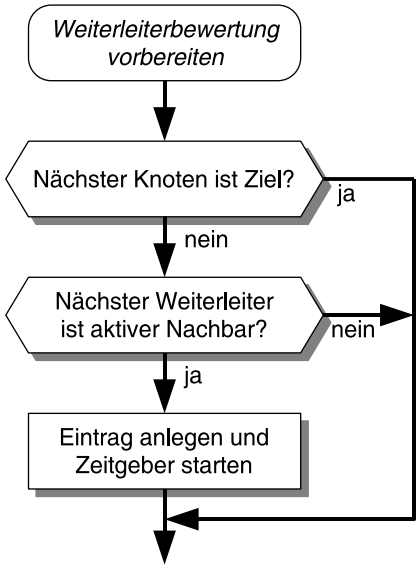


Abbildung 3.25: Vorbereitung der Weiterleitungsbeobachtung

Nicht abgebildet sind die Abläufe beim Absenden eines selbst erzeugten Pakets. Auch hier wird kurz vor der Aussendung des Pakets zum ersten Weiterleiter die Beobachtung in der im Folgenden beschriebenen Weise vorbereitet.

3.6.5.2 Vorbereitung der Beobachtung

Zur Vorbereitung der Beobachtung des nächsten Weiterleitungsvorgangs eines Pakets (siehe Abbildung 3.25) muss dieses in die Liste beobachteter Pakete aufgenommen werden. Sinnvoll ist dies nur, wenn das Paket nach der soeben selbst durchgeführten Weiterleitung oder – im Fall eines selbst erzeugten Pakets – dem Aussenden tatsächlich nochmal weitergeleitet werden muss. Deshalb wird zunächst überprüft, ob der nächste Knoten nicht schon das Ziel des Pakets ist. Die zweite Bedingung für die Aufnahme ist, dass der nächste Weiterleiter sich in der Nachbarschaft des Beobachters befindet (zur Begründung siehe Abschnitt 3.6.4). Sind beide Bedingungen erfüllt, so wird ein Eintrag für das Paket angelegt. Er enthält einen Hashwert über alle Bestandteile des Pakets, die bei der Weiterleitung nicht verändert werden, sowie die Adresse des Schicht-2-Empfängers.

3.6.5.3 Durchführung der Weiterleiterbewertung

Abbildung 3.26 stellt die Abläufe dar, die für jedes empfangene oder mitgehörte Paket stattfinden und weiter oben als Durchführung der Weiterleiterbewertung zusammengefasst wurden.

Der erste Schritt dabei ist, anhand der Liste beobachteter Pakete festzustellen, ob das Paket bereits früher beobachtet bzw. selbst gesendet wurde – nur in diesem Fall kommt eine positive Bewertung in Frage (zur Begründung siehe Abschnitt 3.6.4). Der über das beobachtete Paket berechnete Hash-Wert wird in der Liste gesucht, außerdem muss der Schicht-2-Absender des beobachteten Pakets mit dem im Eintrag gespeicherten Schicht-2-Empfänger übereinstimmen. Wird ein Eintrag gefunden, so wird eine positive Beobachtung in dem Vertrauensprofil registriert, dessen Schlüssel der Schicht-2-Absenderadresse zugeordnet ist.

Unabhängig davon, ob eine positive Bewertung stattgefunden hat, ist das Ziel der nachfolgenden Schritte, die weitere Beobachtung des Pakets vorzubereiten, falls es nochmals weitergeleitet werden

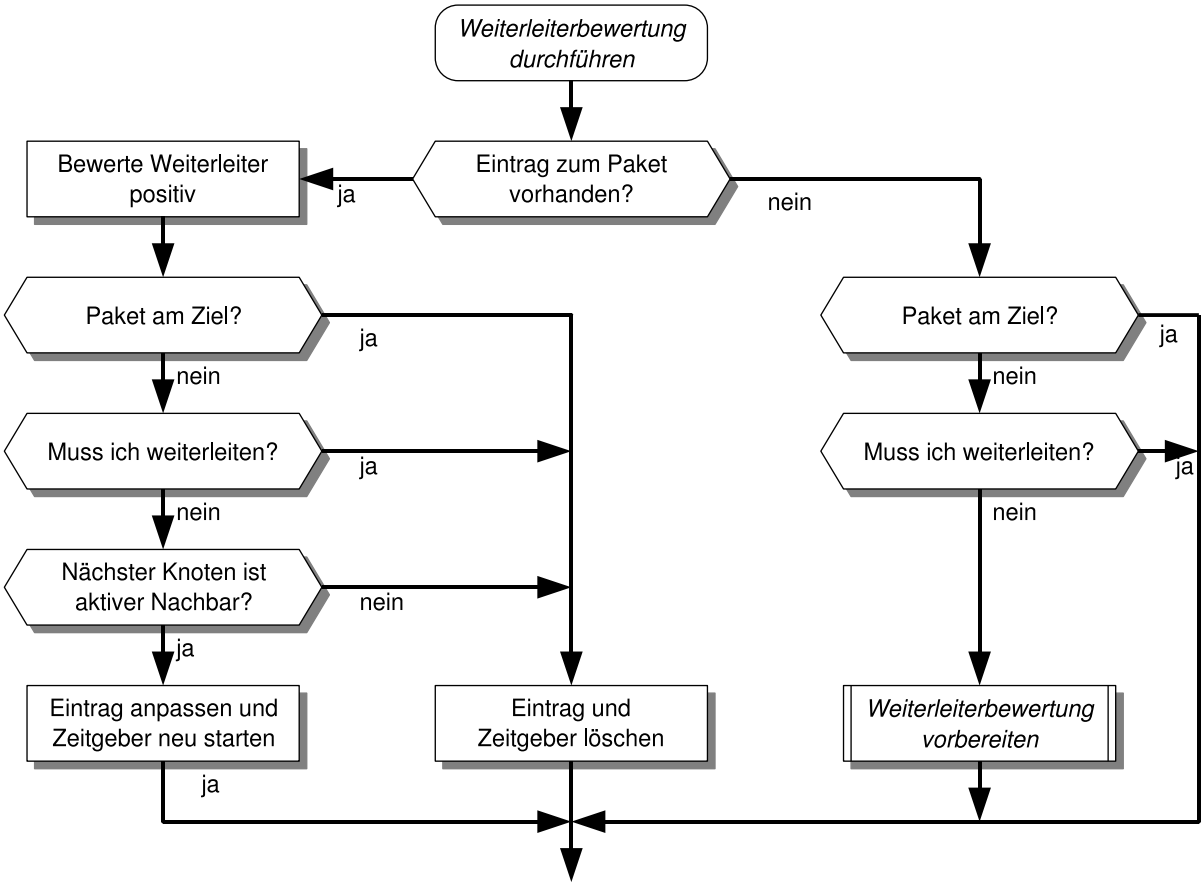


Abbildung 3.26: Ablauf der Weiterleitungsbeobachtung

muss. Die Bearbeitung ähnelt deshalb der aus Abbildung 3.25, mit zwei Unterschieden: Erstens muss zusätzlich unterschieden werden, ob der bearbeitende Knoten selbst weiterleiten muss, oder ob es sich um ein mitgehörtes Paket handelt; im ersten Fall wird der Eintrag in der Liste beobachteter Pakete gelöscht, da das eigene Verhalten ja nicht bewertet zu werden braucht. Der zweite Unterschied besteht darin, dass in manchen Fällen schon ein Eintrag vorhanden war, der nur noch angepasst zu werden braucht.

Schließlich muss noch der Fall eines ablaufenden Zeitgebers behandelt werden; Abbildung 3.27 tut dies schematisch. Eine negative Bewertung wird nach Ablauf des zum anscheinend nicht weitergeleiteten Paket gehörigen Zeitgebers nur dann durchgeführt, wenn angenommen werden kann, dass der säumige Weiterleiter sich noch in der Nachbarschaft des Beobachters aufhält und das Ziel des Pakets erreichbar ist. Diese beiden Fragen werden anhand der Nachbarschaftsinformation und der Weiterleitungsinformation geklärt. Bezüglich letzterer ist zu beachten, dass hier auch bei Einsatz eines reaktiven Wegfindungsverfahrens keine Aktivitäten des Wegfindungsprotokolls ausgelöst werden dürfen, da der entstehende Aufwand in Form zusätzlicher Belastung des Netzes für den Zweck einer einfachen Beobachtung nicht gerechtfertigt wäre. Sind die Bedingungen für die negative Bewertung erfüllt, so wird aus der im Eintrag gespeicherten Schicht-2-Empfängeradresse mit Hilfe der oben genannten Abbildung die zugehörige Schlüsselkennung ermittelt. Ist dies erfolgreich, so wird in dem durch die Schlüsselkennung identifizierten Vertrauensprofil eine negative Beobachtung registriert.

Der Eintrag in der Liste beobachteter Pakete wird abschließend gelöscht, da davon ausgegangen wird, dass das Paket entweder verloren gegangen ist oder den vom Beobachter wahrnehmbaren Bereich des Netzes verlassen hat.

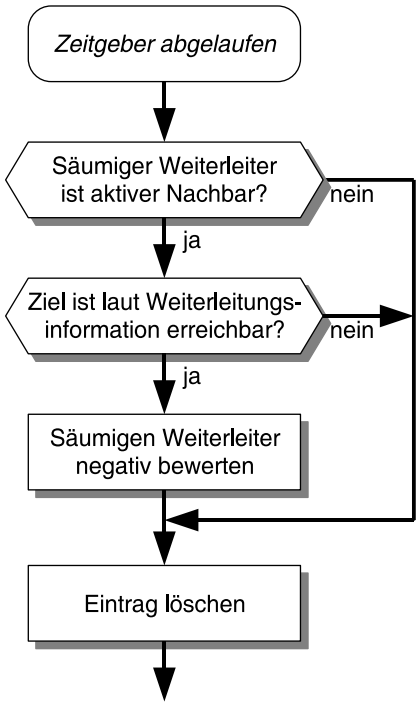


Abbildung 3.27: Negative Bewertung bei Ablauf des Zeitgebers

3.6.6 Beobachtung und Bewertung bei Anfrage-Antwort-Protokollen

Viele oberhalb der Vermittlungsschicht angesiedelte Dienste (und auch einige der im hier vorgestellten Konzept eingesetzten) verwenden Protokolle nach dem Anfrage-Antwort-Paradigma: Ein Dienstnehmer sendet eine Anfragenachricht an einen Dienstgeber und erhält darauf eine Antwortnachricht, womit der Dienstleistungsvorgang abgeschlossen ist. Die korrekte Durchführung solcher Protokolle lässt sich recht einfach immer nach demselben Schema überwachen, deshalb können sie gut in die Beobachtung des Verhaltens anderer Knoten einbezogen werden.

Das Verfahren zur Beobachtung und Bewertung ist grundsätzlich wie folgt: Wird eine Anfrage beobachtet, so legt der Beobachter einen Eintrag in einer lokalen Liste beobachteter Vorgänge an. Dort wird neben der Identität von Dienstnehmer und Dienstgeber auch Information zur Identifikation des Dienstleistungsvorgangs gespeichert, anhand derer Anfrage und Antwort einander zugeordnet werden können. Häufig verwenden Anfrage-Antwort-Protokolle irgendeine Form von Transaktionskennung, die sowohl in der Anfrage als auch in der Antwort enthalten ist, und die der Dienstnehmer verwendet, um zu erkennen, zu welcher seiner Anfragen eine erhaltene Antwort gehört. Diese Transaktionskennung kann dann auch von Beobachtern verwendet werden. Beim Anlegen des Eintrags wird außerdem ein Zeitgeber gestartet, der nach einer Zeitspanne abläuft, innerhalb derer die Dienstleistungsvorgänge des betrachteten Dienstes in aller Regel abgeschlossen sein sollten. Wird in der Folge eine Antwort zu einer registrierten Anfrage beobachtet, so wird der Dienstgeber positiv bewertet und Eintrag und Zeitgeber werden gelöscht. Läuft dagegen der Zeitgeber ab, bevor eine Antwort beobachtet werden konnte, so wird der Dienstgeber negativ bewertet (und der Eintrag ebenfalls gelöscht).

Die Speicherung von Information über beobachtete Anfragen ist aus zwei Gründen erforderlich: Einerseits natürlich, um negative Bewertungen vergeben zu können, wenn keine Antwort erfolgt, andererseits aber auch, um bei der Beoachtung einer Antwort vor einer positiven Bewertung sicherstellen, dass es sich um eine Antwort auf eine tatsächlich gestellte Anfrage handelt. Damit soll verhindert werden, dass ein Knoten, der unaufgefordert wertlose Nachrichten versendet, die wie Antworten aussehen, dafür positive Bewertungen erhält. Der Beobachter kann frei entscheiden, wieviel Speicherplatz

er für solche Informationen verwenden möchte, denn er ist nicht gezwungen, jeden beobachtbaren Vorgang auch tatsächlich zu beobachten und zu bewerten, er kann dadurch lediglich die Genauigkeit seiner Einschätzungen anderer Knoten erhöhen.

Eine wichtige Frage, die noch beantwortet werden muss, ist, welche Knoten als Beobachter in Frage kommen. Kriterium dafür ist die Notwendigkeit, neben der Anfrage auch mit sehr hoher Wahrscheinlichkeit die zugehörige Antwort beobachten zu können. Wenn das betrachtete Protokoll oberhalb der Vermittlungsschicht angesiedelt ist, die Protokollnachrichten also durchs Netz weitergeleitet werden, ist dieses Kriterium nur für einen Teil derjenigen Knoten erfüllt, welche die Anfrage beobachten können. Denn obwohl die Anfrage von all ihren Weiterleitern und deren Nachbarn beobachtet werden kann, ist es in Anbetracht der Dynamik der Topologie von Ad-hoc-Netzen recht unsicher, ob die kurze Zeit später in umgekehrter Richtung übertragene Antwortnachricht genau die selben Knoten passiert. Selbst wenn sich die Netztopologie in der Zwischenzeit nicht ändert, gibt es keine Garantie dafür, dass der Rückweg aus denselben Knoten besteht – gerade bei den in Ad-hoc-Netzen häufig verwendeten bedarfsgesteuerten Wegfindungsverfahren können hier leicht unterschiedliche Wege gewählt werden. Auf jeden Fall als Beobachter geeignet sind alle Nachbarn des Dienstgebers; sie sehen – solange sie Nachbarn bleiben – alle Antworten des Dienstgebers, und können damit alle von ihnen beobachteten Anfragen an diesen bewerten.

Ein weiterer Knoten, der neben der Anfrage auch jede erfolgte Antwort sieht, ist der Dienstnehmer selbst. Es liegt für diesen eigentlich sehr nahe, den Dienstgeber positiv zu bewerten, wenn seine Anfrage beantwortet wurde. Dabei ist allerdings zu beachten, dass es, wenn eine Antwort ausbleibt, nicht feststellbar ist, ob dies die Schuld des Diensterbringers ist, oder ob Anfrage oder Antwort vielleicht unterwegs nicht korrekt weitergeleitet wurden. Als Lösung für dieses Problem wird für das hier vorgestellte Konzept der folgende Ansatz verwendet: Der Dienstnehmer selbst wertet nur positiv, da negative Bewertungen zu unsicher wären. Als Ausgleich für dadurch wegfallende negative Bewertungen des Dienstgebers werten dessen Nachbarn um den Faktor $1 + \frac{1}{n}$ stärker negativ als positiv, wobei n die Anzahl der Knoten in ihrer eigenen Nachbarschaft ist, welche als Näherung für die Anzahl der Nachbarn des Dienstgebers und damit für die Anzahl der Bewertungen ist. Falls der Dienstgeber also tatsächlich n Nachbarn hat, die eine Anfrage an ihn beobachtet haben, so erhält er bei Nichterfüllung insgesamt $n+1$ negative Bewertungen. Die Nachbarn haben damit sozusagen die negative Bewertung durch den Dienstnehmer mit übernommen.

3.7 Verwaltung von Einschätzungen und Vertrauen

Alle Information über das beobachtete Verhalten anderer Teilnehmer sowie über das in diese gesetzte Vertrauen bezüglich bestimmter Vorgänge wird lokal in den jeweiligen *Vertrauensprofilen* der betreffenden Teilnehmer festgehalten. Die in dieser Datenstruktur enthaltenen Vertrauensmaßzahlen werden angepasst, um positive oder negative Beobachtungen zu registrieren, und sie werden als Grundlage von Zugangsentscheidungen abgefragt. Auch Vorgaben des Benutzers bezüglich Vertrauensvorgaben, etwa aufgrund persönlicher Bekanntschaften mit bestimmten anderen Benutzern, können hier registriert werden.

Neben eigenen Einschätzungen werden im Vertrauensprofil auch Meinungen Dritter über den betreffenden Teilnehmer aufbewahrt, die zur Ermittlung einer Gesamteinschätzung miteinander und mit den eigenen Einschätzungen verknüpft werden können. Meinungen Dritter – auch bezeichnet als Fremdmeinungen – werden bei Bedarf angefordert, also etwa dann, wenn bei der Zugangskontrolle eigene Einschätzungen zum fraglichen Teilnehmer nicht vorhanden oder nicht sicher genug sind. Fremdmeinungen werden in Form von *Meinungszertifikaten* übermittelt und gespeichert.

Im Folgenden wird zunächst auf die Struktur von Vertrauensprofilen (Abschnitt 3.7.1) und auf die Alterung der enthaltenen Meinungen eingegangen (Abschnitt 3.7.2), anschließend werden Meinungszertifikate behandelt (Abschnitt 3.7.3). Die zugehörigen Protokolle zum Austausch von Meinungen werden bei den Ausführungen zur Zugangskontrolle beschrieben (Abschnitt 3.9), da sie mit dieser verzahnt sind.

3.7.1 Vertrauenskategorien und Inhalt des Vertrauensprofils

Ein Vertrauensprofil beschreibt verschiedene Aspekte des Verhaltens eines anderen Teilnehmers bzw. des Vertrauens in diesen, die im Weiteren als *Vertrauenskategorien* bezeichnet werden sollen. Es beinhaltet deshalb eine Reihe von Vertrauensmaßzahlen (siehe Abschnitt 3.5), jeweils eine pro Vertrauenskategorie.

3.7.1.1 Konkrete und virtuelle Kategorien

Neben den im Folgenden als *konkret* bezeichneten Vertrauenskategorien, deren Maßzahlen direkt anhand beobachteten korrekten oder inkorrekten Verhaltens bestimmt und deshalb auch durch Zähler r und s für solche Beobachtungen repräsentiert werden, sollen auch noch *virtuelle* Kategorien verwendet werden. Deren Maßzahlen berechnen sich aus den in bestimmter Weise gewichteten Maßzahlen konkreter Kategorien. Ein Beispiel für eine virtuelle Vertrauenskategorie ist die „Kooperativität“: Damit ein Knoten insgesamt als kooperativ gelten kann, muss er nicht nur bei der Weiterleitung, sondern auch bei einigen anderen grundlegenden Dienstleistungen mitarbeiten, etwa indem er Einschätzungen und Schlüssel weitergibt. Deshalb berechnet sich die Maßzahl für die Kooperativität aus denen aller entsprechenden konkreten Kategorien.

3.7.1.2 Empfehlungsvertrauen

Empfehlungsvertrauen ist Vertrauen bezüglich der Fähigkeit zur Einschätzung Anderer und der ehrlichen Weitergabe solcher Einschätzungen (siehe auch Abschnitt 2.3.5.1). Für das hier entworfene Konzept soll davon ausgegangen werden, dass Empfehlungsvertrauen transitiv ist, dass also ein Teilnehmer A, der Teilnehmer B Empfehlungsvertrauen entgegenbringt und weiß, dass Teilnehmer B dies gegenüber Teilnehmer C tut, daraus ebenfalls Empfehlungsvertrauen gegenüber C ableitet.

Grundsätzlich könnte man zu jeder der verwendeten Vertrauenskategorien eine eigene Kategorie für das zugehörige Empfehlungsvertrauen anlegen. Stattdessen soll hier aber der Einfachheit halber davon ausgegangen werden, dass die Fähigkeit und Willigkeit zur Weitergabe von Empfehlungen nicht von der Vertrauenskategorie abhängt, auf welche sich die jeweilige Empfehlung bezieht. Deshalb wird nur eine einzige Kategorie für das Empfehlungsvertrauen verwendet.

Empfehlungsvertrauen kann nicht in gleicher Weise aus Beobachtungen gewonnen werden, wie die Einschätzungen anderer Kategorien. Deshalb wird es auch in der im folgenden Abschnitt gegebenen Aufstellung von Vertrauenskategorien nicht erwähnt. Stattdessen wird in Abschnitt 3.8 detailliert auf das Verfahren zur Gewinnung von Empfehlungsvertrauen eingegangen.

3.7.1.3 Aufstellung vorkommender Vertrauenskategorien

Im Folgenden werden einige sinnvolle Vertrauenskategorien aufgeführt. Zuerst sollen die konkreten Kategorien genannt werden, die jeweils in Form zweier Zähler für positive und negative Beobachtungen vorliegen. Hierzu ist jeweils angegeben, durch welche Ereignisse diese Zähler erhöht werden.

- *Weiterleitung:* Diese Kategorie enthält das Ergebnis der Beobachtung der Weiterleitung, die in Abschnitt 3.6 detailliert beschrieben wurde. Positiv zählen danach beobachtete weitergeleitete Pakete, wenn eindeutig erkannt werden kann, dass sie tatsächlich weitergeleitet und nicht etwa selbst erzeugt wurden. Negativ zählt die unterlassene Weiterleitung von Paketen, wenn angenommen werden kann, dass die Unterlassung schuldhaft erfolgt. Wird wegen fehlgeschlagener Zugangskontrolle nicht weitergeleitet, so führt die dabei erzeugte Fehlermeldung dazu, dass Beobachter nicht negativ werten.

- *Anfrage-Antwort-Dienste:* Bei allen Diensten, die erbracht werden, indem ein Knoten (der Dienstnehmer) eine einteilige Anfrage an einen anderen Knoten (den Dienstgeber) stellt und darauf eine einteilige Antwort erhält, kann in gleicher Weise bewertet werden: Positiv zählt jede in der eigenen Nachbarschaft beobachtete oder an den Beobachter als Dienstnehmer adressierte Antwort, der eine entsprechende Anfrage voranging. Negativ zählt das Ausbleiben von Antworten auf beobachtete Anfragen, wenn der Dienstgeber in der eigenen Nachbarschaft ist. In Abschnitt 3.6.6 wurde dieses Vorgehen näher beschrieben und begründet. Folgende Anfrage-Antwort-Dienste sind feste Bestandteile des hier entworfenen Konzepts:

  - *Meinungsäußerung*: Bewertung dafür, ob ein Knoten auf Anfrage seine Meinung äußert.
  - *Schlüssellieferung:* Bewertung dafür, ob ein Knoten auf Anfrage den Schlüssel der Quelle eines von ihm weitergeleiteten Pakets liefert (siehe Abschnitt 3.10).
  - *Bürgschaftsprotokoll:* Bewertung dafür, ob Bürgschaftsanfragen beantwortet werden. Dieser Dienst muss etwas anders als die vorigen behandelt werden, weil er möglicherweise Benutzerinteraktion beinhaltet. Es muss deshalb eine wesentlich längere Reaktionszeit vorgesehen werden. Andererseits wird dieser Dienst auch relativ selten benötigt, so dass die Notwendigkeit, über längere Zeit Zustandsinformation über im Ablauf befindliche Dienstnutzungen zu halten, nicht zu exzessivem Speicherbedarf führt.

  Die Bewertung ist unabhängig davon, ob die Antwort dem Anfrager inhaltlich weiterhilft; bewertet wird zunächst nur die Mitarbeit entsprechend des Protokolls. Positiv wird also auch dann gewertet, wenn die Antwort eine Fehlermeldung ist, die den Dienstnehmer darüber informiert, warum der Dienstgeber die Anfrage nicht beantworten kann, bzw. wenn auf eine Bürgschaftsanfrage eine ablehnende Antwort erfolgt. Die Qualität der erhaltenen Antwort kann zusätzlich gesondert bewertet werden. Vorgesehen sind beispielsweise die folgenden Kategorien:

  - *Qualität von Meinungsäußerungen:* Hier handelt es sich um eine subjektive Bewertung dafür, ob die Antwort die Erwartungen des Anfragenden erfüllt hat. Konkret wird dann positiv gewertet, wenn die auf Anfrage gelieferten Meinungen dazu führen, dass eine Gesamteinschätzung ausreichend sicher wird, so dass eine ausstehende Zugangsentscheidung getroffen werden kann. Leicht negativ wird gewertet, wenn die Antwort keine für ausstehende Zugangsentscheidungen relevante Information enthält.
  - *Qualität von Schlüssellieferungen:* Ähnlich wie bei Meinungsäußerungen wird hier bewertet, ob der auf Anfrage gelieferte Schlüssel zur Verifikation der Signatur eines wartenden Pakets geeignet war oder nicht.
  - *Bürgschaftsübernahme:* Knoten, die andere um die Übernahme einer Bürgschaft gebeten hatten, werten hier positiv für übernommene und negativ für abgelehnte Bürgschaften. Kämen Bürgschaftsanfragen nur von neu zum Netz hinzugekommenen Knoten, so würde sich die Zählung von Beobachtungen dafür kaum lohnen. Die Nutzung des Bürgschaftsverfahrens kann aber auch für bereits teilweise „etablierte“ Knoten sinnvoll sein, wenn sie Netzbereiche nutzen wollen, in denen sie noch weitgehend unbekannt sind.

- *Sonstige Dienste:* Auf andere, auch anwendungsbezogene Dienste wird in dieser Arbeit bezüglich Beobachtung und Bewertung nicht näher eingegangen. Soweit die Dienste dem Anfrage-Antwort-Paradigma folgen, was häufig der Fall ist, kann in der Regel wie oben beschrieben bewertet werden. Ansonsten lässt sich allgemein nur sagen, dass beobachtetes korrektes Verhalten, das den Regeln des jeweiligen Dienstes genügt, positiv bewertet werden sollte, und offensichtlich falsches Verhalten oder völlig ausbleibende Reaktion negativ.

Basierend auf den genannten konkreten Vertrauenskategorien können virtuelle Kategorien definiert werden. Momentan ist nur eine solche fest vorgesehen, für bestimmte Anwendungszwecke können aber andere sinnvoll sein und bei Bedarf definiert werden.

- *Kooperativität:* Die Kategorie Kooperativität dient zur Bildung einer Gesamteinschätzung der Kooperativität von Netzknoten. Ihr virtueller Wert wird bei Abfrage aus Werten konkreter Kategorien abgeleitet. Für die in Kapitel 4 durchgeführten Simulationen gingen beispielsweise die Werte folgender konkreter Kategorien mit den jeweils dahinter angegebenen Faktoren für r und s in den Kooperativitätswert ein:

Kategorie	*r-Faktor*	*s-Faktor*
Weiterleitung	1	1
Meinungsäußerung auf Anfrage	1	1
Schlüssellieferung auf Anfrage	0,75	1
Qualität von Schlüssellieferungen	0,5	0,5

  Die Wahl der Faktoren bei den beiden die Schlüssellieferung betreffenden Kategorien bewirkt eine einfache negative Wertung bei unterlassener Antwort, aber die Vergabe von insgesamt 1,25 Kooperativitäts-Bewertungspunkten für jede erfolgte Antwort. Damit wird ausgedrückt, dass durch die erhaltene Antwort insgesamt mehr Information über die Kooperativität des Dienstgebers gewonnen werden kann, als ohne Antwort.

3.7.2 Alterung von Fremdmeinungen

Gespeicherte Fremdmeinungen geben Einschätzungen anderer Knoten bezüglich Dritter zu bestimmten Zeitpunkten in der Vergangenheit wieder. Diese Einschätzungen können sich bei ihren Inhabern im Lauf der Zeit durch weitere Beobachtungen ändern, während die zu jeweils einem bestimmten Zeitpunkt in ein Meinungszertifikat verpackte und auf einem anderen Knoten als Fremdmeinung gespeicherte Form natürlich gleich bleibt, so lange sie nicht durch eine neuere Meinungsäußerung ersetzt wird. Diese mögliche und mit fortschreitender Zeit auch immer wahrscheinlicher werdende Abweichung zwischen tatsächlicher Einschätzung und gespeicherter Fremdmeinung wird ausgeglichen, indem bei der Einbeziehung gespeicherter Fremdmeinungen deren Alter – also die Zeitspanne seit Ausstellung des Meinungszertifikats – berücksichtigt wird: Eine Fremdmeinung wird als um so unsicherer angesehen, je älter sie ist.

Zur Bestimmung der tatsächlich berücksichtigten Meinung $\pi' = (p', c')$ zur Fremdmeinung $\pi = (p, c)$ wird π zunächst auf $\phi_{r,s}$ im Beweisraum abgebildet. Dann werden r', s' berechnet als

$$r' = \begin{cases} r - \lambda t & \text{falls} \quad r > s \wedge r \geq \lambda t \\ r - \lambda t\, r/s & \text{falls} \quad r \leq s \wedge r \geq \lambda t\, r/s \\ 0 & \text{sonst} \end{cases}, \qquad s' = \begin{cases} s - \lambda t\, s/r & \text{falls} \quad r > s \wedge s \geq \lambda t\, s/r \\ s - \lambda t & \text{falls} \quad r \leq s \wedge s \geq \lambda t \\ 0 & \text{sonst} \end{cases}.$$

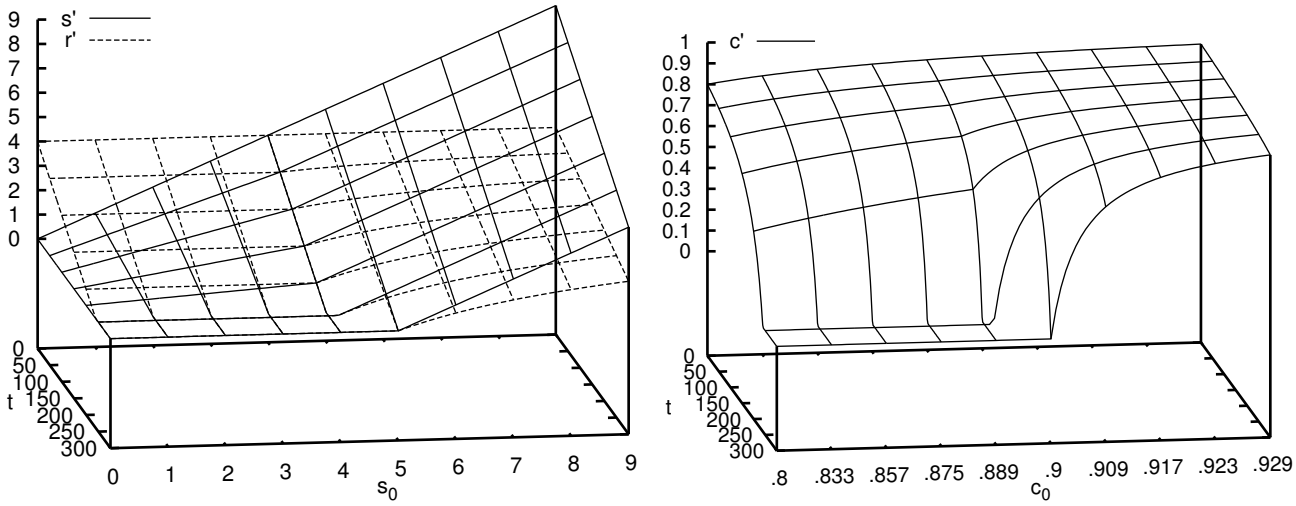


Abbildung 3.28: Zeitabhängige Verringerung von r und s bzw. c bei Startwerten von $r_0 = 4$ und $s_0 = 0 \ldots 9$, $\lambda = \frac{1}{60\,\mathrm{s}}$

Dies bedeutet anschaulich, dass der größere der beiden Werte r und s proportional zur verstreichenden Zeit sinkt, während der kleinere der beiden Werte – falls er ungleich Null ist – genau so schnell (und ebenfalls linear) sinkt, dass beide gleichzeitig Null erreichen. Abbildung 3.28 stellt dies auf der linken Seite für die Startwerte $r_0 = 4$ und s_0 im Bereich $0 \ldots 9$ dar, wobei $\lambda = \frac{1}{60\,\mathrm{s}}$ ist.

Durch Abbildung in den Meinungsraum wird aus $\phi_{r',s'}$ nun $\pi' = (p', c')$ gewonnen. Konstruktionsgemäß gilt dabei stets $p' = p$, falls $0 \neq p \neq 1$ ist. Die Sicherheit c' fällt dagegen mit der verstrichenen Zeit. Abbildung 3.28 stellt dies auf der rechten Seite für dieselben Startmeinungen wie links dar.

3.7.3 Austausch von Einschätzungen

3.7.3.1 Einschätzungszertifikat

Als Antwort auf Anfragen anderer Knoten, etwa bei der Zugangskontrolle (siehe Abschnitt 3.9.3.3) können Einschätzungen in Form so genannter *Einschätzungszertifikate* (auch *Meinungszertifikate*) verschickt werden, die ähnliche Eigenschaften wie Schlüssel-Zertifikate haben: Sie können beliebig weitergegeben und während einer vom Ersteller festzulegenden Gültigkeitszeitspanne beliebig oft von verschiedenen Knoten verwendet werden.

Einschätzungszertifikat
Schlüssel des einzuschätzenden Teilnehmers
Anzahl der übermittelten Kategorien
Liste von Einschätzungseinträgen (einer pro Kategorie)
Zeitstempel
Maximale Gültigkeitsdauer
Signaturfolgenummer und Signatur des Mitteilenden

Jeder Einschätzungseintrag besteht aus einer Kennung für die beschriebene konkrete Vertrauenskategorie sowie der zugehörigen Einschätzung. Diese wird in möglichst platzsparender Form ausgedrückt, nämlich durch Zählerwerte r und s, für die jeweils 2 byte vorgesehen sind. Größere Zählerwerte als 65535 können damit nicht ausgedrückt werden, aber sie bewirken sowieso keine merkliche Änderung der Einschätzung mehr.

Ein Einschätzungszertifikat enthält den vollen Schlüssel des Zielknotens und nicht nur eine Schlüsselkennung in Form eines Hashwerts. Dies hat folgenden Grund: Anhand eines Hashwerts kann der

zugehörige Schlüssel nur identifiziert werden, wenn er bereits vorliegt. Knoten, welche die Übertragung eines Einschätzungszertifikats mithören, aber den Schlüssel nicht besitzen, könnten dieses nicht nutzen. Dadurch, dass der Schlüssel im Einschätzungszertifikat enthalten ist, werden alle mithörenden Knoten in die Lage versetzt, ein Vertrauensprofil für den Zielknoten anzulegen, auch wenn sie dessen Schlüssel vorher nicht kannten. Damit erhöht sich die Nützlichkeit des Einschätzungszertifikats auch für Knoten, die es nicht selbst angefordert haben, und insgesamt wird eine schnellere Verbreitung von Einschätzungen im Netz gefördert.

3.7.3.2 Bürgschaftsanfrage

Mittels einer Bürgschaftsanfrage kann ein Knoten einen anderen bitten, für ihn zu bürgen, so dass er eine genügend gute Einschätzung erhält, um mit dem bürgenden Knoten kommunizieren zu können. Das ist insbesondere dann sinnvoll, wenn ein neuer Knoten einem bestehenden Netz eben deshalb beitritt, weil er mit dem Bürgen kommunizieren möchte. Wenn der Kommunikationswunsch beiderseitig ist, wird der Bürge der Bürgschaftsanfrage in der Regel entsprechen; wenn der potentielle Bürge nicht mit dem Antragsteller kommunizieren möchte, braucht dieser den Netzzugang auch nicht.

Bürgschaftsanfrage
Schlüsselkennung des Anfragenden
Volle Signaturfolgenummer des Anfragenden Signatur des Anfragenden

Mit der Übernahme der Bürgschaft geht der Bürge ein gewisses Risiko ein, da ein Fehlverhalten des neuen Knotens auf ihn zurückfällt. Schon aus diesem Grund muss für die Übernahme einer Bürgschaft ein ausdrückliches Einverständnis des Benutzers vorliegen.

Bürgschaftsanfragen müssen naturgemäß auch dann im Netz weitergeleitet werden, wenn ihr Absender nicht als vertrauenswürdig eingeschätzt wird. Die Rate weitergeleiteter Bürgschaftsanfragen wird deshalb streng begrenzt, um Missbrauch zu verhindern.

Die Antwort auf eine Bürgschaftsanfrage ist ein Einschätzungszertifikat. Während dessen Übertragung zurück zum Anfragenden wird dieses von jedem Weiterleiter und Beobachter, der Einschätzungen des Bürgen vertraut, gespeichert. Dadurch wird die Wahrscheinlichkeit dafür, dass der Anfragende anschließend Pakete zum Bürgen senden kann, ohne an der Zugangskontrolle zu scheitern, stark erhöht.

3.8 Gewinnung von Empfehlungsvertrauen

Empfehlungsvertrauen kann nicht wie Vertrauen anderer Kategorien aus der Beobachtung ablaufender Protokolle gewonnen werden. Zur Gewinnung von Empfehlungsvertrauen muss vielmehr erfasst werden, wie gut von anderen geäußerte Meinungen mit den tatsächlichen Beobachtungen übereinstimmen.

3.8.1 Versuch der beobachtungsorientierten Gewinnung

Man könnte nun zunächst versuchen, bei jeder erfassten Beobachtung neben der zugehörigen Maßzahl für Vertrauen in den beobachteten Teilnehmer auch die Maßzahlen für Empfehlungsvertrauen

in andere Teilnehmer anzupassen, sofern von diesen Teilnehmern Meinungen über den beobachteten Teilnehmer zur passenden Kategorie vorliegen.

Wenn ein Ereignis im Verhalten eines Knotens N_0 beobachtet und der r- bzw. s-Wert der zugehörigen Vertrauensmaßzahl um Δr_0 bzw. Δs_0 erhöht wird, würde dann z. B. außerdem für jede lokal gespeicherte Fremdmeinung $\pi_{N_0}^{N_i} = (p_0^i, c_0^i)$ das eigene Empfehlungsvertrauen $\pi_{N_i} = (p_i, c_i)$ in deren Äußerer N_i modifiziert, indem die zugehörigen r- und s-Werte r_i bzw. s_i um Δr_i bzw. Δs_i wie folgt erhöht werden:

$$\begin{aligned} \Delta r_i &= \Delta' r_i - \min(\Delta' r_i, \Delta' s_i) \\ \Delta s_i &= \Delta' s_i - \min(\Delta' r_i, \Delta' s_i) \end{aligned}$$

mit

$$\begin{aligned} \Delta' r_i &= (\Delta r_0 \quad p_0^i \quad + \Delta s_0 \, (1 - p_0^i)) \; c_0^i \\ \Delta' s_i &= (\Delta r_0 \, (1 - p_0^i) + \Delta s_0 \quad p_0^i \quad) \; c_0^i \end{aligned}$$

$\Delta' r_i$ und $\Delta' s_i$ enthalten beide den Faktor c_0^i, werden also umso größer, je höher die Sicherheit der geäußerten Fremdmeinung ist. Der zweite Faktor gewichtet das Ergebnis Δr_0, Δs_0 der Beobachtung anhand der Position der Fremdmeinung. Ist die Fremdmeinung positiv mit p_0^i nahe 1, so wird bei einer ebenfalls positiven Beobachtung mit Δr_0 nahe 1 und Δs_0 nahe 0 auch $\Delta' r_i$ relativ groß, während $\Delta' s_i$ nahe 0 bleibt. Gleiches gilt bei einer negativen Fremdmeinung (p_0^i nahe 0) und negativen Beobachtung (Δr_0 nahe 0, Δs_0 nahe 1). Trifft dagegen eine positive Fremdmeinung mit einer negativen Beobachtung zusammen, so wird $\Delta' s_i$ groß und Δr_0 klein, ebenso wie bei einer negativen Fremdmeinung und einer positiven Beobachtung.

Die Vorschrift zur Berechnung von Δr_i und Δs_i aus $\Delta' r_i$ und $\Delta' s_i$ bewirkt noch, dass der kleinere der beiden Werte verworfen und vom anderen Wert abgezogen wird. Statt z. B. r_i um 0.6 und s_i um 0.2 zu erhöhen, würde damit r_i nur um 0.4 erhöht und s_i unverändert gelassen. Dieses Vorgehen ähnelt einer als „Uncertainty Maximisation“ bezeichneten, von Jøsang angewendeten Vorschrift [Jøsa01].

Ein Beispiel: Bei einer positiven Beobachtung bezüglich Teilnehmer N_0 mit $\Delta r_0 = 1$, $\Delta s_0 = 0$ und einer vorliegenden Meinung $\pi_{N_0}^X = (0.8, 0.1)$ des Teilnehmers X über N_0 ergäbe sich eine Anpassung $\Delta r_X = 0.8 \cdot 0.1 = 0.08$ und $\Delta s_X = 0.2 \cdot 0.1 = 0.02$ des eigenen Empfehlungsvertrauens in N_0.

Leider spiegeln die mit diesem Verfahren ermittelten Maßzahlen für Empfehlungsvertrauen die Übereinstimmung zwischen Fremdmeinungen und der Gesamtheit der eigenen Beobachtungen in der Regel nur schlecht wider: Wenn das Verhalten des beobachteten Knoten nicht *immer* korrekt oder *immer* inkorrekt ist, wird die Einschätzung des Meinungsäußerers mit dem beschriebenen Verfahren nie $(1, 1)$ erreichen, auch wenn die geäußerte Meinung exakt den Beobachtungen entspricht, denn es werden entweder immer (ohne „Uncertainty Maximisation“) oder zumindest regelmäßig auch negative Beobachtungen registriert. Man erkennt daran, dass Fremdmeinungen nicht anhand einzelner Beobachtungen beurteilt werden dürfen, sondern stattdessen mit der längerfristig aufgrund dieser Beobachtungen ermittelten Einschätzung verglichen werden müssen.

3.8.2 Gewinnung durch Vergleich

Bei der Gewinnung von Empfehlungsvertrauens durch Vergleich von Fremdmeinungen mit eigenen Meinungen sind einige Punkte zu berücksichtigen:

- Der Vergleich sollte nicht nur einmalig – etwa bei Bedarf –, sondern regelmäßig durchgeführt werden, wobei ältere Ergebnisse in abnehmendem Maß berücksichtigt werden. Damit wird widergespiegelt, dass auch Empfehlungsvertrauen nur entstehen kann, wenn ein anderer Teilnehmer über einen längeren Zeitraum stets vertrauenswürdig erscheint.

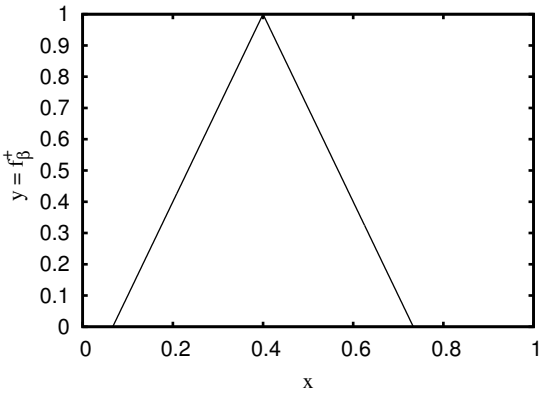


Abbildung 3.29: Verlauf des ersten Entwurfs (3.7) von $f_\beta^+(\pi_A, \pi_B)$ für $p_A = 0.4$, $\beta = 3$ und $p_B = x$

- Das Ergebnis des Vergleichs muss entweder selbst eine Vertrauensmaßzahl sein, oder es muss dazu verwendet werden, eine Vertrauensmaßzahl zu modifizieren. Zu diesem Zweck sollen hier Maßzahlen für Empfehlungsvertrauen ebenso wie solche für andere Kategorien durch zwei positive Größen r und s repräsentiert werden, die bei jedem Vergleich erhöht werden. Bei Bedarf kann dann eine Meinung der Form $\pi = (p, c)$ aus r und s ermittelt werden.

- Wenn mehrere Meinungen eines anderen Teilnehmers vorhanden sind, sind sie alle kumulativ zu berücksichtigen, wenn dessen Vertrauenswürdigkeit ermittelt wird. Sie werden also alle mit den entsprechenden eigenen Meinungen verglichen, und die resultierenden Änderungen an r und s addieren sich.

Zunächst soll eine Bewertungsfunktion $f^+ : \Pi \times \Pi \to [0,1]$ entworfen werden, welche sozusagen die Übereinstimmung zweier Vertrauensmaßzahlen $\pi_A = (p_A, c_A)$ und $\pi_B = (p_B, c_B)$ bewertet. Sie soll bei „völliger Übereinstimmung“ von π_A und π_B den Wert 1 liefern; wenn π_A und π_B dagegen „zu stark voneinander abweichen“, soll sie Null werden (wobei die genaue Bedeutung der in Anführungszeichen gesetzten qualitativen Übereinstimmungsmaße weiter unten noch deutlich wird). Der Wert $f^+(\pi_A, \pi_B)$ soll im Kontext der Gewinnung von Empfehlungsvertrauen von Teilnehmer A in Teilnehmer B zur Erhöhung des r-Werts der Maßzahl des Empfehlungsvertrauens verwendet werden.

3.8.2.1 Entwurf der Bewertungsfunktion f^+

Die Abweichung zweier Einschätzungen voneinander wird in erster Linie durch den Unterschied zwischen den Positionen p_A und p_B beschrieben. Ein erster Ansatz für die Definition von f^+ ist deshalb:

$$f_\beta^+(\pi_A, \pi_B) = \max(0, 1 - \beta\,|p_A - p_B|) \tag{3.7}$$

Der Faktor β legt dabei fest, wie stark sich die Abweichung zwischen den Positionen auf das Ergebnis auswirkt. Abbildung 3.29 zeigt den Verlauf von f^+ für $\beta = 3$, festes π_A und in der Position variiertes $\pi_B = (x, c_B)$. Bei $x = p_A$ ist der Funktionswert wunschgemäß 1, bei steigendem Abstand fällt er ab, bis er bei einem Abstand größer oder gleich $1/\beta$ den Wert 0 erreicht.

Wenn c_A oder c_B klein sind, soll der maximale Funktionswert niedriger ausfallen, als wenn die verglichenen Einschätzungen sehr sicher sind. Dafür soll sich aber bei kleinem c_A oder c_B der Abstand der Positionen weniger stark auf das Ergebnis auswirken. Damit soll ausgedrückt werden, dass bei relativ unsicheren Einschätzungen auch bei abweichender Position eine gewisse Übereinstimmung vorliegt (eben aufgrund der Unsicherheit), wobei diese Übereinstimmung dann allerdings weniger

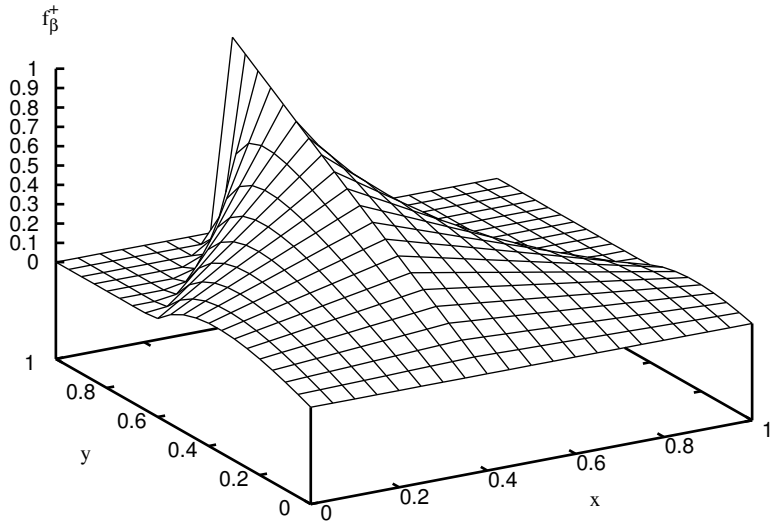


Abbildung 3.30: Verlauf des zweiten Entwurfs (3.8) von $f_\beta^+(\pi_A,\pi_B)$ für $\pi_A = (0.4, 0.99)$, $\beta = 1$ und $\pi_B = (x,y)$

wertvoll ist in dem Sinn, dass Rückschlüsse, die Teilnehmer A von diesem einzelnen Vergleich auf die generelle Korrektheit von Bs Einschätzungen zieht, ebenfalls unsicherer werden. Sind sich beide Teilnehmer dagegen bei ihren Einschätzungen sehr sicher, so ist eine genau übereinstimmende Position sehr wertvoll, während aber auch eine geringe Abweichung der Position als relevante Differenz angesehen werden muss.

Ein neuer Ansatz für die Definition von f^+, der dies berücksichtigt, ist der folgende:

$$f_\beta^+(\pi_A,\pi_B) = \max(0, c_A\,c_B - \beta\,\frac{c_A\,c_B}{1-c_A\,c_B}\,|p_A - p_B|) \qquad (c_A\,c_B \neq 1) \tag{3.8}$$

Der Maximalwert von f^+ ist hier $c_A\,c_B$, und die Steigung der Flanken wird durch den Faktor $\frac{c_A\,c_B}{1-c_A\,c_B}$ ergänzt, der bei $c_A\,c_B \to 0$ gegen null und bei $c_A\,c_B \to 1$ gegen unendlich strebt. Bei einem Abstand größer oder gleich $(1-c_A\,c_B)/\beta$ der Positionen wird der Funktionswert 0. Abbildung 3.30 zeigt den Graphen von f_β^+ bei $\pi_A = (0.4, 0.99)$, $\beta = 1$ und Variation von p_B in x- und c_B in y-Richtung.

Für $c_A\,c_B = 1$ ergibt sich eine Definitionslücke, die durch die Festlegung des Funktionswerts auf

$$f_\beta^+(\pi_A,\pi_B) = \begin{cases} 1 & \text{für} \quad p_A = p_B \\ 0 & \text{für} \quad p_A \neq p_B \end{cases} \qquad \text{bei} \quad c_A = c_B = 1$$

behoben wird. In der Praxis können Einschätzungen mit absoluter Sicherheit ($c = 1$) eigentlich nicht vorkommen, da durch Beobachtungen nie absolute Sicherheit erzielt werden kann. Damit können höchstens (unvorsichtige) Benutzervorgaben zu solchen Einschätzungen führen.

Die Bewertungsfunktion hat nun noch eine Schwäche, durch die sich ein Angreifer einen gewissen Vorteil verschaffen kann, indem er seine Meinungsäußerungen gezielt so gestaltet, dass er bei der Ermittlung von Empfehlungsvertrauen überdurchschnittlich gut bewertet wird. Eine Wahl von p_B in der Mitte des Definitionsbereichs, am besten $p_B = 0.5$, bietet nämlich bei Unkenntnis von p_A bessere Chancen, noch eine positive Bewertung zu erhalten, als eine Wahl am Rand, denn der Bereich, in dem f_β^+ positiv ist, liegt ja aus Sicht von B symmetrisch um p_A und ist damit im Extremfall in der Mitte doppelt so breit wie am Rand.

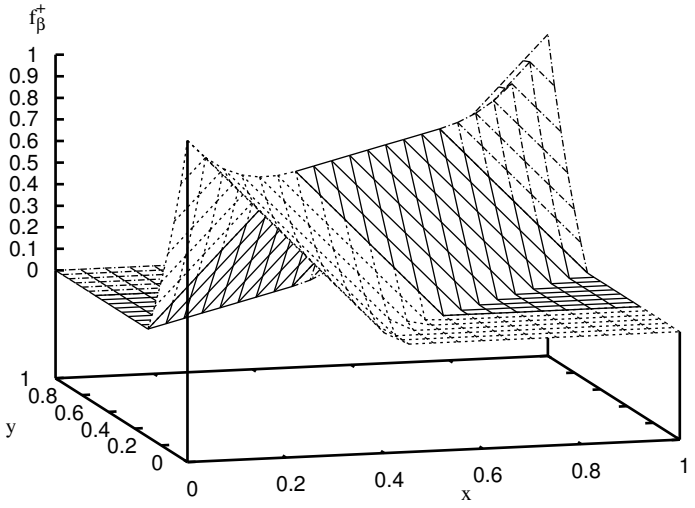


Abbildung 3.31: Verlauf von $f_\beta^+(\pi_A,\pi_B)$ nach (3.9) für $\pi_A=(x,1)$, $\pi_B=(y,0.7)$ und $\beta=1$

Die schlechtere Chance für eine positive Bewertung am Rand kann ausgeglichen werden, indem dort höhere Bewertungen vergeben werden. Dazu wird f_β^+ neu definiert durch

$$f_\beta^+(\pi_A,\pi_B)=\begin{cases}\max(0,\gamma_\beta(c_A,p_B,c_B)\,c_A\,c_B-\beta\,\dfrac{c_A\,c_B}{1-c_A\,c_B}\,|p_A-p_B|) & \text{für} \quad c_Ac_B\neq 1,\\ 1 & \text{für} \quad c_A\,c_B=1, p_A=p_B,\\ 0 & \text{für} \quad c_A\,c_B=1, p_A\neq p_B.\end{cases} \tag{3.9}$$

Der zusätzliche Faktor $\gamma_\beta(c_A,p_B,c_B)$ bewirkt die Anhebung der Funktionswerte in der Nähe des Randes. Er wird so bestimmt, dass der Erwartungswert der Bewertungsfunktion bei gleich verteiltem p_A

$$\int_0^1 f_\beta^+(\pi_A,\pi_B)\,dp_A$$

unabhängig von p_B ist, so dass also keine Wahl von p_B einen Vorteil bei der Bewertung verspricht. Das Integral soll immer den Wert annehmen, den es auch für das nicht um γ erweiterte f^+ hat, wenn p_B so liegt, dass $f_\beta^+(\pi_A,\pi_B)$ bei Variation in p_A rechts und links von p_B noch den Wert null erreicht, wenn also beide Flanken vollständig in den Bereich $[0,1]$ „passen“. Das Integral beschreibt dann die Fläche unter dem dreiecksförmigen Graphen mit Grundseite $g:=2\cdot(1-c_A\,c_B)/\beta$ und Höhe $h:=c_A\,c_B$ und hat damit den Wert $\frac{1}{2}\,g\,h=\frac{1}{\beta}\,c_A\,c_B\,(1-c_A\,c_B)$. Integriert man über (3.9) und setzt das Ergebnis mit dem genannten Wert gleich, so erhält man γ_β mit

$$\gamma_\beta(c_A,p_B,c_B)=\begin{cases}1 & \text{für} \quad p_B\geq\frac{1-c_A\,c_B}{\beta}\ \wedge\ p_B\leq\frac{c_A\,c_B+\beta-1}{\beta},\\ \frac{\sqrt{2}\sqrt{\beta^2\,p_B^2+(1-c_A\,c_B)^2}-\beta\,p_B}{1-c_A\,c_B} & \text{für} \quad p_B<\frac{1-c_A\,c_B}{\beta}\ \wedge\ p_B\leq\frac{c_A\,c_B+\beta-1}{\beta},\\ \frac{\sqrt{2}\sqrt{\beta^2\,(1-p_B)^2+(1-c_A\,c_B)^2}-\beta\,(1-p_B)}{1-c_A\,c_B} & \text{für} \quad p_B\geq\frac{1-c_A\,c_B}{\beta}\ \wedge\ p_B>\frac{c_A\,c_B+\beta-1}{\beta},\\ \frac{\beta^2((p_B-\frac{1}{2})^2+\frac{1}{4})+(1-c_A\,c_B)^2}{\beta\,(1-c_A\,c_B)} & \text{für} \quad p_B<\frac{1-c_A\,c_B}{\beta}\ \wedge\ p_B>\frac{c_A\,c_B+\beta-1}{\beta}.\end{cases}$$

Abbildung 3.31 zeigt den Verlauf von f_β^+, wenn p_A in x- und p_B in y-Richtung variiert werden ($c_A\,c_B=0.7, \beta=1$). Die jeweils verwendeten Fälle bei der Definition von γ_β sind an der Linienausprägung

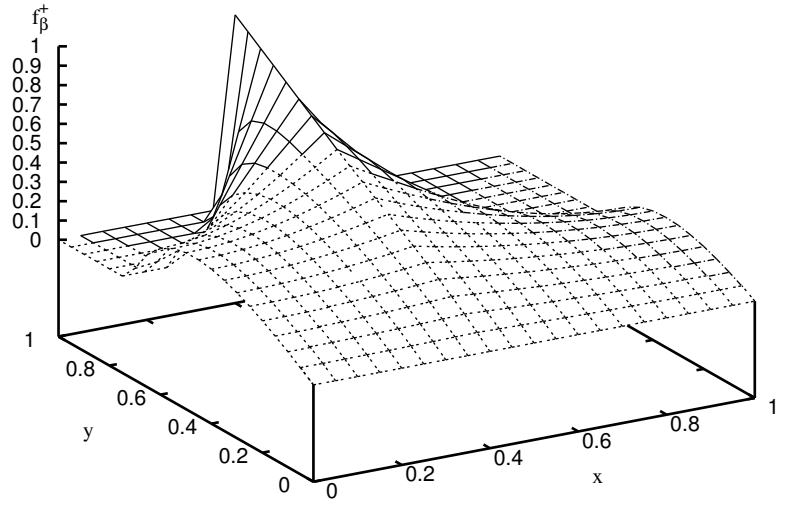


Abbildung 3.32: Verlauf von $f_\beta^+(\pi_A,\pi_B)$ nach (3.9) für $\pi_A=(0.4,0.99)$, $\beta=1$ und $\pi_B=(x,y)$

unterscheidbar: Der erste Fall ist durchgezogen dargestellt, der zweite gleichmäßig gestrichelt und der dritte kurz-lang-gestrichelt; der vierte Fall kommt nicht vor. Man erkennt deutlich, dass bei p_A-Werten nahe 0 oder 1 die Übereinstimmung mit p_B höher bewertet wird, als bei mittleren.

Abbildung 3.32 zeigt den Graphen von f_β^+ bei $\pi_A=(0.4,0.99)$, $\beta=1$ und Variation von p_B in x- und c_B in y-Richtung. Er ähnelt erwartungsgemäß dem aus Abbildung 3.30, aber es ist auch zu erkennen, dass die Funktionswerte für p_B nahe 0 bzw. 1 hier größer sind als dort. γ_β ist beim Übergang zum vierten Fall leider nicht stetig, sondern weist, wenn $c_A\,c_B<1-\beta/2$ ist, je einen kleinen Sprung bei $p_B=(1-c_A\,c_B)/\beta$ und $p_B=(c_A\,c_B+\beta-1)/\beta$ auf. In der Praxis ist dies aber nicht relevant; in Abbildung 3.32 ist es kaum zu erkennen.

Mittels der Bewertungsfunktion f_β^+ kann nun ein Teilnehmer A, der eine eigene Einschätzung π_Z^A des Teilnehmers Z besitzt, anhand einer vorliegenden Fremdmeinung π_Z^B des Teilnehmers B diesen bezüglich Empfehlungen positiv bewerten, und zwar in dem Maße, in dem Bs Meinung mit seiner eigenen Einschätzung übereinstimmt: Er erhöht dazu den B zugeordneten Zähler r für positive Ereignisse um den Wert $f_\beta^+(\pi_Z^A,\pi_Z^B)$.

3.8.2.2 Bewertung der Nicht-Übereinstimmung mittels f^-

Es fehlt nun noch eine entsprechende Möglichkeit der negativen Bewertung bei weniger gut übereinstimmenden Meinungen, also eine Bewertungsfunktion f_β^-, die ein angemessenes Inkrement für den Zähler s für negative Ereignisse liefert. Wir wählen dazu f_β^- als Komplement von f_β^+ in dem Sinn, dass

$$f_\beta^+(\pi_A,\pi_B)+f_\beta^-(\pi_A,\pi_B)=\mu_\beta(p_A,c_A,c_B)\quad\text{für beliebige}\quad\pi_A=(p_A,c_A),\,\pi_B=(p_B,c_B)\in\Omega$$

gelten soll, wobei μ_β der höchste Funktionswert von f_β^+ ist, der bei gegebenem π_A und c_B erreicht werden kann:

$$\mu_\beta(p_A,c_A,c_B):=\max_{x\in[0,1]} f_\beta^+((p_A,c_A),(x,c_B))=f_\beta^+((p_A,c_A),(p_A,c_B)).$$

Damit lässt sich also f_β^- definieren als

$$f_\beta^-(\pi_A,\pi_B):=f_\beta^+(\pi_A,(p_A,c_B))-f_\beta^+(\pi_A,\pi_B).\tag{3.10}$$

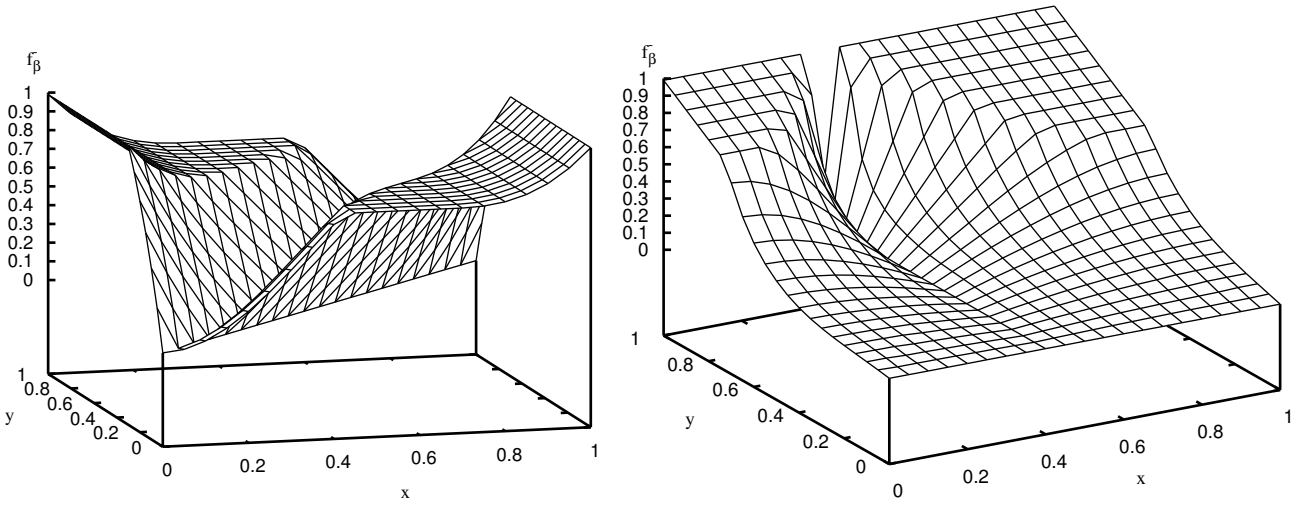


Abbildung 3.33: Verlauf von $f_\beta^-(\pi_A, \pi_B)$ nach (3.10) für $\pi_A = (x, 1)$, $\pi_B = (y, 0.7)$ (links) bzw. $\pi_A = (0.4, 0.99)$, $\pi_B = (x, y)$ (rechts) und $\beta = 1$

Abbildung 3.33 zeigt die Verläufe von f_β^- bei denselben Parametrisierungen, die bei f_β^+ zu den Graphen in den Abbildungen 3.31 bzw. 3.32 führten. Durch Vergleich ist anschaulich zu erkennen, inwiefern f_β^- komplementär zu f_β^+ ist. Es sei ausdrücklich darauf hingewiesen, dass bei jedem Bewertungsvorgang (bezüglich der oben eingeführten Teilnehmer) *sowohl r* um $f_\beta^+(\pi_Z^A, \pi_Z^B)$ *als auch s* um $f_\beta^-(\pi_Z^A, \pi_Z^B)$ erhöht wird. Wenn die Position der Fremdmeinung mit derjenigen der eigenen Einschätzung genau übereinstimmt, ist f_β^- null, so dass effektiv nur positiv bewertet wird; wenn die Positionen zu weit voneinander abweichen, ist f_β^+ null, so dass nur negativ bewertet wird. In allen anderen Fällen wird anteilig sowohl positiv als auch negativ gewertet.

Durch die angegebene Definition von f_β^- ist übrigens auch gewährleistet, dass $\int_0^1 f_\beta^-(\pi_A, \pi_B)\, dp_A$ nicht von p_B abhängt, so dass der oben beschriebene Angriff durch besondere Wahl von p_B auch hier verhindert ist. Außerdem gibt es auch kein bestimmtes c_B, durch welches ein Angreifer sich einen Vorteil verschaffen könnte, indem er es generell in den von ihm geäußerten Meinungen verwendet würde, denn c_B beeinflusst die Ausgabewerte der Bewertungsfunktionen f_β^+ und f_β^- in gleicher Weise; ein höher angegebenes c_B in einer geäußerten Meinung bewirkt also sowohl einen höheren positiven Maximalwert bei Übereinstimmung als auch einen höheren negativen Maximalwert bei fehlender Übereinstimmung.

3.9 Zugangskontrolle

Die Notwendigkeit von Zugangskontrolle in Ad-hoc-Netzen wurde bereits in der Einleitung des Kapitels motiviert (Abschnitt 3.1); im sich daran anschließenden Überblick wurde die Idee des in dieser Arbeit entworfenen Verfahrens grob umrissen (Abschnitt 3.3.5). Das Verfahren soll im Folgenden etwas detaillierter beschrieben werden. Dabei wird hauptsächlich auf den Zugang zum gesamten Netz eingegangen, der verweigert werden kann, indem Pakete des Zugang suchenden Knotens nicht weitergeleitet werden. Der Zugang zu von einzelnen Knoten zur Verfügung gestellten Diensten, die in der Regel auch vom jeweiligen Diensterbringer kontrolliert werden, kann nach denselben Kriterien erfolgen.

3.9.1 Idee des Verfahrens

Für jedes zur Weiterleitung eintreffende Paket wird zunächst anhand der Schicht-3-Absenderadresse der öffentliche Schlüssel des Quellknotens beschafft – dies ist Aufgabe der Schlüsselverwaltung (siehe Abschnitt 3.10). Die Korrektheit der Schlüsselzuordnung wird anhand einer am Paket angebrachten digitalen Signatur des Quellknotens überprüft (in Abschnitt 3.9.2.1 wird begründet, warum dies notwendig ist). Kann kein passender Schlüssel beschafft werden, so schlägt die Zugangskontrolle fehl und das Paket wird nicht weitergeleitet.

Liegt der Schlüssel der Quelle vor, so identifiziert er das zugehörige Vertrauensprofil, und es wird aus den dort gespeicherten Beobachtungen und Fremdmeinungen (siehe Abschnitt 3.7) eine Gesamteinschätzung der Kooperativität des Quellknotens ermittelt (anhand des in Abschnitt 3.5 beschriebenen Verfahrens). Wenn deren Sicherheit ausreichend hoch ist, wird anhand vorgegebener Schwellwerte entschieden, ob der Zugang genehmigt oder abgelehnt wird (Abschnitt 3.9.3.2) .

Falls die lokale Einschätzung der Kooperativität des Quellknotens zu unsicher ist, um eine Entscheidung treffen zu können, werden Meinungen anderer Knoten eingeholt. Dies erfolgt schrittweise in der Reihenfolge zunehmenden Aufwands, wobei nach jedem Schritt kontrolliert wird, ob die inzwischen aquirierten Fremdmeinungen eine genügend sichere Entscheidung erlauben. Zunächst fragt der kontrollierende Knoten seine Nachbarn (Abschnitt 3.9.3.3), dann sukzessive weiter entfernte Knoten, die bei ihm besonders hohes Empfehlungsvertrauen genießen (Abschnitt 3.9.3.4).

Falls mit diesen Maßnahmen keine ausreichend sichere Einschätzung gewonnen wird – und auch in allen anderen Fällen, in denen der Zugang nicht gestattet werden kann (etwa bei fehlendem Quellschlüssel) –, erfolgt als letzter Schritt mit der Ablehnung des Zugangs eine Fehlermeldung an den Quellknoten (Abschnitt 3.9.3.5). Dieser kann dann gegebenenfalls versuchen, die Ablehnungsursache selbst zu beheben, indem er den nötigen Schlüssel oder gespeicherte Fremdmeinungen über sich selbst liefert (Abschnitt 3.9.3.5).

Es sei an dieser Stelle ausdrücklich erwähnt, dass die Zugangskontrolle in der Vermittlungsschicht bestimmungsgemäß nur auf *weiterzuleitende* Pakete angewandt wird. Der *Empfänger* eines Pakets führt keine solche Zugangskontrolle durch, was zur Folge hat, dass die Vermittlungsschichtkommunikation zwischen benachbarten Knoten immer ohne Zugangskontrolle erfolgt. Auf höheren Schichten kann selbstverständlich trotzdem noch eine Zugangskontrolle erfolgen.

3.9.2 Bedrohungsanalyse

Das naheliegendste Ziel von Angriffen gegen die Zugangskontrolle ist deren Umgehung zum eigenen Vorteil, nämlich zur Verbesserung der eigenen Kommunikationsfähigkeit, indem erreicht wird, das eigene Pakete weitergeleitet werden, obwohl noch kein Nachweis für die eigene Kooperativität erbracht wurde (falls sie überhaupt besteht).

Unter Umständen können Versuche zur Verbesserung der Weiterleitungschancen auch zugunsten anderer (etwa „befreundeter“) Knoten unternommen werden. Wahrscheinlicher erscheinen hier aber Angriffe, die zum Ziel haben, die Weiterleitungschancen der Pakete anderer Knoten zu verschlechtern. Diese letztgenannte Art des Angriffs ist (wenn der Angreifer die betroffenen Pakete selbst weiterleitet), allerdings nur sinnvoll, wenn ein Vorteil gegenüber dem schlichten Verwerfen der Pakete erkennbar ist.

Potentiell angreifbare Information, welche die Zugangsentscheidung beeinflusst, kommt sowohl in den der Zugangskontrolle unterworfenen Paketen als auch auf den die Zugangskontrolle durchführenden Knoten vor:

- In den weitergeleiteten Paketen ist Information über ihre Quelle enthalten, deren Vertrauensprofil die Grundlage für die Zugangsentscheidung bildet. Ist die Quelle selbst Angreifer, so kann sie die Pakete gezielt mit falscher Information versehen. Eine nachträgliche Verbesserung oder Verschlechterung der Zugangschancen für bestimmte Pakete durch von der Quelle verschiedene andere Knoten in der Rolle des Angreifers erfolgt ebenfalls durch Modifikation der zugangskontrollrelevanten Information in diesen Paketen.

  Konkret wird der öffentliche Schlüssel und damit die Identität der Quelle anhand der Schicht-3-Absenderadresse eines Pakets bestimmt, wobei die Zugehörigkeit des Schlüssels zum Paketerzeuger anhand der Quellsignatur im Paket überprüft wird. Um nur durch Modifikation (bzw. verfälschte Erzeugung) eines Pakets eine andere Quelle vorzutäuschen müssen also Quelladresse und -signatur ausgetauscht werden. Um eine korrekte Erkennung der Quelle zu verhindern reicht es, durch eine beliebige Modifikation des Pakets zu bewirken, dass die Prüfung der Quellsignatur fehlschlägt. Diese beiden Angriffe werden in Abschnitt 3.9.2.1 näher untersucht. Außerdem spielt die korrekte Abbildung von der Quelladresse auf den Schlüssel eine Rolle; Angriffe zu deren Beeinflussung werden im Abschnitt zur Schlüsselverwaltung (3.10) behandelt.

- Auf den weiterleitenden Knoten sind die Vertrauensprofile selbst gespeichert. Sie umfassen einerseits selbst erfasste Beobachtungen und andererseits Fremdmeinungen, die gegebenenfalls in Form von Meinungszertifikaten von anderen Knoten erhalten wurden. Angriffe, die darauf abzielen, die selbst erfasste Information durch vom Angreifer verfälschte Beobachtungen zu manipulieren, wurden bereits in Abschnitt 3.6 diskutiert und durch die dort einführten Sicherungsmaßnahmen weitgehend unterbunden. Einflussnahme auf gespeicherte Fremdmeinungen kann theoretisch durch Erzeugung zusätzlicher sowie Unterschlagung, Wiederholung oder Modifikation regulärer Nachrichten der Protokolle zur Informationsbeschaffung versucht werden.

  Deshalb werden die Nachrichten dieser Protokolle – wie bei deren Beschreibung in den Abschnitten ab 3.9.3.3 noch deutlich wird – gegen Fälschung und Modifikation (durch Signatur) und gegen Wiederholung (durch Zeitstempel) gesichert, so dass nur die Möglichkeit der Unterschlagung verbleibt. Diese ist nur für Nachrichten von entfernten Knoten durchführbar, die über den Angreifer weitergeleitet werden. Konkret kann der Angreifer weiterzuleitende Meinungszertifikate, die ungünstige Einschätzungen über ihn selbst enthalten, verwerfen, statt sie weiterzuleiten, was ihm allerdings nur dann einen Vorteil bringt, wenn anderswo auch günstigere Einschätzungen über ihn existieren (denn auch ganz ohne bzw. mit neutralen Einschätzungen erhält er keinen Zugang). Da ein Angreifer in aller Regel nur in einen Bruchteil der Wege eingreifen kann, über die Meinungen über ihn ausgetauscht werden, spielt der Angriff praktisch keine Rolle.

Schließlich bleibt noch ein „Angriff“ zu nennen, der erst dadurch entsteht bzw. relevant wird, dass die Durchführung der Zugangskontrolle einen gewissen Aufwand für den durchführenden Knoten darstellt: Möchte ein Knoten diesen Zusatzaufwand einsparen, so kann er einfach auf Zugangskontrolle verzichten, also undifferenziert alle Pakete weiterleiten und darauf vertrauen, dass entweder keine unkooperativen Knoten vorhanden sind oder dass unkooperatives Verhalten durch Zugangskontrollmaßnahmen anderer Knoten verhindert wird. Wenn mehrere Netzteilnehmer diesen eigennützigen Ansatz verfolgen, wird die Wirkung der Zugangskontrolle insgesamt geschwächt. Da allerdings schon ein einzelner den Zugang ablehnender Knoten auf einem Weg durch das Netz ausreicht, um die Kommunikation auf diesem Weg zu unterbinden, ist zu erwarten, dass unkooperative Teilnehmer nur bei einem sehr hohen Anteil indifferenter Teilnehmer sinnvoll zusammenhängend kommunizieren können.

3.9.2.1 Vortäuschen einer anderen Quelle durch Ersetzen der Quellangabe im Paket

Wie in Abschnitt 3.4.4.1 beschrieben kann einerseits der Quellknoten prinzipiell beliebige Identitäten als Quellangabe in erzeugte Pakete eintragen, andererseits kann jeder weiterleitende Knoten unter Verwendung eines eigenen Schlüssels eine neue Signatur für ein weitergeleitetes Paket generieren, ohne dass diese Modifikation von anderen Knoten explizit als solche erkannt wird.

Damit die Signaturprüfung beim nächsten Weiterleiter nicht fehlschlägt, muss dafür gesorgt werden, dass der nächste Weiterleiter bei der Prüfung auch den zur neuen Signatur passenden Schlüssel verwendet. Kann der passende Schlüssel nicht gefunden werden, so schlägt auch die Zugangskontrolle fehl und das Paket wird verworfen. Als Angriff wäre damit die „gefälschte“ Angabe bzw. die Veränderung des weitergeleiteten Pakets wenig sinnvoll, da der Angreifer durch die Manipulation nur dasselbe wie durch simples Nichtabsenden bzw. Verwerfen erreicht (und im zweiten Fall nebenbei noch eine negative Bewertung für nicht erfolgte Weiterleitung erhält).

Liegt der passende Schlüssel vor, so erfolgt die Zugangskontrolle für das modifizierte Paket im weiteren Verlauf nicht mehr anhand des Vertrauensprofils der Quelle, sondern anhand des zum verwendeten Schlüssel des Angreifers gehörigen. Die genauen Auswirkungen der Verwendung einer anderen Identität sind allerdings für den Angreifer meist schwer abzuschätzen, da er dazu die zum ursprünglichen bzw. zum neuen Schlüssel gehörigen lokalen Vertrauensprofile auf allen weiteren Weiterleiterknoten kennen müsste. Bei vielen Wegfindungsverfahren kann er nicht einmal genau feststellen, welche Knoten auf dem weiteren Weg des Pakets liegen, da die Weiterleitungsinformation verteilt gehalten wird und er selbst nur den nächsten Weiterleiter kennt. Lediglich in einem Fall sind die Auswirkungen sicher vorhersehbar: Setzt der Angreifer einen bisher nicht verwendeten Schlüssel ein, so wird das zugehörige leere Vertrauensprofil zum Verwerfen des Pakets führen. Wie oben gilt aber, dass der Angreifer auch gleich selbst verwerfen und dabei zusätzlich einen Sendevorgang einsparen könnte.

Wenn der Angreifer die Möglichkeit hat, eine Identität zu verwenden, die zum Vertrauensprofil eines seit längerer Zeit aktiven Knotens gehört, welcher im Netz als kooperativ bewertet wird, so erhöhen sich wahrscheinlich die Chancen des Pakets, sein Ziel zu erreichen. Wegen der vorgeschrieben Quellsignatur im Paket kann der Angreifer aber nur *eigene* Identitäten verwenden, da er die privaten Schlüssel anderer Identitäten nicht kennt. Damit kann er auch nur auf Vertrauensprofile zurückgreifen, deren Inhalt er durch sein eigenes Verhalten bestimmt hat (denn Vertrauensprofile werden direkt durch Schlüssel identifiziert). Dies bedeutet auch, dass die Verwendung mehrerer verschiedener Identitäten durch einen einzigen Knoten für diesen insofern „teuer“ ist, als er ein Vielfaches an positiven Beobachtungen „erzeugen“ muss, um in allen Profilen als kooperativ gelten zu können.

Ohne Quellsignatur wären sowohl Falschangabe der Quellidentität im Paket durch die Quelle als auch Manipulation der Quellidentität durch Weiterleiter möglich und in manchen Fällen auch sinnvoll:

- Die Quelle selbst in der Rolle des Angreifers könnte beliebige Identitäten anderer Knoten angeben, um bei der Zugangskontrolle deren Vertrauensprofile zu verwenden. Das wäre insbesondere für neue Knoten interessant, die noch sehr wenig Vertrauen im Netz genießen, sowie für unkooperative Knoten, die selbst sehr viele negative Bewertungen gesammelt haben.
- Weiterleiter als Angreifer könnten unter Umständen fremde Pakete benutzen, um ihre eigenen Nutzdaten zu transportieren, indem sie die Quellangabe intakt lassen und die Nutzdaten verändern (wobei Kommunikationspartner der Angreifer eingeweiht sein müssten, um die wahre Herkunft der übermittelten Nutzdaten zu erkennen).

Durch die Quellsignatur als Nachweis sowohl der Inhaberschaft des zugehörigen Schlüssels (und damit des durch den Schlüssel identifizierten Vertrauensprofils) als auch der Integrität der übermittelten Nutzdaten werden diese Angriffe ausgeschlossen.

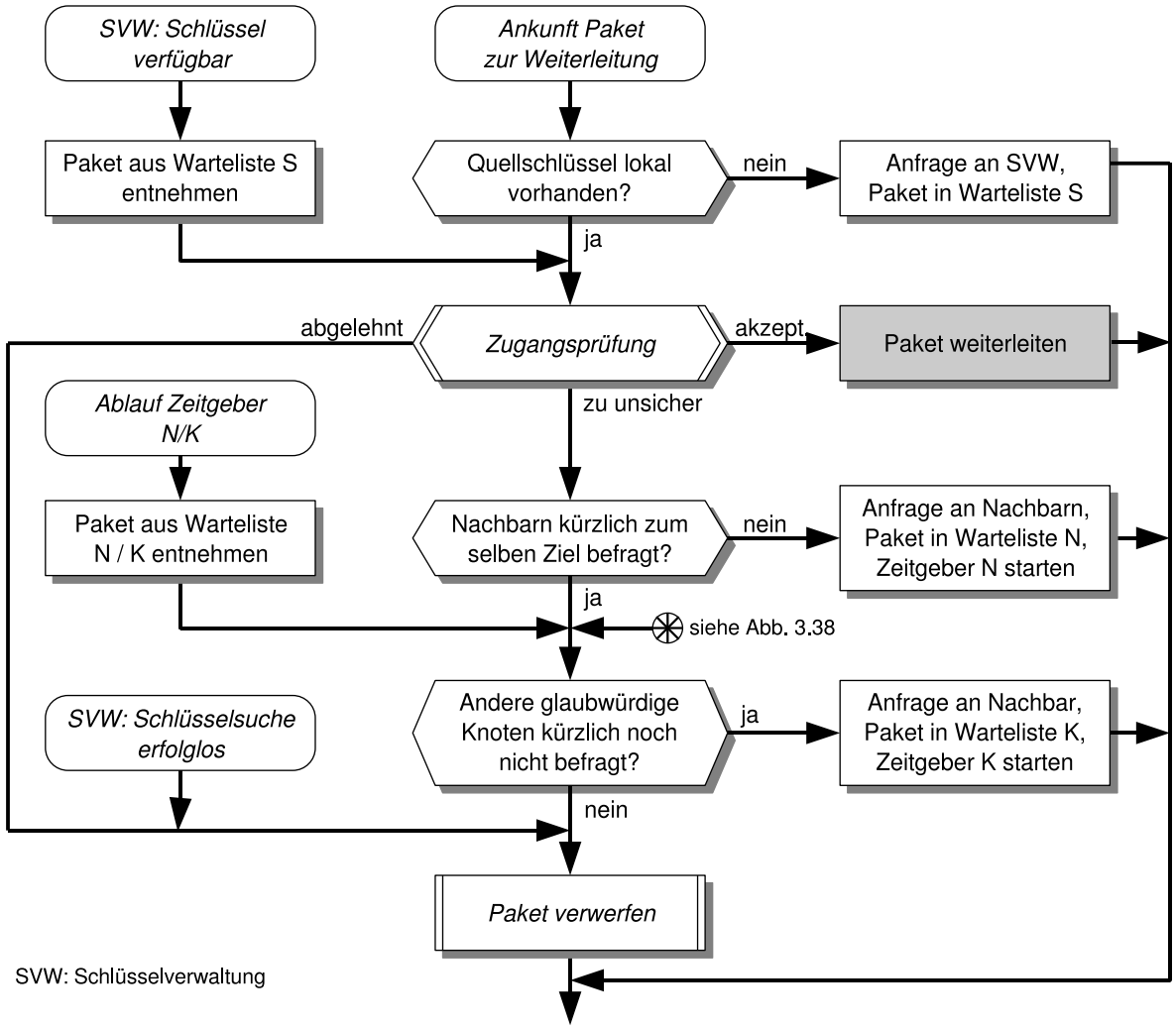


Abbildung 3.34: Verarbeitung weiterzuleitender Pakete bei der Zugangskontrolle

3.9.3 Detaillierte Beschreibung des Verfahrens

3.9.3.1 Identifikation der Quelle

Die Abläufe bei der Bearbeitung weiterzuleitender Pakete sind in Abbildung 3.34 dargestellt; die Integration in die Abläufe der Vermittlungsschicht ist bereits in Abbildung 3.24 gezeigt worden. Als erstes ist die Identität der Quelle festzustellen. Dazu wird überprüft, ob der Schlüssel der Quelle bereits lokal bekannt ist. Der Schlüssel wird dazu in der Liste bekannter Schlüssel anhand der Quelladresse des Pakets gesucht und zur Prüfung der Quellsignatur verwendet. Falls mehrere Schlüssel zur selben Adresse vorliegen, wird durch die Signaturprüfung der richtige ausgewählt. Die genannten Schritte sind eigentlich der Schlüsselverwaltung zuzuordnen, deshalb werden die Datenstrukturen und Operationen erst im zugehörigen Abschnitt 3.10 genauer beschrieben.

Nur wenn ein Schlüssel gefunden wird, der zur Quellsignatur passt, wird direkt mit der Zugangsprüfung fortgefahren. Ansonsten wird das Paket in eine Warteliste gelegt, und die Schlüsselverwaltung wird beauftragt, den passenden Schlüssel zu besorgen. Damit ist die Verarbeitung zunächst abgeschlossen. Erst wenn der benötigte Schlüssel zur Verfügung steht, wird das Paket wieder aus der Warteliste entfernt und es wird mit der Zugangsprüfung fortgefahren.

3.9.3.2 Zugangsprüfung

Mit dem Schlüssel der Quelle steht nun das diesem zugeordnete Vertrauensprofil zur Verfügung, anhand dessen die eigentliche Zugangsprüfung durchgeführt werden kann, wenn es bereits eine ausreichend sichere Einschätzung der Kooperativität des Quellknotens erlaubt.

Die Zugangsprüfung liefert eines der drei Ergebnisse „Zugang genehmigt“, „Zugang verweigert“ oder „Einschätzung zu unsicher“ und erfolgt in mehreren Schritten. Zunächst werden nach dem Widerstandsnetzwerksverfahren (Abschnitt 3.5) alle lokal vorliegenden Beobachtungen und Fremdmeinungen zu einer Gesamteinschätzung der Kooperativität des Quellknotens verknüpft. In die Berechnung der Widerstandswerte gehen dabei Meinungen der virtuellen Kategorie „Kooperativität“ ein, die nach einer Vorschrift wie in Abschnitt 3.7.1.3 beschrieben aus Meinungen konkreter Kategorien ermittelt werden. Das Ergebnis der Verknüpfung liegt dann in Form einer Meinung im Meinungsraum Π vor.

Um nun eine Zugangsentscheidung treffen zu können, werden feste Schwellen benötigt, mit denen die Gesamteinschätzung der Quelle verglichen werden kann. Solche Schwellen lassen sich plausibel im Beweisraum festlegen, was in Textform z. B. wie folgt aussehen könnte:

> „Als vertrauenswürdig genug gilt ein Verhalten, bei dem mindestens dreimal so viele positive wie negative Beobachtungen entstehen. Außerdem müssen mindestens 10 Beobachtungen vorliegen, damit die Bewertung als sicher genug gilt.“

Verallgemeinert man „dreimal“ auf „k-mal“ und 10 auf m Beobachtungen, so lässt sich eine solche Bedingung an eine Vertrauensmaßzahl im Beweisraum in der Form

$$\begin{aligned} r &\geq ks, \\ r+s &\geq m \end{aligned}$$

schreiben. Mit der Abbildung f kann man diese Bedingung in den Meinungsraum Ω als

$$\begin{aligned} b = \frac{r}{r+s+1} &\geq \frac{ks}{r+s+1} = kd, \\ u = \frac{1}{r+s+1} &\leq \frac{1}{m+1}, \end{aligned}$$

und von dort mit der Abbildung g in den Meinungsraum Π als

$$\begin{aligned} p = \frac{b}{b+d} &\geq \frac{kd}{b+d} = k(1-p) \quad \text{entsprechend} \quad p \geq \frac{k}{k+1}, \\ c = 1-u &\geq 1 - \frac{1}{m+1} = \frac{m}{m+1} \end{aligned}$$

übersetzen. Eine positive Zugangsentscheidung nach der oben als Beispiel in textueller Form angegebenen Richtlinie könnte also getroffen werden, wenn für die Gesamtkooperativitätseinschätzung (p,c) der Quelle gilt, dass $p \geq 3/4$ und $c \geq 10/11$ ist.

In Abbildung 3.35 sind auf der linken Seite die Untermengen von Π markiert, die sich durch solche Bedingungen an die Gesamteinschätzung (p,c) ergeben: Liegt c oberhalb einer Schranke c_{min}, so wird eine positive Zugangsentscheidung getroffen, wenn p außerdem rechts von p_{min} liegt.

Für $p < p_{min}$ bei $c \geq c_{min}$ deutet die Gesamteinschätzung auf mangelnde Kooperativität hin, und es kann hier eine negative Zugangsentscheidung getroffen werden; dies wurde auch bei den in Kapitel 4

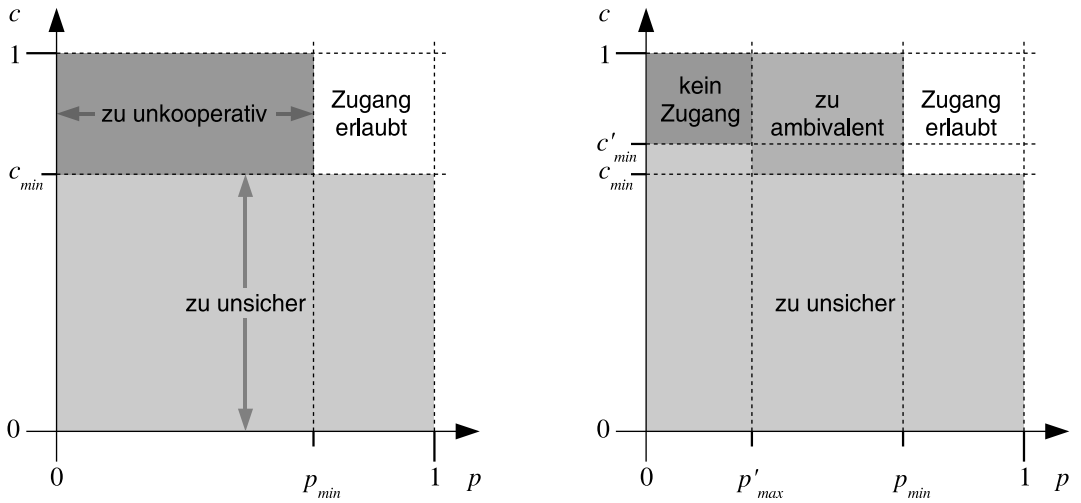


Abbildung 3.35: Zugangskontrollentscheidung abhängig von einer Gesamteinschätzung (p, c) nach einstufigem (links) oder zweistufigem (rechts) Verfahren

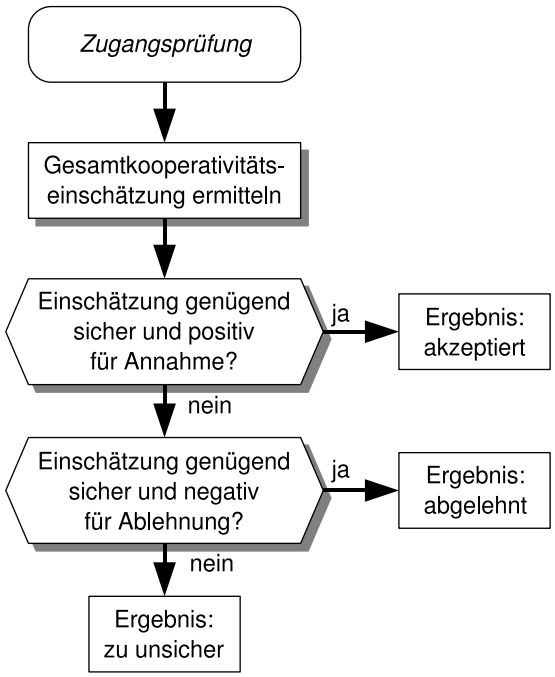


Abbildung 3.36: Detaillierter Ablauf der Zugangsprüfung

beschriebenen Simulationen so gehandhabt. Möchte man stattdessen für die negative Zugangsentscheidung eigene Grenzen festlegen, so kann z. B. bei Fehlschlagen einer *positiven* Zugangsbedingung der eben gezeigten Art analog eine zweite Bedingung an p und c geprüft werden, welche explizit an die *negative* Zugangsentscheidung geknüpft ist. Der Zugang wird dann nur abgelehnt, wenn $p \leq p'_{max}$ ist für eine zu wählende Schranke p'_{max}, wobei außerdem $c \geq c'_{max}$ für ein ebenfalls festzulegendes c'_{max} gelten muss (siehe Abbildung 3.35 rechts). Erfolgt trotz ausreichender Sicherheit durch diese zweite Bedingung keine Ablehnung des Zugangs, so kann ebenso wie bei zu unsicherer Gesamteinschätzung verfahren, also zunächst die Beschaffung zusätzlicher Fremdmeinungen versucht werden. Sind allderdings keine zusätzlichen Fremdmeinungen mehr verfügbar, so sollte der Zugang schließlich doch abgelehnt werden.

Die Abfolge der Schritte bei der Findung einer Zugangsentscheidung (nach dem zuletzt beschriebenen zweistufigen Verfahren) ist zusammenfassend in Abbildung 3.36 dargestellt.

3.9.3.3 Beschaffung von Meinungen der Nachbarn

Wie in Abbildung 3.34 dargestellt führt eine positiv verlaufende Zugangsprüfung zur Weiterleitung, eine negative zum Verwerfen des Pakets. Beim Resultat „Einschätzung zu unsicher“ wird zunächst die erste Stufe des Verfahrens zur Beschaffung zusätzlicher Meinungen betreten, bei der mit relativ geringem Aufwand Meinungen der eigenen Nachbarn eingeholt werden.

Damit in Fällen, wo auch in der Nachbarschaft keine ausreichend sicheren Einschätzungen vorhanden sind, nicht dieselbe Anfrage an die Nachbarn für eine Folge von Paketen einer bestimmten Quelle ständig wiederholt wird, wird im Vertrauensprofil der Quelle vermerkt, wann die Nachbarn des kontrollierenden Knoten zuletzt bezüglich dieser Quelle befragt wurden. Bevor eine neue Anfrage gestellt wird, wird dieser Eintrag jeweils geprüft, und falls nicht eine vorgegebene Zeitspanne seit der letzten Anfrage verstrichen ist, entfällt sie in diesem Fall und es folgt direkt die nächste Stufe der Meinungsbeschaffung (Abschnitt 3.9.3.4).

Wurde also kürzlich noch keine Anfrage zur Quelle des betrachteten Pakets an die Nachbarn gestellt, so erfolgt dies nun durch Aussenden der folgenden Nachricht per Broadcast:

Einschätzungsanfrage
Schlüsselkennung der Quelle
Niederwertiger Teil der Signaturfolgenummer des Anfragenden Signatur des Anfragenden

Als Schlüsselkennung wird der Ausgabewert einer kryptographischen Hashfunktion bei Anwendung auf den öffentlichen Schlüssel der Quelle verwendet. Anhand dieser Kennung können die Nachbarn den Schlüssel und damit das Vertrauensprofil der Quelle identifizieren. Die Anfrage wird signiert, damit etwaiger Missbrauch derartiger Anfragen sicher seinem Verursacher zugeordnet werden kann.

Nach Aussendung der Anfrage wird die Bearbeitung abgebrochen, indem das zu kontrollierende Paket in eine Warteliste gelegt und ein Zeitgeber über eine Zeitspanne gesetzt wird, innerhalb derer Antworten von Nachbarn erwartet werden. Läuft dieser Zeitgeber ab, während das Paket noch in der Warteliste liegt, so wird es daraus entnommen und der zweiten Stufe der Meinungsbeschaffung zugeführt.

Nachbarn, welche die Einschätzungsanfrage erhalten haben und eine nennenswert sichere Einschätzung des fraglichen Knotens besitzen, senden als Antwort ein Einschätzungszertifikat (dessen Aufbau bereits in Abschnitt 3.7.3.1 beschrieben wurde). Falls ihnen außerdem Fremdmeinungen vorliegen, können sie auch diese (ebenfalls in ihrer ursprünglichen Form des Einschätzungszertifikats) ebenfalls

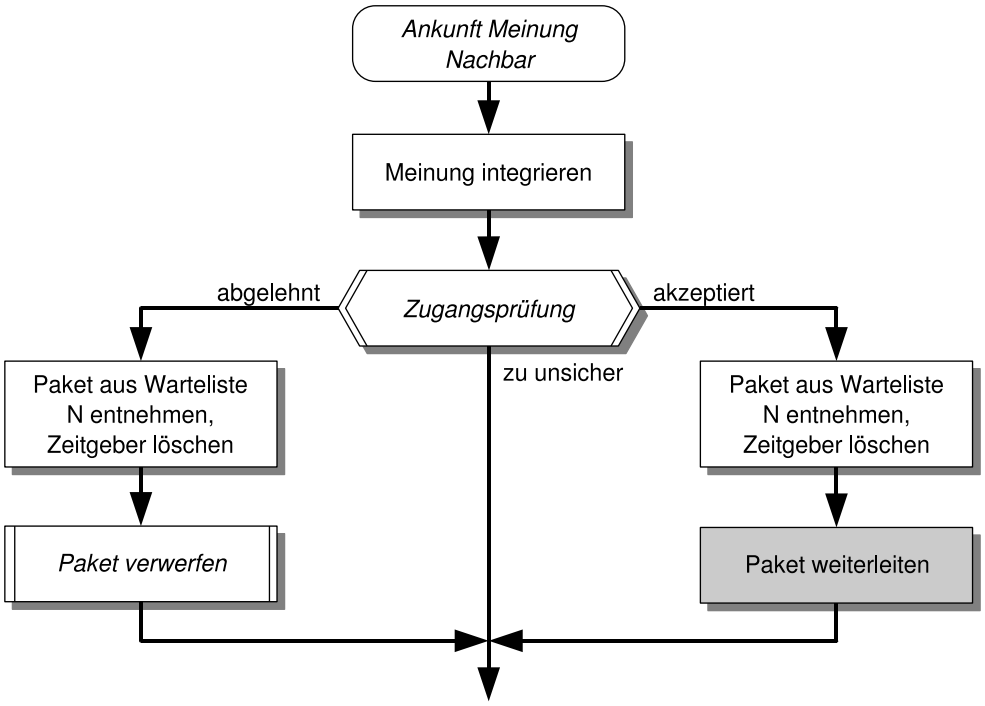


Abbildung 3.37: Verarbeitung der Antwort auf eine Meinungsanfrage an die Nachbarn

an den Anfragesteller weitergeben. Die Verarbeitung eintreffender Antworten ist in Abbildung 3.37 dargestellt: Die neue Meinung wird in die eigene Datenbasis aufgenommen und es wird ein erneute Zugangsprüfung durchgeführt. Liefert diese nun ein eindeutiges Ergebnis, so wird das Paket aus der Warteliste entnommen und entweder verworfen oder weitergeleitet; der zugehörige Zeitgeber wird gelöscht. Ist die Gesamtkooperativitätseinschätzung immer noch zu unsicher, so verbleibt das Paket in der Warteliste und der Zeitgeber läuft weiter.

3.9.3.4 Beschaffung von Meinungen vertrauenswürdiger entfernter Knoten

Liefert auch die Anfrage bei Nachbarn keine ausreichend sichere Gesamtkooperativitätseinschätzung des Quellknotens eines weiterzuleitenden Pakets (oder wurde diese Anfrage gar nicht durchgeführt, weil sie gleich lautend schon kurz vorher aufgetreten war), so bleibt als zweite Stufe der Meinungsbeschaffung noch die Meinungsanfrage bei weiter entfernten Knoten. Damit sich eine solche, mit höherem Übertragungsaufwand verbundene Anfrage lohnt, sollte sie nur an entfernte Knoten gerichtet werden, die beim Anfragenden bereits nennenswertes Empfehlungsvertrauen genießen, denn nur dann können die erhaltenen Antworten die Sicherheit der Einschätzung des fraglichen Quellknotens wesentlich erhöhen.

Um eine Auswahl lohnender Anfrageziele zu erleichtern, pflegt jeder Knoten ständig eine Liste derjenigen anderen Knoten, die bei ihm das höchste Empfehlungsvertrauen genießen. Diese Liste hat eine vorgegebene Maximalgröße, kann aber auch leer sein. Wenn eine Meinungsanfrage an einen der Knoten auf der Liste gerichtet wird, so wird dies zusammen mit dem Zeitpunkt der Anfrage im Vertrauensprofil desjenigen Knotens vermerkt, der den Gegenstand der Anfrage bildet. Anhand dieser Vermerke kann – ähnlich wie oben – verhindert werden, dass wiederholt gleich lautende Anfragen an denselben entfernten Knoten gerichtet werden.

Eine Meinungsanfrage an einen entfernten Knoten kann also nur durchgeführt werden, wenn noch ein Knoten auf der genannten Liste steht, an den dieselbe Anfrage nicht kürzlich schon gerichtet wurde.

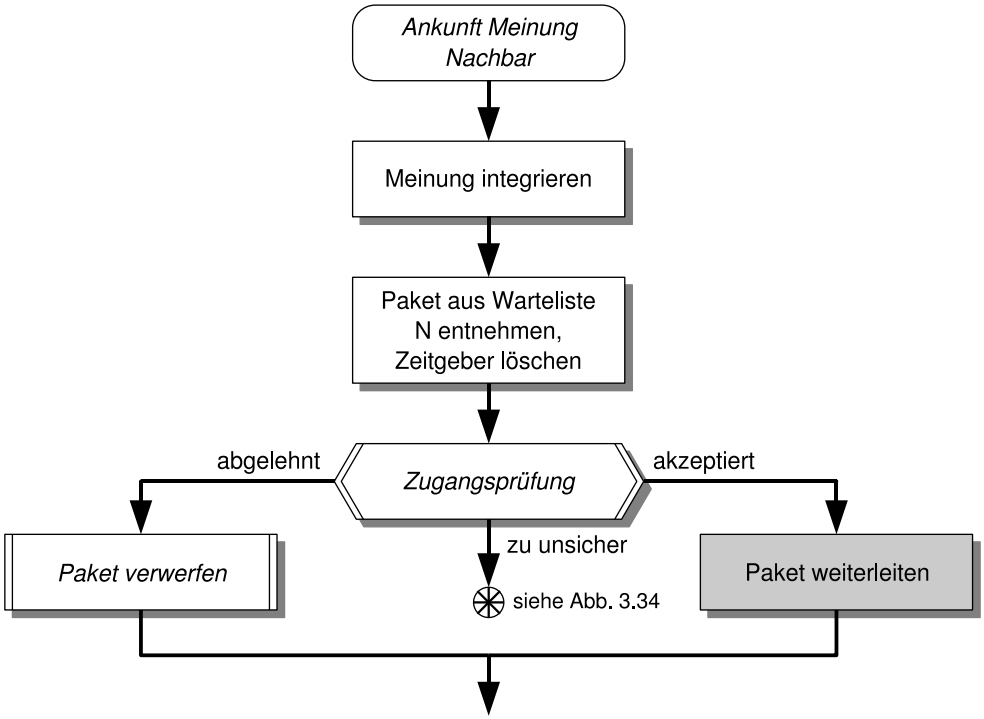


Abbildung 3.38: Verarbeitung der Antwort auf eine Meinungsanfrage an einen nicht benachbarten vertrauenswürdigen Knoten

Gibt es einen solchen Knoten, dann erfolgt die Anfrage an diesen als Unicast im gleichen Nachrichtenformat wie oben (Kasten „Einschätzungsanfrage“). Die Bearbeitung des der Zugangskontrolle zu unterwerfenden Pakets wird nun unterbrochen, indem dieses in eine Warteliste gelegt wird. Ein Zeitgeber wird gestartet, um das eventuelle Ausbleiben einer Antwort aufzufangen. Sollte dieser Zeitgeber ablaufen, ohne dass eine Antwort eingegangen ist, so wird entweder mit dem nächsten Knoten auf der Liste fortgefahren, oder das Paket wird verworfen, falls die Liste erschöpft ist.

Das Vorgehen bei Eingang einer Antwort ist in Abbildung 3.38 dargestellt. Es ähnelt demjenigen bei Antworten von Nachbarn, aber da die Anfrage hier immer nur an einen Knoten gerichtet und also mit Eingang einer Antwort erledigt ist, wird das wartende zu kontrollierende Paket immer sofort aus der Warteliste genommen und der zugehörige Zeitgeber wird gelöscht. Liefert die anschließende Zugangsprüfung ein eindeutiges Ergebnis, so wird das Paket verworfen oder weitergeleitet. Ist die Gesamtkooperativitätseinschätzung noch zu unsicher, so wird an der in Abbildung 3.34 mit einem Stern markierten Stelle fortgefahren und – ebenso wie bei Ablauf des Zeitgebers – entweder eine weitere Anfrage an einen anderen Knoten gestellt, oder das Paket schließlich verworfen, wenn dafür keine anderen entfernten Knoten mehr in Frage kommen.

3.9.3.5 Meldung über Scheitern der Zugangskontrolle

Wenn ein Knoten K während des Vorgangs des Weiterleitens eines Pakets feststellt, dass er den Schlüssel der Quelle des Pakets nicht beschaffen kann oder dass ihm keine positive Bewertung über die Kooperativität der Quelle vorliegt, so verwirft er das Paket und informiert die Quelle mit einer Meldung über das Scheitern der Weiterleitung. Eine solche Meldung wird innerhalb des gesamten Netzes weitergeleitet, unabhängig vom Vertrauen in ihren Ersteller, wobei deren Rate aber streng begrenzt wird.

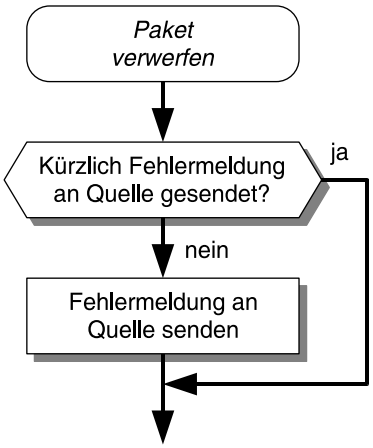


Abbildung 3.39: Ablauf beim Aussenden von Fehlermeldungen

Zugangsfehlermeldung
Grund G: Schlüssel fehlt oder Vertrauen fehlt
Kopie P des verworfenen Pakets
niederwertiger N_K^- Teil der Signaturfolgenummer N_K
Signatur über Nachrichtentyp, G, P, N_K

Anhand der Meldung über das Scheitern der Weiterleitung kann das verworfene Paket auch innerhalb einer Liste gesendeter Pakete identifiziert werden. Dies ist bei dem Knoten erforderlich, der das Paket zuletzt weitergeleitet hatte, damit der zugehörige Zeitgeber angehalten werden kann, so dass die nicht erfolgte Weiterleitung dem verwerfenden Knoten nicht negativ angerechnet wird.

In Abbildung 3.39 ist der Ablauf beim Aussenden von Fehlermeldungen dargestellt. An dem Ablaufdiagramm ist auch zu erkennen, dass die Aussendungsrate für Fehlermeldungen an bestimmte Quellen beschränkt wird.

Der Empfänger der Meldung wird versuchen, entweder durch Übermittlung gespeicherter Meinungszertifikate über sich selbst oder auch mit einer Bürgschaftsanfrage seine Bewertung beim verwerfenden Knoten zu verbessern.

3.9.4 Bootstrapping

Wenn einzelne Knoten, die im Netz noch nicht bekannt sind, diesem beitreten, können sie den Bürgschaftsmechanismus nutzen, um ein gewisses Vertrauen zu erhalten, und damit in die Lage versetzt werden, selbst Netzdienste zu nutzen. Dies funktioniert allerdings nicht, wenn keine geeigneten Bürgen vorhanden sind, weil beispielsweise das Netz gerade erst im Entstehen begriffen ist, also sozusagen alle Knoten neue Teilnehmer sind. In diesem Fall hätte unter Umständen noch kein Teilnehmer Vertrauen zu irgendeinem anderen. Würde jeder Teilnehmer dann alle Dienstleistungen für andere verweigern, könnten auch nie Leistungen beobachtet und somit positive Einschätzungen aufgebaut werden.

Um Abhilfe zu schaffen, wurde ein zusätzlicher Mechanismus vorgesehen, der Situationen wie die eben beschriebene erkennt und Abhilfe schafft, indem er die Zugangskontrolle so lange außer Kraft setzt, bis ein regulärer Betriebszustand erreicht ist. Dieser Mechanismus wird im Folgenden als *Bootstrapping-Mechanismus* bezeichnet, der Zustand der abgeschalteten Zugangskontrolle als *Bootstrapping-Modus*.

Das Kriterium für den Eintritt in den Bootstrapping-Modus sollte so beschaffen sein, dass einerseits das Vorliegen einer Bootstrapping-Situation schnell erkannt werden kann und es andererseits möglichst schwierig ist, entsprechende Bedingungen nur vorzutäuschen, um andere Knoten damit gezielt dazu zu bewegen, Dienstleistungen ohne Zugangskontrolle zu erbringen. Diese Anforderungen erfüllt recht gut das folgende Kriterium:

> Ein Knoten befindet sich dann im Bootstrapping-Modus, wenn er keinen einzigen von ihm als kooperativ eingeschätzten Nachbarn besitzt, bei dem er zusätzlich davon ausgehen kann, dass dieser ihn umgekehrt ebenfalls als kooperativ einschätzt.

Diese Bedingung wird in regelmäßigen Zeitabständen vor Anwendung der eigentlichen Zugangskontrolle geprüft, und zwar etwas häufiger, wenn sich der Knoten im Bootstrapping-Modus befindet (etwa im 1-Sekunden- statt 10-Sekunden-Abstand), weil eine Änderung dann normalerweise wesentlich wahrscheinlicher ist und der Bootstrapping-Modus möglichst schnell verlassen werden sollte.

Die zweite Teilbedingung bewirkt, dass jeder Knoten vor dem Verlassen des Bootstrapping-Modus auch selbst Leistungen erbringt; d. h. es reicht nicht, Leistungen anderer zu beobachten. Damit werden in Bezug auf den einzelnen Knoten die Chancen verbessert, selbst als kooperativ eingeschätzt zu werden und Zugang zu erhalten. Bezogen auf ein gesamtes, möglicherweise soeben neu entstandenes Netz ist es von Vorteil, wenn über jeden Knoten bereits einige Beobachtungen vorliegen.

Um die zweite Teilbedingung prüfen zu können, wird für jeden Nachbarn eine geschätzte Einschätzung des eigenen Verhaltens aus Sicht dieses Nachbarn geführt, welche entsteht, indem eigene Leistungen, die in Anwesenheit des Nachbarn erbracht werden, dort vermerkt werden. Bei der Höhe der Bewertung wird dabei berücksichtigt, dass nicht alle erbrachten Leistungen vom Nachbarn tatsächlich beobachtet und als bewertungsfähig befunden werden müssen.

3.9.4.1 Bedrohungsanalyse

Ein absichtliches böswilliges Erzwingen des Bootstrapping-Modus durch benachbarte Knoten ist nur möglich, wenn *keiner* der benachbarten Knoten beobachtbare Leistungen erbringt, so dass der Knoten im Bootstrapping-Modus kein Vertrauen zu ihnen fassen kann. Ein solcher Angriff ist damit nur in kleinen Netzen unter sehr engen Bedingungen an die Netztopologie durchführbar: Das Netz kann nur aus einem einzelnen Opfer und den zu ihm benachbarten Angreifern bestehen. Gäbe es mehr als ein Opfer, so könnten diese sich gegenseitig bewerten und den Bootstrapping-Modus dann verlassen. Bestünde Kontakt zu weiteren Knoten außerhalb der Nachbarschaft des Opfers, so könnte das Opfer die Weiterleitung von Paketen dorthin beobachten und bewerten. Ein Nutzen für die Angreifer entsteht nur, wenn sie nicht alle benachbart sind, sondern das „in der Mitte“ befindliche Opfer zur Weiterleitung benötigen.

Wenn und solange eine solche Situation vorliegt, gibt es häufig gar keine andere Möglichkeit, ein Netzwerk überhaupt aufrechtzuerhalten, als dass ein einzelner Knoten die gesamte Weiterleitung übernimmt, auch ohne dass ein beabsichtigter Angriff der am Rand befindlichen Knoten vorliegt. In manchen Fällen wäre eine fairere Verteilung der Last möglich, wird aber durch das Wegfindungsverfahren nicht realisiert. Es liegt natürlich im Ermessen des die Hauptlast tragenden Teilnehmers, die durch ihn weitervermittelte Datenmenge zu begrenzen oder die Dienstleistung ganz einzustellen. Da das Problem jedoch nicht auf die Zugangskontrolle zurückgeht, soll es hier nicht weiter erörtert werden.

3.10 Schlüsselverwaltung

Die primäre Aufgabe der Schlüsselverwaltung innerhalb des in dieser Arbeit entworfenen Konzepts ist es, zu gegebenen Schicht-2- oder Schicht-3-Adressen oder Schlüsselkennungen (Ausgabewerte einer kryptographischen Hashfunktion über öffentliche Schlüssel) die Schlüssel der zugehörigen Knoten zu liefern. Dazu wird ein lokaler Zwischenspeicher mit den am häufigsten benötigten Schlüsseln aufgebaut. Immer, wenn ein gesuchter Schlüssel dort nicht vorliegt, wird versucht, diesen durch Anfragen bei anderen Knoten zu beschaffen.

Da die Schlüsselverwaltung sowieso Informationen zur Abbildung von Schicht-2- und Schicht-3-Adressen auf Schlüssel sammeln und pflegen muss, übernimmt sie hier auch noch eine weitere Aufgabe: die Abbildung zwischen Schicht-3- und Schicht-2-Adressen, die für die Weiterleitung und deren Beobachtung erforderlich ist (siehe Abschnitt 3.6.2.1).

Die Schlüsselverwaltung ist zum Aufbau ihrer Datenbasis abhängig von Beobachtungen und von anderen Knoten gelieferten Informationen. Deshalb muss besonders darauf geachtet werden, dass Angreifer die von ihr bereitgestellten Dienste nicht durch zu diesem Zweck konstruierte und ausgesandte Nachrichten oder durch gefälschte Antworten auf Anfragen manipulieren können. Dadurch könnten sie sonst unter Umständen positive Bewertungen für Dienstleistungen anderer Knoten in ihr eigenes oder negative Bewertungen für eigenes Fehlverhalten in fremde Vertrauensprofile „umleiten“, negative Bewertungen ganz vermeiden oder sich Vorteile bei der Zugangskontrolle verschaffen.

3.10.1 Problematik der Zuordnung von Adressen und Schlüsseln

Bei der Abbildung zwischen Adressen und Schlüsseln macht sich ein Charakteristikum von Ad-hoc-Netzen erschwerend bemerkbar: Da es keine besonders autorisierte Instanz zur Vergabe von Adressen gibt, kann der „rechtmäßige Besitz“ von Adressen in Ad-hoc-Netzen nicht nachgewiesen werden. Zwar werden Adressen meist nach einem verteilten Verfahren konfliktfrei vergeben, aber wenn ein Knoten eine andere als die ihm durch dieses Verfahren zugewiesene Adresse verwendet, kann dieser Verstoß von einzelnen anderen Knoten in der Regel nicht einmal erkannt werden[6]. Folgen davon sind:

- Ein Knoten kann mehrere Schicht-2- oder -3-Adressen gleichzeitig verwenden.
- Wenn mehrere Knoten dieselbe Adresse verwenden, ist es sehr schwierig zu entscheiden, welcher der Knoten die beste Legitimation hat.
- Wegen der Ad-hoc-Netzen eigenen Dynamik der Netztopologie kann aus Sicht eines einzelnen Beobachters jederzeit ein neuer Knoten in der Nachbarschaft auftauchen. Dies kann am Auftauchen einer neuen Schicht-2-Adresse in beobachteten Paketen erkannt werden. Andererseits kann das Auftauchen einer neuen Schicht-2-Adresse aber ebenso bedeuten, dass ein vorhandener Nachbarn eine neue Adresse verwendet.
- Wenn ein Paket als Quellangabe eine andere Schicht-3-Adresse trägt als die, die dem Nachbarn zugeordnet ist, der das Paket übermittelt hat, so kann dies entweder bedeuten, dass das Paket von einem beliebig weit entfernten Knoten des Netzwerks stammt, der diese Adresse „regulär“ verwendet, oder, dass ein Nachbarknoten die Adresse verwendet, um den Eindruck zu erwecken, es liege der erste Fall vor.

[6] Theoretisch wäre es zwar denkbar, schon bei der Adressvergabe kryptographische Verfahren einzusetzen, die es ermöglichen, später nachzuweisen, dass eine verwendete Adresse korrekt zugewiesen wurde. Diesen Nachweis bei jeder Verwendung einer Adresse, also zumindest in jedem einzelnen übertragenen Paket zu erbringen, wäre aber in jedem Fall mit erheblichem Zusatzaufwand in Form zusätzlich zu übertragender Information verbunden.

- Ein Angreifer kann gezielt die Adresse eines anderen Knotens verwenden in der Absicht, sich für diesen auszugeben, also den Eindruck zu erwecken, er *sei* der andere Knoten bzw. der andere Knoten habe eine bestimmte Nachricht erzeugt.

Außerdem kann jeder Teilnehmer beliebig viele Schlüssel besitzen, denen jeweils ein eigenes Vertrauensprofil zugeordnet wird. Da ein Knoten die freie Wahl hat, mit welchem Schlüssel er sich identifiziert, kann er auch frei wählen, welches seiner Vertrauensprofile jeweils verwendet werden soll. Wenn die Identifikation mit der Zuordnung einer Adresse verknüpft ist, kann der Knoten zu jedem Schlüssel eine eigene Adresse wählen, die er immer dann verwendet, wenn er sich mit dem zugehörigen Schlüssel identifizieren will.

3.10.2 Bedarf an Schlüsseln bzw. Teilnehmerkennungen

Als Vorbereitung für die Analyse von Bedrohungen gegen die Schlüsselverwaltungsfunktionalität soll zunächst zusammengefasst werden, bei welchen Gelegenheiten sie in Anspruch genommen wird. Außerdem wird jeweils angegeben, welches Verfahren verwendet wird, um aus beobachteten oder empfangenen Paketen deren Sender und Empfänger abzulesen, denn dies beeinflusst wesentlich die Anforderungen an die Sicherung der Schlüsselverwaltung. Die verschiedenen Möglichkeiten wurden in Abschnitt 3.4 vorgestellt und in Tabelle 3.1 (Seite 82) zusammengefasst.

- *Positive Bewertung für erbrachte Weiterleitung:* Am beobachteten oder empfangenen Paket wird dessen letzter Schicht-2-Absender abgelesen und auf einen Schlüssel abgebildet. Das zugehörige Vertrauensprofil wird dann positiv bewertet, falls auch der vorige Weiterleitungsschritt beobachtet worden ist. Diese letzte Bedingung bedeutet nebenbei bemerkt auch, dass die Schicht-2-Absenderangabe des weitergesendeten Pakets mit der Schicht-2-Empfängerangabe des vom Weiterleiter empfangenen Pakets übereinstimmen muss.

- *Negative Bewertung für unterlassene Weiterleitung:* Die Schicht-3-Zieladresse der beobachteten weiterzuleitenden Nachricht wird zunächst auf eine Schicht-2-Adresse abgebildet und mit der Schicht-2-Empfängeradresse im Paket verglichen, um festzustellen, ob eine Weiterleitung erforderlich ist. Wenn dies der Fall ist, wenn also die beiden Schicht-2-Adressen verschieden sind, so wird die Schicht-2-Empfängeradresse aus dem Paket auf einen Schlüssel abgebildet, der dann negativ bewertet wird, wenn innerhalb des erlaubten Zeitraums keine Weiterleitung beobachtet wird.

  In der Praxis wird der für positive *oder* negative Bewertung benötigte Schlüssel sinnvollerweise schon in dem Moment bestimmt, wo die Übertragung eines noch weiterzuleitenden Pakets an den zu bewertenden Weiterleiter beobachtet wird. Er wird dann in der Information zum beobachteten Paket referenziert und in der Folge entweder bei Beobachtung der Weiterleitung für die positive oder nach Verstreichen des erlaubten Zeitraums für die negative Bewertung verwendet.

- *Bewertung bei Anfrage-Antwort-Diensten durch Nachbarn des Diensterbringers:* Zur Erkennung einer an einen Nachbarn gerichteten Anfrage ist zunächst der umgekehrte Test mit Adressabbildung erforderlich wie für die negative Bewertung der Weiterleitung: Wenn die per Abbildung aus der Schicht-3-Zieladresse des Pakets gewonnene Schicht-2-Adresse mit der Empfängerangabe im Paket übereinstimmt, ist es an einen Nachbarn adressiert. Wenn die Nachricht eine Anfrage eines zu beobachtenden Protokolls ist, wird sie – zusammen mit dem Schlüssel, der derselben Adresse zugeordnet ist – registriert. Abhängig davon, ob in der erlaubten Zeitspanne eine Antwort erfolgt oder nicht, wird dieser Schlüssel positiv oder negativ bewertet.

- *Positive Bewertung von Anfrage-Antwort-Diensten durch den Dienstnehmer:* Um erbrachte Dienstleistungen positiv bewerten zu können, benötigt der Dienstnehmer den Schlüssel des Dienstgebers zur Identifikation seines Vertrauensprofils. Im Unterschied zu den vorgenannten Fällen handelt es sich hier im Allgemeinen nicht um einen Nachbarn, was die Schlüsselverwaltung anspruchsvoller macht.

- *Zugangskontrolle:* Bei der Zugangskontrolle ist das relevante Vertrauensprofil das des Schlüssels der Quelle, mit dem die Signatur am Paket erzeugt wurde. Dieser Schlüssel wird anhand der Quelladresse des Pakets ermittelt, was hier jedoch einfacher ist als bei der Bewertung entfernter Dienstgeber (voriger Fall), weil nahezu immer ein Nachbar – nämlich der vorige Weiterleiter – erreicht werden kann, der den Schlüssel bereits kennt. Wie bereits in Abschnitt 3.9.2.1 erläutert wurde, können Signatur und Schlüssel zwar theoretisch von Weiterleitern ausgetauscht werden, eine Motivation dafür ist aber kaum herleitbar.

Neben der Abbildung zwischen Schicht-3- und Schicht-2-Adressen tritt die eigentliche Aufgabe der Schlüsselverwaltung – die Beschaffung von Schlüsseln anhand von Adressen – also in drei Varianten auf, die sich in der Eingabeinformation unterscheiden: Sie kann in der Schicht-2-Adresse eines Nachbarn bestehen oder in der Schicht-3-Adresse der Quelle eines soeben zur Weiterleitung empfangenen Pakets oder in der Schicht-3-Adresse eines entfernten Dienstleisters, der soeben eine Anfrage beantwortet hat.

3.10.3 Idee des Verfahrens

Für die Aufbewahrung von Schlüsseln und Adresszuordnungen wird für alle Aufgaben der Schlüsselverwaltung eine gemeinsame Datenstruktur verwendet, die im Folgenden als Knotendatenbasis bezeichnet wird. Die Verfahren der Beschaffung von fehlenden Schlüsseln dagegen unterscheiden sich bei den drei Varianten.

Die Information zur Abbildung zwischen Schicht-3- und Schicht-2-Adressen und zur Abbildung von Schicht-2-Adressen von Nachbarn auf Schlüssel wird sozusagen proaktiv erfasst: Bei jedem ersten Kontakt mit einem noch nicht bekannten Nachbarn wird ein Schlüsselaustausch mit diesem durchgeführt. Dabei werden im Wesentlichen der eigene Schlüssel und die eigenen Adressen bekanntgegeben, und Schlüssel und Adressen des anderen werden entgegengenommen und in der eigenen Datenbasis gespeichert. Spätere Anfragen werden immer direkt aus dieser Datenbasis beantwortet.

Zur Abbildung von Schicht-3-Quelladressen weiterzuleitender Pakete auf Schlüssel wird, falls noch kein Eintrag zur fraglichen Adresse vorhanden ist, jeweils bei Bedarf eine Anfrage an den Vorgänger bei der Weiterleitung gestellt, der den Schlüssel ja kennen muss, da er selbst soeben Zugangskontrolle dafür durchgeführt hat. Auf dieselbe Weise können grundsätzlich auch Schlüssel entfernter Dienstleister bei Anfrage-Antwort-Diensten ermittelt werden, da der Dienstleister Quelle der Antwortnachricht ist. Bezüglich der Bedrohunganalyse unterscheidet sich dieser Anwendungsfall allerdings vom erstgenannten; auf den Unterschied wird später noch eingegangen.

Während die Anzahl der Nachbarn eines Knotens und damit auch der Speicherbedarf für die ständig präsent gehaltene Information über deren Adressen und Schlüssel beschränkt ist, muss die Anzahl an Einträgen mit Schicht-3-Adressen nicht benachbarter Knoten explizit beschränkt werden. Wenn Platz für neue Einträge benötigt wird, werden dann die am längsten nicht benötigten Einträge gelöscht.

(Unterbundene) Zuordnung	Ziel
eigene Adr. $\not\rightarrow$ eigener Schl.	• selbst trotz Fehlverhaltens negative Bewertungen vermeiden
eigene Adr. $\rightarrow$ fremder Schl.	• anderen Knoten durch gezieltes eigenes Fehlverhalten negative Bewertungen verschaffen
fremde Adr. $\rightarrow$ eigener Schl.	• selbst positive Bewertungen für Leistungen anderer Knoten kassieren
fremde Adr. $\not\rightarrow$ fremder Schl.	• dafür sorgen, dass andere Knoten keine positiven Bewertungen erhalten

Tabelle 3.2: Varianten der Abweichung von der korrekten Zuordnung von Schicht-2-Adressen zu Schlüsseln

3.10.4 Bedrohungen bezüglich der Abbildung von Schicht-2-Adressen auf Schlüssel

Im Folgenden wird analysiert, welche Auswirkungen sich ergeben können, wenn die Schlüsselverwaltungsfunktion falsche Schlüssel liefert, welche Vorteile ein Angreifer davon hätte und welche Möglichkeiten es gibt, solche Fehlfunktionen gezielt herbeizuführen. Zunächst wird die Teilaufgabe der Abbildung von Schicht-2-Adressen auf Schlüssel betrachtet, die bei der Beobachtung und Bewertung des Weiterleitungsverhaltens genutzt wird.

Eine gezielte Beeinflussung der korrekten Funktion der Schlüsselverwaltung beinhaltet immer die Verfälschung von Informationen in der Knotendatenbasis anderer Knoten durch den Angreifer. Bezüglich der Motivation zu solchen Verfälschungen kann man zwei Richtungen feststellen:

- Bezüglich negativer Bewertungen, die mit Hilfe der gespeicherten Information vergeben werden, sind Angreifer hauptsächlich daran interessiert zu bewirken, dass ihre Schicht-2-Adresse nicht mit ihrem (bzw. einem ihrer) eigenen Schlüssel verknüpft wird, so dass sie selbst keine negativen Bewertungen erhalten. Wird die eigene Adresse stattdessen mit einem fremden Schlüssel verknüpft, so können dem zugehörigen Knoten durch eigenes Fehlverhalten negative Bewertungen verschafft werden.

- Bezüglich positiver Bewertungen, die von der Abbildung abhängen, sind Angreifer eher daran interessiert, ihren (bzw. einen ihrer) Schlüssel mit fremden Schicht-2-Adressen zu verknüpfen, um positive Bewertungen auf sich umzuleiten. Entfällt dabei die korrekte Bindung der fremden Adresse an deren Schlüssel, weil beispielsweise nur *ein* Schlüssel an jede Adresse gebunden werden kann oder darf, so erhält der fremde Knoten selbst keine positiven Bewertungen mehr.

Tabelle 3.2 fasst die Möglichkeiten der Verfälschung zusammen, die sich übrigens nicht gegenseitig ausschließen, sondern grundsätzlich alle gleichzeitig versucht werden können. Anhand der schon in Abschnitt 3.10.2 aufgezählten Einsatzgebiete der Schlüsselverwaltung ist zu erkennen, dass alle vier Fälle bei der Bewertung des Weiterleitungsverhaltens sowie bei der Bewertung bezüglich Anfrage-Antwort-Diensten durch Nachbarn des Diensterbringers relevant sind.

Dadurch, dass dieselbe Datenbasis für positive und negative Bewertungen verwendet wird, ergibt sich aber immer auch ein gewisser Nachteil bzw. ein Risiko für den Angreifer:

- Wenn der Schicht-2-Adresse des Angreifers nicht dessen eigener Schlüssel zugeordnet ist, sondern der Schlüssel eines anderen oder gar keiner, so erhält der Angreifer selbst für korrektes

Verhalten (unter dieser Adresse) keine positive Bewertung mehr; gegebenenfalls kommen positive Bewertungen dem anderen Teilnehmer zu. Sinnvoll ist der Angriff also nur, wenn der Angreifer auch keine Leistungen erbringt, für die er normalerweise positiv bewertet würde.

- Wenn einer fremden Schicht-2-Adresse der Schlüssel des Angreifers zugeordnet ist, so erhält der Angreifer negative Bewertungen für Fehlverhalten des zugehörigen Teilnehmers. Dieser andere wird selbst nicht mehr negativ bewertet, falls die Zuordnung seines Schlüssel zur Adresse aufgehoben wurde. Das Verhalten des „Opfers“, also des Knotens, dessen Adresse eine falsche Zuordnung erhält, muss möglichst dauerhaft korrekt sein, so dass keine negativen Bewertungen anfallen. Dazu darf das Opfer den Angriff nicht bemerken, da es sonst durch gezieltes Ändern seines Verhaltens einen „Gegenangriff“ starten könnte.

In den Unterabschnitten 3.10.4.2 bis 3.10.4.5 wird detailliert auf die einzelnen Möglichkeiten der Verfälschung der Zuordnung eingegangen. Vorher werden in 3.10.4.1 aber noch die Auswirkungen der Verwendung verschiedener Adressen beim Senden bzw. zum Empfang von Paketen behandelt, die für einige Angriffe nötig ist.

3.10.4.1 Unterschiedliche Schicht-2-Adressen bei Senden und Empfang

Um positive und negative Bewertungen auf unterschiedliche Vertrauensprofile zu verteilen, müsste ein Angreifer weiterzuleitende Pakete unter einer anderen Schicht-2-Adresse weitersenden als er sie empfangen hat. Damit würde das ausgesendete Paket aber von Beobachtern in ihren Listen in Bearbeitung befindlicher Pakete nicht mehr mit dem empfangenen in Deckung gebracht. Das ursprüngliche Paket erschiene als verloren, was negativ bewertet würde, und das neu ausgesendete Paket erschiene als bisher noch nicht beobachtet, so dass keine positive Bewertung vergeben würde.

Auch wenn ein Angreifer also beliebige Zuordnungen für die Abbildung von Schicht-3- auf Schicht-2-Adressen und von Schicht-2-Adressen auf Schlüssel erzeugen kann, gilt immer: Leitet er nicht trotzdem immer unter der Adresse weiter, unter der er das weiterzuleitende Paket empfangen hat, so entfällt die positive Bewertung und stattdessen fällt eine negative an. Durch das Vorgehen bei der Beobachtung und Bewertung ist damit schon ausreichend sichergestellt, dass positive und negative Bewertungen immer demselben Vertrauensprofil zugerechnet werden.

3.10.4.2 Unterbinden der Zuordnung des eigenen Schlüssels zur eigenen Adresse

Die korrekte Zuordnung des eigenen Schlüssels zur eigenen Schicht-2-Adresse kann trivial verhindert werden, indem kein Schlüsselaustausch mit neuen Nachbarn durchgeführt wird, der eigene Schlüssel also nicht bekanntgegeben wird. Leitet der Angreifer später Pakete nicht weiter oder antwortet er nicht auf Anfragen für andere Dienstleistungen, so kann er nicht negativ bewertet werden.

Als Gegenmaßnahme sollen Knoten, die keinen Schlüsselaustausch mit ihren Nachbarn durchführen, aus dem Netz ausgeschlossen werden. Dazu muss sichergestellt sein, dass Pakete von Knoten, zu denen keine Schlüssel vorliegen, niemals weitergeleitet werden, also auch dann nicht, wenn ansonsten keine Zugangskontrolle angewandt wird, etwa bei Bürgschaftsanfragen. Auch sollen keine Pakete an oder über sie weitergeleitet werden – sie werden damit praktisch vollkommen isoliert, solange sie keine Schlüssel bekanntgeben. Voraussetzung dafür, dass vom Angreifer erzeugte Pakete erkannt und verworfen werden können, ist, dass selbst erzeugte Pakete von der Quelle immer signiert werden. Ist die beschriebene Gegenmaßnahme in Kraft, so ist der Angriff unattraktiv und wirkungslos.

Pakete, die nicht vom Angreifer erzeugt, sondern nur über ihn weitergeleitet wurden, brauchen nicht verworfen zu werden. Dass ein solcher Fall vorliegt, ist aber für den kontrollierenden Knoten nur dann erkennbar, wenn er bereits den vorigen Weiterleitungsschritt desselben Pakets beobachtet hat. Eine positive Bewertung für die Weiterleitung kann ein Knoten ohne Schlüssel selbstverständlich nicht erwarten, schon weil ihm kein Vertrauensprofil zugeordnet werden kann.

3.10.4.3 Zuordnung eines fremden Schlüssels zur eigenen Adresse

Es kommen zwei unterschiedliche Möglichkeiten in Betracht, um eine falsche Zuordnung zwischen einem fremden Schlüssel und einer vom Angreifer verwendeten Schicht-2-Adresse herzustellen:

- *Aktives Zuordnen eines fremden Schlüssels:* Wenn beim Schlüsselaustausch einfach nur der eigene öffentliche Schlüssel mitgeteilt wird, kann eine falsche Assoziation leicht erzeugt werden, indem der Schlüssel eines fremden Teilnehmers mitgeteilt wird.

  Als Gegenmaßnahme, die den Angriff in dieser Form wirksam verhindert, wird beim Schlüsselaustausch ein Challenge-Response-Verfahren verwendet (siehe Abschnitt 2.3.3.1). Dadurch wird sichergestellt, dass nur öffentliche Schlüssel angegeben werden können, zu denen die angebende Partei auch den privaten Schlüssel besitzt. Öffentliche Schlüssel anderer Teilnehmer können also nicht mehr der eigenen Adresse zugeordnet werden.

- *Verwenden einer bereits zugeordneten Adresse:* Wenn der Angreifer die technische Möglichkeit besitzt, seine Schicht-2-Adresse zu ändern, kann er alternativ auch einfach eine fremde Adresse verwenden, der bereits ein Schlüssel zugeordnet ist – bei unsignierten Paketen ist für Dritte in der Regel nicht einmal zu erkennen, dass ein zweiter Sender existiert, der dieselbe Adresse verwendet, geschweige denn, welches Paket von welchem Sender stammt oder welcher der Sender ein Angreifer ist.

  Bei diesem Vorgehen muss sich der Inhaber des Schlüssels in der Nachbarschaft des Angreifers befinden oder sich zumindest kürzlich dort befunden haben. Dies ist notwendig, weil die Schlüssel-Adress-Bindung wegen des im vorigen Punkt verlangten Challenge-Response-Verfahrens nur von ihm selbst erzeugt werden kann und automatisch aufgelöst wird, wenn er nicht mehr als Nachbar wahrgenommen wird. Der Angreifer muss also in Aktion treten, während der ursprüngliche Adressinhaber noch anwesend ist oder nur kurz nachdem er die Nachbarschaft verlassen hat.

Ist ein fremder Schlüssel einer von einem Angreifer verwendeten Schicht-2-Adresse zugeordnet, so bedeutet dies, dass der fremde Teilnehmer negativ bewertet wird, falls Pakete an die Adresse nicht weitergeleitet bzw. Anfragen für andere Dienstleistungen nicht beantwortet werden. Für einen auf dieser Art der Zuordnungsverfälschung beruhenden Angriff sind zwei etwas unterschiedliche Zielsetzungen denkbar:

- Einerseits kann das Ziel die negative Bewertung eines bestimmten anderen Teilnehmers sein. Da nicht vorgesehen ist, dass ein Knoten für selbst (bzw. unter seiner Adresse vom Angreifer) abgesandte Pakete negative Bewertungen erhält, kann der Angreifer nicht gezielt Pakete zum Schaden des Schlüsselinhabers erzeugen. Um negative Bewertungen zu verursachen muss er stattdessen beispielsweise selbst unter einer anderen Adresse Pakete zur Weiterleitung oder andere Anfragen an die fragliche Adresse senden (unter entsprechendem Energieaufwand).

Ist der Inhaber des fremden Schlüssel selbst noch in der Nachbarschaft des Angreifers, so führt er die an seine Adresse gerichteten Dienstleistungsaufträge aus, und hat damit nicht mehr Aufwand als der Angreifer. Negative Bewertungen erhält er nur, wenn er nicht mehr anwesend ist; andererseits muss er aber kürzlich anwesend gewesen sein und die Beobachter dürfen seine Abwesenheit noch nicht bemerkt haben, da sie sonst nicht mehr werten. Insgesamt ist der Angriff recht schwierig durchzuführen, rein destruktiver Natur und dabei sehr wenig effektiv, deshalb werden keine weiteren Gegenmaßnahmen ergriffen.

- Andererseits kann die negative Bewertung auch nur Nebeneffekt sein, wenn das eigentliche Ziel darin besteht, den im vorigen Abschnitt beschriebenen Angriff (Unterbinden der Zuordnung des eigenen Schlüssels zur eigenen Adresse) auszuführen und zu verschleiern, indem ein falscher Schlüssel zugeordnet wird. Der Angreifer möchte hier also Pakete absenden, die weitergeleitet werden, ohne aber selbst Pakete weiterleiten oder negative Bewertungen in Kauf nehmen zu müssen (wobei er selbstverständlich auch auf die Möglichkeit verzichtet, positive Bewertungen zu erhalten).

  Wenn selbst erzeugte Pakete unter einer Schicht-2-Adresse ausgesandt werden, der ein fremder Schlüssel zugeordnet ist, können sie nicht mit diesem Schlüssel signiert werden; der Inhaber des fremden Schlüssels gilt also nicht als Quelle. Stattdessen muss der Angreifer einen eigenen Schlüssel verwenden (und der Zugangserfolg des Pakets hängt damit vom zugehörigen eigenen Vertrauensprofil ab). Für Beobachter und den tatsächlich ersten Weiterleiter entsteht der Eindruck, das Paket werde durch den Angreifer, welcher durch die gefälschte Zuordnung als Inhaber des fremden Schlüssels gilt, lediglich weitergeleitet (allerdings ohne Wertung, da der vorige Schritt nicht beobachtet wurde).

  Der tatsächliche erste Weiterleiter akzeptiert das Paket, weil der verwendeten Schicht-2-Adresse ein Schlüssel zugeordnet ist[7]. Der echte Inhaber des vom Angreifer verwendeten Schlüssels ist der einzige Knoten, der den Angriff feststellen kann. Aus den oben genannten Gründen ist die Wahrscheinlichkeit auch recht hoch, dass er das Paket des Angreifers mithört. Er sollte den Angreifer (identifiziert durch den der verwendeten Quelladresse der Schicht 3 zugeordneten Schlüssel) negativ bewerten.

3.10.4.4 Zuordnung des eigenen Schlüssels zu einer fremden Adresse

Eine inkorrekte Zuordnung des eigenen Schlüssels zu einer fremden, also bereits von einem anderen Knoten verwendeten Adresse kann durch einen Schlüsselaustausch unter der fremden Adresse hergestellt werden. Diese Möglichkeit kann grundsätzlich auch nicht verbaut werden, da die Rechtmäßigkeit der Nutzung einer Adresse nicht nachweisbar ist (siehe Abschnitt 3.10.1).

Als Folge der falschen Zuordnung könnten fälschliche positive Bewertungen für Weiterleitung vergeben werden: Der ursprüngliche Inhaber der Adresse leitet weiter, aber der Angreifer wird – stattdessen oder zusätzlich – positiv bewertet.

Als erste Gegenmaßnahme wird vorgeschrieben, dass immer nur genau einmal positiv gewertet werden darf. Wenn ein zweiter Schlüssel zu einer Adresse auftaucht, wird sofort ein erneuter Schlüsselaustausch versucht, und sollten daraufhin mehrere widersprüchliche Antworten erfolgen, so wird die positive Bewertung für die fragliche Adresse für eine gewisse Zeit ausgesetzt, bevor der Schlüsselaustauschversuch wiederholt wird. (Negative Bewertungen brauchen nicht ausgesetzt zu werden,

[7] Dass der Angreifer den Schlüssel nicht selbst besitzt, kann am Paket nicht festgestellt werden, da Pakete auf Schicht 2 nicht signiert werden. Auf diese Maßnahme wurde wegen des erheblichen Mehraufwands an Rechenleistung bei der Weiterleitung verzichtet.

hier können einfach alle in Frage kommenden Schlüssel bewertet werden. Jeder der Adressanwärter kann an die Adresse gerichtete Pakete weiterleiten und die negative Bewertung damit verhindern.)

Der Angriff ist durch diese Maßnahme sehr viel weniger erfolgversprechend, denn immer wenn der ursprüngliche Adressinhaber mit seinen Nachbarn bereits Schlüssel ausgetauscht hat, führt ein zusätzlicher Schlüsselaustausch mit der selben Adresse durch den Angreifer dazu, dass entweder die ursprüngliche Zuordnung sofort durch den noch anwesenden Adressinhaber wiederhergestellt wird oder – wenn der Angreifer in dem erneuten Schlüsselaustausch auf seiner falschen Zuordnung beharrt – gar nicht mehr gewertet wird. Im zweiten Fall hat der Angreifer damit ebenfalls keinen Vorteil mehr.

Der Angreifer müsste also schon in Aktion treten, wenn bei einem seiner Nachbarn noch kein Schlüssel des Opfers vorliegt, etwa kurz bevor dieser Nachbar in Reichweite des Opfers kommt. Da der Nachbar damit dann eine Schlüsselzuordnung zur Adresse des Opfers hätte, fiele ihm auch bei Empfang des ersten Pakets des Opfers nicht auf, dass es sich tatsächlich um einen ihm noch nicht bekannten Knoten handelt. Allerdings fiele dem Opfer trotzdem auf, dass es einen neuen Nachbarn hat, und so würde der Schlüsselaustausch, der zur Doppelzuordnung führt, eben durch das Opfer initiiert. Um das zu verhindern, müsste der Angreifer auch dem Opfer vorher eine falsche Schlüsselzuordnung für die Adresse des neuen Nachbarn unterschieben. Damit bemerken beide getäuschten Knoten zunächst nichts von der Täuschung und bewerten evtl. fälschlicherweise den Angreifer anstelle des jeweils anderen. Sobald allerdings in der Folge einer der beiden derartig fehlinformierten Knoten eine Nachricht als Quelle absendet, erkennt der andere anhand der bei ihm fehlschlagenden Signaturprüfung eine Unstimmigkeit, die er durch einen (vermeintlich) erneuten Schlüsselaustausch zu beheben versucht.

Insgesamt ist es zwar bedauerlich, dass Angreifer eine Möglichkeit erhalten, die positive Bewertung anderer Teilnehmer durch bestimmte dritte Teilnehmer für eine gewisse Zeit zu unterbinden. Andererseits haben die Angreifer dadurch kaum einen Vorteil, während ihnen durchaus Aufwand entsteht.

3.10.4.5 Unterbinden der Zuordnung fremder Schlüssel zu den Adressen der Inhaber

Die Zuordnung fremder Schlüssel zu den Adressen ihrer Inhaber kann nur durch anhaltendes Stören des Schlüsselaustauschs zwischen anderen Knoten auf physikalischer Ebene, also durch Störsignale, verhindert werden. Damit nicht durch kurze Störung lang anhaltende Wirkung erzielt werden kann, werden fehlgeschlagene Schlüsselaustauschversuche in regelmäßigen Abständen wiederholt.

Der Zweck eines Angriffs könnte nur darin bestehen, zu verhindern, dass ein bestimmter anderer Knoten positive Bewertungen für von ihm erbrachte Leistungen erhält. Insofern entspricht die Wirkung dieses für den Angreifer aufwändigen, rein destruktiven Angriffs derjenigen des im vorigen Abschnitt beschriebenen.

3.10.5 Bedrohungen bezüglich der Abbildung von Schicht-3-Adressen auf Schlüssel

Eine Zuordnung zwischen einem Schlüssel und einer Schicht-3-Adresse kann auf zwei verschiedenen Wegen entstehen: Entweder durch einen Schlüsselaustausch mit einem Nachbarn oder durch eine Schlüsselmitteilung vom vorigen Weiterleiter (bzw. von der Quelle).

Benötigt wird die Abbildung von Schicht-3-Adressen auf Schlüssel bei der Zugangskontrolle zur Identifikation des Vertrauensprofils, aufgrund dessen die Zugangsentscheidung zu treffen ist. Bei der Bewertung entfernter Dienstgeber durch Dienstnehmer ist die Abbildung erforderlich um festzustellen, welches Vertrauensprofil positiv zu bewerten ist.

3.10.5.1 Bedrohungen bei der Zugangskontrolle

Die Zugangskontrolle erfolgt immer aufgrund des Vertrauensprofils, das zu dem Schlüssel gehört, mit dem auch die Quellsignatur in der zu kontrollierenden Nachricht erstellt wurde. Jede wie auch immer erzeugte falsche Zuordnung wird also daran erkannt, dass die Signaturprüfung vor der Zugangskontrolle fehlschlägt. Das betreffende Paket wird in diesem Fall nicht weitergeleitet. Alle Angriffe, die darauf abzielen, durch die Verwendung fremder Vertrauensprofile etwa Zugangschancen für eigene Pakete zu verbessern, sind damit ausgeschlossen.

Die Zugangschancen für fremde Pakete zu verschlechtern ist in gewisser Weise möglich, indem nach deren Weiterleitung eine eventuell vom nächsten Weiterleiter gestellte Anfrage nach dem Quellschlüssel mit einem falschen Schlüssel beantwortet wird. Das entsprechende Paket wird daraufhin vom nächsten Weiterleiter abgelehnt. Da die Antwort auf die Schlüsselanfrage signiert sein muss, kann der nächste Weiterleiter den Angreifer allerdings negativ bewerten. Außerdem entsteht dem Angreifer ein erheblicher Mehraufwand: Er muss zunächst das Paket weiterleiten, dann eine Schlüsselanfrage beantworten und schließlich mit hoher Wahrscheinlichkeit noch eine Fehlermeldung in Richtung Quelle weiterleiten. Von einem Vorteil für den Angreifer kann also kaum die Rede sein.

3.10.5.2 Bedrohungen bei der Bewertung entfernter Dienstgeber

Eine positive Bewertung eines Dienstgebers durch einen Dienstnehmer erfolgt, wenn dieser eine Antwort auf eine Anfrage erhalten hat. Zu bewerten ist dann die Quelle der Antwortnachricht, die mit dem Ziel der Anfragenachricht übereinstimmen sollte. Wenn der Schlüssel der Quelle beim Dienstnehmer nicht bekannt ist, kann genau wie im Fall eines nicht bekannten Quellschlüssels bei der Zugangskontrolle auch hier eine Anfrage an den vorigen Weiterleiter gerichtet werden, der den Schlüssel für die Zugangskontrolle benötigt hat.

Anhand der Quellsignatur im Paket kann überprüft werden, ob der Schlüssel tatsächlich dem Erzeuger des Pakets gehört. Allerdings ist damit nicht automatisch sichergestellt, dass dieser Erzeuger tatsächlich der Dienstgeber ist: Jeder Weiterleiter kann grundsätzlich ein neues Paket desselben (oder anderen) Inhalts erzeugen, mit einem eigenen Schlüssel signieren und statt des Originalpakets weitersenden. Wenn er vom nächsten Weiterleiter oder dem Dienstnehmer eine Anfrage nach dem Quellschlüssel erhält, liefert er dann ebenfalls seinen eigenen Schlüssel, so dass die Signaturprüfung erfolgreich verläuft.

Der die Originalantwort modifizierende Angreifer erhält hierbei wegen der Änderung des Paket-Hashwerts richtigerweise von Beobachtern negative Bewertungen für Nichtweiterleitung. Eine positive Bewertung des angeblichen Dienstgeberschlüssels sollte so moderat ausfallen, dass diese im Angriffsfall anfallenden negativen Bewertungen im Mittel nicht ausgeglichen werden, so dass sich der Angriff für den Angreifer nicht lohnt.

Besser noch wäre eine Ende-zu-Ende-Sicherung des Dienstleistungsvorgangs, etwa indem der Dienstnehmer der Anfrage eine mit dem ihm schon vorab bekannten Schlüssel des Dienstgebers verschlüsselte Zufallszahl beifügt, welche der Dienstgeber entschlüsselt und mit dem signierten Antwortpaket zurücksendet. Damit wäre sichergestellt, dass die Antwort nur vom Dienstgeber erstellt werden kann. Der zusätzliche Aufwand, der unter Umständen durch die Beschaffung des Dienstgeberschlüssels anfällt, ist allerdings nur gerechtfertigt, wenn die Sicherung auch für die Anwendung sinnvoll ist; allein der Zweck der Bewertung rechtfertigt ihn nicht.

Nachbarschaftseintrag
Öffentlicher Schlüssel
Signaturfolgenummer
Schicht-2-Adresse (bei Nachbarn)
Schicht-3-Adresse
Zeitpunkt des ersten Kontakts (bei Nachbarn)
Empfangszeitpunkt des zuletzt beobachteten Pakets
Empfangszeitpunkt des zuletzt beobachteten signierten Pakets
Verweis auf den Knoten, von dem die Information stammt
Vertrauensprofil

Abbildung 3.40: Inhalt eines Eintrags in der Knotendatenbasis

3.10.6 Knotendatenbasis

Die Knotendatenbasis enthält die in Abbildung 3.40 dargestellten Informationen zu jedem Nachbarn sowie zu weiteren Knoten, soweit Speicherplatz verfügbar ist:

- Der *öffentliche Schlüssel* ist die eigentliche Identität des durch den Eintrag beschriebenen Knotens und Primärschlüssel in der Knotendatenbasis. Es kann nur jeweils einen Eintrag zu jedem Schlüssel geben.

- Die *Signaturfolgenummer* dient zur Überprüfung der Aktualität von Änderungsnachrichten und generell als Schutz vor Wiedereinspielung (siehe Abschnitt 3.4.4.2)

- Die *Schicht-2-Adresse* ist immer dann vermerkt, wenn es sich um einen Nachbarn handelt. Sie wird für die Abbildung von Schicht-3- auf Schicht-2- und von Schicht-2-Adressen auf Schlüssel benötigt und beim Schlüsselaustausch zwischen Nachbarn eingetragen. Für die Suche nach Nachbarn anhand ihrer Schicht-2-Adresse wird ein gesonderter Suchindex gepflegt.

- Die *Schicht-3-Adresse* wird für die Abbildung von Schicht-3- auf Schicht-2- und von Schicht-3-Adressen auf Schlüssel benötigt und entweder beim Schlüsselaustausch zwischen Nachbarn oder auf eine Quellschlüsselanfrage bei der Weiterleitung hin eingetragen.

- Der *Zeitpunkt des ersten Kontakts* mit einem Nachbarn wird gelegentlich benötigt, um festzustellen, ob dieser zu einem bestimmten Zeitpunkt in der Vergangenheit bereits anwesend war (siehe Abschnitt 3.10.8.1).

- Die *Empfangszeitpunkte der zuletzt beobachteten Pakete* ohne bzw. mit Signatur werden beim Anlegen des Eintrags sowie bei jeder erfolgreichen Identifikation eines Vertrauensprofils anhand des Absenders eines beobachteten Pakets gesetzt und dient zur Bestimmung der Aktualität des Eintrags. Nachbarn müssen regelmäßig signierte Nachrichten senden, ansonsten werden sie aus dem Nachbarindex gelöscht und nur noch unter den sonstigen Knoten geführt. Wenn ein neuer Eintrag angelegt werden soll, für den kein Speicherplatz mehr zur Verfügung steht, wird der am längsten nicht benötigte Eintrag gelöscht.

- Anhand eines *Ursprungsverweises* auf den Knoten, von dem die im Eintrag enthaltenen Information stammt, kann insbesondere auch erkannt werden, ob die Schicht-3-Adresse eines nicht (mehr) benachbarten Knoten von diesem selbst (in einem Schlüsselaustausch) oder von einem Dritten angegeben wurde.

3.10.7 Schlüsselaustausch zwischen Nachbarn

3.10.7.1 Regulärer Ablauf

Sobald ein Knoten A ein Paket eines Knotens B beobachtet, dessen Schicht-2-Adresse er bisher nicht in seiner Knotendatenbasis registriert hat, beginnt er mit der Durchführung des folgenden Schlüsselaustauschprotokolls:

1. A sendet folgendes Paket an die Schicht-2-Adresse von B:

Hello-1
Schicht-2-Adressen A_A^2 und A_B^2 von A und B
eigener öffentlicher Schlüssel K_A^+ eigene volle Signaturfolgenummer N_A eigene Schicht-3-Adresse A_A^3

2. B schickt als Antwort folgendes Paket an A:

Hello-2
Schicht-2-Adressen A_B^2 und A_A^2 von B und A
eigener öffentlicher Schlüssel K_B^+ eigene volle Signaturfolgenummer N_B eigene Schicht-3-Adresse A_B^3
Signatur von B über Nachrichtentyp, A_A^2, A_B^2, K_A^+, N_A, K_B^+, N_B

3. A registriert B nach Erhalt der Antwort als Nachbarn mit der Schicht-2-Adresse aus der Absenderangabe (die mit der ursprünglich beobachteten übereinstimmen muss) sowie dem Schlüssel aus dem Hello-2-Paket von B, falls die Signatur korrekt ist.

4. Anschließend schickt A folgendes Paket an B:

Hello-3
Schicht-2-Adressen A_A^2 und A_B^2 von A und B
Signatur von A über Nachrichtentyp, A_A^2, A_B^2, K_A^+, N_A, K_B^+, N_B

5. B registriert A als Nachbarn mit der Schicht-2-Adresse aus der Absenderangabe (die mit der ursprünglich beobachteten übereinstimmen muss) sowie dem Schlüssel aus dem Hello-1-Paket von A, falls die Signatur korrekt ist.

Abbildung 3.41 fasst den regulären Ablauf des Protokoll als Weg-Zeit-Diagramm zusammen.

Die Einbeziehung des Schlüssels und der Signaturfolgenummer des Kommunikationspartners in die eigene Signatur entspricht der Verwendung eines Challenge-Response-Verfahrens. Dieses wird benötigt um sicherzustellen, dass der Nachbar den bekanntgegebenen Schlüssel tatsächlich besitzt (und nicht einen fremden öffentlichen Schlüssel angibt oder eine beobachtete Nachricht wiederholt).

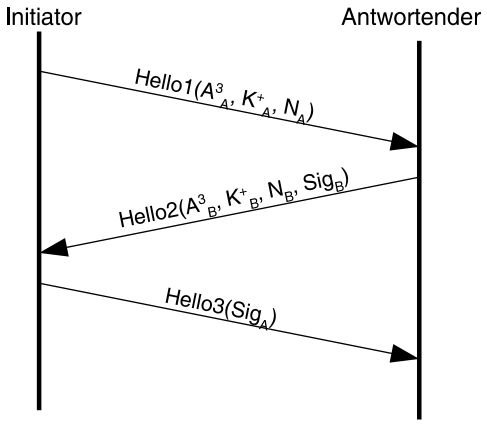


Abbildung 3.41: Regulärer Ablauf des Schlüsselaustauschprotokolls (Bezeichnung siehe Text)

3.10.7.2 Behandlung von Unregelmäßigkeiten im Ablauf

Falls als Antwort auf eine Hello-1-Nachricht statt einer Hello-2-Nachricht ebenfalls eine Hello-1-Nachricht eingeht, so wird davon ausgegangen, dass entweder die eigene Hello-1-Nachricht verloren gegangen ist oder dass beide Parteien gleichzeitig versucht haben, eine Hello-1-Nachricht abzusetzen. In diesem Fall wird als Antwort eine Hello-2-Nachricht verschickt.

In dem Fall, in dem beide Parteien gleichzeitig einen Austausch gestartet haben, kann es dann vorkommen, dass wiederum beide Parteien gleichzeitig mit einer Hello-2-Nachricht „korrigieren“, so dass beide als Antwort wieder eine Hello-2-Nachricht erhalten. Damit haben dann beide Parteien schon alle nötigen Informationen und die erforderliche Signatur, so dass die dritte Protokollnachricht entfallen kann.

Allgemein kann eine Partei den Austausch als beendet betrachten, wenn sie eine Nachricht mit Signatur empfangen und eine versendet hat. Sollte eine Partei keine Nachricht mit Signatur erhalten, weil diese bei der Übertragung verloren gegangen ist, so wiederholt sie nach kurzer Zeit ihre vorige Nachricht, die dann vom Kommunikationspartner nochmals gleichlautend beantwortet wird. In allen anderen Fällen werden unaufgefordert eingehende Hello-2- oder -3-Nachrichten ignoriert.

3.10.8 Abbildung zwischen Schlüsseln und Schicht-3-Adressen

Zur Prüfung der Quellsignatur bei der Zugangskontrolle und positiven Bewertung entfernter Diensterbringer müssen anhand der Schicht-3-Adresse anderer Knoten ihre Schlüssel identifiziert werden.

Wenn ein weiterzuleitendes Paket eintrifft, wird in der Knotendatenbasis nach einem Schlüssel zur Schicht-3-Adresse der Quelle gesucht, der außerdem eine korrekte Verifikation der Paketsignatur erlaubt. Ist diese Suche erfolgreich, so wird der Zeitpunkt der letzten Nutzung des Eintrags angepasst, und die Zugangskontrolle kann anhand des gefundenen Schlüssels durchgeführt werden.

Ist kein Schlüssel vorhanden oder schlägt die Signaturprüfung mit allen gefundenen Schlüsseln fehl, so richtet der kontrollierende Knoten K mit dem folgenden Paket eine Anfrage an seinen Vorgänger V bei der Weiterleitung:

Schlüsselanfrage anhand Schicht-3-Adresse
Schicht-2-Adressen A^2_K und A^2_V von K und V
Quelladresse A^3_Q des weiterzuleitenden Pakets Hashwert H über das weiterzuleitende Paket Signatur über Nachrichtentyp, A^2_K, A^2_V, A^3_Q und H

Der Hashwert H über das zu kontrollierende Paket wird ebenso berechnet wie bei der Beobachtung der Weiterleitung. Da der Vorgänger die Weiterleitung durch K normalerweise beobachten möchte, hat er in seiner Liste in Bearbeitung befindlicher Pakete bereits einen Eintrag mit diesem Hashwert. Diese Liste braucht lediglich um einen Verweis auf den Eintrag des Quellknotens in der Knotendatenbasis ergänzt werden, denn da V in aller Regel vor der Weiterleitung des Pakets an K Zugangskontrolle durchgeführt hat, musste er den Schlüssel der Quelle sowieso identifizieren bzw. beschaffen. Der Hashwert in der Anfrage dient also zur einfachen eindeutigen Identifikation des gesuchten Schlüssels. Die Quelladresse allein würde dafür nicht unbedingt ausreichen, da mehrere Einträge zur selben Schicht-3-Adresse in der Knotendatenbasis erlaubt sind.

Anhand der Signatur kann der Vorgänger feststellen, dass es sich beim Anfragenden um einen registrierten Nachbarn handelt, und er kann diesen eindeutig identifizieren. Er kann damit auch eine Zugangskontrolle für den Dienst der Schlüssellieferung durchführen.

Der Anfragevorgang wird beim Anfragenden identifiziert durch die angegebene Schicht-3-Adresse des Knotens, dessen Schlüssel gesucht wird, und den Vorgänger, an den die Anfrage gesandt wurde. Auf eine zusätzliche Kennzeichnung, etwa durch eine gleichlautende Kennung in Anfrage und Antwort wird verzichtet, weil sie keine Vorteile bieten würde. Solange eine durch Schicht-3-Adresse und Nachbarn identifizierte Anfrage aktiv ist (d. h. solange nicht entweder eine Antwort erhalten wurde oder die erlaubte Antwortzeitspanne abgelaufen ist), werden keine weiteren Anfragen bezüglich derselben Adresse an denselben Nachbarn gesendet, sondern alle Pakete mit der derselben Quelle werden zusammen mit dem ersten (das die laufende Anfrage ausgelöst hat) aufbewahrt und dann abgearbeitet, sobald eine Antwort zur betreffenden Schicht-3-Adresse eintrifft.

Sollte der Vorgänger den gesuchten Schlüssel wider Erwarten nicht kennen, antwortet er mit einer Fehlermeldung. Ansonsten liefert er den Schlüssel mit der folgenden Nachricht an K:

Schlüsselantwort anhand Schicht-3-Adresse
Schicht-2-Adressen A_V^2 und A_K^2 von K und V
Schicht-3-Adresse A_Q^3 Zugehöriger öffentlicher Schlüssel K_Q^+ Volle Signaturfolgenummer N_Q
niederwertiger Teil N_V^- der Signaturfolgenummer N_V Signatur über A_Q^3, K_Q^+, N_Q, N_V

Der Initiator K prüft die Signatur des weiterzuleitenden Pakets mit dem erhaltenen Schlüssel. Falls dies erfolgreich ist, legt er einen Eintrag für den Schlüssel an und der Vorgänger V wird positiv bewertet. Anschließend kann die Zugangskontrolle für das weiterzuleitende Paket durchgeführt werden.

Die Anfrage wird auch gestellt und die Antwort wird ausgewertet, wenn kein Schlüssel des Nachbarn vorliegt, der das auslösende Paket weitergeleitet hat. Wenn der in der Antwort enthaltene Schlüssel dazu geeignet ist, die Signatur am auslösenden Paket zu verifizieren, so wird er in jedem Fall mit der Quelladresse des Pakets assoziiert und das ihm zugeordnete Vertrauensprofil wird für die Zugangskontrolle herangezogen. Wenn auch bei Erhalt der Antwort noch kein Schlüssel des Vorgängers vorliegen sollte, kann er keine positive Bewertung für die erbrachte Dienstleistung erhalten.

3.10.8.1 Negative Bewertung bei unterlassener Schlüssellieferung

Wenn innerhalb einer vorgegebenen Zeitspanne keine Antwort auf eine Schlüsselanfrage eintrifft, soll der säumige Diensterbringer negativ bewertet werden. Da die Anfrage aber einfach an die Schicht-2-Adresse des vorigen Weiterleiters gesendet wurde, ohne dass dessen Schlüssel unbedingt bekannt war (weil dies so ist, wird zunächst auch gar nicht versucht, dem Weiterleiter einen Schlüssel zuzuordnen,

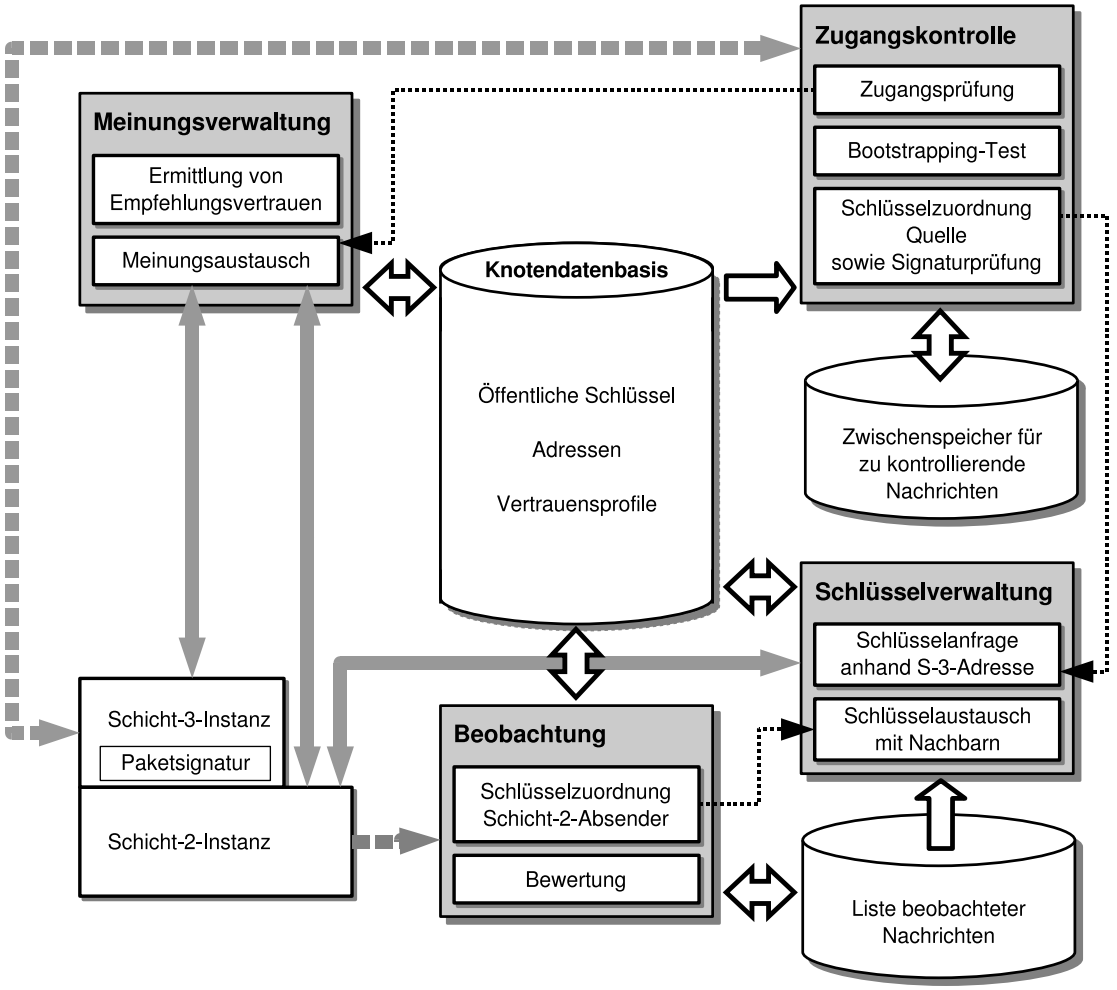


Abbildung 3.42: Architektur und Schnittstellen des entworfenen Zugangskontrollsystems. Breite umrandete Pfeile bedeuten Datenfluss, dünne gestrichelte bezeichnen Kontrollfluss. Die Nutzung von Kommunikationsdiensten ist durch breite graue Pfeile angedeutet, tiefergehende Integration mit den Schichten des Kommunikationssubsystems durch gestrichelte solche.

und in der Anfrage ist entsprechend auch kein Bezug auf den Schlüssel enthalten), muss spätestens bei Ablauf des Zeitgebers ein bewertbarer Schlüssel identifiziert werden.

Dies wird folgendermaßen gelöst: Es werden alle Knoten negativ bewertet, deren Schlüssel im Ablaufmoment der Adresse zugeordnet sind, an welche die Anfrage gerichtet wurde. Ausgenommen werden davon aber Knoten, die noch keine Nachbarn waren, als die Anfrage gestellt wurde. Um diesen Sachverhalt feststellen zu können, wird in der Knotendatenbasis mit jedem Nachbarn der Zeitpunkt des ersten Kontakts mit ihm gespeichert (siehe Abschnitt 3.10.6).

3.11 Gesamtarchitektur

Abschließend soll die Architektur des in der vorliegenden Arbeit entworfenen Zugangskontrollsystems nochmals in ihrer Gesamtheit dargestellt werden. Abbildung 3.42 gibt dazu einen (vereinfachten) Überblick.

Die links unten abgebildeten Schicht-2- und Schicht-3-Instanzen sind diejenigen des um die Zugangskontrolle erweiterten Systems; sie sind hier zwecks Darstellung der Schnittstellen zum Zugangskontrollsystem enthalten. Die von oben an den Schicht-2- und Schicht-3-Instanzen eintreffenden breiten

grauen Pfeile stehen für die reguläre Nutzung der durch die Schichten angebotenen Kommunikationsdienste. So werden etwa für Meinungsanfragen bei entfernten Knoten Schicht-3-Nachrichten dorthin abgesandt, und entsprechend werden solche Meinungsanfragen anderer Knoten entgegengenommen und gegebenenfalls beantwortet. Der Kommunikationsdienst der Schicht-2 wird für Meinungsanfragen bei Nachbarn, den Schlüsselaustausch mit Nachbarn und für Schlüsselanfragen bei der Weiterleitung genutzt.

Neben diesen regulär genutzten Kommunikationsdienstschnittstellen werden zusätzliche Schnittstellen zu den Schicht-2- und Schicht-3-Instanzen benötigt, die in der Abbildung durch seitliche, grau gestrichelt dargestellte Pfeile repräsentiert werden: Erstens müssen zum Zweck der Beobachtung des Verhalten der Nachbarknoten alle empfangenen oder beobachteten Schicht-2-Nachrichten an das Beobachtungsmodul des Zugangskontrollsystems übergeben werden. Und zweitens muss das Zugangskontrollverfahren natürlich in die Abläufe der Schicht 3 eingreifen, so dass Nachrichten erst nach erfolgter und erfolgreicher Zugangsprüfung weitergeleitet oder gegebenenfalls stattdessen Zugangsfehlermeldungen versandt werden. Die für das Zugangskontrollsystem benötigten Paketsignaturen werden im erweiterten System ebenfalls in Schicht 3 erstellt.

Das gesamte Zugangskontrollsystem lässt sich grob in vier Komponenten unterteilen: Beobachtung, Meinungsverwaltung, Zugangskontrolle und Schlüsselverwaltung. Alle diese Komponenten haben Zugriff auf die zentrale Knotendatenbasis, in der Adressen, öffentliche Schlüssel und Vertrauensprofile anderer Knoten gespeichert sind.

Bei der Beobachtung müssen zunächst die Schlüssel der Schicht-2-Absender und -Empfänger beobachteter Nachrichten identifiziert werden, damit positive oder negative Bewertungen den entsprechenden Vertrauensprofilen zugeordnet werden können. Dies geschieht mit Hilfe der Adress-Schlüssel-Zuordnung, die durch den Nachbarschlüsselaustausch in der Knotendatenbasis erzeugt wird. Tritt eine noch unbekannte Schicht-2-Adresse zum ersten mal auf, so wird ein Nachbarschlüsselaustausch angestoßen. Das Verfahren zur Beobachtung und Bewertung der Weiterleitung benötigt und pflegt eine Liste aller derzeit zu beobachtenden Nachrichten. Bewertungen werden in den in der Knotendatenbasis gespeicherten Vertrauensprofilen registriert.

Bei der Zugangskontrolle erfolgt ebenfalls eine Adress-Schlüssel-Zuordnung anhand der Informationen in der Knotendatenbasis, allerdings diesmal betreffend die Schicht-3-Adresse der Quelle zu kontrollierender Nachrichten; liegt die gesuchte Zuordnung nicht vor, so wird eine Schlüsselanfrage anhand der Schicht-3-Adresse angestoßen. Außerdem wird vor Durchführung der eigentlichen Zugangskontrolle noch anhand der Vertrauensprofile der eigenen Nachbarn überprüft, ob die Bootstrapping-Bedingung erfüllt ist, die ein Ausschalten der Zugangskontrolle bewirkt. Bei der Zugangsprüfung selbst wird der aus allen gespeicherten Vertrauensprofilen bestehende Graph auf eine Gesamteinschätzung des Quellknotens hin ausgewertet. Wenn keine ausreichend sichere Einschätzung vorliegt wird eine Meinungsanfrage angestoßen (vorausgesetzt, dies wurde nicht schon kürzlich getan).

Die Durchführung bzw. Beantwortung von Meinunganfragen und die Eintragung relevanter Ergebnisse in die Knotendatenbasis ist die Aufgabe der Meinungsverwaltungskomponente. Außerdem vergleicht sie regelmäßig die gespeicherten Fremdmeinungen mit eigenen Einschätzungen bezüglich derselben Knoten und ermittelt so die Größe des Empfehlungsvertrauens, das anderen Knoten entgegengebracht wird.

Der Schlüsselverwaltung obliegen die bereits angesprochenen Aufgaben der Beschaffung nicht vorliegender Schlüsselzuordnungen zu Schicht-2- bzw. Schicht-3-Adressen. Bei der Beantwortung von Anfragen nach den Schicht-3-Quelladressen selbst weitergeleiteter Nachrichten wird auch die Liste der beobachteten Nachrichten konsultiert, da die Nachricht, welche die Anfrage ausgelöst hat, in der Regel dort noch registriert ist; mit deren Eintrag wird auch ein Verweis auf den Quellschlüssel gespeichert.

3.12 Zusammenfassung

Das in diesem Kapitel vorgestellte Konzept für die Realisierung von Zugangskontrolle in offenen Ad-hoc-Netzen wurde im Hinblick auf ein sehr allgemeines Einsatzszenario entworfen, in welchem die Teilnehmer – Personen mit einem zu drahtloser Kommunikation fähigen Rechner – grundsätzlich vollständig voneinander unabhängig handeln können, also nicht durch übergeordnete Bindungen zur Zusammenarbeit in einer bestimmten Weise verpflichtet sind.

Die eingangs aufgestellten Anforderungen werden dabei voll erfüllt. So werden etwa alle Dienste verteilt erbracht, es ist keine manuelle Konfiguration erforderlich, alle Teilnehmer sind gleichgestellt und es gibt keine hierarchischen Strukturen innerhalb der Teilnehmermenge. Neue Teilnehmer können jederzeit dem Netz beitreten, und trotzdem erhalten aufgrund unkooperativen Verhaltens ausgeschlossene Teilnehmer nicht jederzeit sofort wieder Zugang, indem sie ihre Identität wechseln.

Die Zugangskontrolle wird durch jeden einzelnen Teilnehmer auf der Ebene der Netzwerkschicht durchgeführt, indem Pakete nur weitergeleitet werden, wenn davon ausgegangen werden kann, dass ihr Quellknoten ebenfalls bereit ist, Leistungen für das Netzwerk zu erbringen. Für die sichere Zuordnung von Paketen zu ihren Quellknoten ist es erforderlich, dass jedes Paket von seiner Quelle digital signiert wird.

Grundlage für die Zugangsentscheidung und ein zentraler Bestandteil des Konzepts ist eine in jedem Knoten geführte Datenbasis von Vertrauensprofilen, welche Einschätzungen zur Kooperativität anderer Teilnehmer enthalten. Diese werden durch die automatisierte Beobachtung des Verhaltens benachbarter Teilnehmer bei der Erbringung von Dienstleistungen für das Netzwerk gewonnen. Insbesondere wurde das Verfahren zur Beobachtung und Bewertung des Verhaltens bei der Weiterleitung fremder Pakete genau spezifiziert, möglich ist aber auch die Beobachtung und Bewertung des Verhaltens bezüglich anderer Dienste.

Neben durch eigene Beobachtung gewonnenen Einschätzungen werden in Vertrauensprofilen auch von anderen Teilnehmern geäußerte Meinungen bezüglich Dritter gespeichert. Solche Meinungsäußerungen werden mittels geeigneter Protokolle angefordert, wenn eigene Einschätzungen nicht sicher genug sind, um eine Zugangsentscheidung treffen zu können; sie können nicht nur vom anfordernden, sondern von allen Teilnehmern gespeichert werden, welche die Antwort mithören. Aufgrund der Übereinstimmung zwischen selbst ermittelten Einschätzungen und Fremdmeinungen wird anhand eines dafür entwickelten rechnerischen Verfahrens auch bestimmt, in welchem Maße den Meinungsäußerungen anderer vertraut wird.

Alle bekannten Aussagen von Teilnehmern über andere Teilnehmer (eigene Einschätzungen, fremde Meinungsäußerungen sowie Einschätzungen bezüglich der Vertrauenswürdigkeit fremder Äußerungen) lassen sich in einem Graphen über der Teilnehmermenge zusammenfassen, dem Vertrauensgraphen. Bisher gab es allerdings kein automatisiertes Verfahren zur Ermittlung einer Gesamteinschätzung bezüglich eines bestimmten Teilnehmers, das alle möglichen Wege vom beurteilenden zum beurteilten Knoten im Vertrauensgraph berücksichtigt. Das zu diesem Zweck in der vorliegenden Arbeit entwickelte neuartige Verfahren basiert auf einer Abbildung des Vertrauensgraphen auf ein Netzwerk ohmscher Widerstände und der Ermittlung eines resultierenden Stromes zwischen beurteilendem und beurteiltem Knoten.

Die Identifikation von Teilnehmern innerhalb des Netzwerks und damit auch von Vertrauensprofilen erfolgt bei dem vorgeschlagenen Konzept grundsätzlich direkt anhand öffentlicher Schlüssel. Dadurch kann auf eine aufwändige Zertifizierungsinfrastruktur zur gesicherten Zuordnung zwischen Schlüsseln und Teilnehmeridentitäten verzichtet werden. Trotzdem fallen noch einige Aufgaben im Zusammenhang mit der Verteilung und Verwaltung von Schlüsseln an. So ist das für die Zugangskontrolle

zu verwendende Vertrauensprofil immer dasjenige, dessen Schlüssel zur Verifikation der Quellsignatur am kontrollierten Paket geeignet ist; um diesen passenden Schlüssel aber zu finden, wird eine Abbildungsfunktion zwischen Schicht-3-Adressen und Schlüsseln benötigt. Weiterhin ist für die Bewertung des Verhaltens von Nachbarn auch eine Abbildung von Schicht-2-Adressen auf Schlüssel erforderlich. Diese Aufgaben werden von einer Schlüsselverwaltungskomponente übernommen, welche die benötigten Zuordnungen herstellt und für die weitere Verwendung zwischenspeichert. Dabei kommen besondere hierfür entwickelte Protokolle zum Einsatz, unter anderem eines zum gesicherten Schlüsselaustausch mit Nachbarn.

Kapitel 4

Evaluation

Zur Überprüfung der Funktion und Analyse der Eigenschaften des entworfenen Konzepts wurde es in Form eines Simulationsmodells implementiert. Im Folgenden wird zunächst kurz die dabei verwendete Simulationsumgebung vorgestellt und es werden einige wesentliche Aspekte der Umsetzung des entworfenen Konzepts in ein Simulationsmodell erläutert. Anschließend werden Randbedingungen und Parameter der simulierten Szenarien beschrieben, bevor durch die simulative Analyse einiger grundlegender Eigenschaften der modellierten Netze geeignete Parameterwerte und Vergleichsergebnisse für die nachfolgende Untersuchung des Zugangskontrollsystems gewonnen werden. Ab Abschnitt 4.5 werden sukzessive die Verfahren zur Beobachtung und Bewertung der Weiterleitung, zum Meinungsaustausch und zur Ermittlung einer resultierenden Gesamteinschätzung, zum Bootstrapping und schließlich zur Durchführung der Zugangskontrolle jeweils daraufhin untersucht, ob sie wie gewünscht funktionieren, wie schnell sie die gewünschte Leistung erbringen und welchen Einfluss die Parametrisierung und insbesondere die Wahl des Mobilitätsmodells hat. Den Abschluss des Kapitels bildet eine Zusammenfassung und Bewertung der Ergebnisse.

4.1 Simulationsumgebung

Grundlage für die Implementierung des entworfenen Konzepts in Form eines Simulationsmodells war die für Forschungszwecke frei verfügbare Simulationsumgebung OMNeT++ (Objective Modular Network Testbed in C++), die federführend von András Varga an der Technischen Universität Budapest entwickelt wurde [Varg01]. Die Umgebung erlaubt die objektorientierte Spezifikation eines modular aufgebauten Simulationsmodells und steuert den Simulationsablauf mit Hilfe einer zentralen Warteschlange, anhand derer zeitlich diskrete Ereignisse sukzessive bearbeitet werden.

Die Struktur eines Simulationsmodells wird unter OMNeT++ mit Hilfe einer eigenen Beschreibungssprache festgelegt. Ein Simulationsmodell besteht aus hierarchisch verschachtelten Modulen, was die Repräsentation der logischen Struktur des realen Systems im Modell erlaubt. Module kommunizieren untereinander über Nachrichten, und mit Hilfe solchen Nachrichtenaustauschs erfolgt auch die gesamte Synchronisation zwischen parallel ablaufenden Vorgängen. Nachrichten können entweder direkt an beliebige andere Module gesendet werden oder an in der Modulbeschreibung spezifizierte Ausgangstore. Ausgangstore sind in der Regel mit Eingangstoren anderer Module auf derselben Hierarchieebene verbunden, oder aber mit Ausgangstoren auf der nächsthöheren Hierarchieebene. Solchen Verbindungen zwischen Toren können bestimmte Übertragungseigenschaften (Verzögerung, Datenrate, Fehlerrate) zugewiesen werden. Durch die Verwendung von Toren anstelle direkter Nachrichten

wird die Möglichkeit geschaffen, den Aufbau des Modells aus den Modulen durch Spezifikation der Verbindungen zwischen Toren anzugeben.

Die exakte Funktion einzelner Module der untersten Hierarchiestufe wird spezifiziert, indem ihre Datenstrukturen und Algorithmen in der objektorientierten Programmiersprache C++ implementiert werden. Konkret ist dazu eine neue C++-Klasse vom vorgegebenen `cSimpleModule` abzuleiten und dort eine von zwei virtuellen Funktionen (`handlemessage()` oder `activity()`) neu zu definieren. Diesen beiden Funktionen entsprechen zwei verschiedene Paradigmen bei der Definition des Modulverhaltens:

- Bei der *nachrichtenbasierten* Variante wird in der zugehörigen Funktion (`handlemessage()`) die Reaktion auf jeweils eine empfangene Nachricht festgelegt, die der Funktion zur Bearbeitung übergeben wird. Die Bearbeitung erfolgt hier sozusagen instantan und ohne Unterbrechungen zu genau dem Simulationszeitpunkt, welcher für die Ankunft der Nachricht festgelegt wurde. Pausen im Ablauf können erzeugt werden, indem das Modul sich selbst künstlich um die Wartezeiten verzögerte Nachrichten sendet und die Bearbeitung dann beendet, um sie bei Eintreffen der Selbstnachrichten wieder aufzunehmen.

- Bei der *koroutinenbasierten* Variante wird die zugehörige Funktion (`activity()`) einmalig zu Beginn des Simulationslaufs aufgerufen. Sie implementiert das Verhalten des Moduls während seiner ganzen Lebenszeit. Um zu pausieren oder beispielsweise auf eintreffende Nachrichten zu warten, stehen spezielle Funktionsaufrufe der Simulationsumgebung zur Verfügung, durch welche der Kontrollfluss an die Ablaufsteuerung übergeben wird.

Die unabhängig von der realen Zeit geführte Simulationszeit „vergeht“ ausschließlich während der Übertragung von Nachrichten. Die in Übertragung befindlichen Nachrichten werden dazu in der Reihenfolge ihrer Ankunftszeitpunkte in der erwähnten zentralen Warteschlange verwaltet und der Reihe nach abgearbeitet, indem die Simulationszeit auf den nächsten Ankunftszeitpunkt weitergeschaltet und dann entweder die Empfangsfunktion des Zielmoduls aufgerufen oder die auf die Nachricht wartende Koroutine aktiviert wird. Aus Sicht der Module arbeiten diese parallel, d. h. mehrere Aktionen verschiedener Module können zum selben Zeitpunkt der Simulationszeit stattfinden.

OMNeT++ unterstützt zwei verschiedene Benutzerschnittstellen: Neben einer Kommandozeilenvariante, die gut zur automatischen Abarbeitung von Listen unterschiedlich parametrisierter Simulationsläufe geeignet ist, gibt es auch eine graphische Oberfläche, in welcher die Struktur des Modells sowie die Übertragung von Nachrichten graphisch dargestellt wird. Man kann die Simulation hier auch in Einzelschritten ablaufen lassen und Variablen der Module einsehen und ändern, was während der Entwicklung des Modells und zu dessen Überprüfung sehr vorteilhaft ist.

Neben der Strukturierung der Bestandteile des Modells und der Bereitstellung einer Infrastruktur zum Nachrichtenaustausch stellt OMNeT++ außerdem eine Bibliothek häufig benötigter Hilfsmittel wie etwa Zufallszahlengeneratoren und Klassen zur Sammlung statistischer Daten bereit.

4.2 Implementierung des entworfenen Konzepts innerhalb der Simulationsumgebung

Bei der Modellierung des in dieser Arbeit entworfenen Konzepts unter OMNeT++ bildet das gesamte Ad-hoc-Netz naheliegenderweise die oberste Hierarchiebene. Die zweitoberste Ebene wird durch

Module gebildet, die jeweils einen vollständigen Teilnehmerknoten des Ad-hoc-Netzes repräsentieren; alle Teilnehmerknoten befinden sich also auf derselben Hierarchieebene. Diese Modellierung liegt nahe, da in dieser Arbeit „flache“ Ad-hoc-Netze betrachtet werden, in denen alle Knoten gleichberechtigt sind und in denen eine Strukturierung, die einzelnen Knoten oder Gruppen von Knoten besondere Aufgaben oder Rechte zuordnen würde, vermieden werden soll.

Denkbar wäre nun eine weitere Strukturierung eines Teilnehmerknotens in mehrere Module gewesen. Dieser Ansatz wurde in hinführenden Arbeiten [Banh01, Hof02] auch erprobt. Dabei zeigte sich allerdings, dass die Kommunikation mittels Nachrichten zwischen den einzelnen Komponenten eines Teilnehmerknotens relativ aufwändig ist, da für jeden komponentenübergreifenden Methodenaufruf jeweils ein Nachrichtenobjekt alloziert und mit den Aufrufparametern bestückt werden muss. Ursprünglich war in OMNeT++ zur Definition des Nachrichteninhalts außerdem vorgesehen, dass wiederum einzelne durch Zeichenketten benannte Parameterobjekte zur Nachricht hinzuzufügen waren[1]. Dies wurde einerseits bei der Implementierung als umständlich und unübersichtlich empfunden, andererseits erschien auch der Speicherplatz- und der rechnerische Aufwand zur Verwaltung der Parameter- und Nachrichtenobjekte überhöht. Deshalb wurde die innere Strukturierung der Teilnehmerknoten ausschließlich mit Hilfe der Sprachmittel von C++ realisiert, so dass die Komponenten durch einfache Funktionsaufrufe kommunizieren können. Einige dieser Komponenten sind die folgenden (in der ungefähren Reihenfolge, in der sie sich von unten nach oben in die gebräuchliche Schichtenarchitektur von Kommunikationssystemen einordnen lassen):

- *Physikalischer Knoten und Übertragungsmedium:* Jeder Knoten verwaltet seine eigene Position und passt diese in regelmäßigen Intervallen nach den Vorgaben des für ihn gewählten Mobilitätsmodells an. Eine übergeordnete allwissende Instanz außerhalb des eigentlichen Modells (konkret: eine statische Funktion der Knotenklasse, die Zugriff auf alle Knotenpositionen hat) stellt nach jedem Bewegungsschritt fest, welche Knoten innerhalb ihrer gegenseitigen (hier symmetrisch angenommenen; siehe dazu auch die Bemerkungen in Abschnitt 4.3) Sendereichweite liegen, sich also gegenseitig „hören“ können. Alle Nachrichten, die ein Knoten innerhalb des Modells an einen seiner Nachbarn aussendet (andere Knoten sind für ihn ja nicht erreichbar), werden auch für alle Nachbarn kopiert – jeweils mit gleicher Ankunftszeit. Damit wird das geteilte Medium und das Mithören fremder Übertragungen in der OMNeT++ eigenen ereignisorientierten Weise modelliert. Auf eine explizite Bitübertragungsschicht wird verzichtet; die tatsächliche Übertragung von Nachrichten zu benachbarten Knoten übernimmt OMNeT++ in der im vorigen Abschnitt beschriebenen Weise.

- *Sicherungsschicht und Medienzugriff:* Auf die Einbringung von Übertragungsfehlern, die von OMNeT++ durch das zufallsgesteuerte Setzen eines entsprechenden Kennzeichens an übertragenen Nachrichten unterstützt wird, wurde zunächst verzichtet, da kein besonders starker Einfluss auf das zu untersuchende System zu erwarten ist. Die Integration von Mechanismen zur Fehlererkennung und -behebung hätte die Implementierung der zu untersuchenden Protokolle merklich verkompliziert.

  Die Sicherungsschicht übernimmt neben der hier unnötigen Sicherung gegen Übertragungsfehler üblicherweise auch noch eine weitere wichtige Aufgabe, nämlich die Regelung des Zugriffs verschiedener Teilnehmer auf ein geteiltes Medium. Ein real verwendetes Medienzugriffsverfahren, das in drahtlosen lokalen Netzen nach dem Standard IEEE 802.11 verwendet wird, wurde in Abschnitt 2.2.4 beschrieben. Für die Simulation sollte einerseits ein möglichst „einfaches“

[1] Mittlerweile wurde OMNeT++ um eine eigene Beschreibungssprache zur Spezifikation von Nachrichtenformaten erweitert. Aus solchen Spezifikationen werden durch einen Übersetzer C++-Klassen generiert, so dass ein direkter Zugriff auf die Nachrichteninhalte möglich ist.

Verfahren verwendet werden, damit die Ergebnisse der Evaluation der in dieser Arbeit entworfenen neuen Protokolle nicht durch Eigenheiten des Medienzugriffsverfahrens überlagert werden. Andererseits sollten die typischen Eigenschaften des für Ad-hoc-Netze charakteristischen geteilten Mediums erhalten bleiben, über welches nicht etwa mehrere Übertragungen am selben Ort gleichzeitig und unabhängig voneinander erfolgen können. Gewählt wurde ein Verfahren, bei dem das Medium so gut wie möglich ausgenutzt und zwischen den darum konkurrierenden Knoten relativ fair aufgeteilt wird: Nach Freiwerden des Mediums darf immer derjenige Knoten zuerst senden, bei dem der Sendewunsch zuerst auftrat. Dieser neue Sender darf dann genau ein Paket aussenden und rückt danach ans Ende der Warteschlange (falls er noch weitere sendebereite Pakete hat). Konflikte, bei denen in realen Netzen Paketverluste auftreten, weil mehrere Knoten gleichzeitig mit dem Senden beginnen, werden bei diesem synthetischen Verfahren vollständig vermieden.

Es sei angemerkt, dass Konflikte bei Mithörenden in der Realität für die eigentliche Übertragung zwischen Sender und Empfänger nicht immer schädlich sein müssen. Reale Verfahren versuchen deshalb gar nicht, solche Konflikte in jedem Fall zu verhindern (unter Anderem deshalb, weil so die verfügbare Bandbreite besser ausgenutzt wird – die Mithörbarkeit zählt im Allgemeinen nicht explizit zu den Entwurfszielen bei Medienzugriffsverfahren). Das hat zur Folge, dass die Beobachtung bei solchen Verfahren etwas schlechter funktioniert als bei dem hier verwendeten – die genauen Auswirkungen wären noch zu untersuchen.

- *Netzwerkschicht:* Die Hauptaufgabe der Netzwerk- oder Vermittlungsschicht ist die Verknüpfung von Verbindungsstrecken der darunterliegenden Schicht (zwischen benachbarten Knoten) zu längeren Wegen zwischen nicht benachbarten Knoten sowie die Weiterleitung von Paketen entlang dieser Wege; erst dadurch entsteht ein Gesamtnetz, in dem jeder Teilnehmer mit jedem anderen kommunizieren kann. In realen Ad-hoc-Netzen einsetzbare Verfahren zur Wegfindung müssen verteilt arbeiten und gehen dabei je nach Ansatz verschiedene Kompromisse ein bezüglich Menge und Detailliertheitsgrad der zu verteilenden Information und der Geschwindigkeit der Verteilung. Für die hier durchgeführte Untersuchung sollte wiederum ein möglichst „allgemeines“ Verfahren gefunden werden, das nicht durch charakteristische Stärken oder Schwächen das Ergebnis beeinflusst. Deshalb werden die zu verwendenden Wege hier durch eine allwissende Zentralinstanz in idealer Weise ermittelt. Als Kriterium für die Güte von Wegen gilt in der Regel die Anzahl der zu passierenden Knoten.

- *Anwendung:* Komponenten der Anwendungsschicht bilden in der Simulation die Nutzung durch reale, von den Teilnehmern eingesetzte Dienste und Anwendungen nach. Ob solche Anwendungen durch das in dieser Arbeit entwickelte Zugangskontrollsystem gestört werden oder in welchem Maß andererseits Störungen durch unkooperative Knoten verhindert werden, ist ein wesentliches Kriterium für die Bewertung des Ansatzes. Das von Anwendunginstanzen erzeugte Verkehrsmuster wird weiter unten in Abschnitt 4.3.2 noch beschrieben.

  Auf eine eigene, zwischen Netzwerk- und Anwendungsschicht gelegene Transportschicht wurde verzichtet, da die Untersuchung möglicher Wechselwirkungen zwischen Mechanismen der Transportschicht und des Zugangskontrollsystem nicht im Fokus dieser Arbeit lag.

Wie aus den obigen Beschreibungen bereits hervorging, wurden die Eigenschaften einiger Komponenten eines Netzknotens in einer idealisierten Form modelliert, die in der Realität kaum oder gar nicht erreichbar ist. Einerseits führt dies natürlich dazu, dass die Ergebnisse nicht direkt für eine Implementierung in einem realen Ad-hoc-Netz gelten, andererseits wird so aber vermieden, dass sie von so vielen Parametern gleichzeitig abhängen – nämlich von der Wahl und Parametrisierung konkreter, sich teilweise erheblich unterscheidender Verfahren auf mehreren Ebenen –, dass sie gar keine

allgemeine Aussagekraft mehr hätten. Die Ergebnisse, die unter Voraussetzung idealer Bedingungen gewonnen werden, beschreiben den bestmöglichen Fall. Bei der Übertragung auf reale Netze müssen die einzelnen zusätzlichen Nachteile, die sich durch nicht ideale Verfahren auf den unterschiedlichen Ebenen ergeben, einberechnet werden. Die Untersuchung konkreter Einsatzszenarien mit typischen Übertragungstechnologien, Medienzugriffs-, Nachbarschaftserkennungs- und Wegfindungsverfahren, bestimmten Transportprotokollen und Anwendungen muss Gegenstand zukünftiger Forschungsarbeit sein.

4.3 Eigenschaften von Einsatzszenarien

Ein Ziel bei der simulativen Untersuchung des entworfenen Konzepts ist es festzustellen, für welche Szenarien des Einsatzes von Ad-hoc-Netzen es in welchem Maße geeignet ist. Dazu müssen verschiedene Einsatzszenarien definiert und gegebenenfalls graduell variiert werden. Im Folgenden werden einige quantitative und qualitative Größen benannt und beschrieben, welche eine formale Charakterisierung von Einsatzszenarien erlauben und in den nachfolgenden Abschnitten herangezogen werden, um Zusammenhänge zwischen Eigenschaften von Einsatzszenarien und bestimmten gemessenen Leistungskenngrößen des untersuchten Systems darzustellen.

- *Teilnehmerdichte*: Die mittlere Teilnehmerdichte ρ_{N_T} für ein Ad-hoc-Netz mit N_T Teilnehmern, die über eine Fläche des Inhalts A verteilt sind, wird definiert als $\rho_{N_T} = N_T/A$. Bei den in diesem Kapitel beschriebenen Simulationsläufen wurde zur Untersuchung des Einflusses der Teilnehmerdichte meistens die Teilnehmerzahl N_T bei gleich bleibender Fläche A variiert.

  Es ist zu erwarten, dass die Teilnehmerdichte einen Einfluss auf die Funktion des untersuchten Zugangskontrollsystems hat, da dieses auf Beobachtung und Bewertung durch Nachbarknoten beruht und die Anzahl der Nachbarknoten jedes Knotens von der Teilnehmerdichte abhängt. Um den Zusammenhang zwischen Teilnehmerdichte und Nachbarzahl zu formulieren, wird außerdem eine weitere Größe benötigt, die bestimmt, welche anderen Knoten benachbart sind: die maximale Übertragungsreichweite. Diese kann in der Realität für jede Knotenpaarung verschieden sein, da sie von der Sendeleistung des Senders und der Empfindlichkeit des Empfängers abhängt; außerdem wird sie durch Hindernisse und Störfelder beeinflusst und ist nicht scharf definierbar, sondern äußert sich durch starkes Ansteigen der Fehlerquote, wenn sie überschritten wird. Nimmt man vereinfachend an, dass es eine für alle Knotenpaarungen gleiche und außerdem zeitlich und örtlich unveränderliche Entfernung r gebe, so dass fehlerfreie Signalübertragung genau bei jeder kürzeren Distanz zwischen zwei Knoten möglich sei, so zählen als Nachbarn genau alle Knoten, die sich innerhalb eines Kreises mit dem Radius r um den eigenen Standpunkt befinden. Die mittlere Nachbarzahl N_N in einem Netz mit N_T Teilnehmern und homogener Teilnehmerdichte ist dann

$$N_N = \rho_{(N_T-1)} \cdot \pi \cdot r^2 = \frac{N_T - 1}{A} \cdot \pi \cdot r^2. \tag{4.1}$$

- *Teilnehmerzahl:* Neben ihrem Zusammenhang mit der Teilnehmerdichte bzw. der Nachbarzahl hat die Teilnehmerzahl bei gleichbleibender Dichte außerdem Auswirkungen auf den Speicherbedarf jedes Knotens für Schlüssel und Vertrauensprofile. Ist weniger Speicherplatz vorhanden, als für die gesamte Informationsmenge benötigt würde, so kann mit wachsender Teilnehmerzahl ein immer geringerer Anteil gespeichert werden, und es werden im Bedarfsfall mehr Anfragen erforderlich. Auch wenn genügend Speicherplatz vorhanden ist, muss jeder Knoten bei größerer Teilnehmerzahl insgesamt mehr Anfragen stellen.

- *Mobilität:* Auf die Modellierung der Art und Weise der Bewegung der Knoten wird weiter unten im Abschnitt zu Mobilitätsmodellen eingegangen. Sowohl die Wahl des Modells als auch die Parameter der Mobilitätsmodelle – etwa die Verteilung von Bewegungsgeschwindigkeiten oder Aufenthaltsdauern – beeinflussen quantitativ messbare Größen wie die Dauer eines Nachbarschaftsverhältnisses oder die Anzahl an Änderungen in der Menge der Nachbarn eines Knotens pro Zeiteinheit. Änderungen an Nachbarschaftsverhältnissen verursachen einerseits einen gewissen Aufwand durch den Schlüsselaustausch zwischen Nachbarn und können sich kurzfristig nachteilig auswirken, wenn Nachbarn, zu denen Vertrauensverhältnisse aufgebaut wurden, durch andere, unbekannte Teilnehmer ersetzt werden. Andererseits kann sich die Zerstreuung der dem eigenen Knoten vertrauenden ehemaligen Nachbarn in entfernte Teile des Netzes auch positiv auswirken, weil damit die Wahrscheinlichkeit erhöht wird, dass entfernte Knoten von ihren Nachbarn positive Einschätzungen über den eigenen Knoten erhalten und damit auch bei der Zugangskontrolle positiv entscheiden.

- *Nutzungsverhalten:* Längerfristige Kommunikationsbeziehungen zwischen gleich bleibenden Teilnehmerpaaren dürften sich anders auswirken als kurzlebige Kommunikationsvorgänge mit ständig wechselnden Kommunikationspartnern, da zu erwarten ist, dass für den Aufbau einer Kommunikationsbeziehung innerhalb des Netzes ein erhöhter Aufwand entsteht, wenn die Knoten auf dem entstehenden Übertragungsweg die Endknoten oder sich gegenseitig noch nicht „kennen“.

Von besonderem Interesse sind weiterhin natürlich die Auswirkungen der Anwesenheit unkooperativer Teilnehmer auf die Funktion des Gesamtsystems. Darauf wird in einem eigenen Abschnitt weiter unten eingegangen.

4.3.1 Mobilitätsmodelle

Zur Modellierung des Mobilitätsverhaltens von Netzteilnehmern wurden zwei unterschiedliche Ansätze mit verschiedenen Eigenschaften verwendet: einerseits das bei der Simulation von Ad-hoc-Netzen sehr häufig eingesetzte *Random-Waypoint-Modell* und andererseits ein in hinführenden Arbeiten [Hof02] entwickeltes *Konferenzmodell*.

4.3.1.1 Random-Waypoint-Modell

Beim Random-Waypoint-Modell entsteht das Bewegungsmuster jedes Teilnehmers dadurch, dass er sukzessive eine Folge zufällig gewählter Wegpunkte ansteuert:

1. Der Teilnehmer wählt zufällig einen Zielpunkt im Simulationsgebiet (gleichverteilt) und ebenfalls zufällig eine Geschwindigkeit und bewegt sich im folgenden Verlauf gleichmäßig auf das Ziel zu.

2. Wenn das Ziel erreicht ist, wartet der Knoten dort für eine zufällig gewählte Zeitspanne und fährt anschließend wieder mit Schritt 1 fort.

Die Verteilungen der Geschwindigkeiten und Wartezeiten sind (abgesehen von der Form und Größe des Simulationsgebiets) die einzigen Parameter des Modells. Sie haben wesentlichen Einfluss auf die im vorigen Abschnitt genannten quantitativen Mobilitätskenngrößen.

Beim Random-Waypoint-Modell sind Aufenthaltswahrscheinlichkeit und Teilnehmerdichte innerhalb des Simulationsgebiets homogen verteilt: Die Wahrscheinlichkeit dafür, dass sich ein bestimmter Knoten oder eine bestimmte Anzahl beliebiger Knoten in einem bestimmten Teilgebiet aufhält, ist unabhängig von dessen bzw. deren Lage innerhalb des Simulationsgebiets. Diese Gleichförmigkeit macht das Modell sozusagen „simulationsfreundlich“, denn eine geringe Varianz in der Verteilung der Eingangsgrößen liefert auch bei relativ kurzen Simulationszeiten eine deutlich erkennbare Verteilung vieler Messgrößen. Die geringe Anzahl von Parametern vereinfacht ebenfalls die Auswertung und sorgt außerdem häufig für eine gute Vergleichbarkeit mit fremden Ergebnissen. Nachteilig ist dabei natürlich, dass die Ergebnisse nicht direkt auf reale Szenarien übertragbar sind, da angenommen werden muss, dass das Bewegungsmuster des Random-Waypoint-Modells von bei realen Teilnehmern vorkommenden Mustern recht weit entfernt ist.

Voraussetzung für die Homogenität der Teilnehmerdichte ist, dass das Simulationsgebiet keinen Rand besitzt, denn sonst fällt erstens für Knoten nahe dem Rand ein Teil des Gebiets weg, in dem sich potentielle Nachbarn aufhalten könnten, und zweitens ist die Aufenthaltswahrscheinlichkeit in der Mitte schon deshalb höher, weil sozusagen alle Wege durch die Mitte führen (alle Punkte auf dem Weg zwischen zwei beliebigen Aufenthaltsorten liegen etwa in einem kreisförmigen Simulationsgebiet immer weiter vom Rand entfernt als der randnähere der beiden Aufenthaltsorte). Um ein randloses Simulationsgebiet zu erhalten, wurde ein gekrümmter zweidimensionaler Raum – die Oberfläche eines Torus – zugrunde gelegt, der ein quadratisches Gebiet genau so aufnehmen kann, dass dessen rechte und linke bzw. dessen obere und untere Kante zusammenfallen.

4.3.1.2 Konferenzmodell

Beim Konferenzmodell ist das Bewegungsmuster der Teilnehmer grob dem von menschlichen Teilnehmern einer Konferenz, einer Messe oder einer ähnlichen Veranstaltung nachempfunden:

- Es gibt eine vorgegebene Anzahl von Veranstaltungsorten auf dem Simulationsgebiet. Jeder Veranstaltungsort ist dabei ein kreisförmiges Gebiet, dessen Position und Radius anfangs zufällig bestimmt werden und dann fest bleiben.
- An jedem Veranstaltungsort findet zu jeder Zeit je eine Veranstaltung statt, die eine bestimmte Dauer hat und ein bestimmtes Interessengebiet abdeckt. Beides wird zu Beginn jeder Veranstaltung zufällig festgelegt. Sobald an einem Ort eine Veranstaltung endet, beginnt dort wieder eine neue zufällig bestimmte Veranstaltung.
- Teilnehmer haben eine anfangs zufällig bestimmte Menge von Interessen und besuchen dazu Veranstaltungen. Jeder Teilnehmer hat entweder bereits eine Veranstaltung ausgewählt, oder er ist dabei, eine neue Veranstaltung zu bestimmen.
  - Wenn er eine Veranstaltung gewählt hat und diese noch nicht zu Ende ist, überprüft der Teilnehmer in jedem Zeitschritt zunächst, ob er sich innerhalb des Radius des gewählten Veranstaltungsortes befindet. Wenn ja, wartet er dort bewegungslos das Ende der Veranstaltung ab. Wenn nein, wählt er zufällig eine Geschwindigkeit und bewegt sich damit auf das Zentrum des Veranstaltungsortes zu.
  - Hat er noch keine Veranstaltung gewählt oder ist die gewählte Veranstaltung zu Ende, versucht der Teilnehmer in jedem Schritt einmal, eine neue Veranstaltung zu wählen. Dazu bestimmt er jeweils zufällig eine Veranstaltung und wählt diese dann aus, wenn ihr Interessengebiet in seiner eigenen Interessenmenge enthalten ist. Ist das nicht der Fall, so wartet er bis zum nächsten Schritt bewegungslos.

Beim Konferenzmodell wird versucht, ein etwas realitätsnäheres Bewegungsmuster zu erzeugen als beim Random-Waypoint-Modell, so dass die Ergebnisse eher auf reale Szenarien übertragbar sind und insbesondere auch Schwierigkeiten entdeckt werden können, die sich durch weniger homogene Teilnehmerverteilungen ergeben. Deshalb wird auch ein Simulationsgebiet mit Rand verwendet (rechteckig oder oval). Da die Ergebnisse jedes Simulationslaufs stark vom Aufbau der jeweils erzeugten „Konferenz“ abhängen, sind sehr viele Läufe erforderlich, wenn man stichhaltige statistische Aussagen erhalten möchte.

4.3.2 Nutzungsmodell

Für die Untersuchung der Eigenschaften des Zugangskontrollsystems unter Ausschluss etwaiger Einflüsse durch Abhängigkeiten zwischen Nutzdatenpaketen, wie sie durch bestimmte reale Anwendungen und Protokolle der Anwendungs- und der Transportschicht verursacht werden, wurde das folgende Verfahren zur Erzeugung von Nutzdatenpaketen verwendet: Jeder Knoten sendet einzelne Pakete an zufällig (gleichverteilt) gewählte Knoten im Netz, wobei er zwischen je zwei Paketen eine Pause zufällig gewählter Länge ($\exp_\lambda$-verteilt) einlegt. Die Paketerzeugung ist damit ein Poisson-Prozess. Die durchschnittliche Nachrichtenrate beträgt $1/\lambda$ Nachrichten pro Sekunde.

Dieses Nutzungsmodell stellt gewissermaßen eine Hochlast-Situation für das Zugangskontrollsystem dar, da die Quellknoten der zu kontrollierenden Pakete ständig wechseln und damit ständig neue Informationen beschafft werden müssen. Die Größe der Datenpakete spielt für das Zugangskontrollverfahren keine Rolle und wurde konstant gewählt.

4.4 Eigenschaften der simulierten Netze

Im Folgenden werden zunächst einige grundlegende Eigenschaften der betrachteten Netze untersucht. Die Ergebnisse sollen einerseits dazu dienen, „interessante“ Parametrisierungen für die Analyse der Funktion des entworfenen Konzepts zu finden, und andererseits Vergleichswerte für die dort ermittelten Messwerte liefern.

Zunächst soll der Einfluss der Teilnehmerdichte auf die grundlegende Funktion des Netzwerks betrachtet werden. Dazu wurden Netzwerke ohne Zugangskontrolle mit unterschiedlicher Teilnehmerzahl in einem gleichbleibend großen, randlosen quadratischen Gebiet mit einer Kantenlänge von 600 Metern simuliert. Die Teilnehmermenge blieb dabei während jedes Versuchs konstant. Die Teilnehmer bewegten sich unabhängig voneinander nach dem Random-Waypoint-Modell mit gleichverteilt zufälligen Geschwindigkeiten zwischen 1 und $10\,\frac{\mathrm{m}}{\mathrm{s}}$ und gleichverteilt zufälligen Wartezeiten zwischen 1 und 30 s. Als Nutzlast dienten zufällig adressierte Einzelpakete mit einer Rate von $1,25$ Nachrichten pro Sekunde (Zeitabstände $\exp_\lambda$-verteilt mit $\lambda = 1\,/\,1,25$). Als Übertragungsreichweite wurden 180 m angenommen (auch bei den folgenden Versuchen, soweit nicht anders angegeben).

Abbildung 4.1 zeigt die Erfolgsquote bei der Übermittlung von Anwendungsnachrichten, also den Wert des Quotienten

$$\frac{\text{Anzahl korrekt am Ziel angekommener Nachrichten}}{\text{Anzahl abgesandter Nachrichten}},$$

in Abhängigkeit von der Teilnehmerzahl. Gestrichelt ist außerdem die Weiterleitungsfehlerquote eingezeichnet, welche angibt, wie viele der der Netzwerkschicht zur Übermittlung übergebenen Nachrichten aufgrund eines fehlenden Wegs zum Ziel verworfen werden mussten. Da diese Art von Fehler

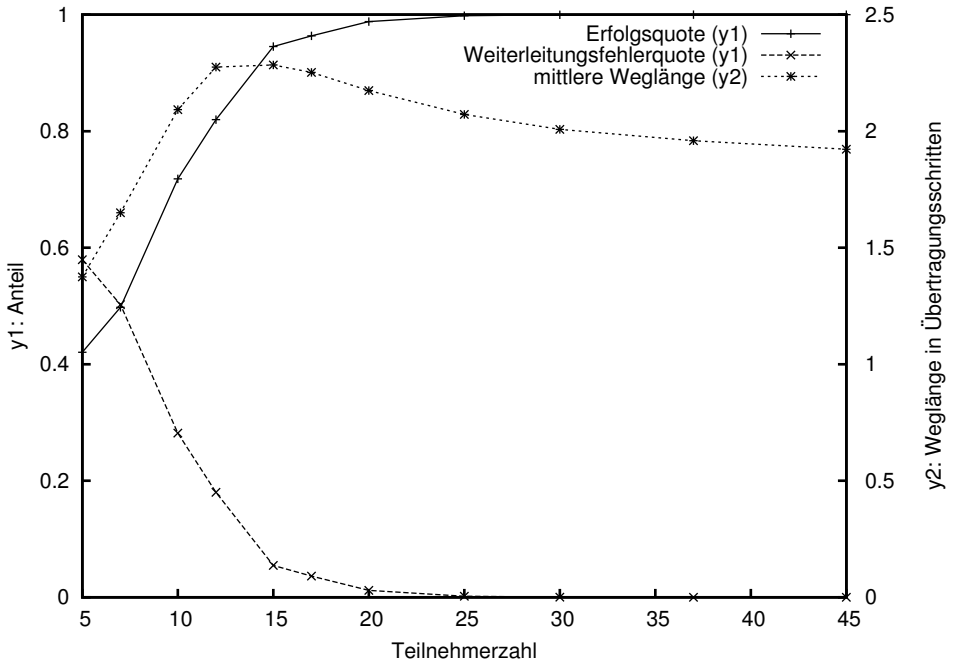


Abbildung 4.1: Erfolgsquote und Weiterleitungsfehlerquote (y1-Achse links) und mittlere Weglänge (y2-Achse rechts) in Abhängigkeit von der Teilnehmerzahl beim Random-Waypoint-Modell

im hier betrachteten Szenario den einzig möglichen Grund für die Nichtankunft einer Anwendungsnachricht darstellt, und da hier ausschließlich Anwendungsnachrichten übermittelt werden, ergänzen sich die oben erwähnte Erfolgsquote und die Weiterleitungsfehlerquote hier immer zu eins.

An den Graphen ist deutlich zu erkennen, dass bei zu geringer Teilnehmerdichte ein großer Teil der Nachrichten nicht zugestellt werden kann, weil kein Weg zwischen Absender und Empfänger existiert; das Netz zerfällt hier häufig in Teilnetze, zwischen deren Teilnehmermengen dann keine Funkverbindung besteht. Mit steigender Teilnehmerdichte erhöht sich die Erfolgsquote, im Beispiel erreichen bei 15 Teilnehmern 95% der Anwendungsnachrichten ihr Ziel, ab etwa 25 Teilnehmern sind es 99%.

Punktiert ist in Abbildung 4.1 außerdem die mittlere Weglänge für erfolgreiche Nachrichtenübertragungen eingezeichnet. Durch das verwendete „allwissende“ Wegfindungsverfahren sind übrigens fast alle Übertragungen, die überhaupt begonnen werden, auch erfolgreich. Nur falls ein existierender Weg *während der Übertragung einer Nachricht* durch Knotenbewegung abreißt, kann eine Nachricht unterwegs verloren gehen, weil sie nicht mehr weitergeleitet werden kann. Die Zeitspanne der Übertragung einer Nachricht (auch bezeichnet als deren Laufzeit) ist recht kurz, wie an Abbildung 4.2 zu erkennen ist: maximal etwa 2 ms. Reale Wegfindungsverfahren haben häufig unvollständige Informationen, weil Knotenbewegungen *während der Verbreitung der Topologieinformation* sie schon verfälschen. Da die Netztopologie sich häufig ändert, aber der Aufwand für die Aktualisierung der auf die Knoten verteilten Topologieinformation begrenzt werden soll, wird die Verbreitung neuer Information meist absichtlich etwas verzögert. Dort wird deshalb häufiger der Fall auftreten, dass bei Beginn der Übertragung die Existenz eines Weges angenommen und erst nach einigen Weiterleitungsschritten festgestellt wird, dass dieser nicht mehr zum Ziel führt.

Von niedrigen zu höheren Teilnehmerzahlen hin steigt in Abbildung 4.1 die mittlere Weglänge (bei fallender Fehlerquote) zunächst an, weil durch die höhere Teilnehmerdichte auch größere Entfernungen überbrückt werden können und weniger Nachrichten wegen fehlenden Wegs verworfen werden müssen. Bei 12 Teilnehmern erreicht die Weglänge im Beispiel (nahezu) ihren Maximalwert. Danach fällt sie wieder etwas ab, da die steigende Teilnehmerdichte nun direktere Wege erlaubt, d. h.

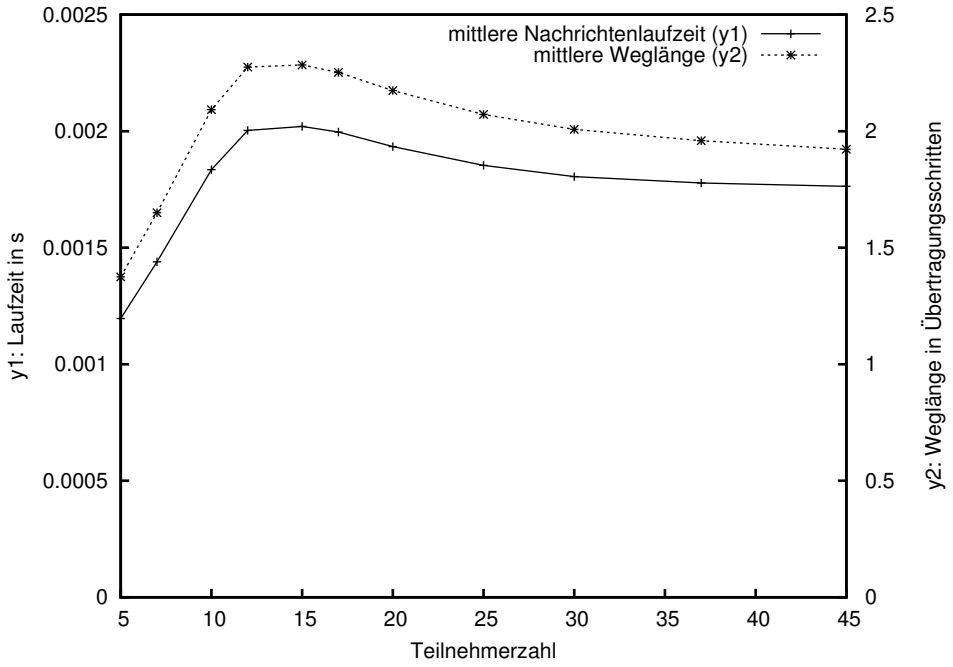


Abbildung 4.2: Mittlere Nachrichtenlaufzeit und mittlere Weglänge in Abhängigkeit von der Teilnehmerzahl beim Random-Waypoint-Modell

es müssen weniger Umwege in Kauf genommen werden. Im Anhang sind ergänzend die relativen Häufigkeiten bestimmter Weglängen bei den hier verwendeten Teilnehmerzahlen angegeben (Abbildung A.1).

Dass die maximale mittlere Weglänge mit knapp $2,3$ Weiterleitungsschritten relativ kurz ausfällt liegt übrigens daran, dass auch die maximale Entfernung zweier Teilnehmer im verwendeten randlosen Simulationsgebiet ($\frac{1}{2}\sqrt{2} \cdot 600\,\text{m} \approx 424\,\text{m}$) nicht besonders groß ist im Vergleich zur Übertragungsreichweite von 180 m. Im Anhang sind ergänzend noch die mittlere Weglänge und die relativen Häufigkeiten bestimmter Weglängen bei verschiedenen Teilnehmerzahlen für ein gleich großes Simulationsgebiet bei einer geringeren Übertragungsreichweite von 100 m dargestellt (Abildungen A.2 und A.3).

Um zu überprüfen, welchen Einfluss die Knotenmobilität auf die betrachteten Messwerte hat, wurden Versuchsreihen mit verschiedenen Mobilitätsmodellen und Parametrisierungen derselben durchgeführt. Zunächst wurden Geschwindigkeit und Wartezeit beim Random-Waypoint-Modell variiert:

- Variation des Geschwindigkeitsmittelwerts zwischen 0,5 und 15 m/s bei Random-Waypoint-Modell mit normalverteilter Geschwindigkeit (Varianz konstant) und konstanter Wartezeit (je einmal mit Wartezeiten von 1 und 180 s).
- Variation des Mittelwerts der Aufenthaltszeit zwischen 1 und 1800 s bei Random-Waypoint-Modell mit normalverteilter Aufenthaltszeit (konstante Varianz) und konstanter Geschwindigkeit (je einmal mit Geschwindigkeiten von 0,5 und 9 m/s)

Dabei wurde die Erwartung bestätigt, dass Geschwindigkeit und Wartezeit in diesem Szenario praktisch keinen Einfluss auf mittlere Weglänge, Nachrichtenlaufzeit oder Weiterleitungsfehlerquote haben. Bei einem realen Wegfindungsverfahren wäre dies in der Regel aus den oben genannten Gründen anders, da die verteilte Weiterleitungsinformation bei sich häufiger ändernder Netztopologie schneller veralten und damit im Mittel weniger genau zur realen Topologie passen würde, so dass weniger ideale Wege gefunden würden und mehr Nachrichten durch abreißende Wege verlorengingen.

Anzahl unterschiedlicher Interessengebiete	8
Anzahl Veranstaltungsorte	8
Radius eines Veranstaltungsortes in m	gleichverteilt aus [10, 50)
Geschwindigkeit der Teilnehmer in m/s	gleichverteilt aus [0.27, 2.22)
Dauer von Veranstaltungen in s	gleichverteilt aus [60, 240)
Anteil anfangs in Veranstaltungen befindlicher Teilnehmer	60 %

Tabelle 4.1: Parameterwahl beim Konferenzmodell

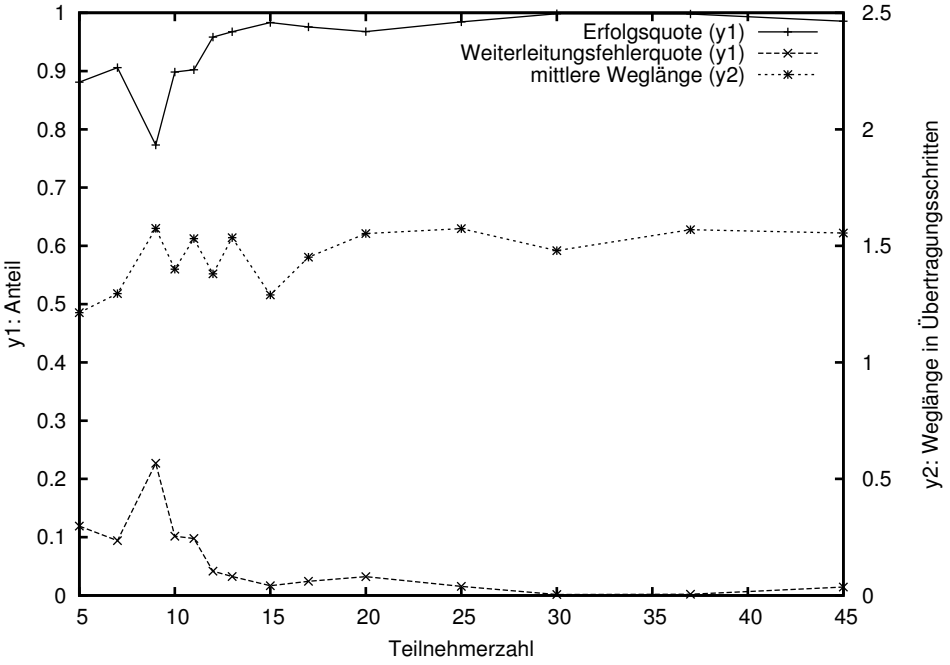


Abbildung 4.3: Erfolgsquote und Weiterleitungsfehlerquote (y1-Achse links) und mittlere Weglänge (y2-Achse rechts) in Abhängigkeit von der Teilnehmerzahlbeim Konferenzmodell

Beim den Experimenten zum Konferenzmodell wurde ein gleich großes Simulationsgebiet verwendet wie beim Random-Waypoint-Modell, wobei es nun allerdings einen Rand besaß, da das Konferenzmodell auf Praxisnähe abzielt und Randlosigkeit nur in theoretischen (oder – bisher ebenfalls fiktiven – weltumspannenden) Netzen erreichbar ist. Auch für die weiteren Parameter wurde versucht, realitätsnahe Belegungen zu finden; so sollte etwa die Bewegungsgeschwindigkeit der Teilnehmer in einem für Fußgänger üblichen Bereich liegen. Tabelle 4.1 gibt die bei den beschriebenen Versuchen gewählten Belegungen an. Wie bereits erwähnt hat die einmalig zu Anfang eines jeden Simulationslaufs zufällig bestimmte Platzierung der Veranstaltungsorte im Simulationsgebiet einen merklichen Einfluss auf die Eigenschaften des entstehenden Netzwerks. Dies ist auch in Abbildung 4.3 deutlich zu erkennen, in welcher zu jeder zu einem Datenpunkt gehörigen x-Koordinate ein Simulationslauf durchgeführt wurde. Neben den großen Schwankungen aller Messgrößen bei niedrigen Knotenzahlen sind auch bei höheren Knotenzahlen noch merkliche Schwankungen vorhanden, wie man etwa an der Erfolgsquote von nur etwa 98,5 % bei 45 Knoten sieht. Eine außerdem durchgeführte Serie gleich parametrisierter Versuche mit je 25 Knoten ergab dort eine mittlere Erfolgsquote von 98,66 % und eine mittlere Weglänge von 1,5 Übertragungsschritten.

Die mittlere Weglänge ist deutlich kürzer als die beim Random-Waypoint-Modell ermittelte. Hier zeigt sich eine bereits genannte Eigenschaft des endlichen Simulationsgebiets: Die Aufenthaltswahrscheinlichkeit für Teilnehmer ist in der Mitte des Simulationsgebiets höher, insbesondere bei geringen Geschwindigkeiten. Durch die Häufung der Teilnehmer in der Mitte werden die Wege zwischen zufällig gewählten Teilnehmerpaaren im Mittel kürzer.

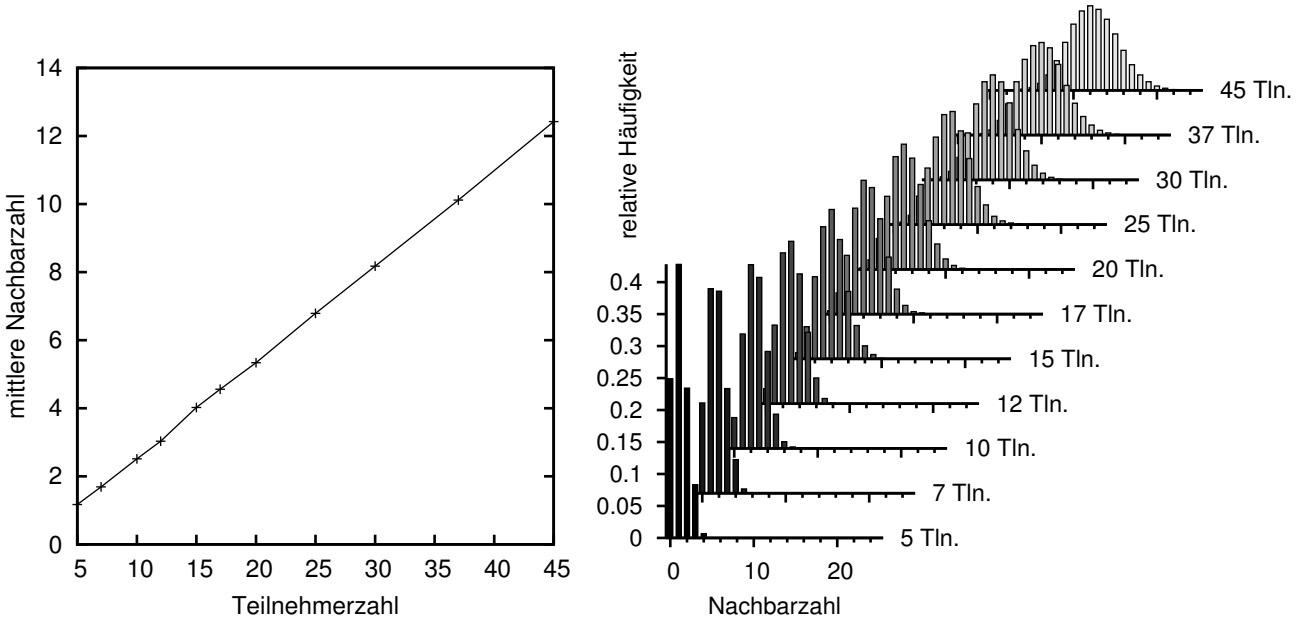


Abbildung 4.4: Mittlere Nachbarzahl (links) und relative Häufigkeiten bestimmter Nachbarzahlen (rechts) in Abhängigkeit von der Teilnehmerzahl bei gleichbleibend großem Simulationsgebiet

4.5 Beobachtung und Bewertung der Weiterleitung

Im Folgenden wird untersucht, ob und unter welchen Bedingungen die in Abschnitt 3.6 beschriebene Methode zur Beobachtung des Weiterleitungsverhaltens benachbarter Knoten ihre Funktion erfüllen kann. Sie soll einen wesentlichen Beitrag zur Beurteilung der Kooperativität anderer Knoten leisten und muss dazu eine ausreichende Menge an Information über das Verhalten der Nachbarn möglichst schnell liefern.

4.5.1 Nachbarzahl und Dauer von Nachbarschaftsverhältnissen

Die Zahl der Nachbarn eines Knotens bestimmt, wie oft eine erbrachte Leistung oder deren Verweigerung maximal beobachtet werden kann, also wie viele Einschätzungen durch das eigene Verhalten beeinflussbar sind. Je länger zwei Knoten benachbart sind, desto sicherer werden ihre gegenseitigen, aus Beobachtungen gewonnenen Einschätzungen in der Regel werden. Deshalb wird nun zunächst betrachtet, wie Nachbarzahl und Dauer von Nachbarschaftsverhältnissen in simulierten Szenarien verteilt sind.

In Abschnitt 4.3 wurde mit Gleichung 4.1 bereits eine theoretische Formel für den Zusammenhang zwischen Teilnehmerzahl, Gebietsgröße, Übertragungsreichweite und mittlerer Nachbarzahl angegeben. Abbildung 4.4 zeigt links die per Simulation unter den auch zu Beginn von Abschnitt 4.4 verwendeten Bedingungen (Random-Waypoint-Modell mit gleichverteilter Geschwindigkeit zwischen 1 und 10 m/s und gleichverteilter Wartezeit zwischen 1 und 30 s) ermittelte Abhängigkeit zwischen Teilnehmerzahl und mittlerer Nachbarzahl bei gleichbleibender Gebietsgröße; die aufgrund der Formel erwartete Linearität ist deutlich zu erkennen. Die relativen Häufigkeiten bestimmter Nachbarzahlen sind im rechten Teil der Abbildung für Teilnehmerzahlen zwischen 5 und 45 Knoten aufgezeichnet.

Um den Einfluss der Mobilität auf die Dauer von Nachbarschaftsverhältnissen und die Auswirkungen auf Beobachtung, Bewertung und (später) Zugangskontrolle in den betrachteten Szenarien zu untersuchen, wurden einige Versuche mit unterschiedlichen Mobilitätsmodellen und Parametrisierungen

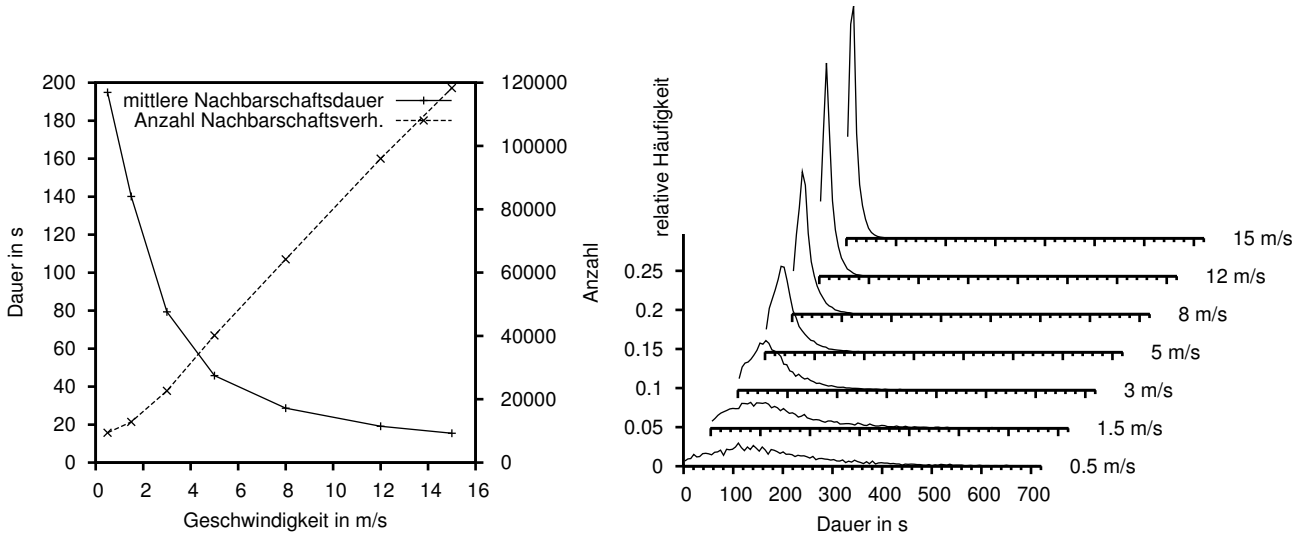


Abbildung 4.5: Mittlere Dauer von Nachbarschaftsverhältnissen (links) und Häufigkeit bestimmter Dauern (6-Sekunden-Intervalle; rechts) in Abhängigkeit von der Geschwindigkeit bei Aufenthaltsdauer 1 s

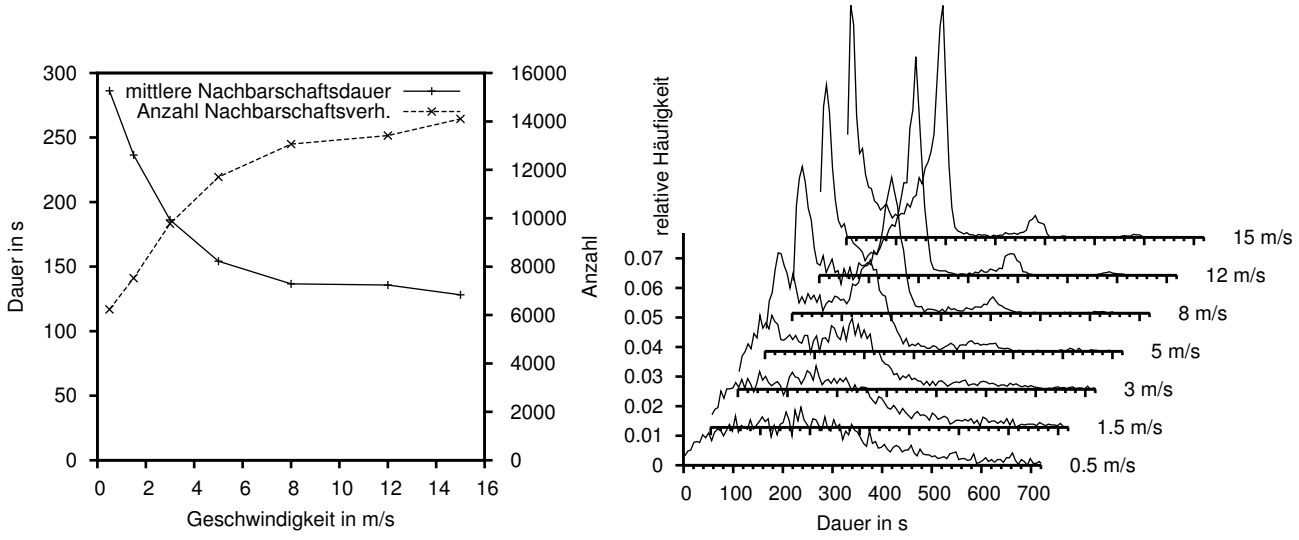


Abbildung 4.6: Mittlere Dauer von Nachbarschaftsverhältnissen (links) und Häufigkeiten bestimmter Dauern (rechts) in Abhängigkeit von der Geschwindigkeit bei Aufenthaltsdauer 180 s

durchgeführt. Dabei wurde ein Netzwerk mit 25 Teilnehmern verwendet, die sich zunächst nach dem Random-Waypoint-Modell in einem randlosen quadratischen Gebiet (Kantenlänge 600 m) bewegten, wobei Geschwindigkeit und Wartezeit jeweils normalverteilt mit geringer Varianz und unterschiedlichen Mittelwerten erzeugt wurden. Die mittlere Nachbarzahl blieb dabei erwartungsgemäß konstant bei etwa dem auch rechnerisch ermittelten Wert ($\frac{24}{600^2\,\mathrm{m}^2} \cdot \pi \cdot 180^2\,\mathrm{m}^2 \approx 6,79$), und auch die ermittelte Verteilung der Nachbarzahlen entsprach jeweils recht gut der Normalverteilung (ähnlich wie in Abbildung 4.4).

Bei sehr kurzen Aufenthaltszeiten (Mittelwert 1 s) ergibt sich ein regelmäßiger Zusammenhang zwischen der Knotengeschwindigkeit und der Dauer von Nachbarschaftsverhältnissen. Dies ist in Abbildung 4.5 zu erkennen, in welcher links die mittlere Nachbarschaftsdauer sowie die dazu umgekehrt proportionale Anzahl von Nachbarschaftsverhältnissen über der gewählten mittleren Knotengeschwindigkeit aufgetragen ist. Auf der rechten Seite sind die relativen Häufigkeiten des Auftretens bestimmter Nachbarschaftsdauern für unterschiedliche Knotengeschwindigkeiten aufgetragen; die er-

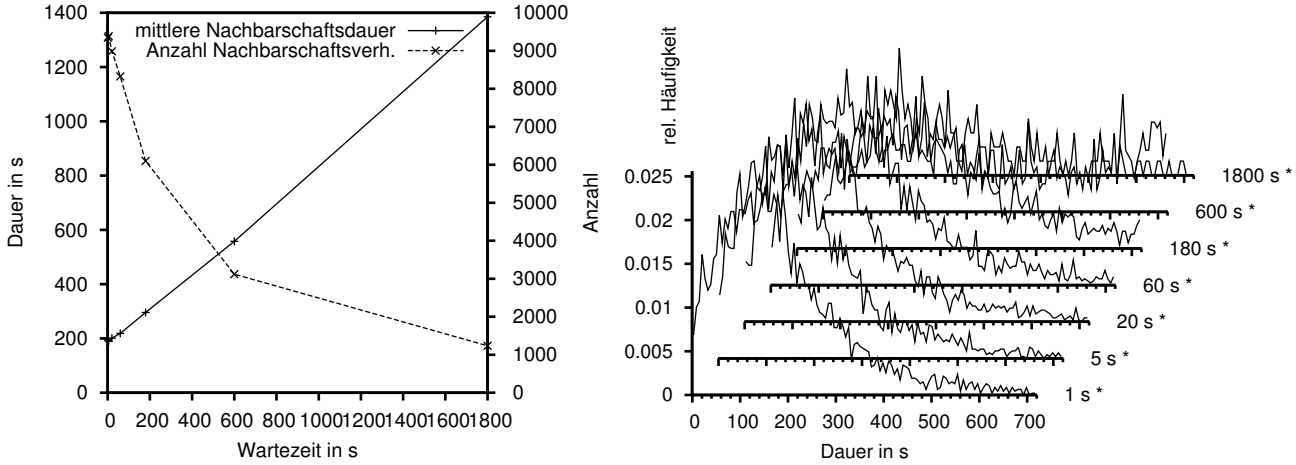


Abbildung 4.7: Mittlere Dauer von Nachbarschaftsverhältnissen (links) und Häufigkeiten bestimmter Dauern (rechts) in Abhängigkeit von der Aufenthaltsdauer bei niedriger Geschwindigkeit.

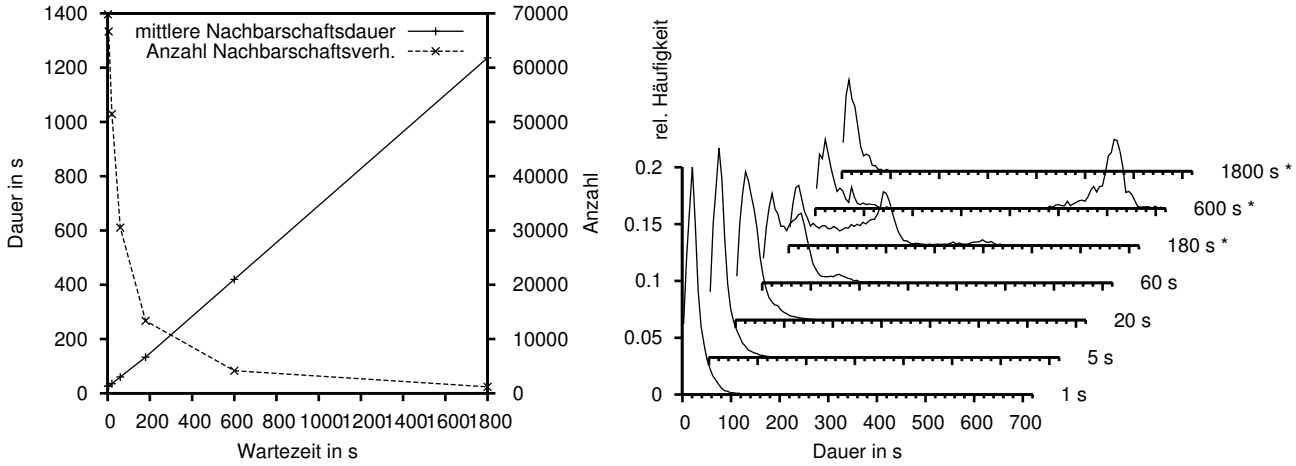


Abbildung 4.8: Mittlere Dauer von Nachbarschaftsverhältnissen (links) und Häufigkeiten bestimmter Dauern (rechts) in Abhängigkeit von der Aufenthaltsdauer bei hoher Geschwindigkeit

mittelten Nachbarschaftsdauern wurden für die Zählung der Häufigkeiten in aufeinander folgende Zeitintervalle der Länge 6 Sekunden eingeordnet.

Bei längeren Aufenthaltszeiten (Mittelwert 180 s) streut die Dauer von Nachbarschaftsverhältnissen stärker und ist im Mittel größer. An der in Abbildung 4.6 rechts dargestellten zugehörigen Verteilung der relativen Häufigkeiten bestimmter Dauern sind wegen der hier hier gewählten geringen Varianz deutliche Häufungen bei Vielfachen der Aufenthaltszeit zu erkennen.

Variiert man statt der Geschwindigkeit die Aufenthaltsdauer, so erhält man die in den Abbildungen 4.7 (Geschwindigkeit um 0,5 m/s) und 4.8 (Geschwindigkeit um 9 m/s) links gezeigten Verläufe der mittleren Nachbarschaftsdauer. Man sieht, dass der Mittelwert nahezu linear mit der Aufenthaltsdauer ansteigt. Rechts sind jeweils wieder Folgen von Häufigkeitsverteilungen dargestellt, wobei allerdings bei größeren Aufenthaltszeiten ein beträchtlicher Teil der auftretenden Dauern außerhalb der gezeigten Skala liegt (gekennzeichnet jeweils durch einen Stern hinter der Angabe der Aufenthaltszeit). Die etwas unübersichtliche rechte Seite der Abbildung 4.7, die hier nur eine Tendenz veranschaulichen soll, ist im Anhang nochmals etwas größer dargestellt (Abbildung A.4).

Insgesamt lässt sich zusammenfassen, dass beim Random-Waypoint-Bewegungsmodell größere Geschwindigkeiten zu einer Häufung kurzer Nachbarschaftsdauern und größere Aufenthaltszeiten zu einer stärkeren Streuung führen.

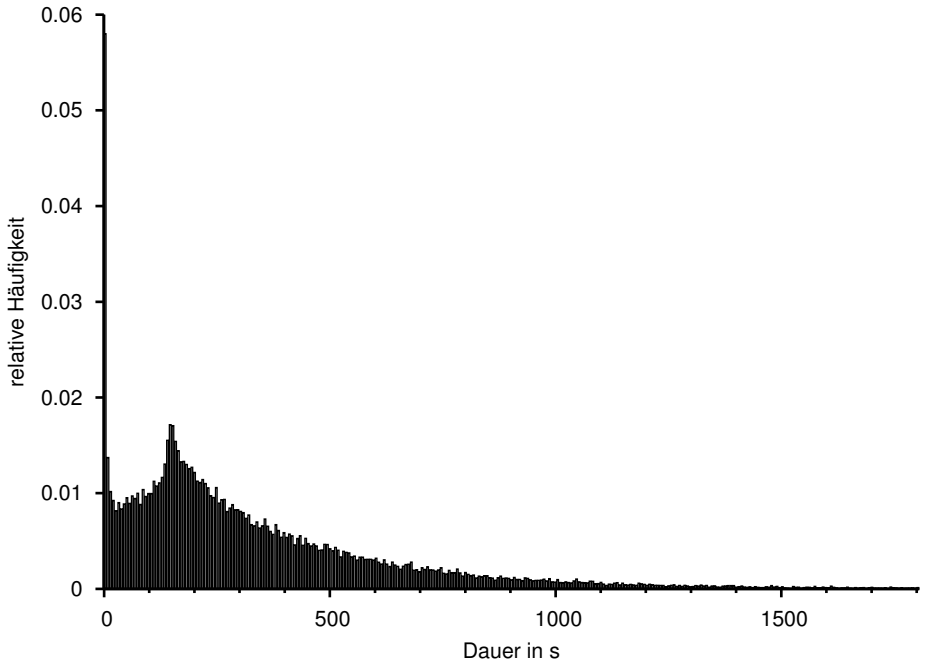


Abbildung 4.9: Über zehn Versuche (je 3 h Simulationszeit) gemittelte Häufigkeitsverteilung der Nachbarschaftsdauer (6-Sekunden-Intervalle) bei Bewegung nach dem Konferenzmodell

Für das Konferenzmodell wurden wieder die in Tabelle 4.1 genannten Parameterwerte verwendet. Die in Abbildung 4.9 gezeigte Verteilung der Dauer von Nachbarschaftsverhältnissen zeigt, dass diese Dauer hier wesentlich mehr streut als beim Random-Waypoint-Modell: Es gibt viele lange haltende Nachbarschaftsverhältnisse. Einerseits geht dies darauf zurück, dass Teilnehmerpaare oder -gruppen zufällig dieselbe Folge von Veranstaltungen wählen und so längere Zeit benachbart bleiben. Andererseits äußert sich auch hier wieder die erhöhte Teilnehmerdichte in der Mitte des Simulationsgebiets.

Auffallend ist eine starke Häufung der niedrigsten erfassten Nachbarschaftsdauer (kleiner 6 s), die hauptsächlich durch sich parallel bewegende Teilnehmer zustande kommt, deren Entfernungen durch leichte Schwankungen der Bewegungsgeschwindigkeiten gerade um die Übertragungsreichweite pendeln.

4.5.2 Bewertungsbedingung

Damit eine erbrachte Weiterleitungleistung auch bewertet werden kann, müssen nicht nur Beobachter vorhanden sein, sondern es muss auch jeweils die in Abschnitt 3.6.4 geforderte Zusatzbedingung erfüllt sein: Der Beobachter muss auch den vorigen Weiterleitungsschritt beobachtet haben, damit er sicher sein kann, dass es sich um eine Weiterleitung und nicht um die Aussendung einer eigenen Nachricht handelt.

Um empirisch zu untersuchen, ob die Zusatzbedingung die Bewertungsmöglichkeiten zu stark einschränkt, wurde im Versuch zunächst die Anzahl aller vergebenen positiven Weiterleitungsbewertungen erfasst und in Verhältnis zur Zahl der insgesamt durchgeführten Weiterleitungsvorgänge gesetzt. Abbildung 4.10 zeigt links das Ergebnis, aufgetragen über der Teilnehmerzahl des verwendeten Netzes; zum Vergleich ist dort außerdem die ermittelte mittlere Nachbarzahl eingetragen. Man sieht, dass die mittlere Zahl der Bewertungen pro weitergeleiteter Nachricht näherungsweise proportional zur Teilnehmer- und etwa halb so groß wie die mittlere Nachbarzahl ist, abgesehen vom unteren Teilnehmerzahlbereich, wo sie etwas höher ist. Bei der für einen nahezu von Weiterleitungsfehlern freien

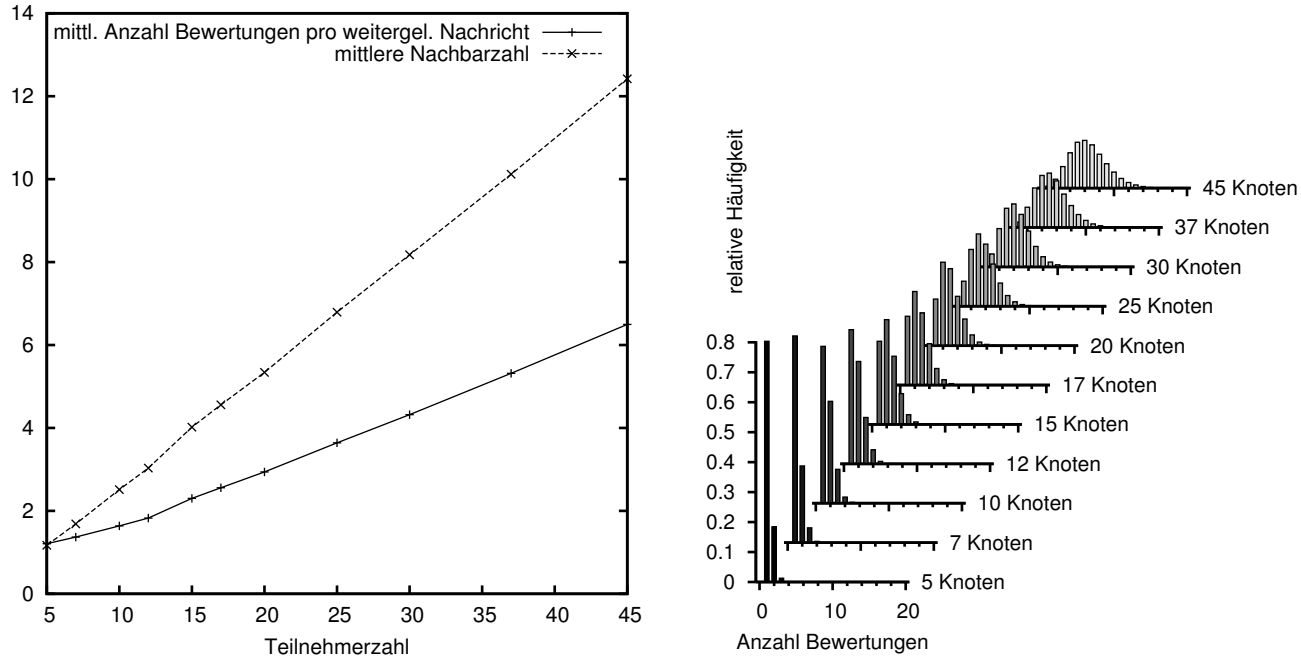


Abbildung 4.10: Anzahl (positiver) Bewertungen pro weitergeleiteter Nachricht (links) sowie Häufigkeiten bestimmter Anzahlen positiver Bewertungen bei weitergeleiteten Nachrichten (rechts)

Netzbetrieb erforderlichen Zahl von 25 Teilnehmern (siehe dazu Abbildung 4.1) wurden immerhin 3,64 positive Bewertungen pro erfolgter Weiterleitung vergeben. Es erscheint durchaus plausibel, aufgrund einer solchen Zahl von bewerteten Beobachtungen zügig Einschätzungen zur Kooperativität der beobachteten Teilnehmer zu gewinnen. Hier wurden nur positive Bewertungen betrachtet, weil die verwendeten Modellteilnehmer niemals absichtlich Leistungen verweigerten. Wäre dies der Fall, käme jeweils eine ähnliche Zahl an Negativbewertungen zustande.

Neben der Gesamtzahl aller Bewertungen wurde außerdem für jede einzelne Nachricht eine Bewertungsquote als Quotient aus Bewertungsanzahl und Beobachterzahl erfasst. Die über alle weitergeleiteten Nachrichten des jeweiligen Versuchs gemittelte Bewertungsquote ist in Abbildung 4.11 über der Teilnehmerzahl aufgetragen.

Dass die gemittelte Bewertungsquote gerade bei sehr niedrigen Teilnehmerzahlen erhöht ist, erscheint zunächst paradox und kommt wie folgt zustande: Bei sehr geringer Teilnehmerdichte ist die Wahrscheinlichkeit dafür, dass ein Teilnehmer genau zwei Nachbarn hat, deutlich höher als diejenige für mehr als zwei Nachbarn (die mittlere Nachbarzahl liegt hier im Versuch auch unterhalb 2). Situationen mit weniger als zwei Nachbarn spielen für die Bewertungsquote keine Rolle, da ein Knoten mit weniger als zwei Nachbarn nicht weiterleiten und damit auch keine Bewertungen erhalten kann.

Leitet ein Teilnehmer in einer Situation mit zwei Nachbarn eine Nachricht weiter, so wird dies nahezu sicher von seinem Vorgänger (der meist auch Quelle ist) beobachtet und bewertet, während der Nachfolger in der Regel nicht wertet, da er den vorigen Übertragungsschritt nicht beobachten konnte; wäre dies der Fall, so hätte der Vorgänger die Nachricht direkt an den Nachfolger übertragen können (und wegen des hier verwendeten idealen Wegfindungsverfahrens hätte er dies auch mit Sicherheit getan). Die Bewertungsquote ist damit in solchen Fällen praktisch immer 1/2, liegt also deutlich höher als der Mittelwert bei allen Teilnehmerzahlen der Abbildung 4.11. Wegen der erwähnten hohen Wahrscheinlichkeit der Zwei-Nachbar-Situation ergibt sich damit bei niedrigen Teilnehmerdichten eine erhöhte mittlere Bewertungsquote.

Bezüglich des Einflusses der Teilnehmermobilität auf die in diesem Abschnitt beschriebenen Zusammenhänge ist festzustellen, dass sie lediglich von der Teilnehmerdichte des betrachteten Netzwerks

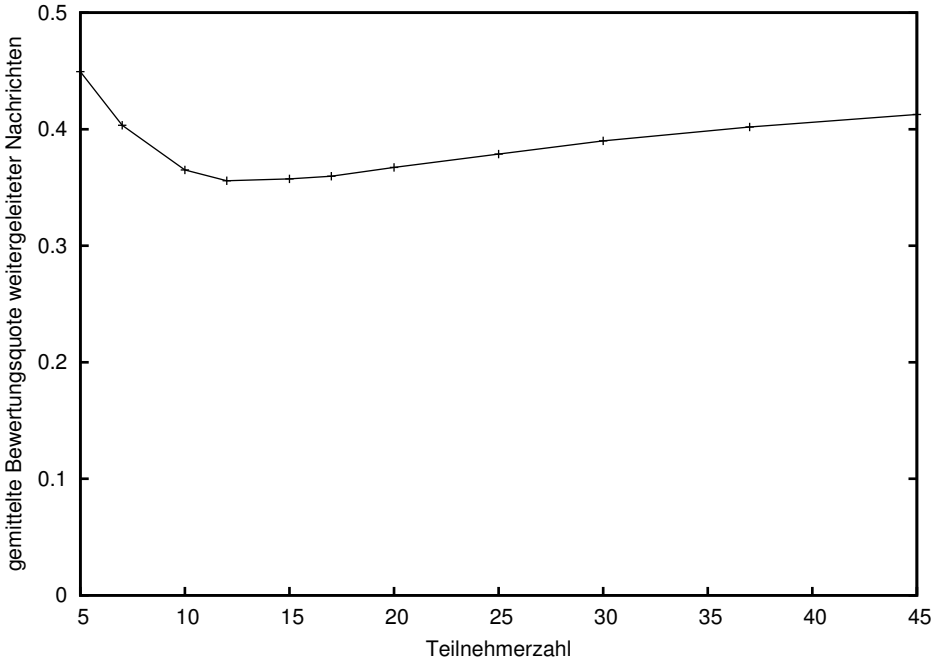


Abbildung 4.11: Gemittelte Bewertungsquote weitergeleiteter Nachrichten in Abhängigkeit von der Teilnehmerzahl

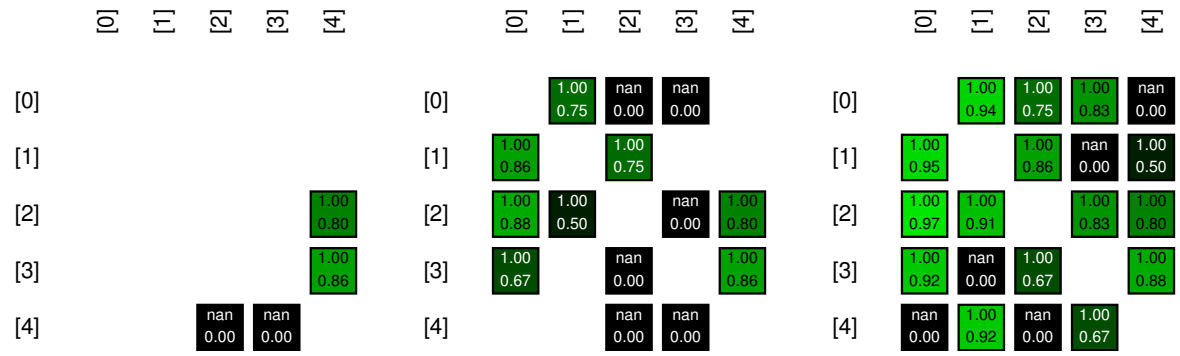


Abbildung 4.12: Einschätzungsmatrix bezüglich Kooperativität in einem Netz mit 5 Teilnehmern nach 30, 60 und 120 s

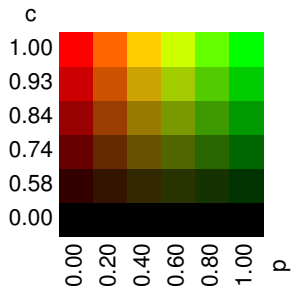


Abbildung 4.13: Farbskala für Einschätzungsmatrizen

abhängen. Da diese bei Verwendung des Random-Waypoint-Modell homogen ist, ergeben sich auch durch Variation der Verteilungen für Geschwindigkeiten und Aufenthaltsdauern keine Änderungen. Beim Konferenzmodell ist die mittlere Teilnehmerdichte dagegen nicht überall im Simulationsgebiet gleich groß, sondern die Teilnehmer konzentrieren sich auf Veranstaltungsorte und Verbindungswege dazwischen, so dass sie sich meist in Gebieten erhöhter Teilnehmerdichte aufhalten. Entsprechende Versuche zeigen, dass sich dadurch regelmäßig eine etwas höhere Bewertungsquote ergibt als nach dem oben gezeigten Zusammenhang zur Teilnehmerzahl der Fall wäre.

4.5.3 Meinungsbildung aufgrund der Weiterleitungsbeobachtung

Um die durch Bewertung des Weiterleitungsverhaltens entstehenden Einschätzungen zu visualisieren, lassen sich alle möglichen Einschätzungen jedes Knotens über jeden anderen zu einem bestimmten Zeitpunkt in einer „Einschätzungsmatrix“ anordnen. Dies wurde in Abbildung 4.12 für ein Netzwerk mit fünf Knoten zu drei verschiedenen Zeitpunkten getan. Jedes Kästchen in einer solchen Matrix steht für eine bei dem am linken Rand durch eine Nummer in eckigen Klammern identifizierten Knoten vorliegende Einschätzung über den am oberen Rand angegebenen anderen Knoten. Jedes Kästchen enthält zwei Zahlenwerte: der obere gibt die Position der Einschätzung an, der untere ihre Sicherheit; beim Sicherheitswert 0 ist die Position undefiniert, was durch das Kürzel „nan“ („not a number“) angezeigt wird. Die durch ein Kästchen repräsentierte Einschätzung wird außerdem auch durch seine Farbe wiedergegeben: Die Position bestimmt den Farbton auf einer Skala zwischen rot (p = 0) und grün (p = 1), die Sicherheit bestimmt die Helligkeit zwischen schwarz (c = 0) und hell-farbig (c = 1); eine Farbskala ist in Abbildung 4.13 dargestellt.

Einschätzungen mit Sicherheitswert 0 sind eigentlich natürlich bedeutungslos; sie werden hier nur angezeigt, weil das entsprechende Vertrauensprofil beim zugehörigen Knoten schon angelegt wurde, was immer beim Schlüsselaustausch mit neuen Nachbarn geschieht. Da nur Nachbarn bewertet werden, mit denen auch ein Schlüssel ausgetauscht wurde, und wegen der Symmetrie des Schlüsselaus-

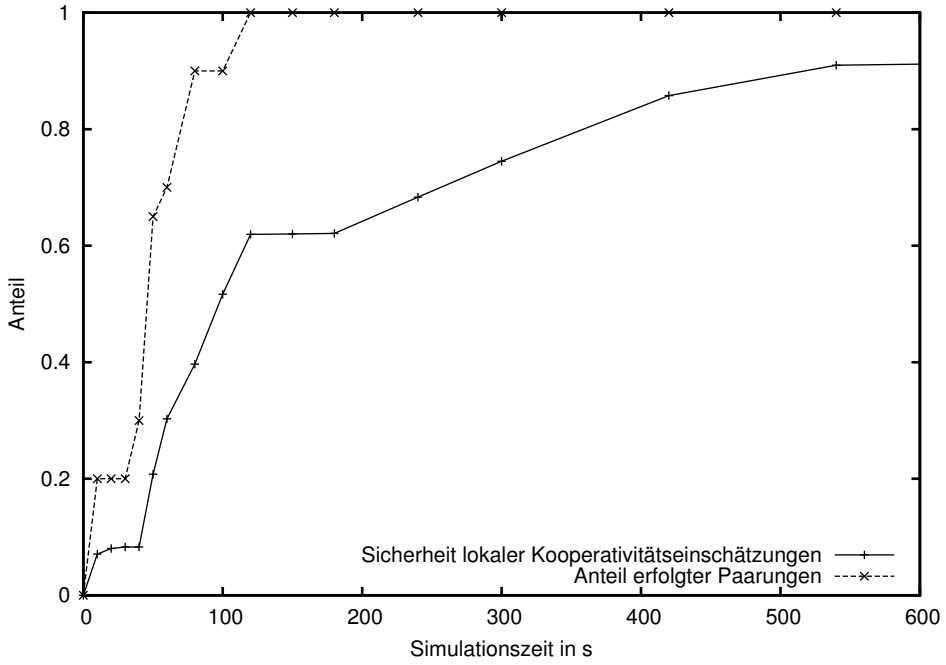


Abbildung 4.14: Durchschnittliche Sicherheit aller möglichen Einschätzungen sowie Anteil der angelegten Vertrauensprofile in einem Netz mit 5 Knoten im Verlauf der Zeit (alle einander unbekannt bei t=0)

tauschs ist auch die Belegung (nicht die Werte) der Einschätzungsmatrixbilder für die selbstermittelten Einschätzungen immer symmetrisch zur links-oben/rechts-unten-Diagonale. In der durch die ganz links in Abbildung 4.12 dargestellte Matrix beschriebenen Situation, welche 30 Sekunden nach Beginn der Simulation mit 5 einander anfänglich unbekannten Knoten vorlag, sind oder waren offenbar nur die Knoten [2] und [4] sowie [3] und [4] benachbart. Während [2] und [3] bereits Beobachtungen zu [4] gesammelt haben, ist das umgekehrt nicht der Fall; vermutlich hat [4] Nachrichten zwischen [2] und [3] weitergeleitet. Nach weiteren 30 Sekunden liegt die zur in der Mitte abgebildeten Matrix gehörige Situation vor, in der bereits Begegnungen in der Mehrzahl der zehn möglichen Paarungen stattgefunden haben und einige zusätzliche Bewertungen erfolgt sind. Nach insgesamt 120 Sekunden ist jeder Knoten jedem anderen begegnet, und es liegen – wie rechts abgebildet – in den meisten Fällen auch schon gegenseitige Einschätzungen mit nicht verschwindender Sicherheit vor. Dass sämtliche Einschätzungen positiv sind, liegt in diesem Fall daran, dass alle Knoten tatsächlich kooperativ waren und alle von ihnen geforderten Leistungen erbracht haben.

Als einheitliches Maß für den Fortschritt der Entwicklung von Einschätzungen in einem ganzen Netz kann man den Mittelwert über die Sicherheitskomponente aller möglichen Einschätzungen bilden; noch nicht existierende Einschätzungen werden dabei mit Sicherheit 0 einbezogen. Abbildung 4.14 zeigt den Anstieg dieses Wertes über die ersten 600 simulierten Sekunden in demselben Netzwerk wie oben.

Um darzustellen, wie sich unkooperatives Verhalten in der Einschätzungsmatrix niederschlägt, wurde für einen weiteren Versuch ein Netz mit 23 völlig kooperativen und zusätzlich 5 weniger kooperativen Knoten betrachtet. Die letzteren verweigerten jeweils einen bestimmten Anteil (10, 25, 50, 75 und 90 %) der von ihnen erwarteten Dienstleistungen. Abbildung 4.15 zeigt die sich ergebende Einschätzungsmatrix nach 30 bzw. 120 s. Die Positions- und Sicherheitswerte sind dort wegen der aus Platzgründen durchgeführten Verkleinerung nicht mehr gut lesbar, aber schon an der Farbverteilung ist deutlich zu erkennen, dass die teilweise unkooperativen Knoten, welche die ersten fünf Kennungen [0] bis [4] tragen, überwiegend entsprechend ihrer tatsächlichen (mangelhaften) Kooperativität eingeschätzt wurden.

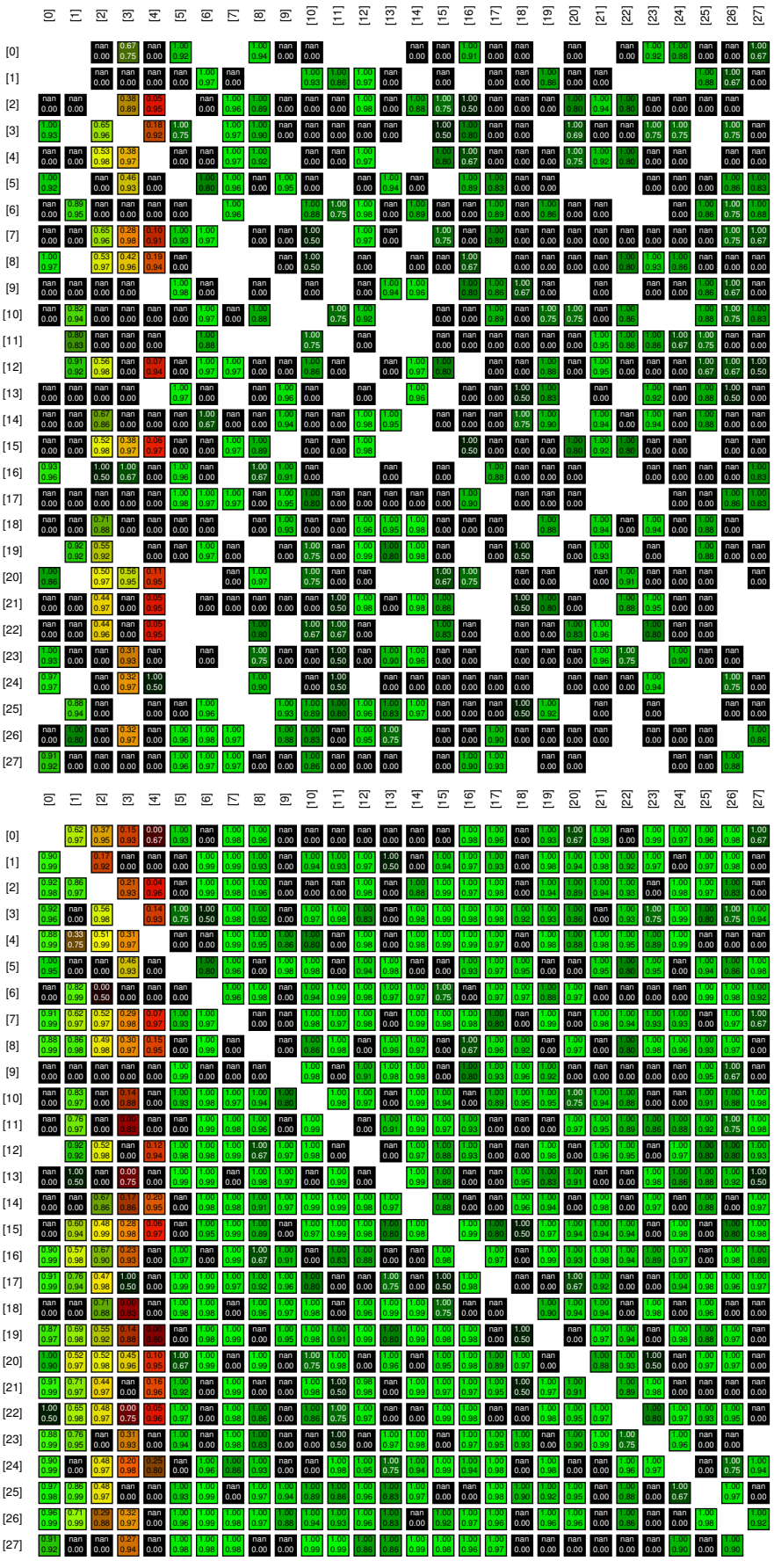


Abbildung 4.15: Einschätzungsmatrix bezüglich Kooperativität nach 30 bzw. 120 s in einem Netz mit 28 Teilnehmern, von denen die ersten 5 Anteile von 10, 25, 50, 75 und 90 % aller Leistungen verweigern

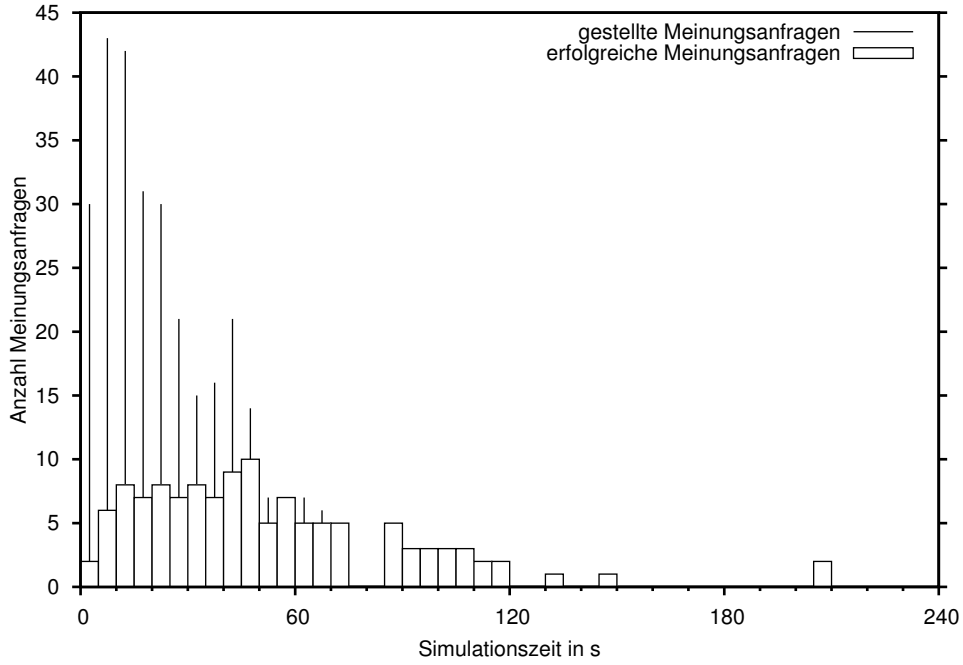


Abbildung 4.16: Anzahl gestellte (Linien) und erfolgreiche (Säulen) Meinungsanfragen jeweils innerhalb von 5-Sekunden-Intervallen

Dass nicht alle Belegungen symmetrisch sind, liegt hier übrigens daran, dass bei dem zugrundeliegenden Versuch bereits der Meinungsaustausch möglich war. Zwar beeinflussen Fremdmeinungen nicht die hier abgebildeten lokalen Einschätzungen, aber zur Speicherung der Fremdmeinungen wird ein Vertrauensprofil angelegt, wodurch das entsprechende Kästchen der Einschätzungsmatrix nicht mehr leer, sondern stattdessen schwarz ist (Sicherheit 0).

4.6 Meinungsaustausch und Gesamtvertrauen

Meinungsaustausch, Ermittlung von Empfehlungsvertrauen und die Entwicklung des Gesamtvertrauens werden hier gemeinsam betrachtet, weil sie voneinander abhängig sind: Fremdmeinungen können bei der Berechnung des resultierenden Gesamtvertrauens nur in dem Maße berücksichtigt werden, wie derjenige, der sie geäußert hat, Empfehlungsvertrauen genießt. Empfehlungsvertrauen wiederum wird mit der Zeit zu Teilnehmern entwickelt, die stets Meinungen äußern, die mit der eigenen Einschätzung übereinstimmen.

Meinungsanfragen werden immer dann gestellt, wenn eigene Einschätzungen nicht sicher genug sind, was natürlich insbesondere dann sehr häufig der Fall ist, wenn ein Netz sich neu aus einander unbekannten Teilnehmern formiert. Abbildung 4.16 zeigt, wieviele Meinungsanfragen in einem solchen Netz in jeweils aufeinanderfolgenden 5-Sekunden-Intervallen auftraten. Die Anzahl gestellter Anfragen wird jeweils als dünne Linie dargestellt, die dickeren Säulen geben an, wieviele Anfragen insofern erfolgreich waren, als sie (nach Auswertung mit Hilfe eines Widerstandnetzwerks) zur Gewinnung einer für die Zugangskontrolle ausreichend sicheren Gesamteinschätzung führten. Anfangs werden recht viele Anfragen gestellt, von denen allerdings nur ein geringer Teil zufriedenstellend beantwortet werden kann, da ja alle Knoten anfangs keine Einschätzungen über andere haben. Eine fehlgeschlagene Anfrage wird frühestens nach 2 Sekunden wiederholt, falls dann wieder Bedarf auftritt. Innerhalb der ersten 80 Sekunden des Netzbetriebs werden genügend Einschätzungen erstellt, so dass alle weiteren Meinungsanfragen erfolgreich verlaufen. Nach etwa 200 Sekunden sind die Ver-

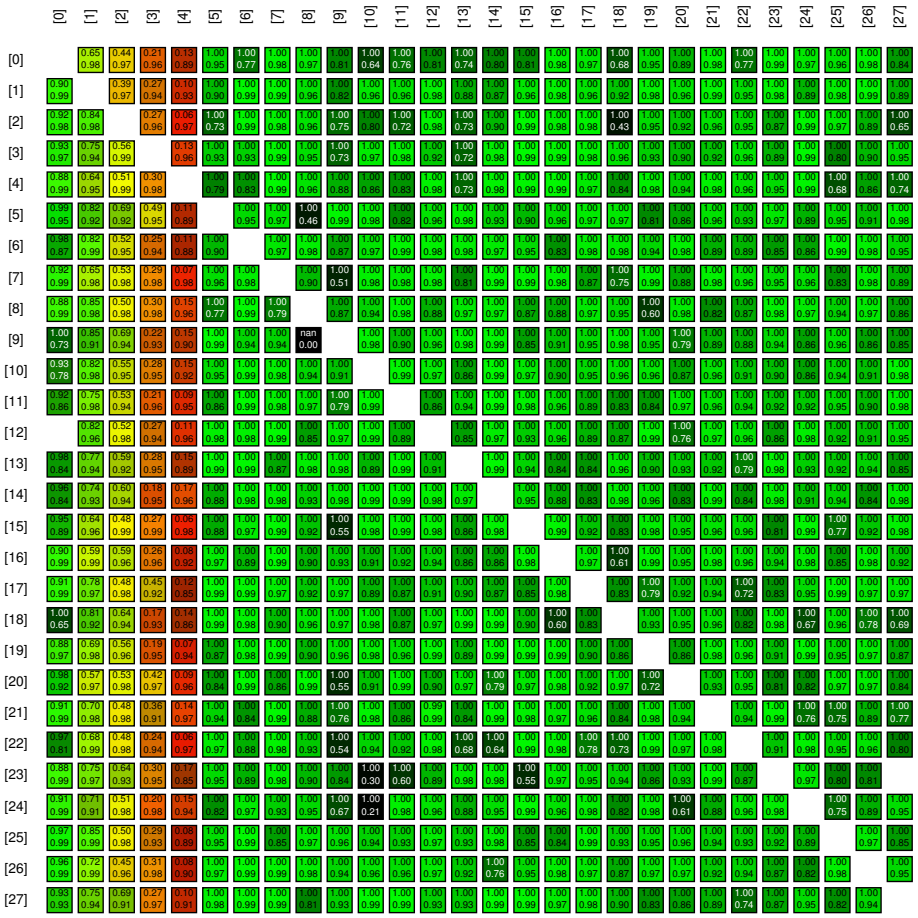


Abbildung 4.17: Gesamteinschätzung bezüglich Kooperativität nach 120 s in dem auch Abbildung 4.15 zugrundeliegenden Netzwerk

trauensprofile auf allen Knoten so gut ausgestattet, dass innerhalb der restlichen Simulationsdauer von hier 1600 Sekunden keine weiteren Anfragen mehr nötig sind.

Die aus eigenen Einschätzungen und auf Meinungsanfragen hin erhaltenen und gespeicherten Fremdmeinungen mittels des Widerstandsnetzwerksverfahrens ermittelte Gesamteinschätzung besitzt im Mittel eine höhere Sicherheit als eine nur auf eigenen Beobachtungen basierende lokale Einschätzung. In Abbildung 4.17 ist eine Einschätzungsmatrix für die Gesamteinschätzungen angegeben, die in dem auch Abbildung 4.15 zugrundeliegenden Netz nach 120 Sekunden vorlagen, also zum selben Zeitpunkt, zu dem die lokalen Einschätzungen den in Abbildung 4.15 rechts dargestellten entsprachen. Im Vergleich ist zu erkennen, dass viele Lücken in den lokalen Einschätzungen bei den Gesamteinschätzungen geschlossen sind.

Es ist zu erwarten, dass die Mobilität der Knoten Einfluss darauf hat, wie schnell Einschätzungen im Netz verteilt werden können. Deshalb wurden einige Versuche mit den bereits früher verwendeten Variationen der Mobilität in Netzen mit 25 Teilnehmern durchgeführt und dabei der Anstieg des Mittelwerts über die Sicherheitskomponente aller möglichen lokalen Einschätzungen bzw. Gesamteinschätzungen beobachtet. Die Ergebnisse sind gesammelt in Abbildung 4.18 zu finden. Auf der linken Seite werden jeweils die lokalen, selbst ermittelten Einschätzungen der Knoten bewertet, auf der rechten Seite die sich aus eigenen Einschätzungen und gespeicherten Fremdmeinungen ergebenden Gesamteinschätzungen.

Für die Ergebnisse der oberen Reihe wurde beim Random-Waypoint-Modell der Mittelwert der Bewegungsgeschwindigkeit der Teilnehmer variiert. Bei sehr niedrigen Geschwindigkeiten um 0,5 m/s

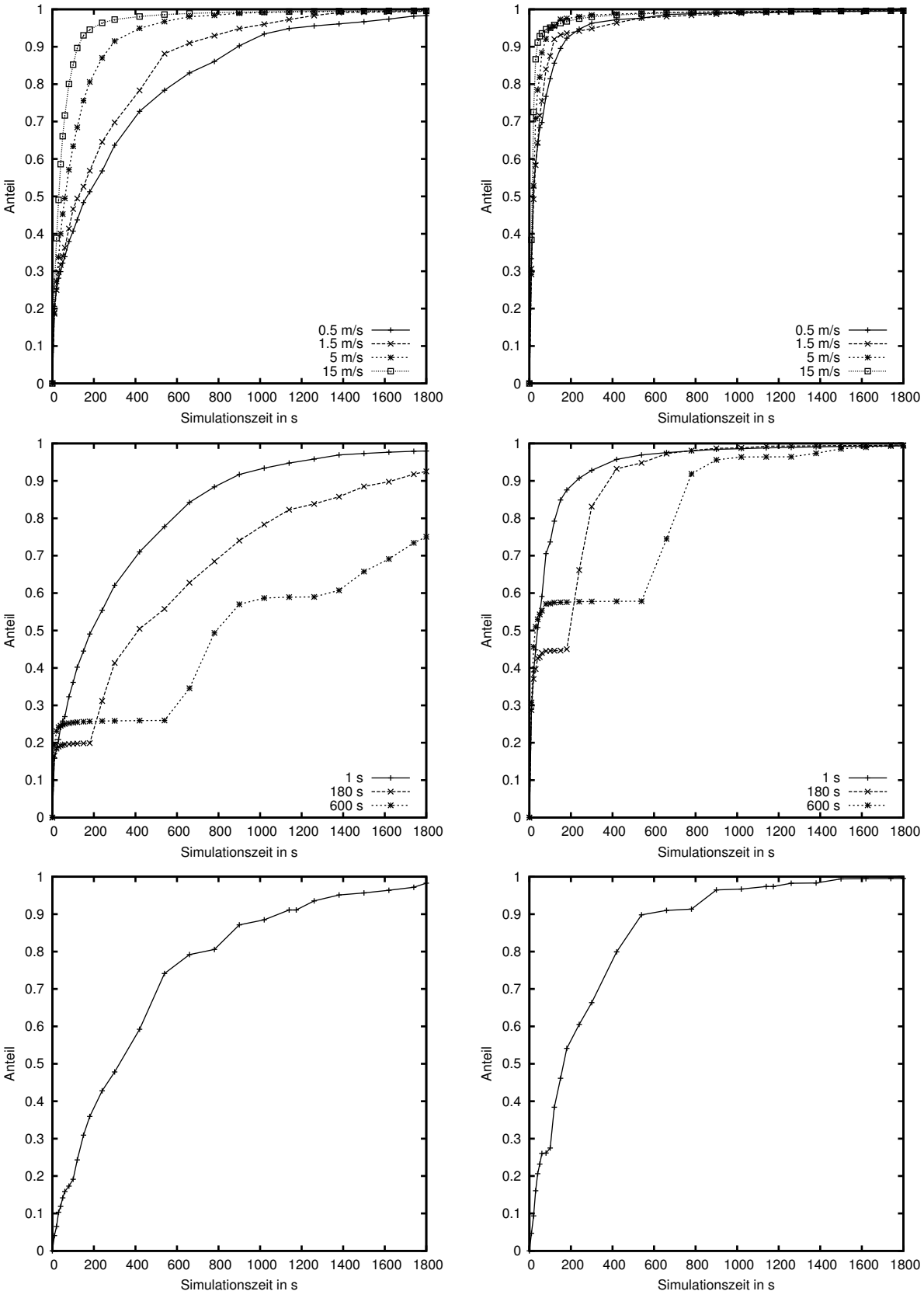


Abbildung 4.18: Verlauf der gemittelten Sicherheit der lokalen Kooperativitätseinschätzungen (links) und der Gesamtkooperativitätseinschätzungen (rechts) beim Random-Waypoint-Modell für verschiedene Bewegungsgeschwindigkeiten (Wartezeit 1 s; obere Reihe) und für verschiedene Wartezeiten (Geschwindigkeit 0.5 m/s; mittlere Reihe) sowie beim Konferenzmodell (untere Reihe)

ändern sich die Nachbarschaftsverhältnisse nur sehr langsam. Das Verhalten der anfänglichen eigenen Nachbarn kann schon nach kurzer Zeit mit großer Sicherheit eingeschätzt werden, deshalb steigt die mittlere Sicherheit der lokalen Einschätzungen anfangs steil an. Weitere Beobachtungen derselben Nachbarn erhöhen die Sicherheit dann aber nur noch unwesentlich, und weil die Beobachtung anderer Knoten erst sukzessive mit der langsamen Durchmischung der Knoten möglich wird, steigt der Sicherheitsmittelwert in der Folge auch recht langsam an. Mit höheren Geschwindigkeiten erfolgt auch die Durchmischung zunehmend schneller, und entsprechend schnell kann jeder Teilnehmer eigene Einschätzungen über andere gewinnen.

Die Möglichkeit des Meinungsaustauschs liefert – wie auf der rechten Seite zu sehen ist – für die aus eigenen Einschätzungen und Fremdmeinungen zusammengesetzte Gesamteinschätzung schon wesentlich schneller eine recht hohe mittlere Sicherheit. Der Unterschied zwischen verschiedenen Bewegungsgeschwindigkeiten ist hier wesentlich geringer als bei lokalen Einschätzungen. Hierzu ist noch anzumerken, dass die für die Versuche verwendete Implementierung insofern eingeschränkt war, als Meinungsanfragen nur an Nachbarn und nicht an entfernte Knoten gerichtet wurden. Außerdem wurden auf Meinungsanfragen hin nur eigene Einschätzungen weitergegeben, keine gespeicherten Fremdmeinungen. Das Hinzufügen dieser Möglichkeiten dürfte zu einer weiteren Beschleunigung der Verbreitung von ausreichend sicheren Meinungen im Netz und damit einem schnelleren Anstieg der mittleren Sicherheit der Gesamteinschätzungen führen.

Den Ergebnissen der mittleren Reihe liegt ebenfalls das Random-Waypoint-Bewegungsmodell zugrunde, wobei die Bewegungsgeschwindigkeit auf 0,5 m/s festgelegt war und die Wartezeit variiert wurde. Da alle Teilnehmer zu Beginn des Simulationslaufs zuerst die Wartephase absolvierten, sind deutliche Stufungen im Ansteigen der mittleren Sicherheitswerte festzustellen: Wenn die Einschätzungen der Nachbarn so sicher geworden sind, dass weitere Beobachtungen momentan keinen merklichen Vorteil mehr bringen, stagniert der Anstieg so lange, bis die Teilnehmer sich wieder bewegen und neue Nachbarschaftsbeziehungen zustande kommen.

Die Ergebnisse der unteren Reihe kamen bei Bewegung nach dem Konferenzmodell zustande. Einerseits ist die Bewegungsgeschwindigkeit dort recht gering, weshalb die Entwicklung der mittleren Sicherheit lokaler Einschätzungen derjenigen beim Random-Waypoint-Modell mit geringer Geschwindigkeit ähnelt. Andererseits bewirkt die weniger ausgeprägte Durchmischung der sich interessenorientiert bewegenden Teilnehmer anscheinend auch, dass fehlende Einschätzungen auch durch Meinungsanfragen bei Nachbarn häufig nicht erhalten werden können, so dass die mittlere Sicherheit der Gesamteinschätzungen ebenfalls recht langsam ansteigt. Hier könnte die Durchführung von Meinungsanfragen bei nicht benachbarten Knoten und die Weitergabe von Fremdmeinungen Abhilfe schaffen; dies muss noch näher untersucht werden.

4.7 Bootstrapping

In Abschnitt 3.9.4 wurde der Bootstrapping-Mechanismus beschrieben, der insbesondere in neu zusammengestellten Netzen die anfängliche Gewinnung von Einschätzungen erleichtern soll, indem er die Zugangskontrolle vorübergehend abschaltet. Dies ist sinnvoll, weil ansonsten in einer Situation, in der keinerlei Einschätzungen vorliegen, wegen Scheiterns der Zugangskontrolle keine Leistungen erbracht werden und damit auch keine Beobachtungen getätigt und keine Einschätzungen gewonnen werden können. Wichtig ist, dass der Bootstrapping-Mechanismus die Zugangskontrolle nicht unterläuft, sondern nur genau in den vorgesehenen Situationen aktiviert und nach kurzer Zeit wieder beendet wird.

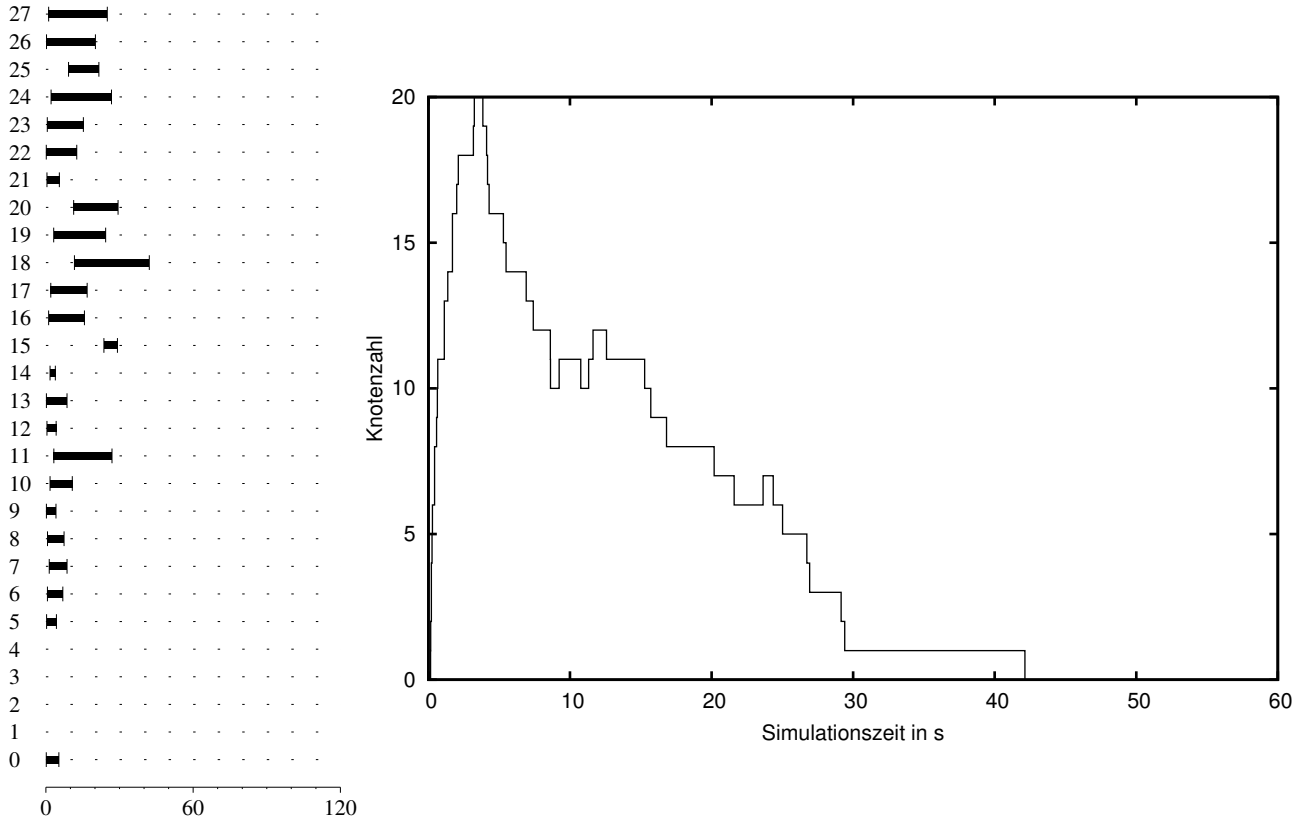


Abbildung 4.19: Im Bootstrapping-Modus verbrachte Zeitintervalle aller Teilnehmerknoten (links) sowie Anzahl von Knoten im Bootstrapping-Modus (rechts) des auch Abbildung 4.15 zugrundeliegenden Netzwerks

Bei den in den vorigen Abschnitten beschriebenen Versuchen zur Meinungsbildung war der Bootstrapping-Mechanismus anfangs ebenfalls aktiv. In Abbildung 4.19 sind links für jeden Teilnehmer des auch schon Abbildung 4.15 zugrundeliegenden Netzwerks die Intervalle innerhalb des Simulationszeitraums gekennzeichnet, in denen der Bootstrapping-Modus aktiviert war. Auf der rechten Seite ist die Anzahl der jeweils im Bootstrapping-Modus befindlichen Knoten über der Zeit dargestellt. Man erkennt, dass der Bootstrapping-Modus jeweils nur für kurze Zeit und kurz nach der Formierung des Netzwerks aktiv war; im nicht dargestellten Teil des insgesamt 1800 s umfassenden Simulationszeitraums war er stets inaktiv.

Versuche mit den verschiedenen bzw. verschieden parametrisierten Mobilitätsmodellen haben gezeigt, dass der Bootstrapping-Mechanismus immer wie gewünscht funktioniert (siehe dazu Abbildungen A.5 und A.6 im Anhang). Bei niedrigeren Bewegungsgeschwindigkeiten bleiben die Knoten im Mittel etwas länger im Bootstrapping-Modus als bei höheren. Lange Aufenthaltszeiten führen manchmal dazu, dass einzelne Knoten längere Zeit in einer Position bleiben, in der sie nicht weiterleiten können, und wenn sie kurz vorher in den Bootstrapping-Modus gewechselt haben, bleibt dieser Zustand lange erhalten. Dies stellt kein Problem dar, da der Zustand ohne Weiterleitungsmöglichkeit auch nicht ausgenutzt werden kann.

4.8 Zugangskontrolle

Besonders wichtig bei der Evaluierung des gesamten Zugangskontrollsystems, das auf den in den vorangegangenen Abschnitten untersuchten Mechanismen aufbaut, sind die beiden folgenden Fragestellungen:

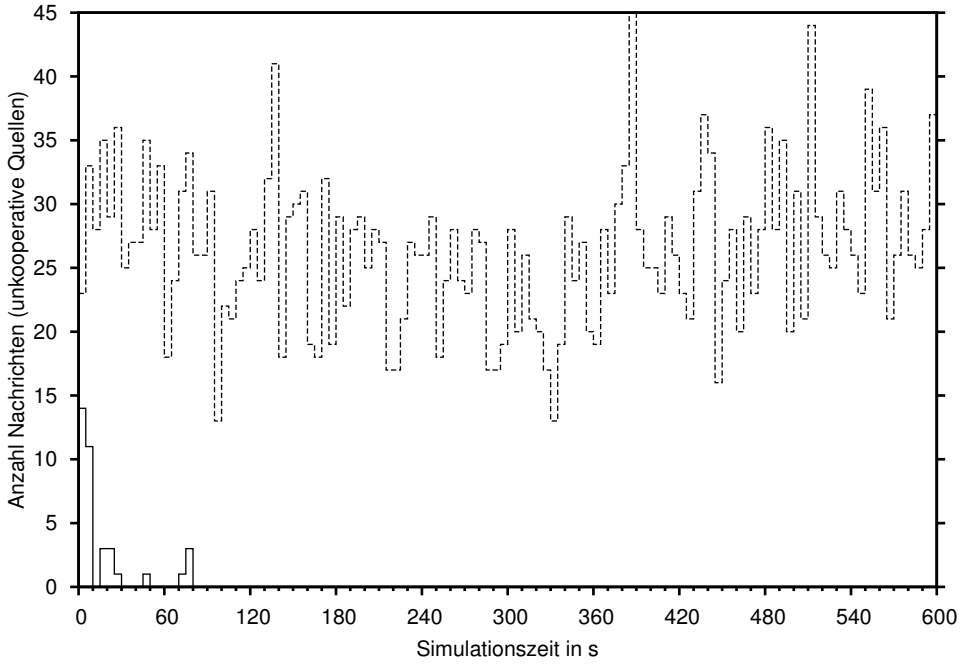


Abbildung 4.20: Anzahlen abgesandter (gestrichelt) und angekommener (durchgezogen) weiterzuleitender Nachrichten aus unkooperativen Quellen innerhalb aufeinanderfolgender 5-Sekunden-Intervallen in einem Netz mit 23 kooperativen und 6 unkooperativen Teilnehmern

- Werden unkooperative Teilnehmer, die selbst keine Leistungen erbringen wollen, wirksam von der Benutzung des Netzes und damit der Ausnutzung der Leistungsbereitschaft anderer Teilnehmer abgehalten?
- Inwiefern beeinträchtigt und belastet das Zugangskontrollsystem die kooperativen Teilnehmer?

Um eine Antwort auf die erste Frage zu finden, wurde untersucht, wie viele der von unkooperativen Teilnehmern abgesandten Nachrichten tatsächlich ihr Ziel erreichen. Betrachtet werden dabei nur Nachrichten, die tatsächlich auch weitergeleitet werden müssen, weil sie an ein zur Quelle nicht benachbartes Ziel adressiert sind. Dass Nachrichten an Nachbarn ihr Ziel erreichen, kann durch das Zugangskontrollsystem weder verhindert werden, noch wäre dies erforderlich, da ja keine Leistung durch andere als den Absender erbracht werden muss.

Abbildung 4.20 basiert auf einem Netz mit 23 kooperativen und 6 vollständig unkooperativen Teilnehmern und zeigt zunächst als gestrichelte Linie die jeweilige Anzahl der von unkooperativen Teilnehmern abgesandten weiterzuleitenden Nachrichten in aufeinanderfolgenden 5-Sekunden-Intervallen, beginnend bei der Entstehung des Netzes aus einander unbekannten Teilnehmern. Als durchgezogene Linie ist die jeweilige Anzahl tatsächlich am Ziel angekommener Nachrichten eingezeichnet. Man sieht, dass direkt nach Entstehung des Netzes einige Nachrichten ihr Ziel erreichen, was darauf zurückzuführen ist, dass sich viele Knoten im Bootstrapping-Modus befinden und keine Zugangskontrolle anwenden. Die letzte nicht kontrollierte Nachricht eines unkooperativen Teilnehmers wurde im Beispielnetz nach knapp 78 s weitergeleitet; alle weiteren Nachrichten innerhalb der Simulationsdauer von 1800 s (dargestellt sind nur die ersten 900 s) wurden durch die Zugangskontrolle abgefangen. Versuche mit höheren Anzahlen unkooperativer Knoten bestätigten die Vermutung, dass der Anteil der anfänglich durchgelassenen Nachrichten und die zeitliche Dauer der Durchlässigkeitsphase nicht von der Anzahl unkooperativer Knoten abhängen (siehe Abbildung A.7 im Anhang).

Für Abbildung 4.21 wurden Netze mit je 23 kooperativen und 2 völlig unkooperativen Teilnehmern verwendet, die sich nach dem – unterschiedlich parametrisierten – Random-Waypoint-Modell beweg-

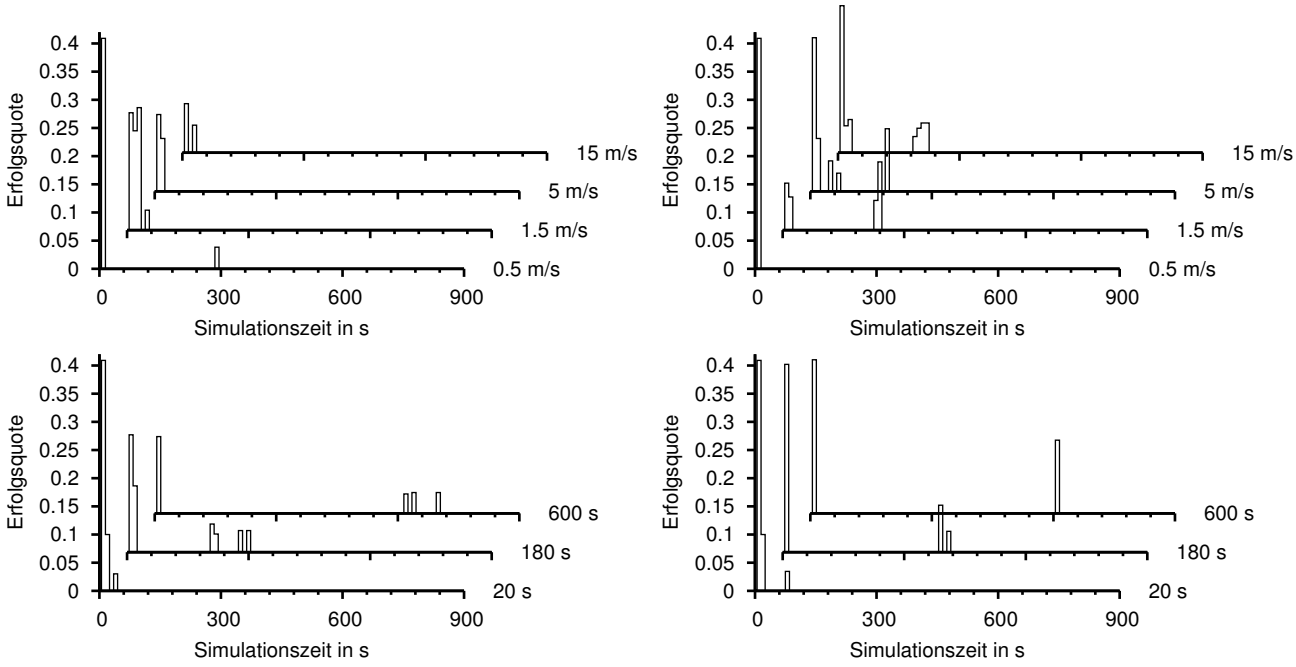


Abbildung 4.21: Weiterleitungs-Erfolgsquote unkooperativer Teilnehmer bezogen auf aufeinanderfolgende 10-Sekunden-Intervalle in Netzen mit unterschiedlich mobilen Teilnehmern.
Obere Reihe: Geschwindigkeit 0,5 / 1,5 / 5 / 15 m/s, Wartezeit 1 s (links) bzw. 180 s (rechts).
Untere Reihe: Geschwindigkeit 0,5 m/s (links) bzw. 9 m/s (rechts), Wartezeit jeweils 20 / 180 / 600 s.

ten. Dargestellt ist die Erfolgsquote bei der Übertragung von Nachrichten unkooperativer Teilnehmer zu nicht benachbarten Zielen, also der Quotient aus den Anzahlen angekommener weitergeleiteter und abgesandter weiterzuleitender Nachrichten, wobei nur Nachrichten betrachtet wurden, die nicht aus anderem Grund (kein Weg zum Ziel) verworfen wurden. Die Zählung erfolgte innerhalb von Intervallen von 10 s Dauer. In allen Fällen ist die Erfolgsquote insgesamt sehr gering, wobei jeweils zu Beginn der Simulationszeit, direkt nach Aufbau des Netzwerks, einige Nachrichten ihr Ziel erreichen, weil das Bootstrapping aktiv ist. Simuliert wurden 30 Minuten; jenseits der in der Abbildung dargestellten ersten 900 Sekunden war die Erfolgsquote unkooperativer Teilnehmer jeweils gleich null.

Für die Serie links oben wurden bei vernachlässigbar kurzer Aufenthaltszeit vier unterschiedliche Bewegungsgeschwindigkeiten vorgegeben (0,5 / 1,5 / 5 / 15 m/s). Lediglich bei der niedrigsten Geschwindigkeit waren zwei Teilnehmer längere Zeit in einer Position, in der sie nicht weiterleiten konnten und begannen ihre Bootstrapping-Phase deshalb erst später; so konnte nach etwa 300 s eine Nachricht eines unkooperativen Knotens passieren. In den drei übrigen Serien spielt die jeweilige längere Aufenthaltszeit eine Rolle: Rechts oben war sie um 180 s lang (bei wie oben variierten Geschwindigkeiten), in der unteren Reihe wurden jeweils Aufenthaltszeiten um 20, 180 und 600 s verwendet (bei Geschwindigkeiten um 0,5 (links) bzw. 9 m/s (rechts)). Längere Aufenthaltszeiten begünstigen – insbesondere, wenn sie wie hier bei allen Knoten zum Simulationsbeginn gleichzeitig stattfinden – Situationen, in denen einzelne Teilnehmer sich von Anfang an längere Zeit in Positionen aufhalten, wo sie nicht weiterleiten können. Erst wenn sie diese verlassen, beginnen sie eine Bootstrapping-Phase, und dann leiten sie für kurze Zeit Pakete unkooperativer Teilnehmer weiter.

Mit einer so geringen Anzahl von Nachrichten, die zu nicht genau vorhersehbaren Zeiten jeweils einen einzelnen Weiterleitungsschritt weit über einen anderen Knoten übertragen werden können, kann keine sinnvolle Kommunikationsbeziehung unterhalten und das Gesamtnetz nicht wesentlich belastet werden. Für die betrachteten Szenarien ist die Zugangskontrolle damit insgesamt als sehr wirksam zu werten.

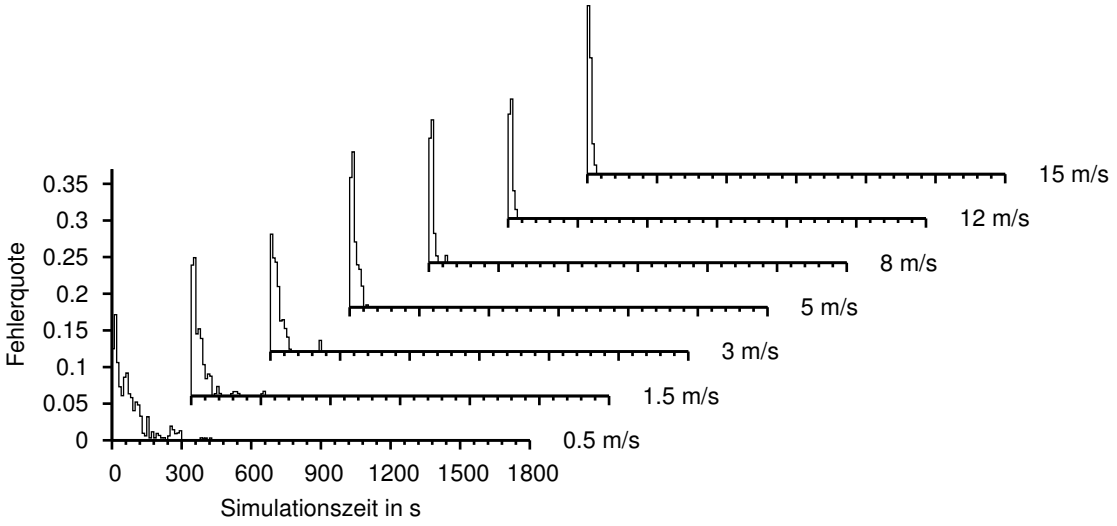


Abbildung 4.22: Zugangsfehlerquote in 10-Sekunden-Intervallen bei 25 Teilnehmern und Random-Waypoint mit Wartezeit 1 s und verschiedenen Geschwindigkeiten

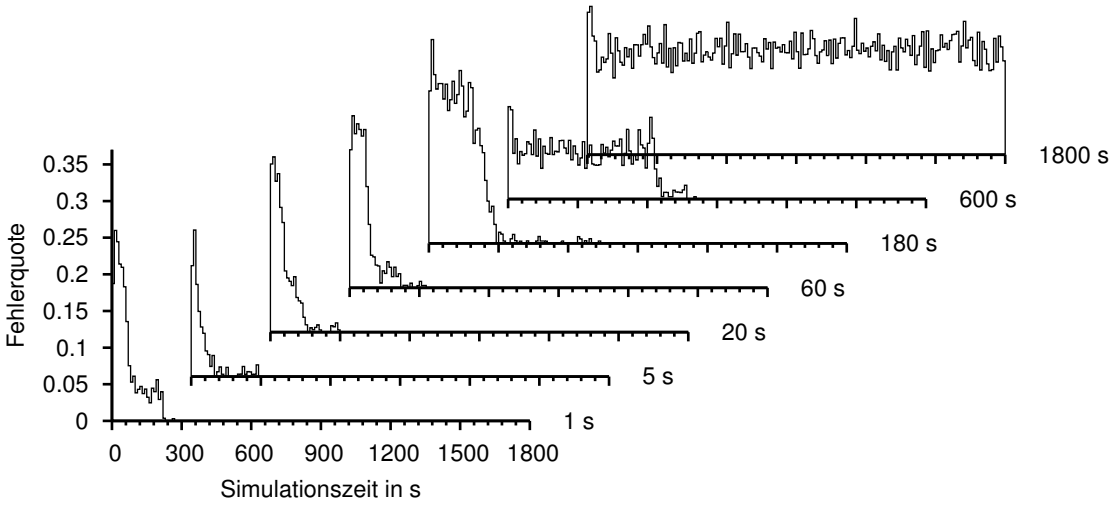


Abbildung 4.23: Zugangsfehlerquote in 10-Sekunden-Intervallen bei 25 Teilnehmern und Random-Waypoint mit Geschwindigkeit $0,5$ m/s und verschiedenen Wartezeiten

Zur Klärung der zweiten wichtigen Frage, derjenigen nach der Beeinträchtigung kooperativer Teilnehmer, wurden Versuche in Netzen mit je 25 völlig kooperativen Teilnehmern durchgeführt. Betrachtet wird die Zugangsfehlerquote, also der Quotient aus Anzahl der Zugangsfehler und Anzahl der zu übertragenden Nachrichten, wieder jeweils bezogen auf 10-Sekunden-Intervalle. Idealerweise sollten in Netzen mit ausschließlich kooperativen Knoten natürlich keine Zugangsfehler auftreten. Dem in Abbildung 4.22 dargestellten Ergebnis liegt Bewegung nach dem Random-Waypoint-Modell mit sehr kurzen Aufenthaltszeiten (um 1 s) zugrunde. Hier treten jeweils kurz nach Formierung des Netzes zu Beginn der Simulationszeit einige Verluste auf, die aber relativ schnell auf ein tolerierbares Maß zurückgehen und nach kurzer Zeit ganz verschwinden. Sie werden dadurch verursacht, dass ausreichend positive Einschätzungen noch nicht weit genug im Netz verbreitet sind. Bei der niedrigsten Geschwindigkeit von $0,5$ m/s dauert dies bis etwa 5 Minuten, wobei aber auch die nach einer Minute erreichte Fehlerquote von 5% schon eine sinnvolle Nutzung des Netzes ermöglicht. Bei höheren Geschwindigkeiten verschwinden die Zugangsfehler größtenteils bereits innerhalb der ersten Minute.

Um einiges ungünstiger wirken sich längere Aufenthaltszeiten aus, insbesondere wenn sie gleichzeitig von allen Teilnehmern begonnen werden; dies war ja auch schon bei Untersuchung der Mei-

nungsbildung festgestellt worden. Durch lange Aufenthalte wird eine Durchmischung der Teilnehmer verzögert, und Meinungen verbreiten sich langsamer. Für Abbildung 4.23 wurden bei sehr niedriger Bewegungsgeschwindigkeit (um $0,5$ m/s) verschieden lange Aufenthaltszeiten verwendet. Es ist deutlich zu erkennen, dass die Zugangsfehlerquote von ihrem anfänglichen hohen Niveau (um 20%) erst abfällt, nachdem die initiale Wartezeit verstrichen ist. Das Anfangsniveau ist unterschiedlich hoch und hängt von der zufälligen Verteilung der Teilnehmer im Simulationsgebiet ab.

Es ist zu erwarten, dass die Mechanismen des Meinungsaustauschs mit entfernten Teilnehmern sowie der Weitergabe gespeicherter Fremdmeinungen, die den Modellteilnehmern hier nicht zur Verfügung standen, geeignet sind, Meinungen wesentlich schneller im ganzen Netz zu verbreiten und damit die anfängliche Verlustphase stark zu verkürzen. Außerdem muss angemerkt werden, dass die Situation des gleichzeitigen Netzeintritts von 25 zunächst unbeweglichen Teilnehmern, die einander unbekannt sind, aber trotzdem sofort mit jedem anderen Teilnehmer kommunizieren möchten, nicht unbedingt einem erwarteten Nutzungsszenario entspricht.

4.9 Zusammenfassung und Bewertung

Durch die Implementierung des in der vorliegenden Arbeit entwickelten Konzepts in einer Simulationsumgebung konnte die Funktionsfähigkeit des Ansatzes überprüft und die Leistungsfähigkeit der einzelnen Bestandteile sowie des Gesamtkonzepts untersucht werden.

Bei der Implementierung der Netzknoten wurden in den nicht die Zugangskontrolle betreffenden Schichten und Komponenten (Funkübertragung, Medienzugriff, Wegfindung) idealisierte Verfahren eingesetzt, um eine Beeinflussung der Ergebnisse durch spezielle Eigenschaften konkreter Verfahren zu vermeiden. Bei der Übertragung der Ergebnisse auf reale Netze müssen Nachteile der dann tatsächlich verwendeten Verfahren gegenüber den idealisierten einberechnet werden.

Die Implementierung selbst erfolgte in der Programmiersprache C++. Der die Zugangskontrolle betreffende Teil ist im Wesentlichen von der Simulationsumgebung unabhängig, d. h. er verwendet keine besonderen Funktionen der Umgebung. Eine Umsetzung dieses Teil von der Simulationsumgebung auf ein reales Betriebssystem dürfte deshalb mit relativ geringem Aufwand machbar sein, wenn das Betriebssystem die benötigten Schnittstellen anbietet. Unproblematisch ist dies sicher bezüglich der erforderlichen Möglichkeiten zum Absenden, Empfangen und Mithören von Nachrichten sowie zur Bestimmung der Systemzeit und zum Warten auf bestimmte Zeitpunkte; diese Schnittstellen bieten gängige Betriebssysteme in ähnlicher Weise an wie die Simulationsumgebung. Etwas mehr Aufwand müsste wohl für die benötigten Eingriffsmöglichkeiten in den Weiterleitungsprozess betrieben werden: Jedes weiterzuleitende Paket muss zunächst dem Zugangskontrollsystem präsentiert werden und anschließend entweder verworfen, zeitweise zurückgehalten oder regulär weitergeleitet werden. Unter einem der quelloffenen Unix-Derivate (Linux, BSD-Varianten) wäre sicher auch diese Betriebssystemänderung ohne Weiteres realisierbar.

Für die Simulation der Teilnehmermobilität wurden zwei unterschiedliche Bewegungsmodelle verwendet: einerseits das verbreitete und damit auch eine gewisse Vergleichbarkeit gewährleistende Random-Waypoint-Modell, das eine gleichmäßig über das Simulationsgebiet verteilte Aufenthaltswahrscheinlichkeit der Teilnehmer mit nur geringen Schwankungen bewirkt und dadurch relativ „glatte“ Ergebnisse liefert, und andererseits ein Konferenz-Modell, mit welchem versucht wird, das Bewegungsmuster von Teilnehmern einer Konferenz oder Besuchern einer Messe nachzuempfinden; es ist vermutlich realitätsnäher, verursacht aber eine recht inhomogene und von der jeweils erzeugten „Konferenz“ abhängige Verteilung der Teilnehmer über das Simulationsgebiet und erfordert deshalb mehr Aufwand zur Gewinnung statistisch verlässlicher Messergebnisse. Beim Random-Waypoint-Modell

wurden die Zufallsverteilungen für Geschwindigkeiten und Aufenthaltszeiten variiert, während beim Konferenzmodell ein fester Parametersatz verwendet wurde, weil der Raum der Möglichkeiten sonst zu groß geworden wäre. Bezüglich der Teilnehmerdichte wurde ermittelt, dass bei etwa 25 Teilnehmern auf einem Simulationsgebiet von $600 \times 600\,\mathrm{m}^2$ ein sinnvoller Netzbetrieb ohne allzu häufige Verluste durch Isolation von Teilnehmern möglich ist. Diese Parametrisierung wurde deshalb in der Folge häufig eingesetzt, sofern nicht ein möglicher Einfluss der Teilnehmerdichte überprüft werden sollte.

Bezüglich des Verfahrens zur Beobachtung und Bewertung des Weiterleitungsverhaltens benachbarter Netzteilnehmer wurde untersucht, bei wie vielen der benachbarten Beobachter eines Weiterleitungsvorgangs die aus Sicherheitsgründen geforderte Zusatzbedingung für die Bewertung erfüllt ist, welche vorschreibt, dass auch der jeweils vorangegangene Weiterleitungsschritt bereits beobachtet worden sein muss. Es zeigte sich, dass im Mittel pro weitergeleiteter Nachricht etwa halb so viele Beobachtungen verwendbar sind wie Nachbarn anwesend; bei sehr geringer Teilnehmerdichte liegt der Anteil sogar etwas höher. Bei 25 Teilnehmern und homogener Aufenthaltswahrscheinlichkeit ergaben sich beispielsweise durchschnittlich $3,6$ Bewertungen pro Nachricht bei durchschnittlich $6,8$ anwesenden Nachbarn. Ein Einfluss der Mobilität auf diesen Zusammenhang war nicht festzustellen.

Die Gewinnung lokaler Einschätzungen anhand der Weiterleitungsbeobachtung funktioniert recht gut, sowohl für kooperative sowie auch für unkooperative Knoten: Die ermittelten Einschätzungen nähern sich schnell der tatsächlichen Verhaltensvorgabe an. Anhand der Untersuchung des Anstiegs der mittleren Sicherheit aller möglichen Einschätzungen konnte allerdings festgestellt werden, dass längere Aufenthaltszeiten der Knoten zu einer Stagnation im Anstieg führen, was auch naheliegt, da nach der weitgehenden „Erforschung“ einer gleichbleibenden Gruppe von Nachbarn kaum noch ein Informationsgewinn möglich ist. Ein negativer Einfluss durch eine relativ kurze Dauer von Nachbarschaftsverhältnissen (die kürzeste sich aus den verwendeten Parametrisierungen ergebende mittlere Nachbarschaftsdauer lag bei etwa 15 s) auf die Gewinnung von Einschätzungen über Nachbarn konnte nicht nachgewiesen werden.

Die zusätzliche Verwendung des Meinungsaustauschs bewirkt, dass die aus eigenen Einschätzungen und Fremdmeinungen per Widerstandsnetzwerkverfahren ermittelten Gesamteinschätzungen noch schneller an Sicherheit gewinnen als aufgrund von Beobachtungen ermittelte lokale Einschätzungen. Insbesondere erhalten die einzelnen Teilnehmer durch den Meinungsaustausch schnell ausreichend sichere Gesamteinschätzungen zu einem wesentlich größeren Anteil der übrigen Knoten als ohne Meinungsaustausch. Auch hier zeigte sich allerdings, dass lange, insbesondere von allen Knoten gleichzeitig verbrachte Aufenthaltszeiten zu einer Stagnation des Anstiegs der mittleren Sicherheit aller möglichen Gesamteinschätzungen führten. Der Grund dafür ist hauptsächlich darin zu suchen, dass der Meinungsaustausch bei der vorliegenden Implementierung auf Nachbarn und auf eigene Einschätzungen beschränkt war, d. h. es wurden keine entfernten Teilnehmer um Meinungsäußerung gebeten und auf Anfragen hin keine Fremdmeinungen weitergegeben.

Diese Schwäche der verwendeten Implementierung äußerte sich auch bei der Untersuchung der Zugangskontrolle. Hier traten in Situationen mit stark eingeschränkter Mobilität durch die Zugangskontrolle relativ hohe Verluste an Nachrichten sich eigentlich völlig kooperativ verhaltender Teilnehmer auf, weil sich die von den Nachbarn der betroffenen Teilnehmer ermittelten positiven Einschätzungen nicht ausreichend im Netz verbreiten konnten. Bei guter Durchmischung der Teilnehmer aufgrund gleichmäßiger Bewegung dagegen verschwanden die in neu aus einander unbekannten Teilnehmern zusammengesetzten Netzen anfänglich auftretenden Verluste sehr schnell.

In allen Fällen wurden Nachrichten sich unkooperativ verhaltender Teilnehmer schon nach kurzer Zeit, d. h. nachdem die meisten Teilnehmer ihre Bootstrapping-Phase abgeschlossen hatten, zuverlässig abgelehnt. Der Zweck der Zugangskontrolle wurde also im vollem Maße erfüllt.

Kapitel 5

Zusammenfassung und Ausblick

Neben mobilen Kleinstrechnern zunehmender Leistungsfähigkeit, die auf Mobiltelefonen und den so genannten persönlichen digitalen Assistenten (PDAs) schon seit einiger Zeit im Umlauf sind, gewinnen auch mobile Allzweckrechner, die mittlerweile meist schon im Auslieferungszustand mit drahtlosen Kommunikationsschnittstellen ausgestattet sind, stark an Verbreitung. Gefördert wird der Trend zur drahtlosen Netzanbindung auch dadurch, dass DSL-Modems (Digital Subscriber Line), die zur Nutzung der beliebten und mittlerweile auch recht preisgünstig verfügbaren breitbandigen Netzzugänge im Heimbereich erforderlich sind, häufig bereits mit drahtlosen Kommunikationsschnittstellen ausgestattet sind, um so eine aufwändige Verkabelung zu vermeiden. Durch die Verbreitung mobiler Rechner mit drahtloser Kommunikationstechnik steigt die Dichte potentieller Teilnehmer an Ad-hoc-Netzen, und es wird dadurch wahrscheinlich, dass solche Netze, die bisher hauptsächlich zu Forschungszwecken aufgebaut werden, in Zukunft tatsächlich in vielen Situationen anstelle infrastrukturbasierter Netze zum Austausch von Daten, zum kooperativen Arbeiten, gemeinsamen Spielen oder für andere Formen der Kommunikation genutzt werden können.

Mit dem Übergang offener Ad-hoc-Netze aus dem experimentellen Stadium in die Nutzung für alltägliche Zwecke treten für die Nutzer ihre jeweiligen Interessen in den Vordergrund und verdrängen dabei das Bewusstsein für die – weiterhin bestehende – Notwendigkeit, das Netzwerk für alle Teilnehmer in fairer Weise verfügbar zu halten. Der Missbrauch der sozusagen durch alle Teilnehmer gemeinsam aufgebauten mobilen Infrastruktur durch Einzelne muss dann sinnvollerweise durch ein Zugangskontrollsystem verhindert werden. Anders formuliert motiviert ein Zugangskontrollsystem die Teilnehmer dazu, Leistungen für die Allgemeinheit zu erbringen, die für den Erhalt der Verfügbarkeit des Netzwerks unverzichtbar sind; die Motivation entsteht daraus, dass unter der Kontrolle des Zugangskontrollsystems nur durch eigene Mitarbeit das Recht zur Nutzung des Netzes erwirkt werden kann.

Eine besondere Herausforderung bei der Entwicklung eines Zugangskontrollsystems für offene Ad-hoc-Netze entsteht dadurch, dass keine zentrale Kontroll- oder Verwaltungsinstanz installiert werden kann, sondern ein vollkommen verteilt arbeitendes Verfahren gefunden werden muss, bei dem alle Teilnehmer gleichberechtigt sind. Weiterhin darf nicht vorausgesetzt werden, dass Teilnehmer sich an durch das Verfahren vorgegebene Regeln halten, sondern das Verfahren muss insbesondere dann wirksam sein, wenn einzelne Teilnehmer dies nicht tun. Unter allen möglichen, von den Teilnehmer frei wählbaren Verhaltensweisen muss sozusagen die einzig erfolgversprechende und erfolgreiche diejenige sein, die durch das Zugangskontrollverfahren vorgegeben ist. Die theoretische Möglichkeit, richtlinientreues Verhalten der Teilnehmer etwa durch die Voraussetzung manipulationssicherer richtlinienientreuer Komponenten in allen Geräten zu erzwingen wird hierbei als unrealisierbar ausgeschlossen.

Aus der Offenheit der betrachteten Netze ergibt sich außerdem eine wichtige Anforderung bezüglich der Zulässigkeit von Annahmen über Teilnehmeridentitäten: Da Teilnehmeridentitäten auf maschineller Ebene sehr leicht neu erzeugt werden können und einzelne Teilnehmer auch durchaus unbemerkt jeweils mehrere Identitäten gleichzeitig verwenden können, darf niemals allein aufgrund von Anzahlen gleich lautender Meinungen oder Voten geschlossen werden, dass ein Konsens mehrerer oder gar aller realer Teilnehmer vorliege. Auch der Versuch, maschinelle Identitäten etwa mittels gegenseitiger Zertifizierung an Identitäten menschlicher Benutzer zu binden, bietet ohne zentrale Kontrollautorität keine Sicherheit. Als Konsequenz daraus müssen insbesondere sicherheitskritische Verfahren für offene Ad-hoc-Netze so entworfen sein, dass sich durch Mehrfachidentitäten keine Vorteile für deren Inhaber ergeben.

5.1 Ergebnisse

In der vorliegenden Arbeit wurde ein Konzept entwickelt, welches die genannte Zielsetzung erfüllt, indem es ein vollständig verteilt arbeitendes Zugangskontrollsystem für offene Ad-hoc-Netze zur Verfügung stellt, in dem alle Teilnehmer gleichberechtigt sind und kooperatives Verhalten motiviert wird, ohne dass generelle Richtlinien dafür künstlich von außen vorgegeben werden.

Eine gleichberechtigende Verteilung des Zugangskontrollsystems auf alle Teilnehmer wird erreicht, indem jede weiterzuleitende Nachricht auf jedem Zwischensystem kontrolliert und nur dann weitergeleitet wird, wenn der jeweils kontrollierende Knoten die Quelle der Nachricht als zur Mitarbeit bereiten Teilnehmer einschätzt. Ein Einhalten einer einheitliche Linie bezüglich der Zugangskontrollpraxis wird dadurch erwirkt, dass die Nachbarn des Kontrollierenden seine Zugangsentscheidungen beobachten und dabei gegebenenfalls ihre Einschätzungen bezüglich des Beobachteten anpassen; wenn ein Teilnehmer also bei der Zugangskontrolle beispielsweise willkürlich restriktiv handelt, muss er damit rechnen, später ebenso behandelt zu werden.

Die Beobachtung des Verhaltens benachbarter Teilnehmer ist auch ansonsten das Mittel der Wahl, um Einschätzungen bezüglich der Kooperativität anderer Teilnehmer zu gewinnen, also bezüglich ihrer Bereitschaft, Leistungen zugunsten der Allgemeinheit zu erbringen. Die dabei ermittelten Einschätzungen bilden die Grundlage für Zugangsentscheidungen. Wichtig ist, dass – im Gegensatz zum Vorgehen bei vielen anderen Ansätzen – Beobachtungen richtigen Verhaltens ebenso registriert werden wie solche falschen Verhaltens. Nur so kann ein möglichst genaues Bild vom Verhalten anderer Teilnehmer gewonnen werden. Ein besonderer Schwerpunkt liegt auf der Beobachtung des Verhaltens bei der Weiterleitung von Paketen anderer. Hierbei wird – ebenfalls ein Unterschied zu vielen anderen Ansätzen – unabhängig vom Wegfindungsprotokoll verfahren. Der Vorteil ist, dass erstens nicht für jedes Wegfindungsverfahren ein eigenes Beobachtungsverfahren spezifiziert werden muss, und dass zweitens nicht nur beim Aufbau neuer Wege, sondern auch bei der insgesamt weit mehr Aufwand verursachenden Nutzung dieser Wege beobachtet werden kann.

Wenn zur Durchführung der Zugangskontrolle Einschätzungen über Teilnehmer benötigt werden, deren Verhalten selbst noch nicht beobachtet werden konnte, so können mit Hilfe eines Meinungsaustauschprotokolls Einschätzungen anderer eingeholt werden. Solchermaßen angeforderte wie auch aus zufällig beobachteten Meinungsaustauschvorgängen entnommene Einschätzungen werden – soweit Platz vorhanden ist – dauerhaft aufbewahrt. Anhand ihrer wird durch Vergleich und Bestimmung der Übereinstimmung mit eigenen Einschätzungen auch ermittelt, inwieweit Meinungsäußerungen anderer in Zukunft vertraut werden soll.

Zur Repräsentation von Einschätzungen kommt keine der sonst gebräuchlichen relativ ausdrucksschwachen eindimensionalen oder gar diskreten Metriken zum Einsatz, sondern stattdessen wurde

eine zweidimensionale Vertrauensmetrik verwendet, in welcher neben der Positionierung zwischen positiven und negativen Bewertungen zusätzlich erfasst werden kann, wie sicher sich der Bewertende bei der Bewertung ist. Diese zwar bereits existierende, aber vorher noch nicht im Umfeld automatisierter Verhaltensbewertung eingesetzte Metrik wurde um eine neue, besser zur Verarbeitung der Einschätzungen geeignete Einschätzungsdarstellung erweitert.

Völlig neu entwickelt wurde ein Verfahren zur Verknüpfung von Einschätzungen, welches auf einer Analogie zwischen Vertrauensgraphen und Widerstandnetzwerken beruht. Mit diesem Verfahren wird es erstmals möglich, beliebige Graphen von Vertrauensbeziehungen zwischen Teilnehmern in einheitlicher Weise auszuwerten und dabei eine resultierende Gesamteinschätzung bezüglich bestimmter Teilnehmer aus der Sicht bestimmter anderer Teilnehmer zu gewinnen. Durch das Verfahren werden *alle* bekannten Vertrauensbeziehungen jeweils in dem Maß berücksichtigt, wie sie für den auswertenden Teilnehmer vertrauenswürdig und relevant sind. Die allgemeine Anwendbarkeit des Verfahrens und die umfassende Berücksichtigung aller bekannten Informationen hebt das Verfahren stark gegenüber bisher verwendeten ab; dort werden meist nur einstufige Vertrauensbeziehungen berücksichtigt, die Menge der Teilnehmer, deren Einschätzungen überhaupt berücksichtigt werden, muss manuell konfiguriert werden, oder es wird allein aufgrund der Anzahl von Meinungsäußerungen in bestimmter Richtung entschieden, ohne dass die individuelle Vertrauenswürdigkeit der einzelnen Teilnehmer berücksichtigt wird.

In Anerkennung der oben beschriebenen Teilnehmerkennungs-Problematik wurde eine bijektiv eindeutige Zuordnung zwischen Teilnehmerkennungen und Benutzern weder vorausgesetzt, noch wurde versucht, eine solche zu erzeugen. Stattdessen wird eine etwas schwächere, dafür aber ohne weitere Hilfsmittel kontrollierbare Bindung verwendet, indem als Teilnehmerkennungen die öffentlichen Teile asymmetrischer Schlüsselpaare eingesetzt werden. Ein Teilnehmer kann und darf sich mehrere solcher Kennungen zulegen, insofern kann aus dem Vorliegen unterschiedlicher Kennungen etwa als Absender zweier Nachrichten nicht geschlossenen werden, dass diese von unterschiedlichen Teilnehmern stammen. Andererseits kann aber ausschließlich der tatsächliche Inhaber einer Teilnehmerkennung Signaturen mit dem privaten Teil des asymmetrischen Schlüsselpaars erzeugen, die dann mit Hilfe des öffentlichen Teils verifizierbar sind. Bei Verwendung solcher Signaturen kann also niemand fremde Teilnehmerkennungen annektieren. Von dieser Eigenschaft wird bei der Zugangskontrolle Gebrauch gemacht: Alle Nachrichten werden von ihren Absendern signiert, und durch eine Signaturprüfung im Rahmen der Zugangskontrolle kann jeweils sicher festgestellt werden, aus welcher Quelle die Nachricht stammt.

Durch die besondere Methode der Identifikation wird auch ein besonders einfaches, aber wirkungsvolles Verfahren zur dezentralen Schlüsselverwaltung ermöglicht, durch welches sichergestellt wird, dass die benötigten Teilnehmerkennungen jeweils an der richtigen Stelle zur Verfügung stehen. Unter Nachbarn wird dazu jeweils bei Kontaktaufnahme ein Schlüsselaustauschprotokoll durchgeführt und eine zeitweise Zuordnung zwischen Schicht-2-Adressen und Schlüsseln hergestellt, anhand derer Beobachtungen den richtigen Teilnehmern zugeordnet werden können; eine ausführliche Bedrohungsanalyse weist nach, dass dabei keine für Angreifer lohnenden Verfälschungen möglich sind. Schlüssel entfernter Teilnehmer werden mittels eines dafür entwickelten Protokolls sukzessive von Vorgängern bei der Weiterleitung angefordert und dann unter Zuordnung zur Schicht-3-Adresse des Inhabers zwischengespeichert. Wegen der Signaturprüfung ist auch hierbei keine Umgehung der Zugangskontrolle möglich.

Anhand einer umfangreichen Implementierung innerhalb einer Simulationsumgebung konnte überprüft werden, dass das Zugangskontrollsystem erwartungsgemäß funktioniert. Simuliert wurden Netze mit zwischen 5 und etwa 45 Teilnehmern, die sich nach verschiedenen Mobilitätsmodellen bewegten. Bezüglich des Verfahrens zur Beobachtung und Bewertung des Weiterleitungsverhaltens benach-

barter Netzteilnehmer wurde dabei festgestellt, dass zuverlässige Einschätzungen über das Verhalten von Nachbarn recht schnell gewonnen werden. Teilnehmer, die sich in vorgegebener Weise unkooperativ verhielten, wurden im Versuch von ihren Nachbarn auch nach kurzer Zeit entsprechend eingeschätzt. Durch den Meinungsaustausch werden die lokal in den jeweiligen Nachbarschaften ermittelten Einschätzungen zügig anderen Netzteilnehmern zur Verfügung gestellt, so dass auch netzweit bald eine ausreichend sichere Grundlage für Zugangsentscheidungen besteht. Es zeigte sich, dass in Netzen mit geringer Mobilität insbesondere auch der Meinungsaustausch mit nicht direkt benachbarten Knoten erforderlich ist, um für eine ausreichende Verbreitung der Einschätzungen zu sorgen. Die Zugangskontrolle lehnte im Versuch nach einer kurzen Zeitspanne, die nach der Formierung eines neuen Netzwerks aus einander unbekannten Teilnehmern zur Gewinnung erster Einschätzungen erforderlich ist, zuverlässig alle Nutzungsversuche unkooperativer Teilnehmer ab.

5.2 Weiterführende Arbeiten

Bei der Durchführung von Versuchen mit größeren simulierten Netzwerken zeigte sich, dass bei der Auswertung der nach kurzer Zeit entstehenden dichten Vertrauensgraphen mittels des dafür neu entwickelten Widerstandsnetzwerkverfahrens ein nicht unerheblicher Berechnungsaufwand anfällt. Dieser ließe sich voraussichtlich sehr stark reduzieren, indem für das Ergebnis nahezu bedeutungslose Vertrauensbeziehungen, die immer in großer Anzahl vorhanden sind, aus zukünftigen Berechnungen herausgenommen würden. Die Bedeutung einzelner Vertrauensbeziehungen lässt sich nach jeder Auswertung sehr gut an den im Widerstandsnetzwerk fließenden Strömen ablesen, wobei sie natürlich abhängig von der Wahl einzuschätzenden Teilnehmers unterschiedlich sein kann. Es wäre also ein Verfahren zu entwickeln, das Beziehungen identifiziert und eliminiert, die sich in allen benötigten Fällen als wenig hilfreich erweisen.

Beim bisherigen Zugangskontrollverfahren wurden feste Schwellwerte vorgesehen für die Entscheidung, ob anhand der jeweiligen Einschätzung eines fraglichen Teilnehmers der Zugang gestattet werden sollte, oder nicht. Insbesondere bei relativ unsicheren Einschätzungen wäre ein denkbarer anderer Weg, zunächst lediglich die Rate erlaubter Dienstnutzungen zu beschränken. Eventuell könnte so die Hürde anfänglich durch neue Knoten zu erbringender Leistungen gesenkt werden, ohne dabei die Gefahr einer übermäßigen Belastung durch unkooperative Teilnehmer in Kauf zu nehmen.

Bezüglich der Evaluation des neuen Verfahrens könnten durch weitere Untersuchungen noch umfassendere Erkenntnisse über die Einflüsse von Parametern und Randbedingungen gewonnen werden. So wurde bisher beispielsweise nicht untersucht, wie sich unterschiedliches Nutzungsverhalten – etwa besonders langlebige oder stark genutzte Kommunikationsbeziehungen innerhalb bestimmter Benutzergruppen – im Zugangskontrollsystem auswirkt. Außerdem wäre es sinnvoll festzustellen, welche Änderungen sich bei Verwendung realer Verfahren beispielsweise für Medienzugang und Wegfindung anstelle der in der Simulation idealisierten Komponenten ergeben.

Anhang A

Zusätzliche Simulationsergebnisse

Die folgenden Abbildungen enthalten zusätzliche Ergebnisse oder großräumigere Darstellungen bereits gezeigter Ergebnisse, auf die in Kapitel 4 aus Platzgründen verzichtet wurde.

- Abbildung A.1: Relative Häufigkeit bestimmter Weglängen zu dem Netz aus Abbildung 4.1.
- Abbildung A.2: Erfolgsquote, Weiterleitungsfehlerquote (y1-Achse links) und mittlere Weglänge (y2-Achse rechts) in Abhängigkeit von der Teilnehmerzahl beim Random-Waypoint-Modell und Sendereichweite 100 m (siehe Seite 168)
- Abbildung A.3: Relative Häufigkeit bestimmter Weglängen bei Teilnehmerzahlen zwischen 10 und 99 beim Random-Waypoint-Modell und Sendereichweite 100 m (siehe Seite 168)
- Abbildung A.4: Relative Häufigkeit bestimmter Nachbarschaftsdauern in Abhängigkeit von der Aufenthaltsdauer bei Bewegung nach dem Random-Waypoint-Modell mit niedriger Geschwindigkeit (rechte Seite der Abbildung 4.7, entzerrt)
- Abbildung A.5: Anzahl von Knoten im Bootstrapping-Modus bei Bewegung nach dem Random-Waypoint-Modell mit verschiedenen Geschwindigkeiten und kurzer (um 1 s; links) bzw. etwas längerer (um 180 s; rechts) Aufenthaltszeit (siehe Seite 183)
- Abbildung A.6: Anzahl von Knoten im Bootstrapping-Modus bei Bewegung nach dem Random-Waypoint-Modell mit verschiedenen Aufenthaltszeiten und niedriger (um 0.5 m/s; links) bzw. höherer (um 9 m/s; rechts) Geschwindigkeit (siehe Seite 183)
- Abbildung A.7: Anteil angekommener weiterzuleitender Nachrichten aus unkooperativen Quellen innerhalb aufeinanderfolgender 10-Sekunden-Intervalle in einem Netz mit 23 kooperativen und zwischen 0 und 12 unkooperativen Teilnehmern (Ergänzung zu Abbildung 4.20)

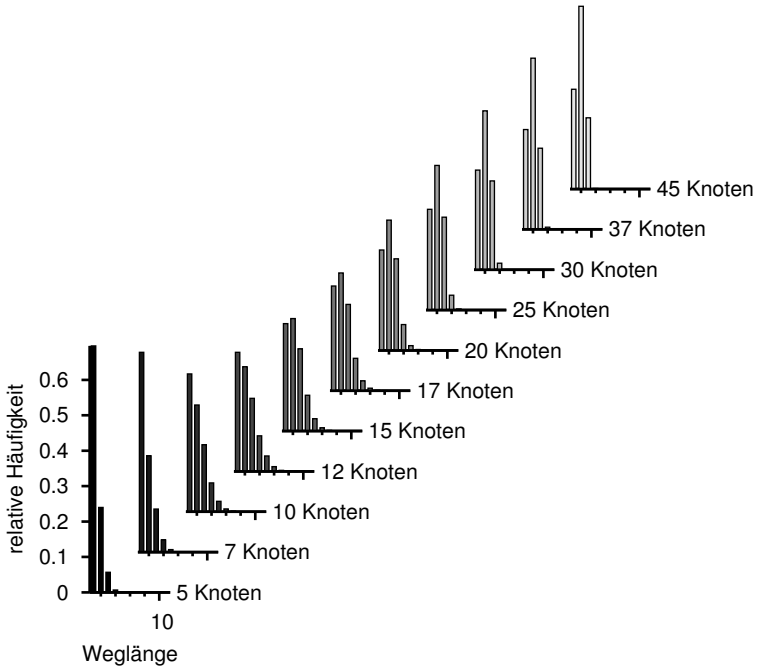


Abbildung A.1: Relative Häufigkeit bestimmter Weglängen

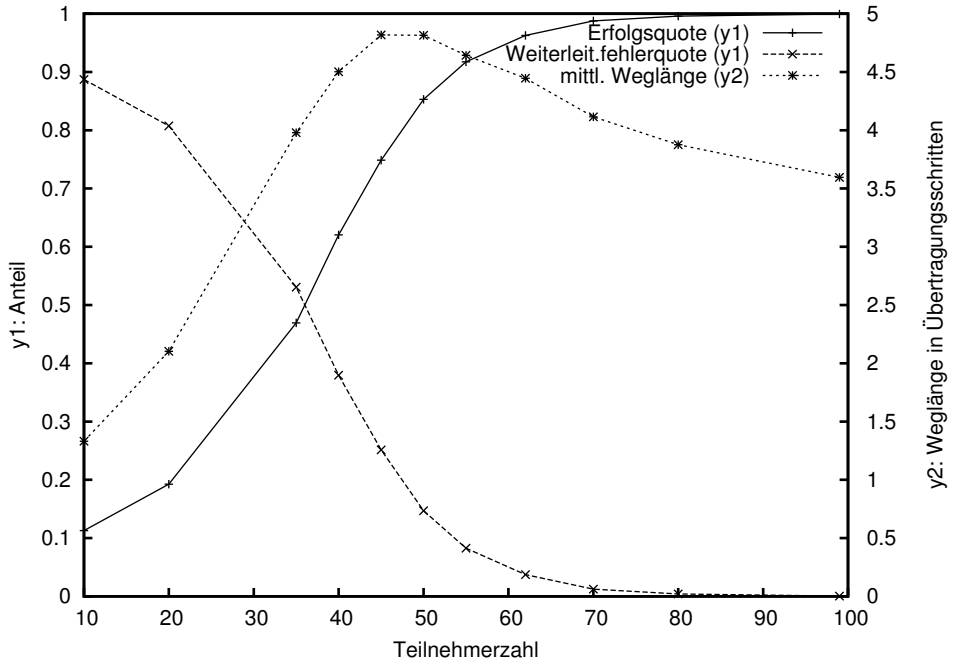


Abbildung A.2: Erfolgsquote und Weiterleitungsfehlerquote (y1-Achse links) und mittlere Weglänge (y2-Achse rechts) in Abhängigkeit von der Teilnehmerzahl beim Random-Waypoint-Modell und Sendereichweite 100 m

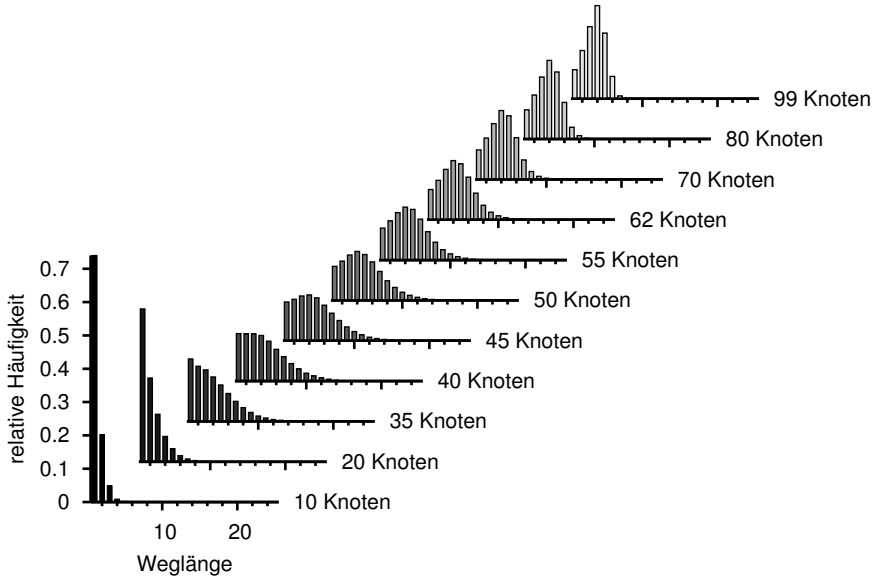


Abbildung A.3: Relative Häufigkeit bestimmter Weglängen bei Teilnehmerzahlen zwischen 10 und 99 beim Random-Waypoint-Modell und Sendereichweite 100 m

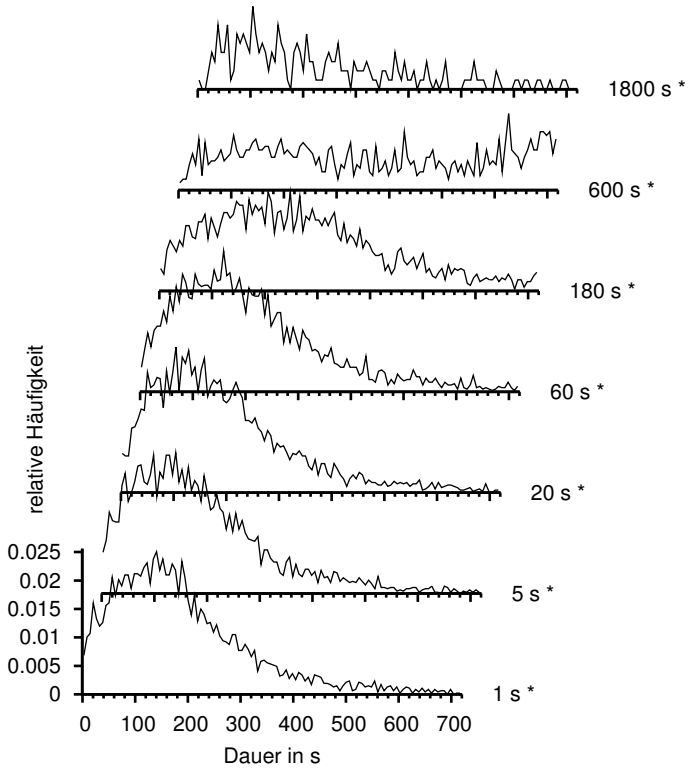


Abbildung A.4: Relative Häufigkeit bestimmter Nachbarschaftsdauern in Abhängigkeit von der Aufenthaltsdauer bei Bewegung nach dem Random-Waypoint-Modell mit niedriger Geschwindigkeit

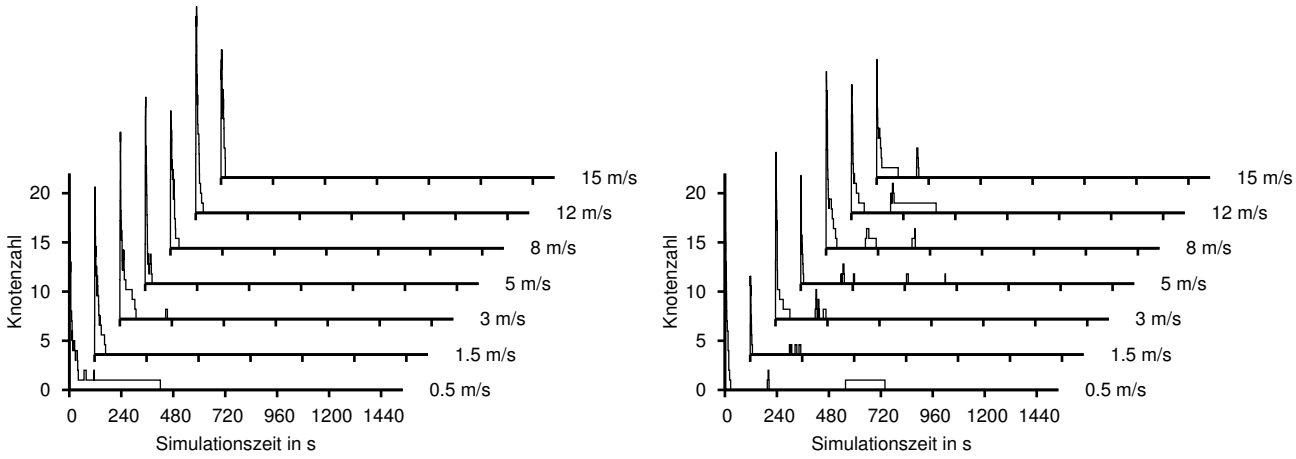


Abbildung A.5: Anzahl von Knoten im Bootstrapping-Modus bei Bewegung nach dem Random-Waypoint-Modell mit verschiedenen Geschwindigkeiten und kurzer (um 1 s; links) bzw. etwas längerer (um 180 s; rechts) Aufenthaltszeit

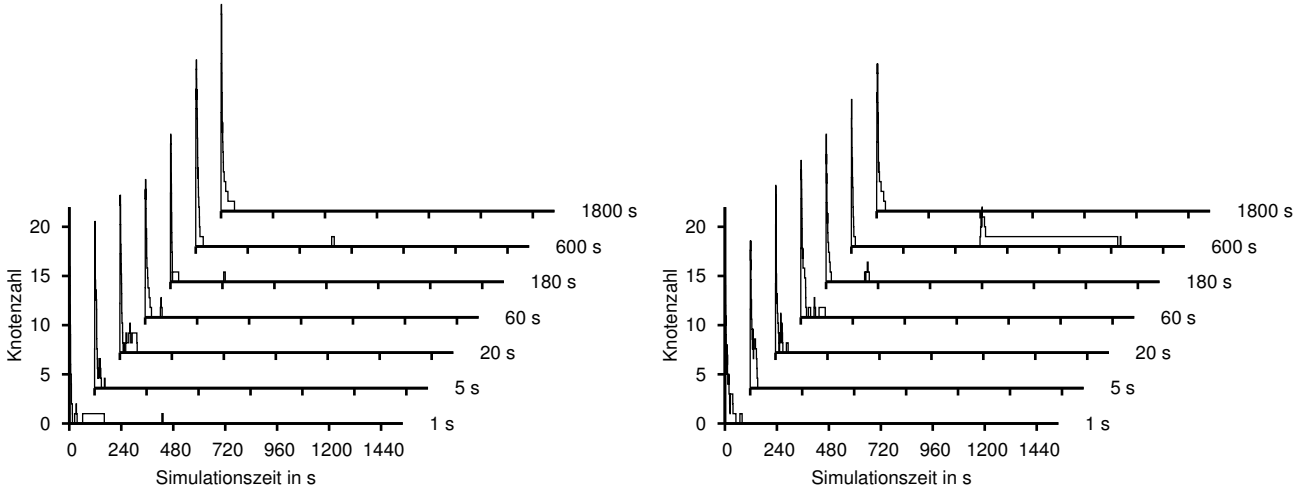


Abbildung A.6: Anzahl von Knoten im Bootstrapping-Modus bei Bewegung nach dem Random-Waypoint-Modell mit verschiedenen Aufenthaltszeiten und niedriger (um 0.5 m/s; links) bzw. höherer (um 9 m/s; rechts) Geschwindigkeit

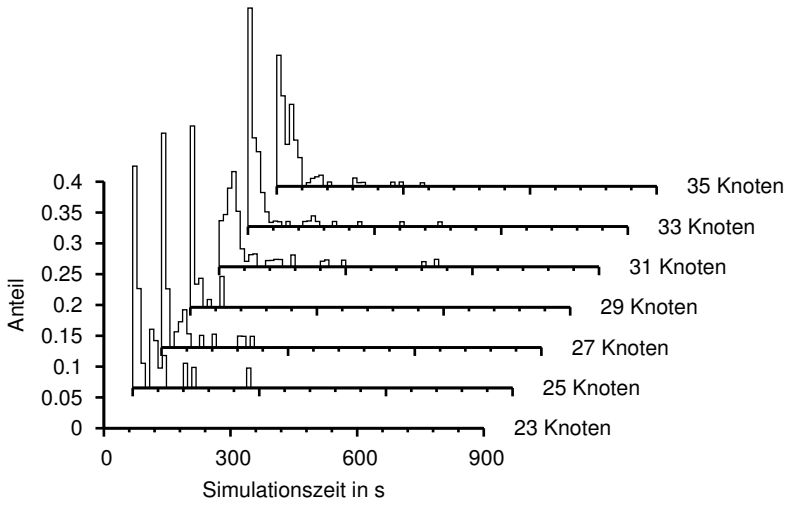


Abbildung A.7: Anteil angekommener weiterzuleitender Nachrichten aus unkooperativen Quellen innerhalb aufeinanderfolgender 10-Sekunden-Intervalle in einem Netz mit 23 kooperativen und zwischen 0 und 12 unkooperativen Teilnehmern

Abkürzungsverzeichnis

ACL	Access Control List
AES	Advanced Encryption Standard
AH	Authentication Header (IP Security)
ANSI	American National Standards Institute
ARP	Address Resolution Protocol
AS	Authentisierungs-Server (Kerberos); Autonomes System (Wegfindung)
BGP	Border Gateway Protocol
BSD	Berkeley Software Distribution (Unix-Variante)
BSS	Basic Service Set
CA	Collision Avoidance
CCA	Clear Channel Assessment
CD	Collision Detection
CONFIDANT	Cooperation of Nodes: Fairness in Dynamic Ad-hoc Networks
CSMA	Carrier Sense Multiple Access
CTS	Clear to Send
DCF	Distributed Coordination Function
DDoS	Distributed Denial of Service
DECT	Digital Enhanced Cordless Telecommunications
DES	Data Encryption Standard
DIFS	DCF Inter-Frame Spacing
DSA	Digital Signature Algorithm
DSL	Digital Subscriber Line
DSR	Dynamic Source Routing
DSS	Digital Signature Standard
DSSS	Direct Sequence Spread Spectrum
ESP	Encapsulating Security Payload (IP Security)
ETSI	European Telecommunications Standards Institute
FHSS	Frequency Hopping Spread Spectrum
FTP	File Transfer Protocol
GHz	Gigahertz
GSM	Groupe Speciale Mobile, Global System for Mobile Communications
HMAC	Keyed-Hashing for Message Authentication
HTTP	Hypertext Transfer Protocol
IBSS	Independent Basic Service Set
ICMP	Internet Control Message Protocol
IDEA	International Data Encryption Algorithm
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force

IGMP	Internet Group Management Protocol
IKE	Internet Key Exchange (IP Security)
IP	Internet Protocol
IPSec	internet Protocol Security
IrDA	Infrared Data Association
ISAKMP	Internet Security Association and Key Management Protocol (IP Security)
ISO	International Organization for Standardization
ISOC	Internet Society
LAN	Local Area Network
LLC	Logical Link Control
MAC	Message Authentication Code; Medium Access Control
MANET	Mobile Ad-hoc Network
MD5	Message Digest Algorithm 5
MHz	Megahertz
MIT	Massachusetts Institute of Technology
OFDM	Orthogonal Frequency Division Multiplex
OMNeT++	Objective Modular Network Testbed in C++
OSI	Open Systems Interconnection
PAN	Personal Area Network
PCF	Point Coordination Function
PDA	Personal Digital Assistant
PGP	Pretty Good Privacy
PKI	Public Key Infrastructure
PLCP	Physical Layer Convergence Protocol
PMD	Physical Medium Dependent
PPP	Point-to-Point Protocol
RC4	Ron's Cipher 4
RC5	Ron's Cipher 5
RFC	Request for Comment
RSA	Rivest-Shamir-Adleman
RTS	Ready to Send
SHA-1	Secure Hash Algorithm 1
SIFS	Short Inter-Frame Spacing
SMTP	Simple Mail Transfer Protocol
SSL	Secure Socket Layer
TCP	Transmission Control Protocol
TGS	Ticket Granting Server
TLS	Transport Layer Security
UDP	User Datagram Protocol
UMTS	Universal Mobile Telecommunications System
URSA	Ubiquitous and Robust Access Control for Mobile Ad-hoc Networks
USA	United States of America
WAN	Wide Area Network
WLAN	Wireless Local Area Network
WWW	World Wide Web

Literaturverzeichnis

[BaBa03] S. Bansal und M. Baker. Observation-based Cooperation Enforcement in Ad Hoc Networks. *ArXiv Computer Science e-prints*, Juli 2003.

[Banh01] Jörg Banholzer. Entwurf und Implementierung einer Simulationsumgebung für mobile Ad-hoc-Netzwerke. Diplomarbeit, Universität Karlsruhe (TH), Februar 2001.

[BBČG+01] L. Blaževic, L. Buttyán, S. Čapkun, S. Giordano, J. Hubaux und J.-Y. Le Boudec. Self-Organization in Mobile Ad-Hoc Networks: the Approach of Terminodes. *IEEE Commun. Mag.*, Juni 2001.

[BDMP99] Roland Bless, Stefan Dresler, Daniel Müller und Frank Pählke. Standardization Efforts for Securing the Internet. In *1999 International Spring Conference: Security Aspects in a Changing Competitive Environment*, München, Mai 1999. Communications Fraud Control Association.

[BHKP+04] Marc Bechler, Hans-Joachim Hof, Daniel Kraft, Frank Pählke und Lars Wolf. A Cluster-based Security Architecture for Ad Hoc Networks. In *Proc. IEEE Conference on Computer Communications (Infocom)*, Hong Kong, März 2004.

[BHMP+02] Marc Bechler, Achim Hauck, Daniel Müller, Frank Pählke und Lars Wolf. Ein Sicherheitskonzept für clusterbasierte Ad-hoc-Netzwerke. In M. Weber und F. Kargl (Hrsg.), *Mobile Ad-hoc-Netzwerke – 1. Deutscher Workshop über mobile Ad-hoc-Netzwerke (WMAN)*, Ulm, März 2002. Gesellschaft für Informatik, S. 135–152.

[Blu01a] Bluetooth Special Interest Group. *Specification of the Bluetooth System, Volume 1: Core, Version 1.1*, Februar 2001.

[Blu01b] Bluetooth Special Interest Group. *Specification of the Bluetooth System, Volume 2: Profile, Version 1.1*, Februar 2001.

[BSSW02] D. Balfanz, D. K. Smetters, P. Stewart und H. C. Wong. Talking To Strangers: Authentication in Ad-Hoc Wireless Networks. In *Proc. Symp. on Network and Distributed System Security (NDSS)*, San Diego, Februar 2002.

[BuLB02] Sonja Buchegger und Jean-Yves Le Boudec. Performance Analysis of the CONFIDANT Protocol (Cooperation Of Nodes: Fairness In Dynamic Ad-hoc NeTworks). In *Proc. ACM Symp. on Mobile Ad Hoc Networking & Computing (MobiHoc)*, Juni 2002.

[BuLB04] Sonja Buchegger und Jean-Yves Le Boudec. A Robust Reputation System for P2P and Mobile Ad-hoc Networks. In *Proc. 2nd Workshop on the Economics of Peer-to-Peer Systems*, Juni 2004.

[BWNHM+03] S. Blake-Wilson, M. Nyström, D. Hopwood, J. Mikkelsen und T. Wright. *Transport Layer Security (TLS) Extensions*. IETF, Juni 2003. RFC 3546.

[CeKa74] V. G. Cerf und R. E. Kahn. A Protocol for Packet Network Interconnection. *IEEE Trans. Commun.* Band COM-22, Mai 1974, S. 627–641.

[CRKGLA89] C. Cheng, R. Riley, S. P. R. Kumar und J. J. Garcia-Luna-Aceves. A Loop-Free Bellman-Ford Routing Protocol without Bouncing Effect. In *Proc. ACM SIGCOMM*, September 1989, S. 224–237.

[DiAl99] T. Dierks und C. Allen. *The TLS Protocol*. IETF, Januar 1999. RFC 2246.

[DiHe76] W. Diffie und M. Hellman. New Directions in Cryptography. *IEEE Trans. Inform. Theory* IT-22(6), November 1976, S. 644–654.

[Dijk59] E. W. Dijkstra. A Note on Two Problems in Connection with Graphs. *Numer. Math.* Band 1, 1959, S. 269–271.

[DoDM99] Elmar Dorner, Stefan Dresler und Daniel Müller. Network-Initiated Hand-Over for Lowering Peak Effects in Cellular Environments. In *Proc. 10th IEEE Workshop on Local Area Networks (LANMAN)*, Sydney, November 1999. IEEE Communcation Society.

[FeSc00] N. Ferguson und B. Schneier. A cryptographic evaluation of IPSec. Technischer Bericht, Counterpane Internet Security Inc., 2000.

[GuGu01] Vipul Gupta und Sumit Gupta. Securing the Wireless Internet. *IEEE Commun. Mag.*, Dezember 2001, S. 68–74.

[HaCa98] D. Harkins und D. Carrel. *The Internet Key Exchange (IKE)*. IETF, November 1998. RFC 2409.

[HeWK04] Q. He, D. Wu und P. Khosla. SORI: A secure and objective reputation-based incentive scheme for ad hoc networks. In *Proc. IEEE Wireless Communications and Networking Conference*, Atlanta, GA, USA, März 2004.

[Hof02] Hans-Joachim Hof. Optimierung und Evaluation eines Sicherheitskonzepts für mobile Ad-hoc-Netze. Diplomarbeit, Universität Karlsruhe (TH), Dezember 2002.

[HuBČ01] J.-P. Hubaux, L. Buttyan und S. Čapkun. The Quest for Security in Mobile Ad Hoc Networks. In *Proc. ACM Symp. on Mobile Ad Hoc Networking & Computing (MobiHoc)*, Long Beach, Oktober 2001.

[HuBu00] J.-P. Hubaux und L. Buttyan. Enforcing Service Availability in Mobile Ad-Hoc WANs. In *Proc. ACM Symp. on Mobile Ad Hoc Networking & Computing (MobiHoc)*, Mai 2000.

[HuPJ02] Yih-Chun Hu, Adrian Perrig und David B. Johnson. Wormhole Detection in Wireless Ad Hoc Networks. Technischer Bericht, Rice University Department of Computer Science, 2002.

[IEE99a] IEEE. *IEEE Standard for Information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks – Specific requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications – Amendment 1: High-speed Physical Layer in the 5 GHz band*, 1999. ISO/IEC 8802-11:1999/Amd 1:2000(E).

[IEE99b] IEEE. *IEEE Standards for Information Technology – Telecommunications and Information Exchange between Systems – Local and Metropolitan Area Network – Specific Requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, 1999. ISO/IEC 8802-11.

[IEE99c] IEEE. *Supplement to 802.11-1999, Wireless LAN MAC and PHY specifications: Higher speed Physical Layer (PHY) extension in the 2.4 GHz band*, 1999. IEEE 802.11b-1999.

[IEE01] IEEE. *IEEE Standard for Information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks – Specific requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications – Amendment 2: Higher-speed Physical Layer (PHY) extension in the 2.4 GHz band – Corrigendum 1*, 2001. IEEE 802.11b-1999/Cor1-2001.

[IEE03a] IEEE. *IEEE Standard for Information technology – Telecommunications and Information Exchange Between Systems – LAN/MAN Specific Requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Spectrum and Transmit Power Management Extensions in the 5GHz band in Europe*, 2003. IEEE 802.11h-2003.

[IEE03b] IEEE. *IEEE Standard for Information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks – Specific requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications – Amendment 4: Further Higher-Speed Physical Layer Extension in the 2.4 GHz Band*, 2003. IEEE 802.11g-2003.

[ISO94] ISO/IEC. *Information technology – Open Systems Interconnection – Basic Reference Model – Part 1: The Basic Model*, 2. Auflage, 1994. International Standard 7498-1, auch ITU-T Recommendation X.200.

[JøIs02] Audun Jøsang und Roslan Ismail. The Beta Reputation System. In *15th Bled Electronic Commerce Conference: e-Reality: Constructing the e-Economy*, Bled, Slovenia, Juni 2002.

[Jøsa96] Audun Jøsang. The right type of trust for distributed systems. In C. Meadows (Hrsg.), *Proc. ACM New Security Paradigms Workshop*, 1996.

[Jøsa98] Audun Jøsang. A Subjective Metric of Authentication. In *Proc. European Symp. on Research in Computer Security (ESORICS)*, LNCS. Springer-Verlag, 1998.

[Jøsa01] Audun Jøsang. A Logic for Uncertain Probabilities. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 9(3), Juni 2001.

[Kant85] Immanuel Kant. *Grundlegung zur Metaphysik der Sitten*. bey Johann Friedrich Hartknoch, Riga. 1785.

[Kauf05] C. Kaufman. *Internet Key Exchange (IKEv2) Protocol*. IETF, Dezember 2005. RFC 4306.

[Kent05a] S. Kent. *IP Authentication Header*. IETF, Dezember 2005. RFC 4302.

[Kent05b] S. Kent. *IP Encapsulationg Security Payload (ESP)*. IETF, Dezember 2005. RFC 4303.

[KeSe05] S. Kent und K. Seo. *Security Architecture for the Internet Protocol*. IETF, Dezember 2005. RFC 4301.

[Kirc45] G. R. Kirchhoff. Ueber den Durchgang eines elektrischen Stromes durch eine Ebene, insbesondere durch eine kreisförmige. *Annalen der Physik und Chemie* LXIV/140(4), 1845, S. 497–513.

[KPPP+06] S. Kumar, C. Paar, J. Pelzl, G. Pfeiffer, A. Rupp und M. Schimmler. How to Break DES for 8,980 Euro. In *Proc. 2nd Workshop on Special-purpose Hardware for Attacking Cryptographic Systems (SHARCS)*, 2006.

[KrBC97] H. Krawczyk, M. Bellare und R. Canetti. *HMAC: Keyed-Hashing for Message Authentication*. IETF, Februar 1997. RFC 2104.

[KrSc04] Daniel Kraft und Günter Schäfer. Distributed Access Control for Consumer Operated Mobile Ad-hoc Networks. In *Proc. IEEE Consumer Communications and Networking Conference (CCNC)*, Las Vegas, Januar 2004.

[LaPW03] B. Lamparter, K. Paul und D. Westhoff. Charging Support for Ad Hoc Stub Networks. *Elsevier Journal of Computer Communications, Special Issue on "Internet Pricing and Charging: Algorithms, Technology and Applications"* 26(13), August 2003.

[LKZL+04] H. Luo, J. Kong, P. Zerfos, S. Lu und L. Zhang. URSA: Ubiquitous and robust access control for mobile ad hoc networks. *IEEE/ACM Trans. Networking* 12(6), Oktober 2004.

[LuLu00] H. Luo und S. Lu. Ubiquitous and Robust Authentication Services for Ad Hoc Wireless Networks. Technischer Bericht TR-200030, Dept. of Computer Science, UCLA, 2000.

[LZKL+02] Haiyun Luo, Petros Zerfos, Jiejun Kong, Songwu Lu und Lixia Zhang. Self-securing Ad Hoc Wireless Networks. In *Proc. 7th IEEE Symp. on Comp. and Communications (ISCC)*, Taormina, 2002.

[MGLM00] S. Marti, T. J. Giuli, K. Lai und M.Baker. Mitigating Routing Misbehaviour in Mobile Ad hoc Networks. In *Proc. 6th International Conf. on Mob. Comp. and Networking (MOBICOM)*, August 2000, S. 255–265.

[MNSS87] S. P. Miller, B. C. Neuman, J. I. Schiller und J. H. Saltzer. Section E.2.1: Kerberos Authentication and Authorization System. Technischer Bericht, M.I.T. Project Athena, Cambridge, Massachusetts, Dezember 1987.

[MSST98] D. Maughan, M. Schertler, M. Schneider und J. Turner. *Internet Security Association and Key Management Protocol (ISAKMP)*. IETF, November 1998. RFC 2408.

[MüSS98] Daniel Müller, Günter Schäfer und Jochen Schiller. An Efficient Authentication Protocol for High Performance Networks. In *Proc. IEEE Global Telecommunications Conference (Globecom)*, Band 2, Sydney, November 1998. Institute of Electrical and Electronics Engineers Inc., S. 886 ff.

[NIS99] National Institute of Standards and Technology, U.S. Department of Commerce. *Data Encryption Standard (DES)*, Oktober 1999. FIPS PUB 46-3.

[NIS01] National Institute of Standards and Technology, U.S. Department of Commerce. *Advanced Encryption Standard (AES)*, November 2001. FIPS PUB 197.

[NYHR05] C. Neuman, T. Yu, S. Hartman und K. Raeburn. *The Kerberos Network Authentication Service (V5)*. IETF, Juli 2005. RFC 4120.

[PaWe02] Krishna Paul und Dirk Westhoff. Context Aware Detection of Selfish Nodes in DSR based Ad-hoc Networks. In *Proc. IEEE Global Telecommunications Conference (Globecom)*, Taipei, Taiwan, November 2002.

[Perk01] C. Perkins. *Ad Hoc Networking*. Addison-Wesley. 2001.

[Perk02] C. Perkins. *IP Mobility Support for IPv4*. IETF, August 2002. RFC 3344.

[Pipe98] D. Piper. *The Internet IP Security Domain of Interpretation for ISAKMP*. IETF, November 1998. RFC 2407.

[Post81] J. Postel. *Internet Protocol*. IETF, September 1981. RFC 791.

[ReSt97] M. K. Reiter und S. G. Stubblebine. Toward acceptable metrics of authentication. In *Proc. IEEE Symp. on Security and Privacy*, 1997, S. 10–20.

[RiSA78] Ronald L. Rivest, Adi Shamir und Leonard M. Adleman. A Method for Obtaining Digital Signatures and Public-Key Cryptosystems. *Communications of the ACM* 21(2), 1978, S. 120–126.

[Schi03] Jochen Schiller. *Mobilkommunikation*. Pearson Education Deutschland, München. 2. Auflage, 2003.

[Schn96] B. Schneier. *Applied Cryptography*. John Wiley. 1996.

[Simp94] W. Simpson. *The Point-to-Point Protocol (PPP)*. IETF, Juli 1994. RFC 1661.

[ThBa04] George Theodorakopoulos und John S. Baras. Trust Evaluation in AdHoc Networks. In *Proc. ACM Workshop on Wireless Security (WiSe)*, 2004.

[Varg01] András Varga. The OMNeT++ Discrete Event Simulation System. In *Proc. European Simulation Multiconference (ESM)*, Prague, Czech Republic, Juni 6–9, 2001.

[WESW98] H. Woesner, J.-P. Ebert, M. Schläger und A. Wolisz. Power-saving Mechanisms in Emerging Standards for Wireless LANs. *IEEE Personal Commun. Mag.* 5(3), 1998.

[ZhHa99] Lidong Zhou und Zygmunt J. Haas. Securing Ad Hoc Networks. *IEEE Network* 13(6), 1999, S. 24–30.

Index

Ad-hoc-Netze bestehen allein aus ihren drahtlos kommunizierenden mobilen Teilnehmerknoten und verzichten auf fest installierte Infrastruktur. Teilnehmer, die sich außerhalb ihrer gegenseitigen Sendereichweite befinden, sind darauf angewiesen, dass dazwischen liegende Knoten Nachrichten für sie weiterleiten. Die aktive Mitarbeit aller Teilnehmer und das Einbringen der eigenen, oft knappen Ressourcen zugunsten Anderer sind damit Voraussetzung für die Funktion solcher Netze. Zum Ausschluss unkooperativer Nutznießer, die selbst nichts beitragen möchten, ist gerade in offenen Ad-hoc-Netzen, in denen die autonomen Teilnehmer einander a priori meist unbekannt sind, eine Zugangskontrolle erforderlich.

Bei dem in der vorliegenden Arbeit vorgestellten, auf alle Netzknoten verteilten Zugangskontrollsystem für offene Ad-hoc-Netze basiert die Zugangsentscheidung auf dem anhand automatischer Beobachtung des Verhaltens von Nachbarn gewonnenen Vertrauen in die Kooperativität der Zugang suchenden Knoten. Durch Meinungsaustausch und Verknüpfung aller bekannten Vertrauensbeziehungen können die Teilnehmer sich schnell gegenseitig einschätzen und unkooperative Knoten sicher ausschließen.

ISBN: 978-3-86644-121-7

www.uvka.de