

Strategisches Verhalten in Systemen mit Interaktions- und Kontaktwahl

zur Erlangung des akademischen Grades eines
Doktors der Ingenieurwissenschaften/

von der Fakultät für Informatik
des Karlsruher Instituts für Technologie (KIT)

genehmigte

Dissertation

von

Björn-Oliver Hartmann

aus Hildesheim

Tag der mündlichen Prüfung: 11.02.2011

Erster Gutachter: Prof. Dr.-Ing. Klemens Böhm

Zweiter Gutachter: Prof. Dr. rer. pol. Dr. rer. nat. Bodo Vogt

Danksagung

Eine Reihe von Menschen hat einen Beitrag zu dieser Arbeit geleistet. Neben ihrer Unterstützung bin ich für die vielen schönen Momente dankbar, die ich mit diesen Menschen erleben durfte.

Mein erster Dank geht an die Gutachter dieser Arbeit: zum einen an Prof. Klemens Böhm, meinen Doktorvater, und zum anderen an Prof. Bodo Vogt. Prof. Klemens Böhm hat mir beigebracht, wissenschaftlich zu arbeiten und hat mich darüber hinaus in meiner Arbeit unterstützt, wissenschaftliche Veröffentlichungen hervorzubringen. Dabei ist er nicht müde geworden, mit mir unsere Veröffentlichungen immer wieder durchzugehen und sie mit mir zu verbessern. Dieser Prozess hat mich sehr bereichert. Von Prof. Bodo Vogt konnte ich vor allem in ökonomischen Fragen und bei der Durchführung verhaltensökonomischer Experimente profitieren.

Mein zweiter Dank geht an Stephan Schosser, der mir zu Beginn meiner Zeit als Doktorand zur Seite gestanden hat und mir durch seine freundschaftliche und professionelle Art den Einstieg in die Welt der Wissenschaft erleichtert hat.

Neben vielen weiteren Studenten, die durch ihre Studien- und Diplomarbeiten einen Beitrag zu dieser Arbeit geleistet haben, möchte ich mich insbesondere bei Anno von Heimburg und Holger Miller bedanken. Mit beiden hat mir die Arbeit Spaß und Freude bereitet und war darüber hinaus inhaltlich bereichernd.

Mein Dank gebührt ebenfalls meinen Kolleginnen und Kollegen am Institut für Programmstrukturen und Datenorganisation – insbesondere natürlich den Kolleginnen und Kollegen des Lehrstuhls für Systeme der Informationsverwaltung – die mir die Arbeit am Institut durch ihre freundliche, hilfsbereite und kollegiale Art verschönert haben. Ein besonderes soziales Highlight am Institut für Programmstrukturen und Datenorganisation war für mich die IPD-Sport-Gruppe, mit der ich viele interessante Sportarten ausprobiert habe und durch deren Veranstaltungen die Arbeit auf eine gesunde und erfrischende Art aufgelockert wurde. Eine weitere angenehme Auflockerung waren die interessanten und teilweise elektrisierenden – dabei aber immer freundlich geführten – Gespräche, die ich mit meinem langjährigen Bürokollegen Andranik Khachatryan führen durfte.

Ein bedeutender Anteil an dieser Arbeit liegt in der Unterstützung und im Verständnis, welches meine Familie und meine Freunde in dieser Zeit für mich aufgebracht haben. Den größten Brocken Verständnis hat dabei meine Frau Maren getragen, die mich an so manchem Abend und so manchem Wochenende bei meiner Arbeit unterstützt hat. Für Dein Verständnis, für Deine Geduld, für Deine Liebe und für so vieles mehr liebe ich Dich!

Inhaltsverzeichnis

1	Einleitung	1
1.1	Verteilte Systeme	1
1.2	Strategien	2
1.2.1	Interaktionswahl	2
1.2.2	Kontaktwahl	2
1.3	Problemdefinition	3
1.4	Strategisches Verhalten	4
1.5	Ziel der Arbeit	4
1.6	Vorgehen	5
1.7	Beiträge der Dissertation	5
1.8	Gliederung der Dissertation	8
2	Verwandte Arbeiten	9
2.1	Studien	9
2.1.1	Studien zur Interaktionswahl	9
2.1.2	Studien zur Kontaktwahl	10
2.1.3	Studien zu Sozialen Netzwerken	11
2.2	Spieltheorie und Anwendungen der Spieltheorie	12
2.2.1	Arbeiten zur Interaktionswahl	12
2.2.2	Arbeiten zur Kontaktwahl	13
2.2.3	Kombination von strategischen Spielen und Kontaktwahl	15
2.3	Verhaltensökonomie	16
2.3.1	Verhaltensökonomische Arbeiten zur Interaktionswahl	16
2.3.2	Verhaltensökonomische Arbeiten zur Kontaktwahl	17
2.3.3	Verhaltensökonomische Arbeiten zur Strategie- und Kontaktwahl	18
2.4	Fazit	19
3	Grundlagen	21
3.1	Verteilte Systeme	21
3.2	Annahmen	25
3.3	Netzwerkstruktur und Kontaktwahl	27
3.4	Ein Modell der Interaktionswahl	28
3.5	Nutzen des Systems	30

Inhaltsverzeichnis

3.6	Wichtige Maße	33
3.7	Methodik	35
4	Anreize zur Vernetzung	37
4.1	Einleitung	37
4.2	Strukturierte Peer-to-Peer Systeme und zusätzliche Kontakte	40
4.3	Methodik	41
4.4	Annahmen	45
4.5	Strategien	46
4.6	Deliberately Poor Connectedness	48
4.6.1	Durchschnittliche Pfadlänge	49
4.6.2	Einfluss unkooperativer Teilnehmer	51
4.6.3	Nutzen der DPC-Strategie	53
4.6.4	Gefahr der DPC-Strategie	54
4.7	Contact-for-Contact Strategie	56
4.8	Analyse der Contact-for-Contact Strategie	58
4.9	Validierung	59
4.9.1	Aufbau der Simulationen	60
4.9.2	Resultate	62
4.10	Diskussion	65
4.11	Fazit	67
5	Gleichgewichte in Systemen mit Interaktions- und Kontaktwahl	69
5.1	Einleitung	70
5.2	Definitionen und Annahmen	72
5.2.1	Netzwerkeigenschaften	72
5.2.2	Strategieraum	74
5.3	Effizienz des Netzwerks	78
5.3.1	Verhalten von Kontakten	79
5.3.2	n -Spieler	80
5.4	Fazit	88
6	Evolution der Interaktions- und Kontaktwahl	89
6.1	Versuchsaufbau	91
6.2	Ergebnisse – Kooperatives Gleichgewicht	95
6.3	Das Entstehen kooperativer Systeme	100
6.4	Fazit	103
7	Das menschliche Verhalten in der Interaktions- und Kontaktwahl	105
7.1	Einleitung	105
7.2	Definitionen und Anpassungen	108
7.3	Hypothesen	109

7.4	Experimentaufbau	112
7.4.1	Experiment-Rahmenwerk	115
7.4.2	Strategie-Spiel	118
7.5	Ergebnisse	118
7.5.1	Kontaktwahl	118
7.5.2	Position vs. Kooperationsgrad	120
7.5.3	Fallenlassen von Freifahrern	121
7.5.4	Kosten für zusätzliche Kontakte	121
7.5.5	Anzahl an Kontakten	124
7.6	Fazit	126
8	PostGame – Eine Studie	127
8.1	Einleitung	127
8.2	Definitionen und Anpassungen	130
8.2.1	Anpassung der Auszahlungsfunktion	130
8.3	Hypothesen	131
8.3.1	Hypothesen zur Kontaktwahl	132
8.3.2	Hypothesen zur Interaktionswahl	132
8.3.3	Hypothesen zur Kombination von Interaktions- und Kontaktwahl	133
8.4	Anforderungen an die Umgebung der Studie	134
8.4.1	Korrektheit	136
8.5	Rahmenwerk des PostGames	136
8.5.1	Dienstmerkmale	136
8.5.2	Architektur	138
8.5.3	Nutzerschnittstelle	140
8.6	Messwerte	143
8.7	Experimentaufbau	143
8.8	Evaluierung	145
8.8.1	Ergebnisse – Hypothesen	145
8.8.2	Ergebnisse – Korrektheitskriterien	150
8.9	Fazit	151
9	Fazit und Ausblick	153
9.1	Zusammenfassung	153
9.2	Ausblick	157
	Appendix	161
A	Deliberately Poor Connectedness	161
A.1	Pfadlänge einer Anfrage	161

Inhaltsverzeichnis

A.2	Beweis – Lemma 4.6.2	165
A.3	Beweis – Lemma 4.8.1	166
B	Gleichgewichte	167
B.1	Beweis – Lemma 5.3.1	167
B.2	Fehlerhafte Reputationssysteme	168
B.3	Optimale Positionen zusätzlicher Kontakte	169
B.4	Nützlichkeit der DSC-Strategie	175
B.5	Beweis – Lemma 5.3.3	176
B.6	Beweis – Lemma 5.3.4	176
C	Experimentdaten	177
C.1	Auszahlung	177
C.2	Zusätzliche Kontakte	179
C.3	Kooperationsgrad	189

Abbildungsverzeichnis

3.1	Zweidimensionales Content-Addressable Network	24
4.1	Einfluss der Teilnehmer mit einer DPC-Strategie auf die C4C-Strategie	63
4.2	Einfluss der Netzwerkgröße auf die C4C-Strategie	63
4.3	Einfluss an Freifahren auf die C4C-Strategie	64
5.1	Ring mit sechs Teilnehmern	74
5.2	Weiterleitungswahrscheinlichkeit ohne zusätzliche Kontakte	82
5.3	Weiterleitungswahrscheinlichkeit mit drei zusätzlichen Kontakten	83
6.1	Einfluss des Parameters r	96
6.2	Interaktionswahl-Strategie im evolutionären Prozess	97
6.3	Kontaktwahl-Strategie im evolutionären Prozess	97
6.4	Erfolgreiche Interaktionswahl-Strategien aus zufälligen Strategien	101
6.5	Erfolgreiche Kontaktwahl-Strategien aus zufälligen Strategien	102
7.1	Initiale Netzwerkstruktur	113
7.2	Software-Client	115
7.3	Kontaktwahl Fenster	117
7.4	Erfolgsrate	123
7.5	Anzahl an Kontakten für die Experimentreihe T_0 und T_1	124
7.6	Anzahl an Kontakten für die Experimentreihe T_3 und T_5	125
8.1	Architektur des PostGame-Rahmenwerks	138
8.2	PostGame Client	140
8.3	Adressraum mit Kontakten	142
8.4	Anzahl an Weiterleitungen relativ zur Adressraum-Distanz der Anfrage	147
8.5	Auszahlung pro Anfrage relativ zur Anzahl an verwendeten Axelrod-Kriterien	147
8.6	Kumulativer Anteil der genutzten Axelrod-Kriterien	148
8.7	Erfolgsrate relativ zur Adressraum-Distanz der Anfrage	149
8.8	Erfolgsrate in Abhängigkeit zur Anzahl an zusätzlichen Kontakten	150
A.1	Teilnehmer, die äquidistant zum Ziel der Anfrage sind	162

1 Einleitung

Verteilte Systeme wie Peer-to-Peer Systeme oder Soziale Netzwerke haben große gesellschaftliche sowie wirtschaftliche Bedeutung. Ihre Teilnehmer können (1) eingehende Anfragen weiterleiten und bearbeiten – oder sie ignorieren, und (2) sie können ihre Kontakte wählen. Sie haben also eine Interaktionswahl- und eine Kontaktwahl-Strategie. Interaktionen von Teilnehmern werden maßgeblich durch Anfragen beziehungsweise die Bearbeitung von Anfragen beeinflusst. Eine Anfrage in einem Sozialen Netz kann z.B. sein: „Kennst Du jemanden, der mir einen Gebrauchtwagen verkaufen kann, wenn ich im Herbst nach Johannesburg ziehe?“ In einem Peer-to-Peer System haben Anfragen oft die folgende Form: „Wer verwaltet das Datenobjekt mit dem Schlüssel ‘9’?“ Bisher war unklar, wie sich Teilnehmer solcher Netze, seien es ihre Auszahlung maximierende Agenten oder menschliche Teilnehmer, verhalten und ob ihr Verhalten zu effizienten, also kooperativen und optimal vernetzten, Systemen führt. Diese Arbeit analysiert das Verhalten unter anderem mit formalen und evolutionären Modellen sowie verhaltensökonomischen Experimenten. Ein zentrales Ergebnis ist, dass menschliche Teilnehmer sich kooperativ verhalten und Netzwerke bilden, die zu einem hohen Gemeinwohl führen, also zu Systemen, in denen die Summe der Auszahlungen der Teilnehmer hoch ist. Es konnte ebenfalls gezeigt werden, dass Kooperation aus zufälligen Strategien entstehen kann, wenn Agenten ihre Strategien mittels evolutionärer Anpassungen verbessern. Das entstehende kooperative und nahezu optimal vernetzte System ist dabei in einem Gleichgewicht. Das Gemeinwohl kann durch in dieser Arbeit vorgeschlagene Strategien weiter erhöht werden. Die erzielten Ergebnisse sind wichtig, da sie zum einen eine Erklärung für kooperatives Verhalten in Systemen mit Interaktions- und Kontaktwahl geben, und da sie es zum anderen Entwicklern solcher Systeme ermöglichen, das Gemeinwohl im System in Richtung Effizienz zu erhöhen.

1.1 Verteilte Systeme

Verteilte Systeme wie Peer-to-Peer Systeme (Peer engl. für „Gleichgestellter“) oder Soziale Netzwerke haben eine große wirtschaftliche und kulturelle Bedeutung. So machen Peer-to-Peer Systeme neben ihrer enormen Verbreitung einen Großteil des Datentransfers im Internet aus [SGG02]. Soziale Netzwerke wie Facebook, Orkut, LinkedIn oder Xing haben Millionen von Teilnehmern [MMG⁺07].¹ Neben diesen

¹vgl. www.facebook.com, www.orkut.com, www.linkedin.com und www.xing.com

technischen verteilten Systemen gibt es eine große Anzahl von nicht-technischen Systemen, die als soziales Netz aufgefasst werden können. Beispiele für solche nicht-technische Soziale Netzwerke sind Freundschaftsbeziehungen in einer Stadt, internationale Firmenbeziehungen oder auch der Informationsfluss in Organisationen.

1.2 Strategien

Soziale Netzwerke und Peer-to-Peer Systeme haben gemeinsam, dass sie im Allgemeinen keine herausgehobene zentrale Kontrollinstanz, den sogenannten *Koordinator*, der die Teilnehmer² des Systems kontrolliert, besitzen. Des Weiteren können ihre Teilnehmer wählen, (1) wie sie mit ihren Kontakten interagieren und (2) welche Kontakte sie haben. Mit anderen Worten, ein Teilnehmer eines solchen Systems hat eine *Interaktionswahl-Strategie* und eine *Kontaktwahl-Strategie*.

1.2.1 Interaktionswahl

Ein wichtiger Bestandteil der Interaktionen von Teilnehmern ist die Anfragebearbeitung. Eine Anfrage kann eine einfache Frage sein wie „Kennst Du jemanden, der mir eine Wohnung vermieten kann, wenn ich im nächsten Frühling in Hamburg bin?“ in einem Sozialen Netzwerk oder eine Daten-Abfrage in einem Peer-to-Peer System wie „Wer hat das Datenobjekt mit Schlüssel ‘123’?“ Üblicherweise interagieren der Teilnehmer, der eine Anfrage erzeugt, und derjenige, der die Anfrage beantwortet, nicht direkt miteinander, sondern über Zwischenknoten. Zwischenknoten sind Teilnehmer, die die Anfrage weiterleiten, bis sie einen Teilnehmer erreicht, der die Anfrage beantworten kann oder bis die Anfrage verworfen wird. Eine Interaktionswahl-Strategie beschreibt, an welche Kontakte eigene Anfragen gesendet werden und von welchen Kontakten welche Anfragen weitergeleitet oder beantwortet beziehungsweise fallengelassen werden, oder abstrakt, ob mit einem Kontakt kooperiert wird oder nicht.

1.2.2 Kontaktwahl

Der Nutzen, den ein Teilnehmer vom System hat, hängt vom Verhalten der Teilnehmer ab, mit denen er interagiert, d.h. vom Verhalten seiner Kontakte und damit von der Netzwerkstruktur. Viele technische strukturierte Systeme haben eine *feste* Netzwerkstruktur [LCP⁺05], bei der die Kontakte eines Teilnehmers fest vorgegeben sind. Diese Kontakte werden auch als *feste Kontakte* bezeichnet.

Beispiele: Im Peer-to-Peer System Chord [SMK⁺01] formen N Teilnehmer einen Ring. Die Kontakte eines Teilnehmers i sind wie folgt. Feste Kontakte sind zum einen diejenigen Teilnehmer, die die Position rechts und links auf dem Ring neben

²Teilnehmer, Spieler, Peer und Subjekt werden in dieser Arbeit synonym verwendet.

Teilnehmer i einnehmen. Zum anderen sind die Teilnehmer, die $2^{i-1} \bmod 2^N$ (mit $1 \leq i \leq N$) Weiterleitungsschritte auf dem Ring entfernt sind, feste Kontakte. In Tapestry und Pastry [LCP⁺05] sind feste Kontakte durch ein Plaxton-Gewebe – also einer Baum-Struktur – definiert. Ein Teilnehmer des Netzes hat auf jeder Ebene des Baumes feste Kontakte. Content-Addressable Networks [RFH⁺01] nutzen einen n -dimensionalen Torus. In diesem sind Teilnehmer mit adjazenten Zonen, also mit solchen Zonen, die eine gemeinsame Kante besitzen, feste Kontakte. ■

Die Entwickler von verteilten Systemen mit fester Netzwerkstruktur haben die Annahme, dass die Teilnehmer ihre Kontakte nicht erweitern beziehungsweise austauschen oder verwerfen. Diese Annahme ist problematisch und die Argumentation ist wie folgt: Teilnehmer sind entweder Menschen oder Software-Artefakte, die von Menschen ausgesucht und kontrolliert werden. Beispielsweise ist für populäre Dateitauschbörsen Protokolle bekannt, dass die Nutzer aus vielen Programmen, die diese Protokolle unterstützen, wählen können. Des Weiteren können stets neue Programme entwickelt werden beziehungsweise existierende Programme im Interesse des Nutzers manipuliert werden. Diese Programme können sich in ihrem Verhalten von den Annahmen der Systementwickler bedeutend unterscheiden. Daher ist es nur folgerichtig anzunehmen, dass sich die Teilnehmer so verhalten, wie es ihnen am nützlichsten ist. Aus diesem Grund ist es natürlich, die Teilnehmer eines Systems ihre Kontakte wählen zu lassen. In anderen Worten, diese Arbeit nimmt an, dass sich die Teilnehmer der oben genannten Systeme nicht so verhalten, wie es ihre Systementwickler angenommen haben, sondern dass sie eine Kontaktwahl-Strategie besitzen.

Kontaktwahl beschreibt, welche Eigenschaften, z.B. welche Position im Netz oder welchen Kooperationsgrad, ein Knoten haben muss, damit er als Kontakt ausgewählt wird. Die Wahl von Kontakten bietet eine strategische Perspektive: Wenn Teilnehmer (ausschließlich) kooperative Kontakte wählen, dann haben unkooperative Teilnehmer Probleme, sich zu vernetzen, und ihre Strategie wird unattraktiv.

1.3 Problemdefinition

Bisher war für Systeme, in welchen Interaktionen über Weiterleitungsketten erfolgen unklar, welches Verhalten Teilnehmer in der Kombination aus Interaktionswahl und Kontaktwahl zeigen und ob die Gesamtheit der unterschiedlichen Strategien zu einem effizienten System führt oder ob z.B. unkooperatives Verhalten sich im System durchsetzt. Teilnehmer eines verteilten Systems erwarten sich von ihrer Teilnahme einen Nutzen. Aus der Sicht des Systementwicklers ist es wünschenswert, wenn der Gesamtnutzen des Systems maximal ist. Ein System ist dann *effizient*, wenn die Summe der Auszahlungen der Teilnehmer maximal ist, oder – angewandt auf diesen Kontext – wenn die Teilnehmer kooperieren und eine Netzwerkstruktur bilden, die eine Anfragebearbeitung mit wenigen Weiterleitungen erlaubt. In einem effizienten System ist die Summe des Nutzens der Teilnehmer maximal, es gilt aber nicht zwangsläufig,

dass daraus folgt, dass der Nutzen eines einzelnen Teilnehmers ebenfalls maximal ist.

Die Klärung dieser Fragestellung ist wichtig und bedeutsam. Ist sie beantwortet, so können Systementwickler erkennen, ob ihre Systeme effizient sind oder nicht. Gegebenenfalls können sie Mechanismen oder Strategien einsetzen, um das Gemeinwohl im System bis zur Effizienz zu erhöhen. Die Lösung der Problemstellung ist wichtig, da Systeme, die nicht effizient sind, die Ressourcen, die durch die Teilnehmer verteilter Systeme bereitgestellt werden, nicht optimal nutzen. Dies ist bei effizienten Systemen der Fall.

1.4 Strategisches Verhalten

Da die Teilnehmer von Peer-to-Peer Systemen und Sozialen Netzwerken autonom über ihre Strategien entscheiden – entweder weil sie aus unterschiedlichen Client-Programmen wählen können, oder weil sie direkt mit anderen Teilnehmern interagieren – ist eine Analyse der Strategien der Teilnehmer nötig. Diese ordnet Aktionen und Ereignissen Kosten beziehungsweise Einnahmen zu, um Strategien vergleichen zu können. Eine strategische Analyse ist aus mehrererlei Hinsicht schwierig durchzuführen.

1. Aus der Literatur ist bekannt, dass im Allgemeinen die Analyse von Interaktionswahl oder Kontaktwahl mit n Spielern schon als Einzelproblem algorithmisch komplex ist [RT02, VR06].
2. Eine formale Analyse wird dadurch erschwert, dass die Teilnehmer nur unvollständiges/ungenaueres Wissen über den Kooperationsgrad anderer Teilnehmer besitzen. Z.B. ist es eine Annahme dieser Arbeit, dass eventuell vorhandene Reputationssysteme – also Systeme zur Bereitstellung von Informationen über den Kooperationsgrad der Teilnehmer eines Systems – mit einem parametrisierbaren Fehler behaftet sind.
3. Erschwerend kommt hinzu, dass die Teilnehmer nur eine eingeschränkte Sicht auf das Netzwerk besitzen. Sie können lediglich den Ausgang ihrer Interaktionen beobachten.

In der Summe machen die Punkte die Analyse der Interaktions- und Kontaktwahl herausfordernd.

1.5 Ziel der Arbeit

Ziel der Arbeit ist es, das Verhalten von Teilnehmern in Systemen mit Interaktions- und Kontaktwahl zu verstehen, und es zur Effizienz des Systems in Verhältnis zu setzen. Falls das Einzelverhalten der Teilnehmer nicht zu effizienten Systemen führt,

sollen Mechanismen oder Strategien entwickelt werden, die das Gemeinwohl erhöhen. Das Erreichen dieses Ziels birgt mehrere Herausforderungen. Zum einen handelt es sich um eine fachübergreifende Thematik, bei der Techniken und Wissen aus der Informatik und der Ökonomie kombiniert werden müssen. Zum anderen muss beachtet werden, dass ihre Auszahlung maximierende Agenten und menschliche Teilnehmer sich unterschiedlich verhalten können. Dies macht die Verwendung von verschiedenen Analysetechniken notwendig.

1.6 Vorgehen

In dieser Arbeit werden unterschiedliche wissenschaftliche Techniken benutzt, um das Problem zu analysieren. Simulationen – sowohl klassische als auch Simulationen aus der evolutionären Spieltheorie – und formale Analysen werden durch verhaltensökonomische Experimente und eine Studie eines Online-Spiels ergänzt, um einen umfassenden Eindruck zu gewinnen.

1.7 Beiträge der Dissertation

Diese Arbeit hat die folgenden wissenschaftlichen Beiträge.

DPC vs. C4C Der Unterhalt von Kontakten bringt in der Regel Kosten mit sich. Daher untersucht diese Arbeit analytisch, wie die Kosten für Senden, Weiterleiten oder Beantworten von Anfragen, der Unterhalt von Kontakten und der Nutzen für das Erhalten einer Antwort, die Wahl von Kontakten beeinflusst. Ebenfalls wird untersucht, welchen intrinsischen Einfluss Teilnehmer, die fremde Anfragen fallen lassen, auf die Kontaktwahl haben.

Ergebnis der Analyse ist, dass unter bestimmten Parameterverhältnissen eine neue Variante des Freifahrerverhaltens (engl. *free riding*) – die sogenannte *freiwillig geringe Vernetztheit* (engl. *Deliberately Poor Connectedness (DPC)*) – lohnender ist als kooperatives Verhalten: Ein Teilnehmer, der eine DPC-Strategie verfolgt, sucht sich einen gut vernetzten Kontakt k und hat ansonsten keine Kontakte. Er spart sich dadurch Kosten für den Unterhalt von Kontakten und hat kaum höhere Kosten: Für seine Anfragen ergibt sich eine Pfadlänge, die nur um eins höher ist als die Pfadlänge der Anfragen von Kontakt k . Da *Deliberately Poor Connectedness* lohnender ist als kooperatives Verhalten, ist zu befürchten, dass sich diese Strategie durchsetzt, wenn keine Gegenmaßnahmen genutzt werden. Dies würde zu keiner effizienten Netzstruktur führen. Daher wurde eine Strategie, genannt *Kontakt-für-Kontakt* (engl. *Contact-for-Contact (C4C)*), entwickelt, welche den Nutzen von *Deliberately Poor Connectedness* so weit verringert, dass ein kooperatives Verhalten nützlicher ist. Die Idee der C4C-Strategie ist es, bei der Anfrageverarbeitung den Nutzen der eigenen Kontakte

KAPITEL 1. EINLEITUNG

für eine gegebene Anfrage ins Verhältnis zu setzen zu dem Nutzen der Kontakte des Teilnehmers, der einem die Anfrage zugesendet hat. Ein Teilnehmer leitet eine Anfrage nur so weit durch den Adressraum des Netzwerkes (also z.B. die Anschriften der Teilnehmer eines Sozialen Netzwerkes), wie es sein Vorgänger getan hat. Dadurch kann ein Teilnehmer nicht in größerem Umfang von den Kontakten anderer profitieren als von seinen eigenen. Eine Kosten-basierte Analyse und Simulationen zeigen, dass der Einsatz der C4C-Strategie den Nutzen von Deliberately Poor Connectedness so weit verringert, dass Teilnehmer, die ihre Auszahlung maximieren, eine kooperative Strategie Deliberately Poor Connectedness vorziehen werden. Dieses Ergebnis ist wichtig, da so eine neue Variante des Freifahrerverhaltens verhindert werden kann. Dies ist insbesondere deshalb bedeutend, da die Effizienz des Systems positiv durch die C4C-Strategie beeinflusst wird.

Gleichgewichte Entwickler technischer Systeme wie z.B. Peer-to-Peer Systeme geben ihren Teilnehmern ein Verhalten, z.B. in Form eines Programms zur Nutzung des Systems, vor. In Sozialen Netzwerken werden oft ebenfalls – meist tradierte – Verhaltensweisen vorgegeben. Um zu verstehen, welche Strategien ihre Auszahlung maximierende Teilnehmer beibehalten und bei welchen sie einen Anreiz haben, ihre Strategie abzuändern, ist eine ökonomische Analyse sinnvoll. Hier ist insbesondere das Gleichgewichtskonzept zu nennen: In einem Gleichgewicht hat kein Teilnehmer einen Anreiz, einseitig von seiner Strategie abzuweichen. Diese Arbeit untersucht, welche Gleichgewichte sich in der Kombination von Interaktions- und Kontaktwahl ergeben können.

Die Ergebnisse sind wie folgt: Zum einen existiert ein Zustand, bei dem kein hinreichender Nutzen mehr gegeben ist und die Teilnehmer das System verlassen. Dieser Fall tritt zum Beispiel ein, wenn Reputationsinformationen zu fehlerhaft sind. Zum anderen existiert ein Gleichgewicht, in dem die Teilnehmer kooperieren und ihre Kontakte gleichmäßig über das System verteilen. Dies führt zu einer nicht-effizienten Netzwerkstruktur. Ein Gleichgewicht mit einem höheren Gemeinwohl ergibt sich, wenn die Teilnehmer eine in dieser Arbeit vorgeschlagene Strategie verfolgen: Wenn Teilnehmer kooperative Kontakte danach wählen, wie schnell – also mit wie wenigen Weiterleitungen – diese in der Lage sind, ihre Anfragen zu verarbeiten, bildet sich eine Netzwerkstruktur, die um Größenordnungen effizienter ist als die im vorherigen Fall beschriebene. Dieses Ergebnis ist bedeutend, da das Gemeinwohl des Systems durch diese relativ einfache und damit mit nur geringen Zusatzkosten behaftete Strategie erheblich verbessert werden kann.

Kooperation aus zufälligen Strategien Mit Hilfe von Simulationen evolutionärer Strategieanpassungen gelingt es in dieser Arbeit zu zeigen, dass Kooperation aus zufälligen Strategien entstehen kann. Dieses Ergebnis ist bedeutsam: Während die Ergebnisse im vorhergehenden Paragraphen besagen, dass die Teilnehmer keinen

1.7. BEITRÄGE DER DISSERTATION

Anreiz besitzen, ein kooperatives Gleichgewicht zu verlassen, zeigen die Simulationsergebnisse, dass evolutionäre Strategieverbesserungen zufälliger Strategien in ein kooperatives Gleichgewicht führen können. Damit ist erstmals eine Erklärung für kooperatives Verhalten in realen Systemen mit Interaktions- und Kontaktwahl gegeben.

Verhalten in der Interaktions- und Kontaktwahl Verhaltensökonomische Experimente haben gezeigt, dass das Verhalten menschlicher Teilnehmer von spieltheoretischen Vorhersagen abweichen kann [CT94, Van02]. Diese Arbeit führt daher verhaltensökonomische Experimente mit insgesamt 120 Subjekten durch, die klären, wie sich menschliche Teilnehmer in der Kombination aus Interaktions- und Kontaktwahl verhalten. Die Ergebnisse sind insofern wichtig, als dass sie die Ergebnisse aus den formalen Analysen bestätigen und neue Erkenntnisse über das menschliche Verhalten hervorbringen. Sie sind bedeutsam, da die meisten verteilten Systeme direkt oder indirekt von menschlichen Teilnehmern beeinflusst werden. Die Ergebnisse der Experimente sind wie folgt:

1. Es gibt einen Zusammenhang zwischen Interaktions- und Kontaktwahl: Kontaktwahl erhöht die Kooperationsrate. Dies ist eine Voraussetzung für effiziente Netzwerke.
2. Teilnehmer lassen Kontakte fallen, die sich ihnen gegenüber unkooperativ verhalten haben. Dies ist eine Erklärung für das in (1) beschriebene Ergebnis.
3. Menschliche Teilnehmer passen ihre Strategien an externe Parameter, z.B. Unterhaltskosten für Kontakte, Weiterleitungskosten, etc. an: Z.B. wählen Teilnehmer umso mehr Kontakte, je billiger der Unterhalt für diese ist.
4. Die durch die Teilnehmer geformten Netzwerke mit einer bestimmten Parameterkonfiguration sind optimal relativ zu entstandenen Netzwerken aus Experimenten mit anderen Parameterkonfigurationen. (Dies gilt auch dann, wenn man die gleichen Parameterkonfigurationen virtuell in die Ergebnisse einrechnet.)

Neben verhaltensökonomischen Experimenten wurde eine Studie an Teilnehmern eines Online-Spiels durchgeführt. Diese bestätigt die oben genannten Ergebnisse. Diese Ergebnisse sind insofern bedeutsam, da sie zeigen, dass ein Systementwickler den Kooperationsgrad eines Systems allein dadurch anheben kann, dass er Kontaktwahl durch die Teilnehmer zulässt. Gleichzeitig kann er eine relativ effiziente Netzwerkstruktur erwarten.

Fazit Die hier zusammengefasste Arbeit erhöht das Verständnis von strategischem Verhalten in dezentralen Systemen, in denen Interaktionswahl und Kontaktwahl in Kombination auftreten. Systementwickler können die Erkenntnisse und Strategien zur Verbesserung des Gemeinwohls auf ihre Systeme anwenden und damit den Nutzen dieser Systeme entscheidend verbessern.

1.8 Gliederung der Dissertation

Kapitel 2 diskutiert verwandte Arbeiten. Kapitel 3 beschreibt die Grundlagen der Arbeit. Es folgen die Beiträge dieser Arbeit in den Kapiteln 4 bis 8: In Kapitel 4 werden die Nützlichkeit der Kontaktwahl und Anreiz-Strategien zur Kontaktwahl analysiert. In Kapitel 5 werden Gleichgewichte in der Kombination aus Interaktions- und Kontaktwahl identifiziert. Es folgen eine Untersuchung evolutionärer Strategieanpassungen (Kapitel 6) und verhaltensökonomische Experimente über das Verhalten menschlicher Teilnehmer (Kapitel 7). Kapitel 8 gibt die Ergebnisse einer Studie über das menschliche Verhalten in einem Online-Spiel mit Interaktions- und Kontaktwahl wieder. Kapitel 9 fasst die Arbeit zusammen und gibt einen Ausblick.

Teile dieser Arbeit wurden ursprünglich in gekürzter Form in [HBKS07, HBV08, HB10b, HB10a, Mil09] veröffentlicht.

2 Verwandte Arbeiten

Um Teile der Arbeit zu motivieren und um sich von anderen Arbeiten abzugrenzen, werden in diesem Kapitel verwandte Arbeiten diskutiert. Das Kapitel beginnt in Abschnitt 2.1 mit Studien, beschreibt dann Arbeiten aus der Spieltheorie sowie Anwendungen der Spieltheorie (Abschnitt 2.2) und endet mit Arbeiten aus der Verhaltensökonomie (Abschnitt 2.3). Die einzelnen Abschnitte sind in der Regel unterteilt in (a) Arbeiten, die sich mit der Interaktionswahl beschäftigen, (b) Arbeiten, die Kontaktwahl untersuchen, und (c) Arbeiten, die sowohl Kontaktwahl als auch Strategiewahl untersuchen. Wie im Folgenden jedoch deutlich werden wird, beschäftigen sich die allermeisten Arbeiten nur mit einem isolierten Thema. Da sowohl diese Arbeit als auch [AK01, BV04, ZL06] zeigen, dass sowohl die Interaktions- als auch die Kontaktwahl die Effizienz eines verteilten Systems beeinflussen, wird in dieser Arbeit argumentiert, dass es nicht ausreichend ist, nur einen Aspekt des Strategieraums – also Interaktionswahl oder Kontaktwahl – in realistischen Systemen zu untersuchen.

2.1 Studien

2.1.1 Studien zur Interaktionswahl

Freifahrerverhalten ist im Internet weit verbreitet und wird durch Studien [AH00, SGG02] belegt.

Adar und Huberman untersuchen Freifahrerverhalten in der Dateitauschbörse Gnutella [AH00]. Sie beobachten, dass 70% der Teilnehmer keine Dateien tauschen.

Saroiu et al. untersuchen die Dateitauschbörsen Gnutella und Napster [SGG02]. Neben der Beobachtung, dass viele Teilnehmer keine Dateien tauschen, stellen sie fest, dass die Teilnehmer falsche Angaben über ihre Fähigkeit, Dateien tauschen zu können, machen, um ihre Unkooperativität zu verschleiern.

Freifahrerverhalten auf Anfrageebene [AH00, CC05, FLSC04, GFJ⁺03] und Gegenmaßnahmen [BB04, FLSC04, JF04, RS07] wurden weithin untersucht.

Feldman et al. [FLSC04] schlagen eine sogenannte „reciprocative decision function“, also eine reziproke Strategie, als Gegenmaßnahme gegen unkooperatives Verhalten auf Anfrageebene vor. „Reziprok“ bedeutet hier, dass sich ein Teilnehmer unkooperativ gegenüber unkooperativen Teilnehmern verhalten soll und kooperativ gegenüber kooperativen Teilnehmern.

KAPITEL 2. VERWANDTE ARBEITEN

Jurca und Faltings [JF04] beschreiben einen anreizkompatiblen Mechanismus für ehrliches Feedback. *Feedback* ist eine Bewertung eines Teilnehmers über eine Aktion seiner Interaktionspartner. Die Idee hinter dem in [JF04] beschriebenen Mechanismus ist, dass beide an einer Aktion beteiligten Parteien eine Bewertung über den Ausgang der Aktion abgeben und dass das Feedback nur dann gewertet wird, wenn beide Parteien die gleiche Bewertung abgeben. Wenn die Bewertung der beiden Parteien unterschiedlich ausfällt, so muss eine Partei unehrlich berichtet haben. Der Ansatz beruht auf einer zentralen Infrastruktur und Seitenzahlungen: Beides ist nicht realistisch in verteilten Systemen. Der Entwurf eines verteilten Systems vermeidet zentrale Flaschenhalse oder zentrale Fehlerstellen grundsätzlich. Daher sind Knoten mit einer hervorgehobenen Rolle unakzeptabel. Da für zentrale Zahlungssysteme dasselbe Argument gilt, bleibt noch die Möglichkeit verteilter Zahlungssysteme. Diese haben sich bisher – möglicherweise auf Grund der durch sie entstehenden zusätzlichen Kosten – nicht in verteilten Systemen durchgesetzt.

Resnick und Sami [RS07] beschreiben einen manipulationsresistenten Algorithmus, der eingesetzt werden kann, um Feedback zu bewerten beziehungsweise um Feedback zu beurteilen. Teilnehmer, die nicht mit der Bewertung eines anderen Teilnehmers zufrieden sind, können das entsprechende Feedback bewerten.

Buchmann und Böhm [BB04] zeigen, dass Teilnehmer an Anfragen angehängtes Feedback nutzen können, um den Kooperationsgrad anderer Teilnehmer in der Interaktionswahl zu schätzen. Allerdings hat Feedback auch einige Nachteile. Schosser et al. [SBV07] zeigen, dass menschliche Teilnehmer kein Feedback in strukturierten Peer-to-Peer Systemen verwenden. Eine Erklärung dafür ist, dass sie ihre lokalen Beobachtungen ihrer Kontakte in der Aussagekraft höher bewerten und für manipulationsresistenter halten als Feedback. Des Weiteren sei angemerkt, dass Feedback Kosten verursacht: Es müssen entweder Nachrichten zur Verteilung des Feedbacks versendet werden, oder es müssen bestehende Nachrichten verlängert werden, um Feedback anzuhängen.

2.1.2 Studien zur Kontaktwahl

Christin und Chuang [CC05] führen eine Kosten-/Nutzen-Analyse von verschiedenen Overlay-Netzwerken durch. Es wird untersucht, ob die Overlay-Netze eine Struktur haben, die effizient ist, und ob die Teilnehmer einen Anreiz haben, diese Struktur beizubehalten, oder ob sie einen Anreiz haben, ihre Strategie zu ändern. [CC05] zeigt, dass es in vielen Systemen einen Anreiz gibt, von dem effizienten Systemzustand abzuweichen. Nur in Systemen, in denen das Weiterleiten einer Anfrage keine Kosten verursacht, entspricht die effiziente Netzwerkstruktur des Stern-Netzwerkes einem Gleichgewicht.

Chun et al. [CFSK04] analysieren, wie Agenten, also ihre Auszahlung maximierende Teilnehmer, Netzwerke formen. Jeder Teilnehmer hat Einkommen durch direkt und indirekt verbundene Teilnehmer, während Kontakte Kosten verursachen. Ein Re-

sultat der Arbeit ist, dass Netzwerke robuster gegen den Ausfall von Knoten oder gegen Attacken sind, wenn die Standardabweichung in der Anzahl an Kontakten klein ist: Wenn alle Teilnehmer die gleiche Anzahl von Kontakten (oder eine ähnliche Anzahl von Kontakten) haben, führt der Ausfall von k aus n Teilnehmern zu dem Ausfall von k/n Prozent der Netzwerkstruktur. Daraus folgt, dass in Systemen, in denen der Ausfall von Teilnehmern nicht ausgeschlossen werden kann, das Stern-Netzwerk ungeeignet ist. Dies gilt ebenfalls für Systeme, bei denen nicht ausgeschlossen werden kann, dass Teilnehmer sich unkooperativ verhalten. Ein Teilnehmer, der keine Anfragen bearbeitet, kann mit einem ausgefallenen Teilnehmer gleichgesetzt werden. Das Einkommen, das durch einen verbundenen Teilnehmer erzeugt wird, ist in [CFSK04] fest, d.h. [CFSK04] lässt die Untersuchung der Interaktionswahl aus.

Lugo und Jiménez [LJ06] nutzen Steuern und Subventionen, um den Kooperationsgrad und die Anzahl der Kontakte pro Teilnehmer zu beeinflussen. Dieser anreizkompatible Mechanismus ist in verteilten Systemen ungeeignet, da keine Koordinatorfreie Infrastruktur zur Erhebung von Steuern und zur Auszahlung von Subventionen in verteilten Systemen bekannt ist.

2.1.3 Studien zu Sozialen Netzwerken

[Mil67, DMW03, KB78] untersuchen Soziale Netzwerke und ihre Leistungsfähigkeit.

Milgram [Mil67] zeigt, dass Teilnehmer eines Experiments in den USA fähig waren, einen Brief an eine ihnen unbekannte Person mit nur ca. sechs Weiterleitungsschritten zuzustellen. Dieser Effekt wird Kleine-Welt-Effekt (engl. small-world effect) genannt und ist eine Eigenschaft von Sozialen Netzwerken. Er bedeutet, dass jeder Mensch über eine kleine Weiterleitungskette mit jedem anderen Menschen verbunden ist.

Killworth und Bernard untersuchen in [KB78] den Kleine-Welt-Effekt. Sie lassen Teilnehmer Listen über andere Teilnehmer ausfüllen, in denen sie begründen, warum sie wem eine Anfrage geben würden. Dabei standen die folgenden Kategorien zur Auswahl: Name, Standort, Beruf, Rasse oder Ethnie, Beziehung oder weiteres. 45% der Teilnehmer wurden wegen ihrer Position, 47% wegen ihres Berufs ausgewählt.

Dodds et al. [DMW03] wiederholen das Kleine-Welt-Effekt Experiment [Mil67] mit E-Mails statt mit Briefen. Im Einzelnen wurden einige Teilnehmer (aus über 60.000 Teilnehmern) gebeten, eine E-Mail an eine Person zu schicken, von der sie die Adresse nicht kannten. Sie sollten dabei die E-Mail lediglich an ihre Kontakte weiterleiten. [DMW03] bestätigt die Ergebnisse aus [Mil67] und stellt darüber hinaus fest, dass stark vernetzte Personen dabei keine wichtige Rolle spielen. Weiterhin stellen sie fest, dass vor allem Geschäftskontakte für die Aufgabe genutzt wurden.

A / B	B kooperiert	B defektiert
A kooperiert	1 Jahr / 1 Jahr	10 Jahre / 0 Jahre
A defektiert	0 Jahre / 10 Jahre	6 Jahre / 6 Jahre

Tabelle 2.1: Auszahlungsmatrix des Gefangenendilemmas

2.2 Spieltheorie und Anwendungen der Spieltheorie

2.2.1 Arbeiten zur Interaktionswahl

Spieltheoretische Modelle [Axe06, FM86, NS05, Now06] und Anwendungen der Spieltheorie [FLSC04] können genutzt werden, um zu verstehen, welche Strategien Teilnehmer in der Interaktionswahl wählen.

Das Gefangenendilemma [Axe06] modelliert eine Konfliktsituation: Zwei Gefangene, die gemeinsam eine Straftat begangen haben, sitzen nach der Verhaftung in unterschiedlichen Zellen. Unabhängig voneinander werden sie verhört. Schweigen beide, so bekommen sie aufgrund der Indizien eine geringe Gefängnisstrafe (z.B. ein Jahr). Beschuldigt ein Gefangener den anderen (und dieser schweigt), dann geht ersterer aufgrund der Kronzeugenregelung straffrei aus. Der andere hingegen bekommt eine hohe Strafe (z.B. zehn Jahre). Beschuldigen sich beide gegenseitig, bekommen sie eine mittlere Strafe (z.B. 6 Jahre). Die Auszahlungsmatrix eines Gefangenendilemmas ist in Tabelle 2.1 zu sehen. Das Dilemma des Gefangenen ist nun, dass er nicht wissen kann, wie sich sein Kollege verhält. Er muss daher annehmen, dass dieser diejenige Aktion wählt, die ihm am meisten nützt, und das ist in jedem Fall, dass der andere ihn verrät. Folglich muss auch er seinen Kollegen verraten. Aus Sicht der Gefangenen ist dieses Ergebnis zwar ein Gleichgewicht, aber kein effizienter Zustand, da – wenn sie beide geschwiegen hätten – sie im effizienten Zustand eine insgesamt kleinere Strafe hätten verbüßen müssen.

Während es im einfachen Gefangenendilemma nur ein Gleichgewicht gibt, nämlich defektieren [BEG06], kommen im wiederholt gespielten Gefangenendilemma mehrere Gleichgewichte vor. Das Folk-Theorem [FM86] über wiederholte Spiele impliziert, dass wenn die Wahrscheinlichkeit für weitere Runden im Spiel hoch genug ist, dass Kooperation aus sogenannten Trigger-Strategien entstehen kann. Ein Spieler mit einer Trigger-Strategie defektiert unnachgiebig, sobald der Mitspieler das erste mal defektiert hat [NS05]. Damit kann im wiederholten Gefangenendilemma der effiziente Zustand, d.h. Kooperation, ein Gleichgewicht sein.

Das Gefangenendilemma kann als Modell für die Interaktion der Endpunkte einer Weiterleitungskette angesehen werden. Diese interagieren in einem großen Netz nur mit geringer Wahrscheinlichkeit wiederholt. In dieser Situation macht es – folgt man dem einfach gespielten Gefangenendilemma – keinen Sinn, Anfragen zu bear-

2.2. SPIELTHEORIE UND ANWENDUNGEN DER SPIELTHEORIE

beiten. Allerdings sei angemerkt, dass Kontakte wiederholt miteinander interagieren. Feldman et al. [FLSC04] nutzen das Gefangenendilemma, um das Verhalten der Endpunkte einer Weiterleitungskette zu modellieren. Da die Endpunkte in großen Netzen in der Regel nur einmal miteinander kommunizieren, würde das auf sie angepasste Gefangenendilemma zu der Vorhersage führen, dass die Endpunkte defektieren, d.h. nicht miteinander kooperieren [FLSC04]. Nowak et al. [NS05, Now06] zeigen, dass Kooperation im wiederholten Gefangenendilemma durch unterschiedliche Maßnahmen entstehen kann: Durch indirekte Reziprozität und Reputation [NS05] sowie durch Netzwerk-Reziprozität. *Indirekte Reziprozität* bedeutet, dass Teilnehmer i nur dann mit Teilnehmer j kooperiert, wenn j sich bereits gegenüber anderen Teilnehmern kooperativ verhalten hat. Reputation in [NS05] hängt von den Beobachtungen eines Teilnehmers selbst sowie von den Meinungen von Dritten ab – also von Feedback.

Für die Endpunkte einer Weiterleitungskette kann also aus den bisher erörterten Arbeiten geschlossen werden, dass – ohne weitere Maßnahmen – Defektion zu erwarten ist, während für die wiederholt interagierenden Kontakte Kooperation möglich ist. Da die Endpunkte einer Weiterleitungskette jedoch über ihre Kontakte interagieren, ist es nicht trivial, zu bestimmen, ob und unter welchen Umständen die Teilnehmer des Systems kooperieren oder defektieren.

In den Modellen von [Axe06, NS05, Now06] existieren keine Weiterleitungsketten. Diese sind jedoch in den meisten realen Systemen vorhanden. Gottlob et al. zeigen in [GGS03], dass die analytische Untersuchung von Spielen mit Zwischenknoten schwierig ist.

Roughgarden und Tardos [RT02] analysieren die Effizienz von Systemen, in denen Teilnehmer autonom eine Interaktionswahl-Strategie verfolgen, mit der Effizienz von zentralen Systemen: Die Latenz der Anfragebearbeitung ist nahe dem Optimalzustand, selbst dann wenn die Knoten nach eigenem Interesse handeln. Es kann daraus gefolgert werden, dass keine zentrale Instanz nötig ist, um Anfragen durch ein System zu leiten. Im Gegensatz zu der hier vorgestellten Arbeit, wird in [RT02] davon ausgegangen, dass die Netzwerkstruktur fest ist.

Valiant und Roughgarden zeigen in [VR06], dass die Analyse von n -Spieler Problemen schwierig ist, wenn die (sich verändernde) Netzwerkstruktur dabei berücksichtigt wird. Alle bisher genannten Ansätze untersuchen Kontaktwahl nicht. Es sei jedoch angemerkt, dass aus [BV04] bekannt ist, dass die Netzwerkstruktur den Ausgang eines Spiels beeinflusst. Daher werden im folgenden Arbeiten diskutiert, die Netzwerkbildung untersuchen.

2.2.2 Arbeiten zur Kontaktwahl

Eine der ersten Arbeiten auf dem Gebiet der Netzwerkbildung (aka. Kontaktwahl) ist das *Connections Model* [JW96]. Im Connections Model haben Teilnehmer positiven Nutzen, wenn sie mit anderen Teilnehmern direkt oder indirekt – also über mehrere

KAPITEL 2. VERWANDTE ARBEITEN

Zwischenstationen – verbunden sind. Je weniger Zwischenstationen auf dem kürzesten Pfad zwischen zwei Knoten vorhanden sind, desto mehr profitieren die beiden Teilnehmer von einander. Verbindungen, also Kontakte, verursachen Kosten. Sei N die Anzahl an Teilnehmern im System und δ die Einnahmen, die durch einen Kontakt verursacht werden. Seien δ^2 die Einnahmen, die durch einen Teilnehmer, der über eine Zwischenstufe erreicht werden kann, erzeugt werden (also die Einnahmen durch einen Kontakt eines Kontakts). Ein Kontakt erzeugt Kosten c . Dann sind im Connections Model die drei folgenden Netzwerk-Architekturen effizient:

1. Das *komplette Netzwerk* (engl. *complete network*) ist effizient, wenn gilt, dass die Verbindungskosten kleiner sind als der Wertverlust durch einen zusätzlichen Zwischenknoten ($c < \delta - \delta^2$).
2. Bei einem *Stern-Netzwerk* (engl. *star network*) ist ein Teilnehmer mit allen anderen Teilnehmern direkt verbunden und es existieren sonst keine Verbindungen. Das Stern-Netzwerk ist effizient, wenn die Kontaktkosten höher sind als der Wertverlust durch einen zusätzlichen Zwischenknoten und wenn gleichzeitig gilt, dass sich Verbindungen überhaupt lohnen. Es muss also die folgende Formel gelten:

$$\delta - \delta^2 < c < \delta + \frac{N - 2}{2} \cdot \delta$$

3. In allen anderen Fällen ist das *leere Netzwerk* (engl. *empty network*) effizient. Dieses besitzt keine Verbindungen.

Andere Netzwerkarchitekturen sind im Connections Model nicht effizient.

Im Netzwerkmodell von Bala und Goyal [BG00b] erzeugen Knoten mit denen ein Teilnehmer (indirekt oder direkt) verbunden ist Gewinne. „Verbunden sein“ wird durch zwei Modelle unterschiedlich aufgefasst. Im „one-way-flow“-Modell sind Verbindungen gerichtet und nur der Knoten, der eine Verbindung vorschlägt (und seine Vorgänger im Graphen) profitieren davon. Im „two-way-flow“-Modell sind Verbindungen ungerichtet, und beide Knoten einer Verbindung (und ihre Vorgänger im Graphen) profitieren von einer Verbindung. Direkte Verbindungen erzeugen Kosten bei dem Teilnehmer, der die Verbindung vorgeschlagen hat. Bala und Goyal zeigen, dass das Rad (one-way-flow Modell) oder der zentral gesponserte Stern (two-way-flow Modell) sowie das leere Netzwerk strikte Nash-Gleichgewichte sind. Ein zentral gesponserter Stern ist ein Stern, bei dem alle Verbindungen vom Zentrum des Sterns vorgeschlagen worden sind. Wie in [Jac03] werden auch in [BG00b] keine Interaktionswahl-Strategien untersucht.

Epstein et al. [EFM07] zeigen, welche Topologien *starke Gleichgewichte*, also solche, in denen auch Gruppen von Teilnehmern ihre Auszahlungen nicht verbessern können, darstellen. Sie untersuchen dabei ein Netzwerkmodell, bei dem jeder Spieler eine Quelle (d.h. den Ursprung des Informationsflusses) und eine Senke (d.h. das

2.2. SPIELTHEORIE UND ANWENDUNGEN DER SPIELTHEORIE

Ziel des Informationsflusses) kennt, mit denen er sich verbinden will. Epstein et al. zeigen, (a) dass für eine einzelne Quelle und Senke stets ein starkes Gleichgewicht existiert und (b) dass für eine Quelle und n Senken für eine Menge von parallelen Graphen starke Gleichgewichte existieren.

Die in diesem Abschnitt diskutierten Arbeiten grenzen sich von dieser Arbeit dadurch ab, dass sie neben der Netzwerkbildung nicht die Interaktionswahl untersuchen. So ist zum Beispiel das Weiterleiten von Anfragen ein wichtiger Bestandteil der Interaktionswahl. Dieser Aspekt wird von den Arbeiten außer Acht gelassen.

2.2.3 Kombination von strategischen Spielen und Kontaktwahl

Es ist nicht ausreichend, Interaktionswahl und Kontaktwahl isoliert zu untersuchen. Dies wird im Verlauf dieses Abschnitts deutlicher werden, in dem Arbeiten diskutiert werden, die in ihren Modellen die Kombination aus Kontaktwahl und ökonomischen Spielen untersuchen.

Berninghaus und Vogt [BV04] untersuchen Kontaktwahl in Kombination mit zwei 2x2 Spielen. In einem 2x2 Spiel können zwei Spieler aus jeweils zwei Aktionen wählen. In [BV04] können sich Teilnehmer einer Gruppe von Teilnehmern aussuchen, mit welchem anderen Teilnehmer sie das 2x2 Spiel spielen wollen. [BV04] zeigt, dass die Netzwerkstruktur von den Verbindungskosten und der Verteilung an Strategien im System abhängt. Folglich sollten Kontaktwahl und die Analyse ökonomischer Spiele gemeinsam untersucht werden. Oder in anderen Worten: Es ist unklar, ob wissenschaftliche Ergebnisse, die aus Systemen mit ausschließlich festen Kontakten abgeleitet wurden, gelten, wenn die Teilnehmer der Systeme ihre Kontakte wählen können.

Zhang und Liu [ZL06] untersuchen die Dynamik von Netzwerken, die aus Teilnehmern mit unterschiedlichen Strategien bestehen. Im Einzelnen können die Teilnehmer die folgenden Strategien nutzen:

1. Sie wählen jeweils den Teilnehmer als Kontakt, der in der Vergangenheit am kooperativsten bei der gemeinsamen Verarbeitung von Anfragen gewesen ist.
2. Sie wählen den Teilnehmer, mit dem sie die höchste Auszahlung generiert haben.
3. Es wird der Teilnehmer als Kontakt gewählt, der am kooperativsten gewesen ist.
4. Sie wählen den Teilnehmer, der die meisten Kontakte besitzt.

Das Ergebnis der in [ZL06] durchgeführten Analyse ist, dass Strategie (3) am besten für das Gemeinwohl des Systems ist.

Goyal und Vega-Redondo [GVR05] zeigen, dass die Kosten zum Unterhalt von Kontakten die Strategien eines gespielten Koordinationsspieles beeinflussen. Bei diesem Koordinationsspiel haben Teilnehmer einen Gewinn, wenn sie die gleiche Aktion durchführen und einen Verlust, wenn sie unterschiedliche Strategien verwenden.

Abramson und Kuperman [AK01] zeigen, dass das Verhalten von Teilnehmern sich ändert, wenn sich die Netzwerkstruktur ändert. Dies zeigt, dass Interaktionswahl und Kontaktwahl gemeinsam untersucht werden sollten.

Viele Arbeiten berücksichtigen nicht, dass ein Teilnehmer erfolgreich den Kooperationsgrad anderer Teilnehmer einschätzen muss, damit er auf unkooperatives Verhalten reagieren kann. In Situationen, in denen der Kooperationsgrad anderer Teilnehmer nicht eingeschätzt werden kann, ist von vielen Freifahrern auszugehen. Aus Sicht eines kooperativen Teilnehmers erzeugen diese Teilnehmer keinen Nutzen und können mit ausgefallenen Knoten verglichen werden. Bala und Goyal [BG00a] untersuchen Kontaktwahl in der Gegenwart von unzuverlässigen Teilnehmern. In vielen Situationen ist es nach [BG00a] optimal, wenn sich „superconnected networks“ bilden – also Netzwerke, die mehr Verbindungen enthalten, als jene, die unter der Abwesenheit von unzuverlässigen Knoten optimal wären. Ein Teilnehmer benötigt also redundante Verbindungen, d.h. zusätzliche Kontakte, in unzuverlässigen Netzwerken.

2.3 Verhaltensökonomie

2.3.1 Verhaltensökonomische Arbeiten zur Interaktionswahl

Schossler et al. [SBV07] nutzen Laborexperimente, um zu zeigen, dass die Teilnehmer eines verteilten Informationssystems zur Anfragebearbeitung in der Interaktionswahl *indirect partner interaction* – also indirekte Kommunikation über Kontakte – nutzen. D.h. sie senden ihre Anfragen über eine Kette von Kontakten. Dies erhöht den Kooperationsgrad im System, da die Kontakte wiederholt miteinander interagieren und aufeinander angewiesen sind.

Keser und van Winden untersuchen, welchen Einfluss wiederholtes Interagieren im Vergleich zu einfachem Interagieren hat [KvW00]. Sie untersuchen Gruppen von Teilnehmern in einem Öffentliche-Güter-Spiel [BEG06]. Das Ergebnis der Untersuchung ist, dass die Anzahl von Freifahrern beim einfachen Interagieren signifikant höher liegt als beim wiederholten Interagieren. Dies gibt bereits einen Hinweis darauf, dass wiederholte Interaktionen zwischen Kontakten zu Kooperation führen kann.

Ule et al. [USRC09] untersuchen, welchen Einfluss Bestrafung unkooperativer Teilnehmer und Belohnung kooperativer Teilnehmer auf den Kooperationsgrad von Teilnehmern im Helping Game hat. Beim Helping Game entscheidet ein Spieler (der Spender), ob er einem anderen (zugelosten) Spieler (dem Empfänger) hilft oder nicht. Helfen erzeugt Kosten beim Spender und Einnahmen beim Empfänger, wobei die

Einnahmen die Kosten übertreffen. Beim Experiment wird berücksichtigt, dass Teilnehmer großer Systeme sich oft nicht kennen und nur einmal miteinander interagieren. Die Autoren finden heraus, dass in Systemen, in denen es nicht möglich ist, Teilnehmer zu bestrafen, Freifahrerverhalten am lohnendsten ist. Ist es jedoch möglich, Teilnehmer zu bestrafen, so führt die Strategie, die kooperative Teilnehmer belohnt, zur höchsten Auszahlung.

Grimm und Mengel untersuchen in einem Gefangenendilemma-Szenario wie die Auswahl von Gruppen den Kooperationsgrad beeinflusst [GM09]. Die Experimentteilnehmer können im Experiment zwischen zwei Gruppen wählen. In der ersten ist der Gewinn des Defektors klein, in der zweiten Gruppe groß. Die Erwartung an den Experimentaufbau ist, dass kooperative Teilnehmer eher die erste Gruppe bevorzugen, während unkooperative Teilnehmer die zweite Gruppe wählen. Die Autoren beobachten, dass je größer der Unterschied im Gewinn für die Defektoren zwischen den Gruppen ist, desto klarer trennen sich kooperative und unkooperative Teilnehmer auf die Gruppen auf.

2.3.2 Verhaltensökonomische Arbeiten zur Kontaktwahl

[JK08, RW90] untersuchen wie gegebene (feste) Netzwerkstrukturen die strategischen Spiele der Teilnehmer beeinflussen.

Judd und Kearns [JK08] zeigen für gegebene Netzwerkstrukturen, dass die Spieler sich unterschiedlich verhalten in Abhängigkeit der gegebenen Netzwerkstruktur. Insbesondere beeinflusst die Netzwerkstruktur die Effizienz des Systems.

Raub und Weesie [RW90] zeigen, dass die Position eines Knotens im Netzwerk beeinflusst, wie schnell sich seine Reputation im Netzwerk verbreitet. Dies hat einen Einfluss auf die Ergebnisse der Interaktionen. Daher wird gezeigt, dass die Netzwerkstruktur Interaktionswahl-Strategien beeinflusst.

Weitere Arbeiten untersuchen, wie Menschen Netzwerke bilden (ohne dabei die Interaktionswahl der Teilnehmer zu berücksichtigen).

Vanin [Van02] untersucht in einer Studie das menschliche Verhalten im Connections Model [Jac03]. Es werden dabei Gruppen von bis zu vier Studenten genutzt. Das Ergebnis der Studie ist, dass in manchen Gruppen Stern-Netzwerke beobachtet werden konnten, in anderen dagegen nicht.

Falk und Kosfeld [FK03] untersuchen das menschliche Verhalten im Netzwerkmodell von Bala und Goyal [BG00b]. Das Ergebnis ihrer Labor-Experimente ist, dass menschliche Teilnehmer im one-way-flow Modell Strategien verwenden, die zu den vorhergesagten strikten Nash-Gleichgewichten führen, dass dies aber im two-way-flow Modell nicht der Fall ist. Eine Erklärung, die Falk und Kosfeld für dieses Verhalten haben, ist, dass die Subjekte im one-way-flow Modell mit großer Dynamik lernen, während das im two-way-flow Modell nicht der Fall ist. Des Weiteren zeigen die Experimente in [FK03], dass Verbindungskosten einen positiven Effekt für das Auftreten strikter Nash-Gleichgewichte im one-way-flow Modell haben, während

KAPITEL 2. VERWANDTE ARBEITEN

Verbindungskosten einen negativen Effekt im two-way-flow Modell besitzen. Falk und Kosfeld erklären dies mit auftretenden Ungleichgewichten in den Auszahlungen der Teilnehmer.

Berninghaus et al. [BEOV07] untersuchen die Bildung von Stern-Netzwerken in kleinen Gruppen von je sechs Subjekten. Die Basis ihres Experiments ist das two-way-flow Modell von Bala und Goyal [BG00b]. Da Berninghaus et al. die Dynamik der Netzwerkbildung untersuchen, wenden sie zwei Szenarien an: Im ersten ist die Zeit diskret, d.h. die Subjekte spielen in Runden, im zweiten ist die Zeit kontinuierlich. Die Resultate der verhaltensökonomischen Experimente sind wie folgt: Im diskreten Fall führten die Strategien der Teilnehmer in 50% der Experimente zu Stern-Netzwerken. Im kontinuierlichen Fall konnten Stern-Netzwerke häufiger beobachtet werden. Des Weiteren tauschten die Subjekte die zentrale Position des Sterns durch. Dies führt zu ausgeglicheneren Auszahlungen und vermeidet die von Falk und Kosfeld beobachteten Ungleichgewichte in den Auszahlungen.

Goeree et al. [GRU09] untersuchen ebenfalls das Netzwerkmodell von Bala und Goyal [BG00b]. Sie bestätigen die Ergebnisse von Falk und Kosfeld, zeigen jedoch, dass Stern-Netzwerke gebildet werden, wenn einzelne Teilnehmer einen höheren Wert für das System besitzen. Führen Verbindungen zu einem bestimmten Teilnehmer zu (deutlich) höheren Einnahmen, so entstehen Stern-Netzwerke mit einer hohen Wahrscheinlichkeit.

2.3.3 Verhaltensökonomische Arbeiten zur Strategie- und Kontaktwahl

Neben verhaltensökonomischen Arbeiten, die entweder die Strategiewahl der Teilnehmer oder die Kontaktwahl der Teilnehmer untersuchen, existieren Arbeiten, die die Kombination von Strategiewahl und Kontaktwahl untersuchen [Ule05, BEO10, RU02, CD08, CB10].

Ule und Riedl untersuchen die Kombination aus Gefangenendilemma und Kontaktwahl [Ule05, RU02]. In dem in [Ule05, RU02] genutzten Netzwerk-Spiel wählen die Spieler pro Runde (a) ihre Kontakte und (b) ihre Aktion im Gefangenendilemma. Ein Spieler wählt dabei für alle seine Kontakte die gleiche Aktion im Gefangenendilemma. Das Ergebnis der Untersuchung ist, dass Teilnehmer eines solchen Szenarios einen hohen Kooperationsgrad zeigen. Die Autoren schließen aus diesem Ergebnis, dass Kontaktwahl ein geeigneter Mechanismus zur Steigerung der Kooperativität der Teilnehmer ist. Im Gegensatz zu dem in dieser Arbeit vorgestellten Ansatz, werden keine Weiterleitungen von Anfragen betrachtet. Interaktionen im System betreffen stets nur zwei Teilnehmer. Dies gilt auch für die im Folgenden vorgestellten Arbeiten.

Berninghaus et al. untersuchen Hawk-Dove Spiele und Kontaktwahl [BEO10]. Ein Hawk-Dove Spiel ist ein 2x2 Spiel, bei dem die Auszahlungsmatrix aus Tabelle 2.2

	Hawk	Dove
Hawk	d / d	a / c
Dove	c / a	b / b

Tabelle 2.2: Auszahlungsmatrix von Hawk-Dove Spielen; es gilt $a > b > c > d$

gilt. Ein Ergebnis der Untersuchung ist, dass die Subjekte auf neue strategische Konstellationen zuerst durch Veränderung ihrer Kontakte reagieren. Dies unterstreicht die Bedeutung der gemeinsamen Analyse von Strategie- und Kontaktwahl. Erneut sind die Interaktionen lediglich auf Interaktionen zwischen zwei Teilnehmern beschränkt.

Corbae und Duffy [CB10] sowie Corten und Buskens [CD08] untersuchen die Kombination aus Koordinationsspielen und Kontaktwahl. Beide Arbeiten zeigen, dass die kombinierte Analyse der Strategie- und Kontaktwahl zu anderen – d.h. realistischeren – Ergebnissen führt als die Analyse der Einzelgebiete.

2.4 Fazit

Neben der Vielzahl von Aussagen in diesem Kapitel sollen zwei erneut hervorgehoben werden, da sie im Folgenden als Motivation herangezogen werden.

Viele Untersuchungen stützen sich entweder auf die Analyse der Interaktionswahl-Strategien oder auf die Analyse von Kontaktwahl-Strategien. Wie Berninghaus und Vogt in [BV04] und Goyal und Vega-Redondo in [GVR05] zeigen, ist dies im Allgemeinen nicht ausreichend. Daher werden in dieser Arbeit Interaktionswahl-Strategien und Kontaktwahl-Strategien in Kombination untersucht.

Das Verständnis von Strategien, die durch Teilnehmer, die ihre Auszahlung maximieren, genutzt werden, ist fundamental für die Analyse verteilter Systeme. Neben der Analyse von Teilnehmern, die ihre Auszahlung maximieren, ist aber das Verständnis menschlichen Verhaltens entscheidend, da zum einen Menschen oft die handelnden Parteien in einem System sind oder da sie zum anderen Software auswählen, kontrollieren oder sogar in ihrem Interesse manipulieren, die für sie das Handeln im System übernimmt. Wie in [CT94] und [Van02] gezeigt wird, verhalten sich menschliche Teilnehmer nicht in jeder Situation so wie Teilnehmer, die ihre Auszahlung maximieren. Daher ist das Verständnis menschlichen Verhaltens und das Verständnis von Teilnehmern, die ihre Auszahlung maximieren, gleichermaßen wichtig.

Es sei an dieser Stelle angemerkt, dass obwohl Arbeiten existieren, die strategische Spiele und Kontaktwahl kombinieren, es ein Alleinstellungsmerkmal dieser Arbeit ist, dass die Interaktionswahl in Kombination mit der Kontaktwahl untersucht wird. Teilnehmer vieler Systeme wie Soziale Netzwerke oder Peer-to-Peer Systeme können ihre Interaktionswahl-Strategie wählen und entscheiden, wie und ob sie Anfragen beantworten, weiterleiten oder verwerfen. Daher ist eine Analyse der Interaktionswahl für das Verständnis solcher Systeme entscheidend.

KAPITEL 2. VERWANDTE ARBEITEN

3 Grundlagen

Ziel dieses Kapitels ist es, Grundlagen für die folgenden Kapitel zu legen. So werden Definitionen, Maße und Annahmen, die für alle Kapitel gelten, bereits hier vorgestellt. Abweichungen davon beziehungsweise Definitionen, die nur ein Kapitel betreffen, werden in dem jeweiligen Kapitel behandelt.

Das Kapitel beginnt mit grundsätzlichen Eigenschaften verteilter Systeme, wie zum Beispiel den Spielern, und beschreibt verteilte Systeme beispielhaft an Content-Addressable Networks [RFH⁺01]. Es folgen Annahmen, die dieser Arbeit zugrunde liegen (Abschnitt 3.2). Danach folgen die Aktionen der Spieler (Abschnitte 3.3 und 3.4), sowie deren Auszahlungsfunktion (Abschnitt 3.5). Das Kapitel schließt mit einer Diskussion der in der Arbeit verwendeten Methoden (Abschnitt 3.7). Diese motiviert gleichzeitig die Struktur der folgenden Kapitel.

3.1 Verteilte Systeme

Spieler Ein verteiltes System besteht aus und definiert sich über seine Teilnehmer, die im Folgenden auch als *Spieler* bezeichnet werden. Ein System habe n Teilnehmer. Um die Teilnehmer eines verteilten Systems unterscheiden zu können, wird in dieser Arbeit o.B.d.A. angenommen, dass $N = \{0, \dots, n - 1\}$ eine Menge von Identifikatoren ist, die die Teilnehmer des Netzwerkes unterscheidbar macht. Es sei die Kardinalität einer Menge mit $|\cdot|$ bezeichnet. Z.B. gilt, dass die Menge der Identifikatoren gleich der Menge der Teilnehmer im Netz ist: $|N| = n$. Bezieht sich eine in dieser Arbeit definierte Menge M auf einen Teilnehmer i , so sei sie mit dem Subskript i gekennzeichnet: M_i . Wird das Subskript weggelassen, so bezieht sich die Menge auf alle Teilnehmer.

Datenverwaltung in technischen Systemen Das Bearbeiten von Anfragen verursacht Kosten. Während das Erstellen einer Antwort von der Größe und der Struktur des Netzes unabhängig ist, hängt die Länge der Weiterleitungskette und damit die Gesamtkosten des Weiterleitens von der Größe und der Struktur des Netzes ab. Verteilte Systeme können danach kategorisiert werden, wie Anfragen in ihrem System verarbeitet werden. In einem unstrukturierten System hat ein Anfrager keinen Anhaltspunkt darüber, welcher andere Teilnehmer im System seine Anfrage beantworten kann [ATS04]. Allein aus diesem Grund ist die Anfragebearbeitung in unstrukturierten Systemen nicht effizient. Der Schwerpunkt dieser Arbeit liegt daher

auf strukturierten Systemen. Im Unterschied zu unstrukturierten Systemen haben die Teilnehmer strukturierter Systeme Anhaltspunkte darüber, welcher ihrer Kontakte eine Anfrage beantworten kann beziehungsweise welcher ihrer Kontakte gut geeignet ist, die Anfrage weiterzuleiten. In technischen Systemen erzeugt man dieses Wissen üblicherweise, indem man die Datenobjekte durch eine *verteilte Hashtabelle* (engl. *distributed hash table (DHT)*) verwaltet [CW79]. Dabei weist man jedem Datenobjekt einen Schlüssel zu [ATS04]. Eine übliche Methode ist die Verwendung einer öffentlich bekannten Hash-Funktion [CW79]. Mit Hilfe dieser Hash-Funktion wird jedem Datenobjekt ein *Schlüssel* aus dem durch die Zielmenge der Hash-Funktion definierten *Adressraum* zugeordnet. Der Schlüssel wird ebenfalls als die *Zielkoordinate der Anfrage* oder als das *Ziel der Anfrage* bezeichnet. Teilnehmer eines strukturierten Systems verwalten Abschnitte des Adressraums – also Teile der verteilten Hashtabelle. Alle Datenobjekte mit Schlüsseln, die Teil eines Abschnitts des Adressraums sind, werden durch einen Teilnehmer verwaltet. Ein solcher Abschnitt wird ebenfalls als *Zone* bezeichnet. Zusätzlich ist dem Teilnehmer bekannt, welche Abschnitte des Adressraums seine Kontakte verwalten. Dadurch kann er für eine gegebene Anfrage, d.h. für die Suchanfrage nach einem gegebenen Schlüssel, feststellen, welcher seiner Kontakte verantwortlich für den gesuchten Schlüssel ist beziehungsweise welcher seiner Kontakte den Abschnitt des Adressraums verwaltet, der dem Schlüssel am nächsten gelegen ist. Dadurch ist es den Teilnehmern eines Systems mit einer verteilten Hashtabelle möglich, Anfragen in jedem Weiterleitungsschritt näher an den Teilnehmer, der für den Schlüssel des gesuchten Datenobjektes verantwortlich ist, zu senden.

Content-Addressable Networks In diesem und in weiteren Abschnitten werden *Content-Addressable Networks (CAN)* [RFH⁺01] als Beispiel für strukturierte Peer-to-Peer Systeme genutzt. Die Struktur eines Content-Addressable Networks basiert auf einem n -dimensionalen Torus. Ein *Torus* ist eine Hyperebene, deren Kanten so um den Adressraum herumgebogen sind, dass sie sich wieder berühren. Dadurch hat ein sich auf dem Torus bewegendes Objekt den Eindruck, dass keine Begrenzung des Adressraumes existiert. Im Fall eines Content-Addressable Networks ist der Adressraum ein normierter Torus mit n Dimensionen. Da die Anschauung in höheren Dimensionen oft versagt, wird in Teilen dieser Arbeit ein zwei-dimensionales Content-Addressable Network als Beispiel verwendet, d.h. der Adressraum ist zwei-dimensional: $[0, 1] \times [0, 1]$.

Definition 3.1 (Adressraum-Distanz im CAN)

Sei $\Delta(a, b)$ die euklidische Distanz der Koordinate a und der Koordinate b im Adressraum:

$$\Delta(a, b) = \sqrt{\min(|a.x - b.x|, 1 - |a.x - b.x|)^2 + \min(|a.y - b.y|, 1 - |a.y - b.y|)^2}$$

Ein Teilnehmer eines strukturierten Peer-to-Peer Systems ist verantwortlich für die Verwaltung eines Unterraums des Adressraums. In Content-Addressable Networks verwalten die Teilnehmer Zonen. Eine Zone ist ein Abschnitt des Adressraums der Form $[i, j] \times [l, m]$ mit $0 \leq i < j \leq 1$ und $0 \leq l < m \leq 1$. Sei $zentrum(i)$ das Zentrum der Zone von Teilnehmer i . Teilnehmer mit Zonen, die adjazente Kanten besitzen, also solche Zonen, die eine Kante gemeinsam haben, sind feste Kontakte. Ein Teilnehmer, der eine Anfrage erzeugen möchte, transformiert diese mit Hilfe der Hash-Funktion auf eine Koordinate des Adressraums. Er erzeugt also das Ziel der Anfrage. Der Teilnehmer sendet die Anfrage dann an einen Kontakt, der dem Ziel der Anfrage am nächsten gelegen ist. In dieser Arbeit wird dieses Vorgehen im Gegensatz zum Vorgehen in Content-Addressable Networks aufgeweicht, so dass Kontakte gewählt werden können, die näher am Ziel der Anfrage liegen, als der Erzeuger oder der letzte Weiterleiter der Anfrage selbst. Dies erlaubt es den Teilnehmern, Kontakte nach strategischen Gesichtspunkten, z.B. dem Kooperationsgrad, auszusuchen. Das Weiterleiten wiederholen die Teilnehmer, bis die Anfrage ihr Ziel erreicht. Da das Ziel der Anfrage eindeutig ist, gibt es nur einen Teilnehmer, der die Anfrage beantworten kann. Dieser sendet (nach den Annahmen in [RFH⁺01]) die Antwort zur Anfrage direkt an den Erzeuger der Anfrage.

Beispiel: Abbildung 3.1 zeigt ein zweidimensionales Content-Addressable Network. Rechtecke repräsentieren Zonen. Teilnehmer i ist an dem Datenobjekt interessiert, das durch den Schlüssel k repräsentiert wird. Schlüssel k ist in der Zone von Teilnehmer j . Teilnehmer i kennt Teilnehmer j nicht, aber er weiß, dass Teilnehmer f näher zum Schlüssel gelegen ist als er selbst. Daher sendet i seine Anfrage an Teilnehmer f . Teilnehmer f und weitere Teilnehmer verfahren analog und leiten die Anfrage weiter, bis sie Teilnehmer j erreicht. ■

Datenverwaltung in Sozialen Netzwerken Im Gegensatz zu technischen Systemen, wie zum Beispiel strukturierten Peer-to-Peer Systemen, gibt es in Sozialen Netzwerken in der Regel keinen künstlichen Adressraum. Experimente in Sozialen Netzwerken haben aber gezeigt, dass die Teilnehmer dieser Systeme ebenfalls in der Lage sind, Teilnehmer mit bestimmten Eigenschaften mit einer hohen Effizienz zu finden [Mil67, DMW03, KB78]. So sind die Teilnehmer eines Experiments in den USA in der Lage gewesen, einen Brief an eine ihnen unbekannte Person im gleichen Land mit nur circa sechs Weiterleitungsschritten zuzustellen [Mil67]. Dieser sogenannte Kleine-Welt-Effekt von Sozialen Netzwerken, dass jeder Mensch mit jedem anderen Mensch über eine kleine Weiterleitungskette verbunden ist, spiegelt sich in diesem und anderen Experimenten wider. Als Adressraum, der die Anfragebearbeitung beschleunigt, kann in diesem Beispiel die räumliche Anordnung der Teilnehmer in den Vereinigten Staaten von Amerika gesehen werden. Es kann also auch in Sozialen Netzwerken in der Regel davon ausgegangen werden, dass die Teilnehmer einschätzen können, welcher ihrer Kontakte am besten dafür geeignet ist, eine Anfrage

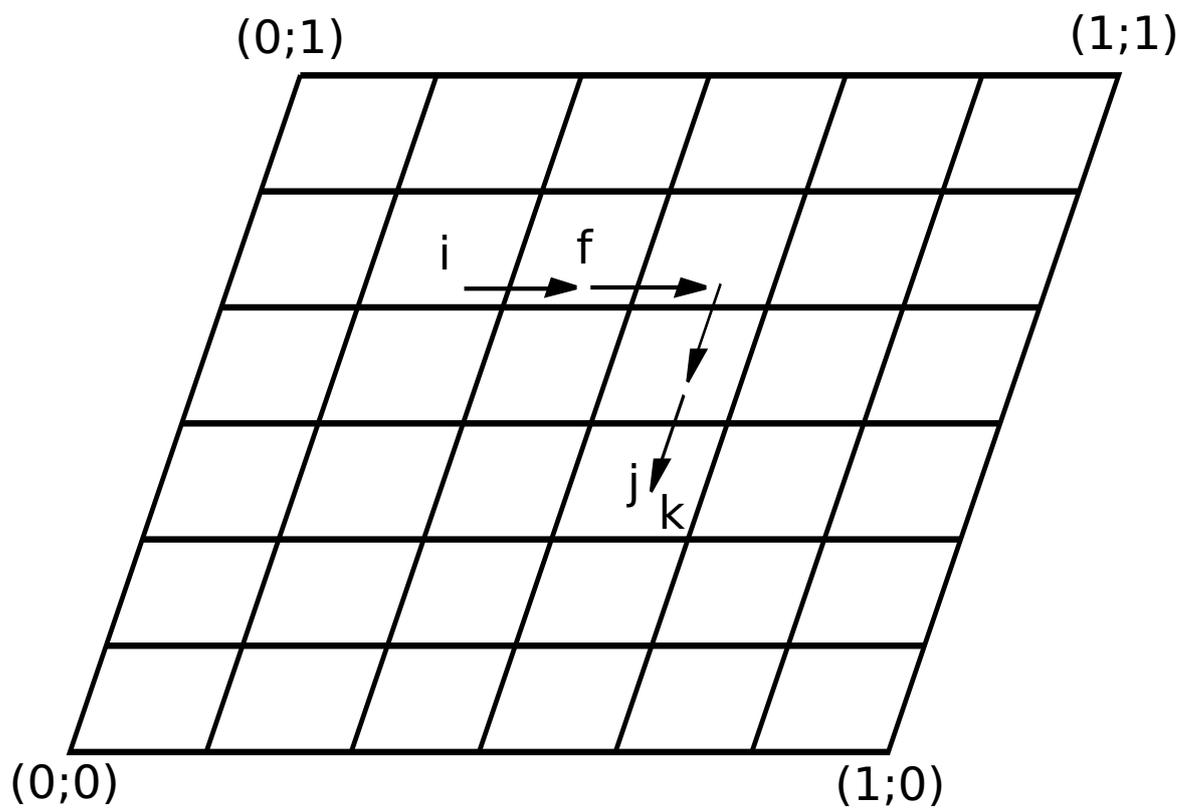


Abbildung 3.1: Zweidimensionales Content-Addressable Network

weiterzuleiten oder zu beantworten.

3.2 Annahmen

Aufgrund der Komplexität des Problems werden in dieser Arbeit einige Annahmen getroffen, um das Problem handhabbar zu machen. Im Folgenden wird motiviert, warum diese Annahmen realistisch beziehungsweise nötig sind.

Geiziges Weiterleiten Die Effizienz eines Systems hängt nicht nur von der Struktur des Netzwerks ab, sondern auch von den Strategien beziehungsweise dem Verhalten seiner Teilnehmer. Beim Weiterleiten von Anfragen an einen Kontakt spielt neben seiner Strategie – also z.B. seinem Kooperationsgrad – auch seine Lage im Adressraum eine Rolle. So kann eine Anfrage nie ihr Ziel erreichen, wenn die Teilnehmer sie nicht näher an ihr Ziel sondern weiter davon weg leiten. Da ein Teilnehmer die Möglichkeit hat, Anfragen fallen zu lassen, wird in dieser Arbeit davon ausgegangen, dass ein Teilnehmer, der eine Anfrage weiterleitet, sie an einen Kontakt weiterleitet, der näher zum Ziel der Anfrage gelegen ist als er selbst. Dieses Verhalten wird auch als *geiziges Weiterleiten* (*engl. greedy forwarding*) bezeichnet. Geiziges Weiterleiten ist nicht unrealistisch, da man das Weiterleiten einer Anfrage an einen Kontakt, der weiter weg vom Ziel der Anfrage ist, als unkooperatives Verhalten auslegen kann, das ebenso behandelt wird wie das Fallenlassen einer Anfrage. Ebenfalls ist unklar, welche Anreize ein solches Teilnehmerverhalten erklären können, da ein Teilnehmer weder seine Reputation noch seine Auszahlung dadurch erhöht, dass er eine Anfrage weg vom Ziel leitet. Geiziges Weiterleiten ist insbesondere dann sinnvoll, wenn Kontakte miteinander kooperieren, da dann ein weit entfernter Kontakt ebenso kooperativ ist wie ein dem Ziel der Anfrage nahe gelegener Kontakt. Wie in Kapitel 5 und in [Sch08] gezeigt wird, kann in der Tat davon ausgegangen werden, dass Kontakte unter bestimmten Umständen miteinander kooperieren.

Nutzen und Auszahlung Für einen Teilnehmer haben Ereignisse und Aktionen einen Nutzen. Da der Nutzen eines Teilnehmers für eine bestimmte Aktion oder ein bestimmtes Ereignis ein individuelles Maß ist, d.h. dass jeder Teilnehmer eine andere Bewertung hat, werden unterschiedliche Strategien in dieser Arbeit auf Basis der Auszahlung durch eine Strategie verglichen. Die Annahme, dass ein Teilnehmer die Aktion wählt, die die höchste zu erwartende Auszahlung erzielt, ist üblich in der ökonomischen Literatur [Qui93, BEG06]. Es ist eine Annahme dieser Arbeit, dass eine Aktion oder ein Ereignis, das eine höhere Auszahlung hat als eine andere Aktion oder ein anderes Ereignis, auch zu einem höheren Nutzen führt.

Uniforme Auszahlungsfunktion Der individuelle Nutzen einer Aktion oder eines Ereignisses für einen Teilnehmer ist schwer zu bemessen. Im Allgemeinen kann nicht ausgeschlossen werden, dass unterschiedliche Teilnehmer – insbesondere menschliche Teilnehmer – unterschiedliche Bewertungen für den Nutzen von Aktionen oder Ereignissen besitzen. Dennoch wird in dieser Arbeit vereinfachend die Annahme getroffen, dass die Auszahlungsfunktion und ihre Koeffizienten für alle Teilnehmer gleich sind. So bringt beispielsweise der Erhalt einer Anfrage für alle Teilnehmer die gleiche Auszahlung.

Gleichverteilte Schlüssel In dieser Arbeit soll der Nutzen unterschiedlicher Strategien verglichen werden. Daher ist es wichtig, einzelne Teilnehmer nicht durch äußere Einwirkungen zu bevorzugen oder zu benachteiligen. Die Schlüssel der Anfragen aller Teilnehmer seien daher gleichverteilt über den Adressraum. Dies schließt übermäßige Last auf Grund ungleichverteilter Schlüssel aus. Es gelte zusätzlich, dass das Ziel einer Anfrage nicht im Adressraum des Erzeugers liegt. (Andernfalls kann der Teilnehmer die Anfrage selbst beantworten und das Erzeugen und Senden einer solchen Anfrage ist sinnlos.) Dennoch sei angemerkt, dass die Last durch den Einfluss der Strategien unterschiedlich sein kann.

Feste und zusätzliche Kontakte Die Netzstruktur eines verteilten Systems bestimmt sich durch die Kontakte der einzelnen Teilnehmer. Diese Arbeit unterscheidet zwei Arten von Kontakten: (a) Kontakte, die ein Teilnehmer wählen kann, und (b) Kontakte, die ein Teilnehmer nicht wählen kann. Im ersten Fall sprechen wir von zusätzlichen Kontakten, während wir im zweiten Fall von festen Kontakten sprechen. In technischen Systemen ist es üblich, dass bestimmte Kontakte vorgegeben – also fest – sind, während andere gewählt werden können. (Unabhängig davon, ob das Wählen zusätzlicher Kontakte ursprünglich vom Systementwickler vorgesehen wurde oder nicht.) In Sozialen Netzwerken gibt es oft ebenfalls feste und zusätzliche Kontakte. Ein Beispiel ist die Kooperation von Firmen. Manche Firmen stehen in einem Besitzverhältnis zueinander. So gehören viele Firmen wiederum anderen Firmen. Die Verbindung zwischen diesen Firmen ist also durch das Besitzverhältnis vorgegeben. Zusätzlich ist es natürlich möglich, mit weiteren Firmen zu kooperieren. Diese Firmen können als zusätzliche Kontakte aufgefasst werden.

Der Unterhalt von zusätzlichen Kontakten verursacht Kosten, die mit der Anzahl an zusätzlichen Kontakten steigen. Da die Anzahl zusätzlicher Kontakte eine strategische Entscheidung der Teilnehmer ist, ist es wichtig, diese Kosten zu berücksichtigen, wenn Strategien untersucht werden sollen. Die Kosten für feste Kontakte hingegen sind konstant und hängen nicht von der Strategie des Teilnehmers ab. Um die Analysen der Strategien der Teilnehmer zu vereinfachen, werden daher die Kosten für feste Kontakte nicht berücksichtigt.

3.3. NETZWERKSTRUKTUR UND KONTAKTWAHL

Zeit Es ist eine Annahme dieser Arbeit, dass die Zeit diskret ist beziehungsweise durch diskrete Zeitpunkte angenähert werden kann. Ziel dieser Arbeit ist es, zu verstehen, ob und welche Strategien in der Interaktions- und Kontaktwahl zu effizienten Systemen führen. Es wird daher nicht darauf eingegangen, wie Teilnehmer ihre Auszahlung dadurch erhöhen können, dass sie Anfragen sortieren oder zu einem späteren Zeitpunkt verarbeiten. Es wird vereinfacht angenommen, dass die Verarbeitung aller Anfragen, die ein Teilnehmer in einer Zeiteinheit erhält, in der nächsten Zeiteinheit durchgeführt wird. Dennoch sei angemerkt, dass die Zeit, die nötig ist, eine Anfrage zu verarbeiten, von der Anzahl an Weiterleitungen abhängt: Für jede Weiterleitung wird eine Zeiteinheit benötigt.

Zonen Alle Zonen haben die gleiche Größe. Andernfalls kommt es zu Ungleichgewichten in der Anfragelast, ohne dass dies von den Strategien der Teilnehmer abhängt. Des Weiteren ist diese Annahme aus folgendem Grund realistisch: Die Größe der Zonen eines Content-Addressable Network unterscheiden sich nur um einen kleinen Faktor [RFH⁺01].

Ebenfalls wird angenommen, dass sich die Zuordnung von Zonen zu Teilnehmern nicht ändert. Diese Annahme erleichtert die formalen Analysen, stellt aber keine besondere Einschränkung des allgemeinen Falles dar, da in dieser Arbeit von den verwalteten Inhalten abstrahiert wird und da deshalb alle Zonen gleichwertig sind. Des Weiteren sind Algorithmen zur Lastbalancierung bekannt, die Vergleichbares erreichen [RLS⁺03].

Technische Fehler Diese Arbeit nimmt an, dass selbst wenn alle Teilnehmer kooperieren, Anfragen wegen technischer Fehler verloren gehen. Daher ist die Wahrscheinlichkeit, dass eine Anfrage verloren geht, stets größer als null.

Einfügen und Manipulieren von Datenobjekten Das Einfügen von Datenobjekten in ein verteiltes System als auch die Problematik des manipulativen Veränderns von Datenobjekten wird in dieser Arbeit nicht betrachtet. Obwohl beide Anliegen wichtige Forschungsfelder sind, liegen sie nicht im Fokus dieser Arbeit und würden den Rahmen der Arbeit sprengen.

3.3 Netzwerkstruktur und Kontaktwahl

Die Struktur des Netzes besteht aus n Teilnehmern und aus den Kontaktbeziehungen zwischen ihnen. Diese Arbeit geht von zwei Arten von Kontakten aus: festen Kontakten und zusätzlichen Kontakten. Ein *fester Kontakt* ist durch das verteilte System vorgegeben. So hat z.B. ein Teilnehmer eines Sozialen Netzwerkes eine Familie, die ihm von außen vorgegeben ist, während er sich seine Freunde aussuchen kann. In

Strukturierten Peer-to-Peer Systemen sind die Kontakte eines Teilnehmers meist vorgegeben. Dennoch ist es möglich, dass der Teilnehmer durch die Anfragebearbeitung weitere Teilnehmer – auch *potentielle Kontakte* genannt – kennenlernt und mit diesen in der Zukunft direkt interagiert. Ein Kontakt, der zusätzlich zu den festen Kontakten gewählt wird, wird als *zusätzlicher Kontakt* bezeichnet. Sei K_i die Menge aller Kontakte von Teilnehmer i . Sei $K_i^+ \subseteq K_i$ die Menge der zusätzlichen Kontakte von Teilnehmer i , und $K_i^f = K_i \setminus K_i^+$ die Menge der festen Kontakte von Teilnehmer i .

Ein Teilnehmer des Systems hat also grundsätzlich die folgenden zwei Aktionen: einen *zusätzlichen Kontakt zu wählen* oder einen *zusätzlichen Kontakt fallenzulassen*.

Um im weiteren Verlauf des Kapitels die Unterhaltskosten für zusätzliche Kontakte zu definieren, benötigen wir die folgenden zwei Hilfsfunktionen:

$$g(i, j) = \begin{cases} \text{wahr} & \text{falls Teilnehmer } j \text{ zusätzlicher Kontakt von Teilnehmer } i \text{ ist} \\ \text{falsch} & \text{sonst} \end{cases} \quad (3.1)$$

$$nr(i, j) = \text{Anzahl der Zeiteinheiten, an denen } g(i, j) \text{ gilt} \quad (3.2)$$

Die Unterhaltszeit für zusätzliche Kontakte des Teilnehmers i sei mit $r(i)$ bezeichnet und ist wie folgt:

Definition 3.2 (Unterhaltszeit für zusätzliche Kontakte)

$$r(i) = \sum_{n \in N} nr(i, n)$$

3.4 Ein Modell der Interaktionswahl

Soziale Netzwerke und Peer-to-Peer Systeme sind Koordinator-freie Systeme, die zum großen Teil als Systeme zur Verwaltung von Wissen und Daten aufgefasst werden können. Anfragebearbeitung hat einen entscheidenden Anteil an der Verwaltung von Wissen und Daten im Bezug auf das Nutzerverhalten. Gegenstand dieses Kapitels ist es, die Aktionen und Ereignisse der Anfragebearbeitung zu motivieren und zu definieren.

Erzeugen einer Anfrage Hat ein Teilnehmer eines Sozialen Netzwerkes oder eines Peer-to-Peer Systems ein Informationsbedürfnis, so hat er die Möglichkeit, eine Anfrage (engl. request) zu *erzeugen*. Sei R_i die Menge der Anfragen, die Teilnehmer i erzeugt hat. Teilnehmer i wird im Folgenden als *Erzeuger* der Anfragen $req \in R_i$ bezeichnet. Die Anfrage beschreibt (a) das Informationsbedürfnis und (b) die Adresse des Erzeugers, d.h. z.B. eine Koordinate aus der Zone des Erzeugers. Da der Teilnehmer das Informationsbedürfnis nicht selbst beantworten kann, hat Teilnehmer i die

3.4. EIN MODELL DER INTERAKTIONSWAHL

Möglichkeit, die erzeugte Anfrage $req \in R_i$ an einen seiner Kontakte zu senden. Sei $R_{i \rightarrow j} \subseteq R_i$ die Menge der Anfragen, die Teilnehmer i erzeugt und die er an seinen Kontakt j sendet. Kontakt j sei als *erster Weiterleiter* (engl. *first forwarder*) oder in der Kurzform als $ff(req)$ bezeichnet.

Beantworten einer Anfrage Erhält ein Teilnehmer eine Anfrage von einem anderen Teilnehmer, so hat er drei Möglichkeiten:

1. Wenn er das Informationsbedürfnis stillen kann, so kann er die Anfrage *beantworten*. Er ist also das *Ziel der Anfrage*. Das Ziel (engl. *destination*) der Anfrage req sei auch mit $dest(req)$ bezeichnet.
2. Wenn er das Informationsbedürfnis nicht stillen kann, kann der Teilnehmer die Anfrage an einen Kontakt *weiterleiten*.
3. Er kann die Anfrage ignorieren, d.h. er kann sie *fallen lassen*.

Möglichkeiten (2) und (3) werden in den folgenden Paragraphen behandelt. Das Beantworten einer Anfrage durch den *Beantworter* bedeutet zum einen, dass er eine Antwort, die das Informationsbedürfnis stillt, erstellt und zum anderen, dass er die Antwort an den Erzeuger der Anfrage sendet. Es wird dabei davon abstrahiert, ob die Anfrage direkt an den Erzeuger der Anfrage geschickt wird oder über Zwischenstationen. Da die Adresse des Erzeugers in der Anfrage enthalten ist, kann der Beantworter die Anfrage grundsätzlich direkt zu ihm senden. Sei $A_i \subseteq R_i$ (A von Antwort (engl. *answer*)) die Menge der Anfragen, die Teilnehmer i erzeugt hat und für die Teilnehmer i eine Antwort erhalten hat. Sei W_j (W von Bearbeiten (engl. *to work on*)) die Menge der Anfragen, die Teilnehmer j erhält und die er beantwortet hat. Sei $W_{i \rightarrow j} \subseteq W_j$ die Menge der Anfragen, die Teilnehmer i an Teilnehmer j weitergeleitet hat und die Teilnehmer j beantwortet.

Weiterleiten einer Anfrage Erhält ein Teilnehmer eine Anfrage, die er nicht beantworten kann, so kann er sie an einen seiner Kontakte weiterleiten. Vereinfachend wird davon ausgegangen, dass ein Teilnehmer nur solche Kontakte auswählt, die näher zum Ziel der Anfrage gelegen sind als er selbst. Diese Annahme wird genauer in Abschnitt 3.2 diskutiert. Sei F_i die Menge der Anfragen, die Teilnehmer i weiterleitet (engl. *forwarding*). Es sei dabei angemerkt, dass das Senden erzeugter Anfragen nicht als Weiterleitung modelliert wird, d.h. es gilt: $R_i \cap F_i = \emptyset$. Sei $F_{i \rightarrow j} \subseteq F_j$ die Menge der Anfragen, die Teilnehmer i an Teilnehmer j weiterleitet. Teilnehmer, die eine Anfrage gemeinsam weiterleiten, werden in dieser Arbeit als *Weiterleitungskette* bezeichnet.

Ignorieren von Anfragen Das Bearbeiten von Anfragen, sei es das Weiterleiten oder das Beantworten von Anfragen, verursacht in der Regel Kosten. Es ist daher denkbar, dass Teilnehmer dadurch Kosten sparen, dass sie Anfragen ignorieren, d.h. fallen lassen.

Einfaches Spiel vs. wiederholtes Spiel In großen Netzwerken, tritt das folgende Szenario auf: Die Endpunkte einer Weiterleitungskette interagieren mit hoher Wahrscheinlichkeit nur ein einziges Mal, spielen also ein sogenanntes one-shot Spiel, während Kontakte mehrfach, also wiederholt, miteinander interagieren. Dieses Szenario ist aus dem folgenden Grund interessant: Die Theorie sagt unkooperatives Verhalten in one-shot Kooperationsspielen voraus [BEG06]. Auf der anderen Seite ist bekannt, dass Kooperation entstehen kann, wenn Kooperationsspiele wiederholt gespielt werden [BEG06]. Diese Arbeit konzentriert sich daher auf Szenarien, bei denen Endpunkte nur einmal miteinander interagieren, während Kontakte wiederholt miteinander interagieren können.

3.5 Nutzen des Systems

Ein Teilnehmer eines verteilten Systems zur Anfragebearbeitung verspricht sich einen Nutzen von der Teilnahme am System. Der Erhalt einer Antwort auf eine Anfrage erzeugt positiven Nutzen. Das Verrichten von Arbeit, wie zum Beispiel das Beantworten oder Weiterleiten von Anfragen, sowie der Unterhalt von Kontakten, erzeugt negativen Nutzen. Wie in Abschnitt 3.2 motiviert, ist es eine Annahme dieser Arbeit, dass der Nutzen einer Aktion oder eines Ereignisses proportional zur Auszahlung durch dieses Ereignis beziehungsweise durch diese Aktion ist. Im Folgenden wird beschrieben, welche Auszahlung ein Teilnehmer hat.

Einkommen Ein Teilnehmer nimmt an einem verteilten System zur Anfragebearbeitung teil, da er sich die Beantwortung von Informationsbedürfnissen verspricht. Positiver Nutzen, d.h. *das Einkommen eines Teilnehmers*, entsteht, wenn der Teilnehmer Antworten zu seinen Anfragen erhält. Seien a die Einnahmen, die ein Teilnehmer durch den Erhalt einer Antwort (maximal) erhält. In manchen Fällen ist eine Antwort auf eine Anfrage umso wertvoller, je eher der Erzeuger der Anfrage sie empfängt. Zum Beispiel kann der Händler, der zuerst von einer Missernte erfährt, meist den höchsten Gewinn erzielen, da er seine Einkaufsstrategie als erster anpassen kann. Aus diesem Grund wird in dieser Arbeit der Wert einer Anfrage diskontiert (das heißt, der Wert einer Anfrage wird abgezinst). Je kleiner Diskontierungsfaktor $\delta \in [0, 1]$ ist, desto schneller verliert eine Antwort ihren Wert. (Für $\delta = 1$ ist keine Diskontierung vorhanden.) Die Diskontierung erfolgt dabei pro Weiterleitungsschritt einer Anfrage. Sei $h(req)$ die Anzahl der Weiterleitungsschritte, die eine Anfrage $req \in A$ benötigt,

um beantwortet zu werden.

Definition 3.3 (Einkommen)

Das Einkommen eines Teilnehmers i besteht aus der Summe der Einnahmen aus den erhaltenen Antworten:

$$Einkommen(i) = a \cdot \sum_{req \in A_i} \delta^{h(req)}$$

Ausgaben Ein Teilnehmer eines verteilten Systems hat Kosten, wenn er Arbeit verrichtet. Im Einzelnen sind dies Interaktionskosten und Kosten für den Unterhalt von zusätzlichen Kontakten.

Interaktionskosten sind Kosten für das Erzeugen von Anfragen q sowie Kosten für das Weiterleiten f und Kosten für das Beantworten w von Anfragen.

Der Unterhalt von zusätzlichen Kontakten verursacht Kosten c pro Zeiteinheit und pro zusätzlichem Kontakt (d.h. pro $r(i)$, vergleiche Abschnitt 3.3).

Die Ausgaben eines Teilnehmers bestehen aus der Summe der oben genannten Einzelkosten:

Definition 3.4 (Ausgaben)

$$Ausgaben(i) = q \cdot |R_i| + f \cdot |F_i| + w \cdot |W_i| + c \cdot r(i)$$

Die Auszahlung eines Teilnehmers ist die Differenz aus den Einnahmen und den Ausgaben. Die Auszahlung wird durch die Auszahlungsfunktion (engl. value function) beschrieben.

Definition 3.5 (Auszahlungsfunktion)

$$\begin{aligned} Auszahlung(i) &= Einkommen(i) - Ausgaben(i) \\ &= a \cdot \sum_{req \in A_i} (\delta^{h(req)}) - q \cdot |R_i| - f \cdot |F_i| - w \cdot |W_i| - c \cdot r(i) \end{aligned}$$

Die Auszahlungsfunktion des Teilnehmers i sei auch kurz mit $z(i)$ bezeichnet. Tabelle 3.1 gibt eine Übersicht über die Abkürzungen.

Es sei bereits an dieser Stelle angemerkt, dass eine negative Auszahlung, d.h. $z(i) < 0$, dadurch von einem Teilnehmer vermieden werden kann, dass dieser das System verlässt. Rein praktisch könnte also auch die folgende Auszahlungsfunktion herangezogen werden:

$$Auszahlung'(i) = \max\{Einkommen(i) - Ausgaben(i), 0\}$$

KAPITEL 3. GRUNDLAGEN

Aktion/Ereignis	Anzahl der Aktionen/Ereignisse	Auszahlungs-Koeffizient
Erzeugen einer Anfrage	$ R $	$-q$
Weiterleiten einer Anfrage	$ F $	$-f$
Beantworten einer Anfrage	$ W $	$-w$
Erhalt einer Antwort	$ A $	$+a$
Unterhalt von zusätzlichen Kontakten	$r(\cdot)$	$-c$

Tabelle 3.1: Kosten und Nutzen von Aktionen und Ereignissen

Um pathologische Systeme auszuschließen, sei in dieser Arbeit angenommen, dass die folgenden Verhältnisse zwischen den Auszahlungskoeffizienten gelten: Einnahmen führen zu positivem Nutzen und Ausgaben zu negativen Nutzen, d.h. es gilt $a, q, f, w, c \geq 0$.

Gemeinwohl Während die Auszahlung eines Teilnehmers den Nutzen beschreibt, den dieser vom System hat, beschreibt das *Gemeinwohl* (engl. *social welfare*) den Gesamtnutzen des Systems. Während die Auszahlung zur Analyse der Strategien herangezogen wird, lässt sich aus dem Gemeinwohl der Wert (engl. *value*) des Systems ableiten. Ist der Wert des Systems maximal beziehungsweise sein Gemeinwohl, so ist das System effizient.

Definition 3.6 (Gemeinwohl)

$$v = \sum_{i \in N} \text{Auszahlung}(i)$$

Definition 3.7 (Effizienz)

Ein System ist dann effizient, wenn das Gemeinwohl v maximal ist.

Eine Strategie $S(i)$ eines Teilnehmers i und seine Auszahlung sind für die Analyse eines verteilten Systems wichtig. Mittels der Auszahlungen und der Strategien der Teilnehmer können Gleichgewichtszustände definiert werden.

Definition 3.8 (Nash-Gleichgewichte)

Ein Nash-Gleichgewicht (kurz: Gleichgewicht) beschreibt einen Zustand, in dem jeder Teilnehmer $i \in N$ seine Auszahlung $\text{Auszahlung}(i)$ nicht dadurch verbessern kann, dass er einseitig von seiner Strategie $S(i)$ abweicht [Nas54].

Eine Definition des Strategieraums findet sich in Kapitel 5.

3.6 Wichtige Maße

Kooperation Die Effizienz eines Systems hängt vom Verhalten der Teilnehmer ab. Ein wichtiges Maß, das das Verhalten eines Teilnehmers beschreibt, ist sein Kooperationsgrad. Dieser beschreibt den Anteil an Anfragen, den der Teilnehmer verarbeitet.

Definition 3.9 (Kooperationsgrad)

Der Kooperationsgrad $C(i)$ von Teilnehmer i ist der Anteil an bearbeiteten Anfragen aus der Menge der erhaltenen Anfragen.

$$C(i) = \frac{|W_i| + |F_i|}{\sum_{\substack{x \in N, \\ x \neq i}} |R_{x \rightarrow i}| + |F_{x \rightarrow i}|}$$

Der durchschnittliche Kooperationsgrad aller Teilnehmer sei mit \bar{C} bezeichnet. Er ist interessant, da sich so bestimmen lässt, ob die Strategien der Teilnehmer kooperativ sind. (Es sei angemerkt, dass der durchschnittliche Kooperationsgrad im Allgemeinen ungleich dem systemweiten Kooperationsgrad ist, der die Anzahl der beantworteten Anfragen ins Verhältnis zur Anzahl aller Anfragen setzt. Dieser wird in dieser Arbeit nicht betrachtet, da die Analyse der Leistung des Systems – wie allgemein üblich [RT02, Jac03] – das Verhältnis des Gemeinwohls zur Effizienz nutzt.)

Definition 3.10 (Durchschnittlicher Kooperationsgrad)

Der durchschnittliche Kooperationsgrad \bar{C} ist die Summe der Kooperationsgrade $C(i)$ geteilt durch die Anzahl der Teilnehmer n .

$$\bar{C} = \frac{\sum_{i \in N} C(i)}{n}$$

In der Analyse der Nützlichkeit kooperativer Teilnehmer ist es wichtig zu wissen, ob Teilnehmer gegenüber ihren zusätzlichen Kontakten ein anderes Verhalten zeigen als gegenüber ihren festen Kontakten.

Definition 3.11 (Kooperationsgrad bezüglich zusätzlicher Kontakte)

Der Kooperationsgrad, den ein Teilnehmer i bezüglich seiner zusätzlichen Kontakte zeigt, ist der Anteil der Anfragen, den Teilnehmer i von seinen zusätzlichen Kontakten erhält und bearbeitet:

$$C_c(i) = \frac{\sum_{j \in K_i^+} (|W_{j \rightarrow i}| + |F_{(i|j \rightarrow i)}|)}{\sum_{j \in K_i^+} (|F_{j \rightarrow i}| + |R_{j \rightarrow i}|)}$$

dabei sei $F_{(i|j \rightarrow i)} \subseteq F_i$ die Teilmenge von F_i , für die gilt, dass ihre Anfragen über zusätzliche Kontakte an Teilnehmer i gesendet wurden. D.h. es gilt: $F_{(i|j \rightarrow i)} = F_i \cap (F_{j \rightarrow i} \cup R_{j \rightarrow i})$.

KAPITEL 3. GRUNDLAGEN

Der Durchschnitt des Kooperationsgrads bezüglich zusätzlicher Kontakte über alle Teilnehmer sei mit \overline{C}_c bezeichnet.

Die *Erfolgsrate* eines Teilnehmers i bezeichnet den Anteil seiner Anfragen, den er beantwortet bekommt. Die Erfolgsrate lässt also Rückschlüsse darüber zu, wie erfolgreich der Teilnehmer darin ist, zusätzliche Kontakte mit einem hohen Kooperationsgrad zu wählen.

Definition 3.12 (Erfolgsrate)

Die Erfolgsrate ist der Anteil an beantworteten Anfragen bezüglich der Anzahl gesendeter Anfragen:

$$\text{erfolgsrate}(i) = \frac{|A_i|}{|R_i|}$$

Anzahl an Weiterleitungen (Hops) Ein Maß, das einen Einfluss darauf hat, wie effizient ein System ist, ist die durchschnittliche Anzahl an Weiterleitungen \bar{h} (kurz: die durchschnittliche Anzahl an Hops). Eine geringere Anzahl an Hops führt zu einer geringeren Weiterleitungslast (falls alle anderen Parameter des Systems gleich bleiben).

Definition 3.13 (Durchschnittliche Anzahl an Weiterleitungen)

$$\bar{h} = \frac{1}{|A|} \cdot \sum_{req \in A} h(req)$$

Definition 3.14 (Durchschnittliche Anzahl an Weiterleitungen des Teilnehmers i)

$$\bar{h}(i) = \frac{1}{|A_i|} \cdot \sum_{req \in A_i} h(req)$$

Die durchschnittliche Anzahl an Weiterleitungen ist als Maß unbrauchbar, um zu bestimmen, ob Teilnehmer Anfragen mit Zielen, die weit vom Erzeuger entfernt liegen, genauso gut beantwortet bekommen wie Anfragen mit Zielen nahe am Erzeuger. Das Maß $\bar{h}_\Delta(i)$ ist die durchschnittliche Anzahl an Weiterleitungen relativ zu der Distanz im Adressraum zum Ziel der jeweiligen Anfrage.

Definition 3.15 (Relative durchschnittliche Anzahl an Weiterleitungen)

$$\bar{h}_\Delta(i) = \frac{1}{|A_i|} \cdot \sum_{req \in A_i} \frac{h(req)}{\Delta(i, \text{dest}(req))}$$

Dabei sei $\Delta(i, j)$ der Euklidische Abstand im Adressraum zwischen den Zentren der Zonen der Teilnehmer i und j , d.h. zwischen $\text{zentrum}(i)$ und $\text{zentrum}(j)$.

$\bar{h}_{\Delta}(i|j)$ ist die durchschnittliche Anzahl an Hops der Anfragen von Teilnehmer i , die er an Teilnehmer j weitergeleitet hat.

Definition 3.16 (Relative durchschnittliche Anzahl an Weiterleitungen von i zu j)

$$\bar{h}_{\Delta}(i|j) = \frac{1}{|A_{i \rightarrow j}|} \cdot \sum_{req \in A_{i \rightarrow j}} \frac{h(req) - 1}{\Delta(j, dest(req))}$$

Durchschnitt Der Durchschnitt eines Maßes x über alle Teilnehmer n sei mit \bar{x} bezeichnet. Zum Beispiel wird der durchschnittliche Kooperationsgrad über alle Teilnehmer mit \bar{C} bezeichnet. Der Durchschnitt eines Maßes $m(i)$ über alle Teilnehmer $i \in N$ sei wie folgt:

$$\bar{m} = \frac{\sum_{i \in N} m(i)}{n}$$

3.7 Methodik

Teilnehmer verteilter Systeme zur Anfragebearbeitung können zum einen Software-Artefakte, die ein Programm ausführen, (d.h. Agenten) sein und zum anderen Menschen.

Agenten folgen einem Programm. Das Programm eines Agenten hat in der Regel zum Ziel, dass der Agent seinen Nutzen maximiert. Dies bedeutet auf die Anfragebearbeitung bezogen, dass der Agent seine Auszahlung maximiert. Es ist also wichtig zu verstehen, wie sich ihre Auszahlung maximierende Agenten in verteilten Systemen verhalten und ob sie einen Anreiz haben, von einer vom System-Entwickler vorgegebenen Strategie abzuweichen oder nicht.

Da Software im Allgemeinen und Agenten im Speziellen von Menschen geschrieben werden und da Menschen Software steuern beziehungsweise direkt im verteilten System agieren, ist es ebenfalls wichtig, das Verhalten von menschlichen Teilnehmern in der Anfragebearbeitung zu verstehen. Wie bereits in Kapitel 2 motiviert, kann im Allgemeinen nicht angenommen werden, dass menschliches Verhalten dem Verhalten von Agenten, die ihre Auszahlung maximieren, entspricht. Ebenso wenig kann das Gegenteil angenommen werden.

Die Methodik dieser Arbeit versucht, die in den beiden vorherigen Paragraphen genannten Ziele zu erreichen. Die Methodik ist wie folgt:

- Kapitel 4 stellt eine Vorstudie dar, die grundsätzlich prüfen soll, ob es sich überhaupt lohnt, zusätzliche Kontakte zu wählen – und unter welchen Bedingungen. Dazu wird eine mathematische Abschätzung getroffen.

KAPITEL 3. GRUNDLAGEN

- In Kapitel 5 wird formal untersucht, ob ihre Auszahlung maximierende Agenten von einer durch den System-Entwickler vorgegebenen Strategie abweichen. Falls dem so ist, soll durch Veränderungen der vorgegebenen Strategie ein Zustand erreicht werden, der ein Gleichgewicht darstellt.
- Kapitel 6 verschärft die Überprüfung aus Kapitel 5 dahingehend, dass ein auch unter Einstreuen zufälliger Strategien stabiler Zustand gesucht wird. Des Weiteren wird überprüft, ob die durch den System-Entwickler vorgeschlagene Strategie auch „natürlich“ hätte entstehen können. „Natürlich“ bedeutet hier, dass die vorgegebene Strategie durch evolutionäre Strategieanpassungen aus zufälligen Strategien entstehen kann.
- Kapitel 7 untersucht menschliches Verhalten in der Anfragebearbeitung. Im Gegensatz zu den vorherigen Kapiteln, werden keine Annahmen – wie die Maximierung der Auszahlung – über das Verhalten der Teilnehmer gemacht. Vielmehr soll das menschliche Verhalten beobachtet und beschrieben werden. Die angewendete Methodik sind verhaltensökonomische Experimente. Diese erlauben es nicht nur, Verhalten zu beobachten, sondern auch zu messen, wie signifikant ein Verhalten ist. Kapitel 8 ergänzt die Ergebnisse aus Kapitel 7: Dort wird eine Studie zur Beobachtung des menschlichen Verhaltens in großen Netzen beschrieben (in Kapitel 7 wurden kleine Netze von 12 Teilnehmern untersucht). Im Gegensatz zu Kapitel 7 handelt es sich nicht um ein Experiment, sondern um eine Studie. Diese lässt zwar die Beobachtung von Verhalten zu, ist aber nicht in der Lage, die Signifikanz eines Verhaltens zu bestimmen. Insofern kann die Studie nur als Ergänzung zu den in Kapitel 7 durchgeführten Experimenten gesehen werden. Vorteil der Studie gegenüber verhaltensökonomischen Experimenten ist, dass in der Studie ein größeres Netzwerk untersucht wird. Dies ermöglicht es, Effekte, die evtl. nur in großen Netzwerken zu beobachten sind, aufzuspüren.

Der Mix der genutzten Methoden soll sowohl das Verhalten von Agenten als auch von Menschen untersuchen. Wie im Folgenden deutlich werden wird, unterscheidet sich das Verhalten von Agenten und Menschen in der Interaktions- und Kontaktwahl kaum.

4 Anreize zur Vernetzung

Teilnehmer verteilter Systeme interagieren in der Regel nicht direkt miteinander, sondern mit Hilfe von Weiterleitungsketten. So gibt ein Teilnehmer eine Anfrage an seine Kontakte weiter, die diese an ihre Kontakte weitergeben können, und so fort. Da der Unterhalt von Kontakten normalerweise Kosten verursacht, gilt es zu klären, wann es vorteilhaft ist, nur wenige Kontakte zu unterhalten und wann dies nicht der Fall ist.

Ziel dieses Kapitels ist es, zu prüfen, ob es grundsätzlich eines Anreizes bedarf, damit Teilnehmer Kontakte auswählen. Dabei soll untersucht werden, unter welchen exogenen Einflüssen Teilnehmer einen natürlichen Anreiz besitzen Kontakte zu wählen und unter welchen Umständen dies nicht der Fall ist.

Wie im Folgenden gezeigt wird, existiert eine neue Form des Freifahrens: Teilnehmer sind bewusst kaum vernetzt (engl. *Deliberately Poorly Connected (DPC)*) und sparen sich die Unterhaltskosten für Kontakte. Im Gegensatz zum klassischen Freifahrerverhalten [FLSC04], das sich auf die Interaktionswahl bezieht, bezieht sich *Deliberately Poor Connectedness* auf die Kontaktwahl. Aus globaler Sicht ist dieses Verhalten problematisch: Da nur wenige Kontakte pro Teilnehmer dazu führen, dass die Weiterlast steigt, kann das Gemeinwohl des Systems sinken. Im Folgenden wird ein formales Modell vorgestellt, das zeigt, dass Teilnehmer, die die DPC-Strategie nutzen, in vielen Fällen eine höhere Auszahlung haben als Teilnehmer, die viele Kontakte besitzen, also im Sinne des Systementwicklers kooperieren. Um diese Situation zu verbessern, d.h. um Effizienzverluste durch *Deliberately Poor Connectedness* zu verhindern, wird eine Anreiz-basierte Strategie vorgeschlagen. Die Idee ist, dass die Teilnehmer sich kooperativer gegenüber Kontakten verhalten, die keine DPC-Strategie verfolgen, als gegenüber Teilnehmern, die die DPC-Strategie nutzen.

4.1 Einleitung

Systementwickler aus der Informatik nehmen zum Teil an, dass die Teilnehmer ihrer Systeme *vollständig kooperativ* sind: Teilnehmer bearbeiten die Anfragen anderer Teilnehmer und unterhalten Kontakte. Dennoch ist es von heutigen Peer-to-Peer Systemen, also einer technischen Klasse von verteilten Systemen zur Anfragebearbeitung, bekannt, dass deren Teilnehmer sich oft unkooperativ verhalten [SGG02, SW02]: 25% aller Teilnehmer der Datei-Tauschbörse Gnutella teilen keine einzige Datei. In diesem Kapitel werden solche (konventionelle) Freifahrer, d.h. solche Teilnehmer, die Anfragen von anderen Teilnehmern fallen lassen, als *Freifahrer auf*

KAPITEL 4. ANREIZE ZUR VERNETZUNG

Anfrageebene bezeichnet. Diese Freifahrer sind nach der Nomenklatur dieses Kapitels *unkooperativ auf der Anfrageebene*. Eine andere Art des Freifahrens hängt mit dem Unterhalt von Kontakten zusammen: In der in [SW02] durchgeführten Untersuchung des Verhaltens der Teilnehmer der Datei-Tauschbörse FastTrack haben 48% aller Teilnehmer nur einen einzigen Kontakt. Dieses Verhalten lässt sich zum einen dadurch erklären, dass viele Nutzer sich nur kurz mit dem Netz verbinden und in dieser Zeit kaum Kontaktbeziehungen aufbauen, zum anderen gibt es eine ökonomische Erklärung: (a) Da die Verwaltung und der Unterhalt von Kontakten Kosten verursachen (z.B. da Datenstrukturen angelegt und aktuell gehalten werden müssen beziehungsweise weil überprüft werden muss, ob ein Kontakt noch erreichbar ist) und (b) da diese Unterhaltskosten mit der Anzahl der Kontakte wachsen, ist es für einen Teilnehmer attraktiv, nur wenige Kontakte zu unterhalten. Falls diese wenigen Kontakte kooperativ sind, hat der Teilnehmer Zugang zum System und kann das gesamte Netz dadurch nutzen, dass er Anfragen an seine (wenigen) Kontakte sendet. In dieser Arbeit nennen wir diese neue Art des Freifahrerverhaltens *Deliberately Poor Connectedness (DPC)* (deutsch: *freiwillig geringe Vernetztheit*). Ohne einen Mechanismus oder eine Strategie, die *Deliberately Poor Connectedness* eindämmt, muss erwartet werden, dass Teilnehmer, die ihren Nutzen maximieren, eine DPC-Strategie wählen. Dadurch entsteht das folgende Problem: Von einer globalen Sicht aus betrachtet – also aus der Sicht des Systementwicklers – führt *Deliberately Poor Connectedness* zu kaum vernetzten Systemen. Dies wiederum führt zu einer erhöhten Weiterleitungslast, was wiederum zu einem geringeren Gemeinwohl und damit zu weniger Nutzen für alle Teilnehmer führen kann. Allerdings zeigen verhaltensökonomische Experimente [SBV07], dass Teilnehmer strukturierter Peer-to-Peer Systeme sich *reziprok* Verhalten, d.h. sie kooperieren nur dann mit Kontakten, wenn diese ebenfalls kooperativ sind und ignorieren Kontakte, die unkooperativ sind (beides auf Anfrageebene). Es ist also unklar, ob sich *Deliberately Poor Connectedness* durchsetzt. Eine Voraussetzung für reziprokes Verhalten ist das Vermögen, den *Kooperationsgrad* von anderen Teilnehmern einzuschätzen, d.h. einzuschätzen, welchen Anteil der Anfragen, die ein Teilnehmer erhält, von ihm bearbeitet werden. In vielen verteilten Systemen sind die Teilnehmer an lokale Informationen gebunden, d.h. sie können lediglich beobachten, welcher Anteil der Anfragen, die sie an einen Kontakt senden, verloren geht. Es ist dabei zu beachten, dass dieser Kontakt nicht derjenige Teilnehmer einer Weiterleitungskette sein muss, der die Anfrage fallen gelassen hat. Dennoch behandeln menschliche Teilnehmer einen Kontakt, bei dem der Anteil der gesendeten Anfragen zu einem großen Teil fallengelassen wird, wie einen unkooperativen Teilnehmer [SBV07]. Die Tatsache, dass viele verteilte Systeme auf lokalen Informationen beruhen, macht die Analyse ihrer Strategien herausfordernd: Ein Teilnehmer kennt den wahren Kooperationsgrad seiner Kontakte (oder anderer Teilnehmer) nicht. Teilnehmer, die aufbauend auf lokalen Informationen den Kooperationsgrad ihrer Kontakte schätzen, machen Fehler und treffen daher teilweise falsche Entscheidungen.

Ziel Das Ziel dieser Arbeit sind effiziente Netzwerke – also solche, in denen die Summe der Auszahlungen der Teilnehmer maximal ist. Mit Ausnahme von pathologischen Situationen, in denen z.B. das Gemeinwohl sinkt, wenn Teilnehmer Anfragen bearbeiten, bedeutet das (a) die Abwesenheit von Freifahrern auf Anfrageebene und (b) die Abwesenheit von Deliberately Poor Connectedness. Während Punkt (a) – also die Abwesenheit von Freifahrern auf Anfrageebene – bereits untersucht wurde (vergleiche Kapitel 2), konzentriert sich dieses Kapitel auf Deliberately Poor Connectedness – also auf die Frage, ob es sich für einen Teilnehmer überhaupt lohnt Kontakte zu wählen und unter welchen Umständen.

Die Analyse von Deliberately Poor Connectedness ist nicht trivial: Es müssen alle Abhängigkeiten zwischen Interaktionswahl und Kontaktwahl berücksichtigt werden. Beispielsweise ist zu vermuten, dass zusätzliche Kontakte nützlich sind, wenn der Anteil an Freifahrern auf Anfrageebene hoch ist: Mit Hilfe von zusätzlichen Kontakten können unkooperative Teilnehmer überbrückt werden. Dadurch kann ein größerer Anteil an Anfragen erfolgreich verarbeitet werden. Daher müssen in dieser Analyse sowohl die Unterhaltskosten für Kontakte als auch die Kosten der Anfragebearbeitung berücksichtigt werden. Die Analyse wird dadurch zusätzlich erschwert, dass die vorkommenden Strategien komplex sind. Teilnehmer interagieren indirekt über Zwischenknoten einer Weiterleitungskette. Daraus folgt, dass die Strategien mehrerer Knoten bestimmen, ob eine Anfrage bearbeitet wird. Daher ist es schwierig die durchschnittliche Pfadlänge oder die Wahrscheinlichkeit, dass eine Anfrage bearbeitet wird, zu berechnen.

Beiträge Die in diesem Kapitel durchgeführte Analyse zeigt, dass Deliberately Poor Connectedness zu höheren Auszahlungen führt als Strategien, bei denen viele zusätzliche Kontakte gewählt werden, d.h. zu höheren Auszahlungen als bei kooperativen Strategien. Im Einzelnen ist Deliberately Poor Connectedness umso vorteilhafter, je besser die Teilnehmer in der Lage sind, Freifahrer auf Anfrageebene zu identifizieren, also je besser die Teilnehmer in der Lage sind, den Kooperationsgrad anderer Teilnehmer einzuschätzen. Wenn Teilnehmer Freifahrer auf Anfrageebene mit hoher Wahrscheinlichkeit erkennen, dann senden sie keine Anfragen mehr an diese und nur wenige Anfragen gehen verloren. Daher ist es nicht vorteilhaft, viele zusätzliche Kontakte zu unterhalten, denn mit nur wenigen zusätzlichen Kontakten ist die Wahrscheinlichkeit, eigene Anfragen erfolgreich beantwortet zu bekommen, hoch. Folglich können die Teilnehmer die Unterhaltskosten für zusätzliche Kontakte dadurch reduzieren, dass sie nur wenige zusätzliche Kontakte wählen und Deliberately Poor Connectedness wird attraktiv. In anderen Worten, Strategien und Mechanismen, die Freifahrer auf Anfrageebene verhindern, reduzieren den Nutzen von Deliberately Poor Connectedness nicht, sondern fördern diese problematische Strategie sogar noch. Dieses Ergebnis ist überraschend, da dieser negative Zusammenhang bisher unbekannt war.

KAPITEL 4. ANREIZE ZUR VERNETZUNG

Ein Beitrag dieser Arbeit ist es, dass eine Strategie gegen *Deliberately Poor Connectedness* entwickelt wurde, die sogenannte *Contact-for-Contact (C4C)* Strategie (zu deutsch: Kontakt-für-Kontakt Strategie). Die Idee der C4C-Strategie ist es, dass ein Teilnehmer, der diese Strategie nutzt, nur eine kleine Untermenge seiner Kontakte denjenigen Teilnehmern zur Verfügung stellt, die – nach seiner Einschätzung – eine DPC-Strategie verwenden. Auf der anderen Seite stellt ein Teilnehmer, der die C4C-Strategie verwendet, kooperativen Teilnehmern (also solchen Teilnehmern, die keine DPC-Strategie verwenden) alle seine Kontakte zur Verfügung. Die C4C-Strategie ist wie folgt: Ein Teilnehmer, der die C4C-Strategie verwendet, leitet eine Anfrage nur so weit durch den Adressraum (relativ zum Ziel der Anfrage), wie es sein Vorgänger in der Weiterleitungskette getan hat. Die Konsequenz ist, dass die Anfragen von Teilnehmern mit einer DPC-Strategie nicht über zusätzliche Kontakte geleitet werden und dass dadurch die Anzahl an Weiterleitungen für diese Anfragen höher ist als für die Anfragen kooperativer Teilnehmer. Eine höhere Anzahl an Weiterleitungen erhöht die Wahrscheinlichkeit, dass Anfragen verloren gehen, da eine Anfrage in jedem Weiterleitungsschritt wegen eines unkooperativen Teilnehmers oder einem technischen Defekt verloren gehen kann. Daraus folgt, dass Teilnehmer, die eine DPC-Strategie verfolgen, eine geringere Auszahlung erhalten als kooperative Teilnehmer. Diese intuitive Argumentation wird im Folgenden durch eine formale Analyse bestätigt. Des Weiteren wird in Simulationen gezeigt, dass die C4C-Strategie auch dann effektiv ist, wenn der Anteil der Teilnehmer, die die C4C-Strategie benutzen, gering (d.h. um die 10%) ist. Die Anwendung der C4C-Strategie benötigt keine komplexen Statistiken, keine zusätzlichen Nachrichten oder komplexe Berechnungen und ist damit nahezu kostenlos. Des Weiteren ist sie robust gegen Weißwaschen (engl. whitewashing) der Reputation, d.h. die C4C-Strategie ist robust gegen Teilnehmer, die das Netz verlassen und wieder in dieses eintreten, um ihre schlechte Reputation zu löschen.

Gliederung des Kapitels Kapitel 4.2 motiviert die Vorteile zusätzlicher Kontakte. Es folgen eine Erläuterung der in diesem Kapitel verwendeten Methodik in Kapitel 4.3 und die Beschreibung zusätzlicher Annahmen (Kapitel 4.4). Nach der Beschreibung der relevanten Strategien in Abschnitt 4.5 folgt die Analyse der DPC-Strategie (Abschnitt 4.6). In Abschnitt 4.7 wird die C4C-Strategie vorgeschlagen, gefolgt von einer Analyse ihrer Effektivität in Abschnitt 4.8. Es folgt die Validierung der Ergebnisse (Abschnitt 4.9) und ein Fazit (Abschnitt 4.11).

4.2 Strukturierte Peer-to-Peer Systeme und zusätzliche Kontakte

Mittels strukturierter Peer-to-Peer Systeme ist es unter anderem möglich, große Mengen von (Schlüssel, Daten)-Paaren zu verwalten. In diesem und in weiteren Abschnit-

ten werden Content-Addressable Networks (vergleiche Abschnitt 3.1) als Beispiel für ein strukturiertes Peer-to-Peer System genutzt. Allerdings sei angemerkt, dass die Ergebnisse dieses Kapitels nicht auf CAN-spezifischen Eigenschaften beruhen. Eine Diskussion dieses Themas findet sich in Abschnitt 4.10.

Zusätzliche Kontakte In einem zweidimensionalen Content-Addressable Network beträgt die durchschnittliche Pfadlänge $\frac{\sqrt{n}}{2}$ Weiterleitungsschritte, wenn alle Teilnehmer kooperieren und in Abschnitt 3.1 erörterten und in [RFH⁺01] beschriebenen Weiterleitungsalgorithmus benutzen. $\frac{\sqrt{n}}{2}$ Weiterleitungsschritte sind im Durchschnitt nötig, da die maximale Anzahl an Weiterleitungen der zweifachen Länge einer Dimension des Content-Addressable Networks entspricht (d.h. \sqrt{n}); Faktor $\frac{1}{2}$ ergibt sich, da eine durchschnittliche Anfrage die halbe Anzahl der Schritte benötigt.

Beispiel: In einem zweidimensionalen Content-Addressable Network mit 10.000 Teilnehmern, die einen Kooperationsgrad von 100% besitzen, leiten 50 ($= \frac{\sqrt{10000}}{2}$) Teilnehmer eine Anfrage im Durchschnitt weiter. ■

Die Anzahl an Weiterleitungen kann durch zusätzliche Kontakte um Größenordnungen gesenkt werden.

Beispiel: Bollobas und Chung [BC88] zeigen, dass die minimale Pfadlänge zwischen zwei Knoten auf einem Ring aus n Knoten mit n zufällig gewählten zusätzlichen Kontakten $\log_2(n)$ Schritte beträgt. Kleinberg [Kle00] zeigt, dass in einem $n \times n$ -Grid mit n^2 zusätzlichen Kontakten, die einer inversen Power-Law Verteilung folgen (d.h. die zusätzlichen Kontakte eines Teilnehmers nehmen exponential ab, je weiter sie von ihm im Adressraum entfernt sind), Anfragen in $O(\log(n) \cdot \log(n))$ Schritten weitergeleitet werden können. ■

Durch zusätzliche Kontakte kann also die Effizienz des Systems erhöht werden. Zusätzliche Kontakte können sich durch die Anfragebearbeitung kennenlernen: Der Beantworter sendet eine Antwort auf eine Anfrage (aus der er die Adresse des Erzeugers kennt) direkt zum Erzeuger. Dieser lernt den Beantworter durch die Antwort kennen. Neben der Verringerung der Weiterleitungslast haben zusätzliche Kontakte einen weiteren Vorteil: Die Anzahl der Teilnehmer, die eine Anfrage bearbeitet, sinkt und damit die Wahrscheinlichkeit, dass eine Anfrage fallengelassen wird.

4.3 Methodik

Um das Vorgehen im Rest des Kapitels zu erläutern, soll kurz die Methodik des Folgenden kommentiert werden. Dieses Kapitel untersucht Deliberately Poor Connectedness in Systemen mit Interaktions- und Kontaktwahl an Hand von zwei Abstraktionsebenen. Die verschiedenen Abstraktionsebenen entstehen durch die zwei verwendeten Methoden: (1) eine formale Analyse und (2) Simulationen. Die formale Analyse, also Methodik (1), erlaubt es, den Erwartungswert für die Auszahlung eines

KAPITEL 4. ANREIZE ZUR VERNETZUNG

Teilnehmers zu berechnen. Die Anzahl zusätzlicher Kontakte ist ein exogener Parameter des der Analyse zugrunde liegenden Modells. In anderen Worten, Methode (1) ignoriert, wie genau sich zwei zusätzliche Kontakte gebildet haben. In Methode (2), d.h. in den Simulationen, wählen Teilnehmer ihre zusätzlichen Kontakte nach einem vorgegebenen Protokoll. Dieses Protokoll wird im Verlauf dieses Abschnitts erläutert werden.

Verhalten Es folgt eine Beschreibung, wie kooperative Teilnehmer in den Simulationen zusätzliche Kontakte wählen. Während die Simulationen in Abschnitt 4.9 erörtert werden, werden die Details der Kontaktwahl bereits hier beschrieben. Dies soll zum einen die Übersichtlichkeit erhöhen und zum anderen schon hier eine Intuition darüber geben, wie Teilnehmer ihre Kontakte wählen. Des Weiteren wurde in den verhaltensökonomischen Experimenten im Zuge dieser Arbeit beobachtet, dass sich menschliche Teilnehmer in der Tat so oder ähnlich verhalten, wie es im Folgenden beschrieben wird.

Simulierte Teilnehmer wählen Kontakte während der Anfragebearbeitung, daher wird die Kontaktwahl hier ebenfalls beschrieben. Die Idee der Kontaktwahl ist wie folgt: Ein Teilnehmer, der einen neuen zusätzlichen Kontakt sucht, kann seinen Anfragen eine „Anfrage für einen zusätzlichen Kontakt“ beifügen. Wenn eine Anfrage beantwortet wird, so kann der Beantworter die „Anfrage für einen zusätzlichen Kontakt“ bestätigen, wenn er ebenfalls an einem neuen zusätzlichen Kontakt interessiert ist. In diesem Fall fügt der Beantworter den Erzeuger zu seinen Kontakten hinzu. Der Erzeuger fügt den Beantworter zu seinen Kontakten hinzu, sobald er die Antwort auf die Anfrage mit der Bestätigung des Beantworters erhält.

Algorithmus 1: Senden einer erzeugten Anfrage

Input : Anfrage req , Teilnehmer i

- 1 Koordinate $k \leftarrow$ Ziel der Anfrage req ;
- 2 Menge an Teilnehmern $K_i \leftarrow$ Menge der Kontakte von i ;
- 3 **foreach** Kontakt $j \in K_i$ *aufsteigend geordnet nach* $\Delta(\text{zentrum}(j), k)$ **do**
- 4 **if** i hält j für kooperativ **then**
- 5 **if** Anzahl zusätzlicher Kontakte $< MAX$ **then**
- 6 füge „Anfrage für einen zusätzlichen Kontakt“ zu req hinzu;
- 7 sende req to j ;
- 8 return;
- 9 **else**
- 10 entferne(j, K_i);

Es folgt eine genauere Beschreibung wie kooperative Teilnehmer Anfragen verarbeiten und Kontakte wählen. Algorithmus 1 beschreibt, wie Teilnehmer erzeugte

Anfragen senden. Zuerst sucht Teilnehmer i den Kontakt, der dem Ziel der Anfrage am nächsten liegt und den er für kooperativ hält (Zeilen 2-4). Dieser sei mit Teilnehmer j bezeichnet. Er entfernt dabei diejenigen zusätzlichen Kontakte aus seiner Kontaktliste, die er für unkooperativ hält (Zeile 10). Es sei angemerkt, dass feste Kontakte nicht entfernt werden, da sie vom System vorgegeben sind. Wenn Teilnehmer i nicht die von ihm ausgewählte maximale Anzahl an zusätzlichen Kontakten überschreitet, dann fügt er der Anfrage eine Anfrage für einen zusätzlichen Kontakt hinzu (Zeilen 5-7). Zuletzt sendet Teilnehmer i die Anfrage an Teilnehmer j (Zeile 7).

Algorithmus 2: Beantworten von Anfragen

Input : Anfrage req , Teilnehmer j

- 1 Teilnehmer $f \leftarrow$ letzter Weiterleiter von req ;
- 2 **if** j hält f für unkooperativ auf der Anfrageebene **then**
- 3 | ignoriere req ;
- 4 | return;
- 5 **if** kann req nicht beantworten **then**
- 6 | leiteWeiter(req);
- 7 | return;
- 8 Teilnehmer $i \leftarrow$ Erzeuger von req ;
- 9 Antwort $a \leftarrow$ erzeuge Antwort für req ;
- 10 **if** Anzahl zusätzlicher Kontakte $<$ MAX **then**
- 11 | **if** req beinhaltet „Anfrage für einen zusätzlichen Kontakt“ **then**
- 12 | füge i Kontakten hinzu;
- 13 | füge „Akzeptiere neuen Kontakt“ zu a hinzu;
- 14 sende a zu i ;

Algorithmus 2 beschreibt, wie kooperative Teilnehmer Anfragen beantworten und Kontaktanfragen annehmen, falls sie Interesse haben. (Unkooperative Teilnehmer lassen die Anfrage mit einer bestimmten Wahrscheinlichkeit fallen.) Als erstes überprüft ein Teilnehmer j , der eine Anfrage erhält, ob sein Vorgänger in der Weiterleitungskette, also der Teilnehmer, von dem er die Anfrage erhalten hat, kooperativ ist (Zeile 2). Ist dies nicht der Fall, dann ignoriert Teilnehmer j die Anfrage und lässt sie fallen (Zeile 3-4). Wenn Teilnehmer j nicht das Ziel der Anfrage verwaltet und sie folglich nicht beantworten kann, dann leitet er die Anfrage weiter (Zeilen 5-7). Andernfalls produziert i eine Antwort (Zeile 9) und sendet diese im Verlauf des Algorithmus an den Erzeuger (Zeile 14). Falls j weniger zusätzliche Kontakte hat als er anstrebt, so überprüft er, ob die Anfrage eine „Anfrage für einen zusätzlichen Kontakt“ enthält. Ist dem so, fügt Teilnehmer j zum einen den Erzeuger der Anfrage, d.h. Teilnehmer i , seinen Kontakten hinzu, zum anderen fügt er der Antwort die Information „Akzeptiere neuen Kontakt“ hinzu, damit i darüber in Kenntnis gesetzt wird,

KAPITEL 4. ANREIZE ZUR VERNETZUNG

dass er zusätzlicher Kontakt von Teilnehmer j ist (Zeilen 10-13).

Algorithmus 3: Erhalt einer Antwort

Input : Antwort a

- 1 Teilnehmer $j \leftarrow$ Beantworter, d.h. Erzeuger von a ;
- 2 **if** *Anzahl zusätzlicher Kontakte* $<$ *MAX* **then**
- 3 **if** a beinhaltet „Akzeptiere neuen Kontakt“ **then**
- 4 füge j Kontakten hinzu;

Sobald ein Teilnehmer i eine Antwort erhält, verhält er sich, wie in Algorithmus 3 beschrieben. Wenn die Antwort einen „Akzeptiere neuen Kontakt“-Hinweis enthält, so prüft Teilnehmer i , ob er die von ihm gewählte maximale Anzahl an zusätzlichen Kontakten bereits überschritten hat (Zeile 2). Falls dem nicht so ist, fügt er seinen Kontakten den Beantworter seiner Anfrage, d.h. Teilnehmer j , hinzu (Zeile 4). Falls Teilnehmer i bereits genügend zusätzliche Kontakte besitzt, so fügt er Teilnehmer j nicht seinen Kontakten hinzu und eine unidirektionale Verbindung entsteht.

Beispiel: Teilnehmer i akzeptiert 20 zusätzliche Kontakte und hat im Moment 19 zusätzliche Kontakte. Daher hängt er seinen Anfragen eine „Anfrage für einen zusätzlichen Kontakt“ an. Nach einiger Zeit empfängt Teilnehmer j die Anfrage von i , fügt i zu seinen Kontakten hinzu und sendet ihm die Antwort zu seiner Anfrage mit dem „Akzeptiere neuen Kontakt“-Hinweis. Falls Teilnehmer i in der Zwischenzeit schon einen weiteren zusätzlichen Kontakt zu seinen Kontakten hinzugefügt hat, so wird Teilnehmer i Teilnehmer j nicht als zusätzlichen Kontakt akzeptieren. Dies ist darin begründet, dass Teilnehmer i sonst seine selbst gewählte maximale Anzahl an zusätzlichen Kontakten überschreiten würde. Daher ist Teilnehmer i zusätzlicher Kontakt von Teilnehmer j , aber j nicht zusätzlicher Kontakt von i . ■

Es sei angemerkt, dass unidirektionale Verbindungen nur vorübergehend existieren, da sie nicht stabil sind: Sei Teilnehmer j zusätzlicher Kontakt von Teilnehmer i , aber Teilnehmer i kein Kontakt von Teilnehmer j . Sendet nun i Anfragen an j , so kann Teilnehmer j diese (a) verarbeiten oder (b) ignorieren. In Fall (a) hat Teilnehmer j Ausgaben jedoch keine Einnahmen bezüglich Teilnehmer i . Im Sinne der Auszahlungsmaximierung ist es rational, i aus der Kontaktliste zu entfernen. In Fall (b) merkt Teilnehmer i , dass j unkooperativ ist und lässt j daher fallen, um seine Verluste zu begrenzen.

Algorithmus 4 zeigt wie kooperative Teilnehmer Anfragen weiterleiten. Das Weiterleiten von Anfragen ist analog zu dem Senden erzeugter Anfragen (vergleiche Algorithmus 1) mit dem Unterschied, dass die Teilnehmer keine „Anfrage für einen zusätzlichen Kontakt“ an ihre Anfragen anhängen.

Vereinfachung der Auszahlungsfunktion Die Auszahlungsfunktion, die in diesem Kapitel zugrunde gelegt wird, entspricht im Wesentlichen der in Kapitel 3.5.

Algorithmus 4: Weiterleiten von Anfragen

Input : Anfrage req , Teilnehmer f

- 1 Koordinate $k \leftarrow$ Ziel der Anfrage req ;
- 2 Menge an Teilnehmern $K_f \leftarrow$ Menge der Kontakte von f ;
- 3 **foreach** Kontakt $j \in K_f$ *aufsteigend geordnet nach* $\Delta(\text{zentrum}(j), k)$ **do**
- 4 **if** f hält j für kooperativ **then**
- 5 leite req an j ;
- 6 return;
- 7 **else**
- 8 entferne(j, K_f);

Um die Auszahlungsfunktion für die formale Analyse möglichst verständlich zu gestalten, wird sie wie folgt vereinfacht: (a) Es wird keine Diskontierung verwendet, d.h. es gilt $\delta = 1$ und (b) die Unterhaltskosten für zusätzliche Kontakte werden vereinfacht. Die Anzahl an Runden, für die Teilnehmer i zusätzliche Kontakte unterhalten muss, d.h. $r(i)$ wird durch die durchschnittliche Anzahl an zusätzlichen Kontakten $\overline{|K_i^+|}$ multipliziert mit der Anzahl an Runden ru , die sich Teilnehmer i im System befindet, approximiert. Es ergibt sich die folgende Auszahlungsfunktion:

$$z(i) = a \cdot |A_i| - q \cdot |R_i| - f \cdot |F_i| - w \cdot |W_i| - c \cdot \overline{|K_i^+|} \cdot ru \quad (4.1)$$

4.4 Annahmen

Zusätzlich zu den allgemeinen Annahmen in Kapitel 3.2 werden in diesem Kapitel weitere Annahmen gemacht, welche die formale Analyse vereinfachen.

Statische Teilnehmermenge Die formale Analyse in den Kapiteln 4.6 und 4.8 macht die Annahme, dass sich die Anzahl an Teilnehmern im System nicht verändert. Diese Annahme wird in den in Kapitel 4.9 durchgeführten Simulationen fallengelassen und die Anzahl an Teilnehmern im System wird variiert.

Die Strategie eines Teilnehmers ändert sich ebenfalls über den Verlauf einer Analyse oder einer Simulation nicht. Während der Anteil der Teilnehmer mit einer bestimmten Strategie in den Simulationen variiert wird und während die Teilnehmer verhaltensökonomischer Experimente (vgl. Kapitel 7) ihre Strategie frei abändern können, sind die Strategien in der formalen Analyse unveränderlich. In Kapitel 6 wird diese Annahme fallengelassen. Dort wird explizit die Dynamik von Strategien in der Interaktions- und Kontaktwahl untersucht.

4.5 Strategien

Das Verhalten eines verteilten Systems wird durch die Strategien der Teilnehmer (und die Netzwerkstruktur) bestimmt. Im Folgenden werden Strategien definiert, die für die Analyse von Deliberately Poor Connectedness wichtig sind.

In einem vollständig kooperativen System verarbeiten die Teilnehmer alle Anfragen, die sie erhalten. In realen Systemen hingegen sind Freifahrer wahrscheinlich [SGG02, CFSK04, FLSC04]. Eine Gegenmaßnahme gegen Freifahrer auf Anfrageebene ist es, sich reziprok zu verhalten, und ihre Anfragen zu ignorieren.

Definition 4.1 (Reziprokes Verhalten auf Anfrageebene)

Ein Teilnehmer verhält sich reziprok auf Anfrageebene, wenn er Anfragen von Kontakten bearbeitet, die seine Anfragen bearbeiten, und Anfragen von Kontakten ignoriert, die seine Anfragen ignorieren.

Da ein Teilnehmer die Strategie seiner Kontakte (und aller anderen Teilnehmer) nicht kennt, ist es praktisch unmöglich, sich reziprok zum echten Verhalten des Kontaktes zu verhalten. Vielmehr kann sich ein Teilnehmer nur reziprok gegenüber dem vermuteten Verhalten seiner Kontakte verhalten. So kann eine Schätzung des Verhaltens des Kontaktes zum Beispiel durch Beobachtungen und Statistik erfolgen. Es ist aus [SBSV06] bekannt, dass menschliche Teilnehmer daher Schwellenwert-Strategien (oft auch als Schwellwert-Strategien bezeichnet) in ihren Strategien gegen unkooperative Teilnehmer verwenden: Wenn ein Teilnehmer beobachtet, dass der Anteil der unbeantworteten Anfragen an den Anfragen, die er an einen Kontakt gesendet hat, einen Schwellenwert überschreitet, so ignoriert der Teilnehmer fortan die Anfragen, die ihm dieser Kontakt zusendet. Diese Schwellenwert-Strategie ist eine Approximation reziproken Verhaltens: Menschliche Teilnehmer machen ihre Kontakte für den Erfolg oder das Scheitern der Anfragebearbeitung einer ganzen Weiterleitungskette verantwortlich. Auch dann, wenn nicht der Kontakt sondern ein weiterer Teilnehmer in der Weiterleitungskette die Anfrage nicht bearbeitet. Dies ist insofern begründet, als dass die Kontakte wiederum verantwortlich sind, (kooperative) Kontakte auszuwählen.

Neben Schwellenwert-Strategien sind *Reputationssysteme* [RKZF00, JIB07] übliche Mittel gegen unkooperative Teilnehmer auf Anfrageebene. Reputationssysteme sammeln Informationen über den Kooperationsgrad jedes Teilnehmers und bieten sie den Teilnehmern an. Es sei angemerkt, dass diese Arbeit der eben genannten Definition von Reputationssystemen aus [RKZF00] folgt. Demnach ist eine Schwellenwert-Strategie noch kein Reputationssystem, da eine Schwellenwert-Strategie allein noch nicht für die Verteilung der Reputationsinformationen sorgt. Damit die im Folgenden durchgeführte Analyse nicht von einer bestimmten Technik zur Reputationsbestimmung abhängt, also von einem bestimmten Reputationssystem, nutzt die Analyse ein *Reputationsorakel* mit einstellbaren Fehlern f_p und f_n . Dieses wird von allen

Teilnehmern genutzt und stellt die (mehr oder minder fehlerhaften) Reputationsinformationen zur Verfügung. f_p ist der Anteil an „false positives“ und f_n der Anteil an „false negatives“. In anderen Worten, das Reputationsorakel klassifiziert Freifahrer auf Anfrageebene in f_p Prozent der Fälle als kooperativ und es klassifiziert auf der Anfrageebene kooperative Teilnehmer in f_n Prozent der Fälle als unkooperativ. Mit dem Reputationsorakel ist es möglich, Reputationsysteme mit unterschiedlicher Qualität im formalen Modell zu berücksichtigen.

Neben reziprokem Verhalten werden im Folgenden vier weitere Strategien definiert, die für die Untersuchung von Deliberately Poor Connectedness von Bedeutung sind: Strategien, die kooperativ oder unkooperativ auf Anfrageebene sind, und orthogonal dazu Strategien, die kooperativ oder unkooperativ auf der Kontaktebene sind.

Definition 4.2 (Kooperation auf Anfrageebene)

Ein Teilnehmer ist kooperativ auf der Anfrageebene, wenn er alle Anfragen bearbeitet, die er von Teilnehmern erhält, die er für kooperativ hält.

Definition 4.3 (Kooperation auf Kontaktebene)

Ein Teilnehmer ist kooperativ auf der Kontaktebene, wenn er mindestens J zusätzliche Kontakte hat. Ein Teilnehmer mit J oder mehr zusätzlichen Kontakten sei mit hinreichend verbunden bezeichnet.

J ist ein exogener, netzwerkabhängiger Parameter (vgl. [BG00a]).

Definition 4.4 (Vollständige Kooperation)

Ein Teilnehmer ist vollständig kooperativ, wenn er kooperativ auf Anfrageebene und kooperativ auf Kontaktebene ist.

Es sei angemerkt, dass in dieser Arbeit unterschiedliche Abstraktionsebenen im Zusammenhang mit dem Begriff der Kooperation verwendet werden: eine intuitive und eine präzise. Während „kooperatives Verhalten“ (wie z.B. in der Einleitung verwendet) intrinsisch ist, unterliegt die Definition von „vollständiger Kooperation“ einem beobachtbaren Verhalten (und ist daher eine präzisere Definition).

Definition 4.5 (Unkooperativ auf Anfrageebene)

Ein Teilnehmer ist unkooperativ auf der Anfrageebene, wenn er Anfragen fallen lässt, ohne dabei reziprok zu handeln. Das bedeutet, dass ein solcher Teilnehmer Anfragen ignoriert, die von Teilnehmern stammen, die er für kooperativ hält.

Sei u_f der Anteil an nicht weitergeleiteten Anfragen und u_a der Anteil an nicht beantworteten Anfragen. Beide Werte seien bezogen auf unkooperatives Verhalten. Es sei angemerkt, dass das hier vorgestellte Modell gemischte Strategien auf Anfrageebene zulässt: Parameter u_f und u_a erlauben die Modellierung von Teilnehmern, die jeden beliebigen Anteil von Anfragen bearbeiten. Dies reicht von Teilnehmern, die keine Anfragen bearbeiten ($u_f = u_a = 1$), bis hin zu Teilnehmern, die alle An-

KAPITEL 4. ANREIZE ZUR VERNETZUNG

fragen bearbeiten ($u_f = u_a = 0$). Neben der Anfrageebene können Teilnehmer sich unkooperativ auf der Kontaktebene verhalten.

Definition 4.6 (Deliberately Poor Connectedness (DPC))

Ein Teilnehmer nutzt eine Deliberately Poor Connectedness Strategie, wenn er $J_d < J$ zusätzliche Kontakte besitzt (sobald das System einen stabilen Zustand erreicht hat; wie in Abschnitt 4.4 angemerkt, werden in diesem Kapitel keine dynamischen Systeme betrachtet).

Die Perspektive eines einzelnen Teilnehmers zeigt, warum die Nutzung einer DPC-Strategie attraktiv ist: Ein Teilnehmer mit einer DPC-Strategie hat nur geringe Kosten zum Unterhalt zusätzlicher Kontakte und die Pfadlänge seiner Anfragen ist nur einen Schritt länger als die Pfadlänge der Anfragen seiner Kontakte. Er kann seine Anfragen stets über seine festen Kontakte weiterleiten.

Von einer globalen Perspektive hingegen ist Deliberately Poor Connectedness problematisch. Ein Teilnehmer mit einer solchen Strategie trägt nicht zur Netzwerkstruktur bei und in der Konsequenz steigt die Weiterleitungslast im System an.

Um die formale Analyse nicht in ihrer Komplexität ausufern zu lassen, wird in dieser angenommen, dass Teilnehmer mit einer DPC-Strategie keine zusätzlichen Kontakte besitzen, d.h. es gilt: $J_d = 0$. Diese Annahme wird in den Simulationen in Abschnitt 4.9 fallen gelassen, um ein umfangreiches Bild zu gewinnen.

Per Definition sei p die Wahrscheinlichkeit, dass die Teilnehmer eine bestimmte Strategie verfolgen. Der obere Index beschreibt dabei die Strategie:

- Der Index „kh“ (**k**ooperativ auf Anfrageebene, **h**inreichend verbunden) bedeutet, dass der Teilnehmer vollständig kooperativ ist.
- Der obere Index „kd“ (**k**ooperativ auf Anfrageebene, **D**PC) steht für Teilnehmer mit einer DPC-Strategie, die kooperativ auf Anfrageebene sind.
- „uh“ (**u**nkooperativ auf Anfrageebene, **h**inreichend verbunden) steht für Teilnehmer, die unkooperativ auf Anfrageebene und hinreichend verbunden sind.
- Index „ud“ (**u**nkooperativ auf Anfrageebene, **D**PC) steht für Teilnehmer, die unkooperativ auf Anfrageebene sind und die eine DPC-Strategie nutzen.

Es folgt:

$$p^{kh} + p^{kd} + p^{uh} + p^{ud} = 1$$

4.6 Deliberately Poor Connectedness

Bisher ist unklar, in welchen Situationen Deliberately Poor Connectedness zu höheren Auszahlungen führt als vollständige Kooperation. In diesem Abschnitt wird formal die zu erwartende Auszahlung für Teilnehmer mit einer DPC-Strategie berechnet

und die erwartete Auszahlung für vollständig kooperative Teilnehmer bestimmt. Dies soll zur Klärung der Frage dienen.

4.6.1 Durchschnittliche Pfadlänge

Die erwartete Auszahlung wird unter anderem dadurch beeinflusst, über wie viele Weiterleitungsschritte eine Anfrage gesendet wird. Daher ist eine genaue Schätzung der durchschnittlichen Pfadlänge wichtig.

Ein exaktes Vorgehen macht es nötig, zuerst eine spezifischere Definition der Pfadlänge vorzunehmen. Sei D_{opt} die durchschnittliche Pfadlänge, wenn keine Anfragen fallengelassen werden (d.h. wenn $u_f = u_a = 0$ gilt). Sei D_{avg} hingegen die tatsächliche Pfadlänge, die auftritt, wenn Anfragen fallengelassen werden. Offensichtlich wird die Pfadlänge geringer, wenn ein Teilnehmer aus der Weiterleitungskette die Anfrage fallen lässt und zwar um mindestens einen Schritt (nämlich mindestens den letzten in der Weiterleitungskette). Es folgt: $D_{avg} \leq D_{opt}$. In der Gegenwart von zusätzlichen Kontakten verändert sich die durchschnittliche Pfadlänge: Je mehr zusätzliche Kontakte die Teilnehmer haben, desto stärker sinkt die durchschnittliche Pfadlänge. Die durchschnittliche Pfadlänge mit ausschließlich festen Kontakten ist hingegen unveränderlich. Im Content-Addressable Network (ohne zusätzliche Kontakte) beträgt sie $\frac{\sqrt{n}}{2}$ [RFH⁺01] (eine Erklärung findet sich in Abschnitt 3.1).

Die Bestimmung der durchschnittlichen Pfadlänge für kooperative Kontakte mit zusätzlichen Kontakten ist nicht trivial: Es ist unklar, (a) wie die Teilnehmer ihre Kontakte im Adressraum verteilen und (b) ob die Teilnehmer ihre zusätzlichen Kontakte überhaupt für Anfragen anderer Teilnehmer nutzen (falls sie entsprechende Kontakte besitzen). Daher wird in der formalen Analyse angenommen, dass vollständig kooperative Teilnehmer ihre Kontakte in einer optimalen Art und Weise verteilen. Sollte sich herausstellen, dass Teilnehmer mit einer DPC-Strategie eine höhere zu erwartende Auszahlung besitzen als vollständig kooperative Teilnehmer, so folgt, dass dies natürlich auch dann der Fall ist, wenn vollständig kooperative Teilnehmer ihre Kontakte nicht optimal verteilen. Mit anderen Worten stellt sich die Frage, ob Deliberately Poor Connectedness erfolgreicher ist als kooperatives Verhalten, selbst wenn wir annehmen, dass die kooperativen Teilnehmer ihre Kontakte optimal verteilen. Falls dem so ist, folgt, dass Deliberately Poor Connectedness wahrscheinlicher ist als Kooperation.

Im optimalen Fall haben alle Teilnehmer zusätzliche Kontakte und nutzen diese. Als eine untere Schranke für die durchschnittliche Anzahl an Weiterleitungen kann die folgende Situation betrachtet werden: Für jede Anfrage teilt ein zusätzlicher Kontakt den verbleibenden Suchraum in zwei Teile. Haben die Teilnehmer $|K^+|$ zusätzliche Kontakte, so folgt, dass Anfragen in $\log_{|K^+|}(n)$ Schritten bearbeitet werden können. Daher wird für den optimalen Fall angenommen, dass die Pfadlänge von

KAPITEL 4. ANREIZE ZUR VERNETZUNG

vollständig kooperativen Teilnehmer D_{best}^{coop} wie folgt ist:

$$D_{best}^{coop} = \log_{|K^+|}(n) \quad (4.2)$$

Es sei angemerkt, dass in der Realität – selbst wenn die Teilnehmer eine optimale Verteilung der zusätzlichen Kontakte besitzen – die Pfadlänge um Größenordnungen höher liegt [Kle00].

Nachdem gezeigt wird, dass Deliberately Poor Connectedness lohnender ist als vollständige Kooperation, wird gezeigt werden, dass die in dieser Arbeit vorgeschlagene C4C-Strategie gegen Deliberately Poor Connectedness effektiv ist. In dem zugehörigen Beweis wäre es falsch anzunehmen, dass der beste Fall für kooperative Teilnehmer eintritt. Im Gegenteil muss vom schlechtesten Fall für kooperative Teilnehmer ausgegangen werden, wenn die Effektivität der C4C-Strategie gezeigt wird: Wenn ein kooperativer Teilnehmer eine Anfrage weiterleitet, dann muss im schlechtesten Fall davon ausgegangen werden, dass die Weiterleitungskette aus Teilnehmern mit der DPC-Strategie besteht. Es muss also davon ausgegangen werden, dass die Anfrage lediglich im ersten Schritt mit Hilfe von zusätzlichen Kontakten weitergeleitet wird. In anderen Worten, ein kooperativer Teilnehmer kann im schlechtesten Fall nur von seinen eigenen zusätzlichen Kontakten profitieren. Falls ein Teilnehmer in der Weiterleitungskette doch seine zusätzlichen Kontakte nutzen sollte, dann erhöht sich die Auszahlung des Erzeugers verglichen mit dem schlechtesten Fall. Wenn vollständig kooperative Teilnehmer die Algorithmen 1 bis 3 aus Abschnitt 4.3 verwenden, dann ist die Pfadlänge der Anfragen kooperativer Teilnehmer im schlechtesten Fall D_{worst}^{coop} wie im folgenden Lemma beschrieben.

Lemma 4.6.1 *Wenn jeder kooperative Teilnehmer über den Adressraum gleichverteilte zusätzliche Kontakte hat, dann ist die Anzahl der Weiterleitungsschritte im schlechtesten Fall D_{worst}^{coop} wie folgt:*

$$D_{worst}^{coop} = \frac{1}{2} \int_0^{\frac{\sqrt{n}}{2}} \left(\omega_0(r) \cdot r + \omega_1(r) \cdot \left(r + \frac{\sqrt{n}}{2} \right) \right) dr \quad (4.3)$$

mit

$$\omega_0(r) = 4 \cdot r \cdot \frac{|K^+| + 1}{n} \cdot e^{-2 \cdot r^2 \cdot \frac{|K^+| + 1}{n}}$$

und

$$\omega_1(t) = (-4t + 2\sqrt{n}) \frac{|K^+| + 1}{n} \left[1 - \omega_0\left(\frac{\sqrt{n}}{2}\right) \right] e^{(2t^2 - 2\sqrt{nt}) \frac{|K^+| + 1}{n}}$$

Appendix A.1 enthält den zugehörigen Beweis. $\omega_0(r)$ und $\omega_1(r)$ sind die Wahrscheinlichkeiten, dass ein Teilnehmer ein oder mehr zusätzliche Kontakte besitzt, die den (Adressraum-) Abstand r zum Ziel der Anfrage besitzen. Durch die Torus-Eigenschaft des Adressraums müssen dabei zwei Fälle unterschieden werden. In Fall (a) formen Punkte mit Abstand r zum Teilnehmer t einen Rhombus. In Fall (b)

4.6. DELIBERATELY POOR CONNECTEDNESS

fehlen die Ecken des Rhombus, aufgrund der Torus-Eigenschaft. $\omega_0(r)$ gilt dabei in Fall (a), bei welchem $0 \leq r \leq \frac{\sqrt{n}}{2}$ gilt, und $\omega_1(r)$ in Fall (b), bei dem $r > \frac{\sqrt{n}}{2}$ gilt. Eine umfangreichere Beschreibung findet sich in Appendix A.1. Es sei angemerkt, dass Gleichung (4.3) im für kooperative Teilnehmer schlechtesten Fall gilt: Selbst, wenn nur der erste Teilnehmer einer Weiterleitungskette zusätzliche Kontakte verwendet und alle weiteren Teilnehmer dies nicht tun (weil sie z.B. eine DPC-Strategie verwenden), gilt Gleichung (4.3).

Natürlicherweise ist die Anzahl an Anfragen, die ein Teilnehmer weiterleitet, abhängig von der Anzahl an Anfragen im System und ihren Pfadlängen. Ist die Anzahl der Anfragen, die ein Teilnehmer im Durchschnitt erzeugt, mit $|\overline{R}|$ gegeben, so erhält man die Anzahl an Weiterleitungen, wenn man $|\overline{R}|$ mit der durchschnittlichen Pfadlänge multipliziert:

$$\overline{F} = \frac{|\overline{R}| \cdot (D_{avg} - 1) \cdot n}{n} = |\overline{R}| \cdot (D_{avg} - 1) \quad (4.4)$$

Dies bedeutet, dass jede Anfrage durch $(D_{avg} - 1)$ Teilnehmer weitergeleitet werden muss. (Die Subtraktion um eins ist nötig, da im Modell die Kosten des ersten Weiterleitungsschrittes bereits durch die Kosten für das Erzeugen der Anfrage modelliert sind.)

4.6.2 Einfluss unkooperativer Teilnehmer

Neben der durchschnittlichen Pfadlänge, ist die Effizienz eines Systems negativ mit dem Anteil an unkooperativen Teilnehmern auf Anfrageebene korreliert: Selbst wenn Anfragen von kooperativen Teilnehmern erfolgreich an unkooperativen Teilnehmern vorbeigeleitet werden, gehen dennoch Anfragen verloren, da das Ziel einer Anfrage ebenfalls ein unkooperativer Teilnehmer sein kann, der die Anfrage fallen lässt. Die Wahrscheinlichkeit, dass Anfragen verloren gehen, ist ebenfalls negativ korreliert mit dem Vermögen der Teilnehmer, zwischen unkooperativen und kooperativen Teilnehmern zu unterscheiden: Wenn die Teilnehmer in der Lage sind, unkooperative Teilnehmer mit einer geringen Fehlerrate zu identifizieren, dann sind sie ebenfalls in der Lage, ihre Anfragen an kooperative Teilnehmer zu senden. Andersherum folgt aus dem Unvermögen unkooperative Teilnehmer zu identifizieren, dass Anfragen fälschlicherweise an unkooperative Teilnehmer geleitet werden und verloren gehen. In der in dieser Arbeit vorgenommenen Analyse wird davon ausgegangen, dass ein Teilnehmer erst das Reputationsorakel (also die Analogie zu einem Reputationssystem; vergleiche Abschnitt 4.5) befragt, bevor er eine Anfrage an einen Teilnehmer sendet. Er erfragt dabei, ob der Teilnehmer kooperativ auf Anfrageebene ist oder nicht. Falls ja, dann sendet er ihm die Anfrage zu, falls nein, dann sucht er sich einen anderen Kontakt. Falls ein Teilnehmer laut dem Orakel nur unkooperative Kontakte besitzt, verwirft er die Anfrage (vergleiche Algorithmus 1). Im Folgenden wird näherungs-

KAPITEL 4. ANREIZE ZUR VERNETZUNG

weise berechnet, wie hoch die Wahrscheinlichkeit ist, dass eine Anfrage weitergeleitet wird.

Die Näherung der Wahrscheinlichkeit P_F , dass eine Anfrage weitergeleitet wird, wird aus vier Teilen berechnet. In den ersten beiden Fällen sendet ein Teilnehmer, der kooperativ auf Anfrageebene ist, die Anfrage, in den beiden letzten Fällen ein Teilnehmer, der unkooperativ auf Anfrageebene ist.

1. Wir nehmen an, dass die Anfrage von einem Teilnehmer, der kooperativ auf Anfrageebene ist ($p^{kh} + p^{kd}$), zu einem Teilnehmer gesendet wird, der ebenfalls kooperativ auf Anfrageebene ist ($p^{kh} + p^{kd}$). Die Anfrage wird nur dann verarbeitet, wenn der Empfänger den Sender nicht fälschlicherweise als unkooperativ einschätzt ($1 - f_n$). Die Gesamtwahrscheinlichkeit beträgt: $(p^{kh} + p^{kd})^2 \cdot (1 - f_n)$
2. Sendet ein auf Anfrageebene kooperativer Teilnehmer ($p^{kh} + p^{kd}$) die Anfrage an einen Teilnehmer, der unkooperativ auf Anfrageebene ist ($p^{uh} + p^{ud}$), so darf (a) der Empfänger den Sender nicht fälschlicherweise als unkooperativ einschätzen ($1 - f_n$) und (b) muss der unkooperative Empfänger die Anfrage trotz seiner Unkooperativität verarbeiten ($1 - u_f$). Die Gesamtwahrscheinlichkeit beträgt: $(p^{kh} + p^{kd}) \cdot (p^{uh} + p^{ud}) \cdot (1 - f_n) \cdot (1 - u_f)$
3. Wir nehmen an, dass die Anfrage von einem Teilnehmer, der unkooperativ auf Anfrageebene ist ($p^{uh} + p^{ud}$), zu einem Teilnehmer gesendet wird, der kooperativ auf Anfrageebene ist ($p^{kh} + p^{kd}$). Der Empfänger muss den unkooperativen Sender fälschlicherweise als kooperativ einschätzen, sonst verarbeitet er die Anfrage nicht (f_p). Die Gesamtwahrscheinlichkeit beträgt: $(p^{uh} + p^{ud}) \cdot (p^{kh} + p^{kd}) \cdot f_p$
4. Sendet ein auf Anfrageebene unkooperativer Teilnehmer ($p^{uh} + p^{ud}$) die Anfrage an einen ebenfalls auf Anfrageebene unkooperativen Teilnehmer ($p^{uh} + p^{ud}$), so muss (a) der Empfänger den Sender fälschlicherweise als kooperativ einschätzen (f_p) und (b) die Anfrage trotz seiner Unkooperativität verarbeiten ($1 - u_f$). Die Gesamtwahrscheinlichkeit beträgt: $(p^{uh} + p^{ud})^2 \cdot f_p \cdot (1 - u_f)$

Insgesamt ergibt sich für die Näherung der Wahrscheinlichkeit P_F die folgende Formel:

$$\begin{aligned}
 P_F = & (p^{kh} + p^{kd})^2 \cdot (1 - f_n) \\
 & + (p^{kh} + p^{kd}) \cdot (p^{uh} + p^{ud}) \cdot (1 - f_n) \cdot (1 - u_f) \\
 & + (p^{uh} + p^{ud}) \cdot (p^{kh} + p^{kd}) \cdot f_p \\
 & + (p^{uh} + p^{ud})^2 \cdot f_p \cdot (1 - u_f)
 \end{aligned} \tag{4.5}$$

Im Fall einer nahezu perfekten Unterscheidung von kooperativen und unkooperativen Teilnehmern durch das Reputationssystem, d.h. es gilt f_n, f_p sind nahe null, geht P_F

4.6. DELIBERATELY POOR CONNECTEDNESS

gegen $((p^{kh} + p^{kd})^2 + (p^{kh} + p^{kd}) \cdot (p^{uh} + p^{ud}) \cdot (1 - u_f))$. Dies bedeutet, dass nur noch Anfragen bearbeitet werden, die von Teilnehmern stammen, die kooperativ auf Anfrageebene sind. In der Konsequenz ist Freifahren auf Anfrageebene unattraktiv und als Strategie unwahrscheinlich.

Die Wahrscheinlichkeit P_A , dass eine Anfrage beantwortet wird, kann analog hergeleitet werden:

$$\begin{aligned}
 P_A &= (p^{kh} + p^{kd})^2 \cdot (1 - f_n) \\
 &+ (p^{kh} + p^{kd}) \cdot (p^{uh} + p^{ud}) \cdot (1 - f_n) \cdot (1 - u_a) \\
 &+ (p^{uh} + p^{ud}) \cdot (p^{kh} + p^{kd}) \cdot f_p \\
 &+ (p^{uh} + p^{ud})^2 \cdot f_p \cdot (1 - u_a)
 \end{aligned} \tag{4.6}$$

Da eine Anfrage über $(D_{opt} - 1)$ Schritte geleitet werden muss, ist die Approximation für Wahrscheinlichkeit $P(D_{opt}, P_F, P_A)$, dass eine Anfrage bearbeitet wird, wie folgt:

$$P(D_{opt}, P_F, P_A) = P_F^{D_{opt}-1} \cdot P_A \tag{4.7}$$

Die durchschnittliche Anzahl $\overline{|A|}$ an Anfragen, die beantwortet werden, ist wie folgt:

$$\overline{|A|} = P(D_{opt}, P_F, P_A) \cdot \overline{|R|} = P_F^{D_{opt}-1} \cdot P_A \cdot \overline{|R|} \tag{4.8}$$

4.6.3 Nutzen der DPC-Strategie

Setzt man die Gleichungen (4.4) und (4.8) in Gleichung (4.1) ein, so erhält man die zu erwartende Auszahlung eines kooperativen Teilnehmers. Sie sei in der Folge mit z_c beschrieben:

$$\begin{aligned}
 z_c &= P(D_{opt}, P_F, P_A) \cdot a \cdot \overline{|R|} - (q \cdot \overline{|R|} + P(D_{avg}, P_F, P_A) \\
 &\cdot (D_{avg} - 1) \cdot f \cdot \overline{|R|} + w \cdot |W| + c \cdot \overline{|K^+|} \cdot r)
 \end{aligned} \tag{4.9}$$

Ohne Strategien gegen Deliberately Poor Connectedness haben Teilnehmer mit einer DPC-Strategie aus zwei Gründen eine höhere Auszahlung: Erstens, sie haben keine Unterhaltskosten für zusätzliche Kontakte. Zweitens, sie haben eine geringere Weiterleitungslast, da sie keine Anfragen über zusätzliche Kontakte erhalten.

Es sei angemerkt, dass eine Formel, die spezifiziert, wie viele Anfragen ein Teilnehmer i empfängt, unnötig komplex und unpraktisch ist: Die Weiterleitungslast hängt (a) von der Anzahl an Kontakten ab, (b) von ihren Strategien (d.h. von dem Anteil der Anfragen, den sie fallenlassen), (c) davon, wie viele Anfragen sie erzeugen und wie viele Anfragen sie selbst erhalten, und (d) davon, wie viele Anfragen die Kontakte zu Teilnehmer i leiten. Daher wird in dieser Arbeit eine Näherung verwendet: Teilnehmer mit einer DPC-Strategie leiten um die $(\frac{|K^f|}{|K^+| + |K^f|} \cdot \overline{F})$ Anfragen weiter. Das folgende Beispiel macht die Intuition dahinter deutlich.

Beispiel: Sei ein stark vernetztes System angenommen. Kooperative Kontakte haben vier feste und im Durchschnitt 40 zusätzliche Kontakte. Teilnehmer d hingegen hat lediglich vier feste Kontakte (und keine zusätzlichen). Es folgt, dass Teilnehmer d nur etwa $\frac{4}{44} = \frac{1}{11}$ der Anfragen als Weiterleitungen empfängt wie ein kooperativer Teilnehmer. ■

Es wird angenommen, dass diese Näherung angemessen ist, da von einem statistischen Standpunkt aus betrachtet Kriterien (b) bis (d) für alle Teilnehmer eines großen Systems gleich sind. Kriterium (a) – die Anzahl an zusätzlichen Kontakten – wird vom Teilnehmer selbst gewählt. Dies ist Teil der Näherung.

Es sei angemerkt, dass die DPC-Strategie auch einen Nachteil für ihre Anwender hat: Teilnehmer mit einer DPC-Strategie erhalten geringfügig weniger Antworten als kooperative Teilnehmer. Ihre Anfragen haben eine etwas höhere Pfadlänge und in der Konsequenz eine höhere Wahrscheinlichkeit, auf einen Freifahrer auf Anfrageebene zu treffen, der die Anfrage fallen lässt. Dennoch ist dieser Effekt nicht bedeutend: Wenn ein Teilnehmer mit einer DPC-Strategie seine Anfrage an einen kooperativen festen Kontakt sendet, so hat die Anfrage nur eine Weiterleitung mehr als die Anfragen des festen Kontakts. Die Wahrscheinlichkeit $P_d(D_{opt}, P_F, P_A)$, dass eine Anfrage, die von einem Teilnehmer mit einer DPC-Strategie erzeugt wurde (und der kooperativ auf der Anfrageebene ist), beantwortet wird, ist:

$$P_d(D_{opt}, P_F, P_A) := P(D_{opt} + 1, P_F, P_A) \quad (4.10)$$

Insgesamt folgt, dass die zu erwartende Auszahlung eines Teilnehmers mit einer DPC-Strategie z_d wie folgt ist:

$$z_d = P_d(D_{opt}, P_F, P_A) \cdot a \cdot \overline{|R|} - (q \cdot \overline{|R|} + \frac{|K^f|}{|K^+| + |K^f|} \cdot P(D_{opt}, P_F, P_A) \cdot (D_{avg} - 1) \cdot f \cdot \overline{|R|} + w \cdot \overline{W}) \quad (4.11)$$

Um zu analysieren, in welchen Situationen eine Kombination aus Deliberately Poor Connectedness und Kooperation auf Anfrageebene zu höheren Auszahlungen führt als vollständige Kooperation, wird im Folgenden die zu erwartende Auszahlung von Teilnehmern mit der entsprechenden Strategie verglichen.

4.6.4 Gefahr der DPC-Strategie

Im Folgenden wird die Auszahlung eines Teilnehmers mit einer vollständig kooperativen Strategie mit der Auszahlung eines Teilnehmers mit einer DPC-Strategie verglichen. Wie sich zeigen wird, ist Deliberately Poor Connectedness in vielen Situationen lohnender als vollständige Kooperation. Der Vergleich der Auszahlungen wird in Relation des Gewinns a durch den Erhalt einer Anfrage zu den Bearbeitungskosten und Kosten für zusätzliche Kontakte durchgeführt. Dieser Vergleich ist dadurch motiviert,

4.6. DELIBERATELY POOR CONNECTEDNESS

dass für verschiedene Anwendungen und Systeme der Gewinn für eine Antwort (im Verhältnis zu den Kosten) unterschiedlich sein kann. Insbesondere kann a sehr klein sein, z.B. in Internet-Anwendungen.

Der Vergleich der zu erwartenden Auszahlungen aus dem vorhergehenden Abschnitt führt zum folgenden Lemma:

Lemma 4.6.2 *Es sei die Wahrscheinlichkeit, dass eine Anfrage weitergeleitet wird kleiner als 100%, d.h. $P_F < 1$. Ein Teilnehmer, der eine Deliberately Poor Connectedness Strategie bei gleichzeitiger Kooperation auf der Anfrageebene verfolgt, hat eine höhere Auszahlung als ein Teilnehmer, der vollständig kooperativ ist, immer dann, wenn die folgende Ungleichung gilt:*

$$a < \frac{\frac{|K^+|}{|K^+| - |K^f|} \cdot (D_{avg} - 1)}{(1 - P_F)} \cdot f + \frac{1}{(1 - P_F)} \cdot c \cdot |K^+| \quad (4.12)$$

Appendix A.2 enthält den Beweis. Lemma 4.6.2 hat zwei Implikationen: Erstens, bei hohen Weiterleitungskosten f oder hohen Unterhaltskosten für zusätzliche Kontakte c führen nur hohe Gewinne a für den Erhalt einer Anfrage dazu, dass vollständige Kooperation lohnender ist als Deliberately Poor Connectedness (mit Kooperation auf der Anfrageebene). In diesem Fall sind der erste und der zweite Summand auf der rechten Seite der Ungleichung (4.12) hoch und in der Konsequenz müssen die Gewinne für Antworten a ebenfalls hoch sein. Diese Implikation ist natürlich, da viele zusätzliche Kontakte zu einer relativ gesehen höheren Weiterleitungslast führen. Im Gegensatz dazu ist die zweite Implikation unerwartet und überraschend: Wenn alle Teilnehmer Anfragen mit einer großen Wahrscheinlichkeit weiterleiten ($P_F \rightarrow 1$), dann ist Deliberately Poor Connectedness lohnender als ein kooperatives Verhalten; wenn $P_F \rightarrow 1$ gilt, dann sind beide Nenner in Ungleichung (4.12) nahe null und folglich wird die rechte Seite der Ungleichung gegen Unendlich wachsen. Dies ist problematisch: Die Situation tritt dann auf, wenn es den Teilnehmern eines verteilten Systems zur Anfragebearbeitung gelingt, den Kooperationsgrad anderer Teilnehmer korrekt einzuschätzen, also immer dann, wenn die Klassifikation der Freifahrer auf Anfrageebene (z.B. durch ein Reputationssystem) erfolgreich ist. Dieses Ergebnis ist überraschend, da nicht zu erwarten gewesen ist, dass notwendige Gegenmaßnahmen gegen Freifahrerverhalten auf Anfrageebene einer neuen Art des Freifahrens – nämlich Deliberately Poor Connectedness – Vorschub leistet. Das folgende Beispiel zeigt für übliche Parameterwerte und in der Abwesenheit von Mechanismen oder Strategien gegen Deliberately Poor Connectedness, dass DPC-Strategien zu einer höheren Auszahlung führen als vollständige Kooperation. Dies ist der Fall, solange der Gewinn für den Erhalt einer Anfrage nicht bedeutend größer ist als die Kosten des Weiterleitens oder die Kosten des Unterhalts für zusätzliche Kontakte.

Beispiel: Es seien 1.000.000 Teilnehmer im System. 95% von ihnen seien kooperativ auf Anfrageebene und 5% bearbeiten keine Anfragen. Des Weiteren nutzen

KAPITEL 4. ANREIZE ZUR VERNETZUNG

50% der Teilnehmer Deliberately Poor Connectedness. In anderen Worten sei also die folgende Verteilung an Strategien angenommen: $p^{kh} = 0,475$; $p^{kd} = 0,475$ sowie $p^{uh} = 0,025$; $p^{ud} = 0,025$. Teilnehmer, die keine DPC-Strategie nutzen, haben 40 zusätzliche Kontakte ($|\bar{K}^+| = 40$). Das durch das Reputationsorakel dargestellte Reputationssystem habe eine false-positive Rate von 0,03 und eine false-negative Rate von 0,05. Weiterleitungskosten für eine Anfrage und Kosten für den Unterhalt eines zusätzlichen Kontaktes seien beide eins (d.h. $c = f = 1$). Es folgt aus Lemma 4.6.2, dass der Gewinn einer Anfrage a mindestens 600 mal höher sein muss als die Weiterleitungskosten oder die Unterhaltskosten für Kontakte, damit vollständige Kooperation zu einer höheren Auszahlung führt als Deliberately Poor Connectedness. Andernfalls ist Deliberately Poor Connectedness wahrscheinlicher als Kooperation. ■

Wenn Deliberately Poor Connectedness lohnender ist als andere Strategien, dann folgt daraus, dass die Teilnehmer eine DPC-Strategie verfolgen und dass nur wenige (oder keine) zusätzlichen Kontakte im System existieren. Mit Ausnahme des Falls, in dem nur ein leeres Netzwerk effizient ist, führt ein solches Netz nicht zu einem effizienten System (vergleiche Kapitel 2).

Wie in den Ergebnissen der verhaltensökonomischen Experimente in Kapitel 7 zu sehen ist, nutzen einige menschliche Teilnehmer in der Tat Deliberately Poor Connectedness und haben keine oder nur sehr wenige zusätzliche Kontakte.

4.7 Contact-for-Contact Strategie

Deliberately Poor Connectedness führt zu kaum vernetzten Systemen. Teilnehmer dieser Systeme haben eine höhere Weiterleitungslast als nötig. Ebenfalls ist problematisch, dass Deliberately Poor Connectedness zu nicht effizienten Systemen führt. Um eine Gegenmaßnahme gegen Deliberately Poor Connectedness zu haben, wird in dieser Arbeit die sogenannte C4C-Strategie (Contact-for-Contact Strategie) vorgeschlagen. Nutzen hinreichend viele Teilnehmer diese Strategie, so stellt sie sicher, dass ein Teilnehmer im gleichen Maß von den zusätzlichen Kontakten anderer profitiert wie von seinen eigenen: Ein Teilnehmer leitet eine Anfrage nur so weit durch den Adressraum (relativ zum Ziel der Anfrage), wie es sein Vorgänger in der Weiterleitungskette getan hat. Folglich können Anfragen von Teilnehmern mit einer DPC-Strategie nicht von den zusätzlichen Kontakten anderer profitieren. Dies führt zu einer großen Pfadlänge für Anfragen von Teilnehmern mit einer DPC-Strategie, was wiederum zu einer geringeren Auszahlung führt. Durch die C4C-Strategie geben kooperative Teilnehmer Teilnehmern mit einer DPC-Strategie einen Anreiz, sich stärker zu vernetzen und einen Beitrag zur Netzwerkstruktur zu leisten. Algorithmus 5 beschreibt die C4C-Strategie: Bevor ein Teilnehmer f eine Anfrage req weiterleitet, überprüft f wie weit sein Vorgänger in der Weiterleitungskette die Distanz der Anfrage zum Ziel verkürzt hat. Dazu berechnet er (a) die Distanz zwischen seinem

4.7. CONTACT-FOR-CONTACT STRATEGIE

Algorithmus 5: Weiterleiten einer Anfrage – C4C-Strategie

Input : Anfrage req , Teilnehmer f

- 1 Koordinate $t \leftarrow$ Ziel der Anfrage req ;
- 2 Teilnehmer $p \leftarrow$ der letzte Teilnehmer der req zu f geleitet hat;
- 3 Radius $\varrho \leftarrow \Delta(\text{zentrum}(p), t) - \Delta(\text{zentrum}(f), t)$;
- 4 **if** $\varrho < 0$ **then**
- 5 \perp return;
- 6 Teilnehmermenge $K_f^+ \leftarrow$ Menge der zusätzlichen Kontakte von f ;
- 7 **foreach** Kontakt $i \in K_f^+$ **aufsteigend geordnet nach** $\Delta(\text{zentrum}(i), t)$ **do**
- 8 **if** $\Delta(\text{zentrum}(f), \text{zentrum}(i)) < \varrho$ **then**
- 9 \perp leite req zu Kontakt i ;
- 10 \perp return;
- 11 leite req zum dem, dem Ziel der Anfrage am nächsten gelegenen, festen Kontakt von f , der kooperativ ist;

Vorgänger p (engl. predecessor) und dem Ziel der Anfrage t (engl. target) und zieht (b) seine eigene Distanz zum Ziel der Anfrage von der in (a) berechneten Distanz ab (Zeile 3). Falls Teilnehmer f weiter vom Ziel der Anfrage entfernt ist als sein Vorgänger p , dann verwirft er die Anfrage (Zeilen 4-5). Hat die Distanz nicht zugenommen, so wählt Teilnehmer f einen Kontakt mit den folgenden Eigenschaften: Der Kontakt ist der dem Ziel der Anfrage am nächsten gelegene zusätzliche Kontakt, der nicht weiter als ϱ von ihm entfernt ist (Zeilen 6-10). ϱ ist der Radius der C4C-Strategie bezüglich einer gegebenen Anfrage und bezeichnet die in Zeile 3 berechnete Distanz. Falls kein zusätzlicher Kontakt mit einer solchen Eigenschaft existiert, dann leitet Teilnehmer f die Anfrage zu einem (kooperativen) festen Kontakt, der näher am Ziel der Anfrage liegt als er selbst. (Falls kein solcher fester Kontakt existiert, verwirft er die Anfrage.)

Es war ein Entwurfsziel bei der Entwicklung der C4C-Strategie, dass kooperative Teilnehmer die Anfragen von Teilnehmern mit einer DPC-Strategie nicht einfach verwerfen. Dies würde es z.B. unmöglich machen, einem solchen System beizutreten, da ein Teilnehmer, der einem System beitrifft, keine zusätzlichen Kontakte besitzt und auch keine kennenlernen könnte, da seine Anfragen fallengelassen werden würden.

Die Verwendung von Deliberately Poor Connectedness sorgt dafür, dass ein solcher Teilnehmer beim Weiterleiten seiner Anfragen diese ihrem Ziel kaum näher bringt. Falls die festen Kontakte eines Teilnehmers mit einer DPC-Strategie die C4C-Strategie benutzen, so werden sie nur ihre festen Kontakte für seine Anfragen verwenden. In der Konsequenz werden die Pfadlängen für Anfragen von Teilnehmern mit einer DPC-Strategie so groß wie in einem System ohne zusätzliche Kontakte. Durch die Verwendung der C4C-Strategie erzeugen kooperative Teilnehmer eine Si-

tuation, in der nur die Teilnehmer, die durch zusätzliche Kontakte zur Netzwerkstruktur beitragen, auch von ihr einen Nutzen haben. In dieser Situation werden Teilnehmer, die ihre Auszahlung maximieren, zusätzliche Kontakte wählen, wie im nächsten Abschnitt in einer formalen Analyse gezeigt werden wird. Des Weiteren zeigt diese Analyse, dass vollständige Kooperation lohnender ist als Deliberately Poor Connectedness, wenn die kooperativen Teilnehmer die C4C-Strategie nutzen. Dies ist der Fall, obwohl der schlechteste Fall für kooperative Teilnehmer angenommen wird: Selbst wenn der erste Teilnehmer, dem kooperative Teilnehmer ihre Anfragen weiterleiten, eine DPC-Strategie verwendet, haben kooperative Teilnehmer eine höhere Auszahlung als Teilnehmer mit einer DPC-Strategie.

4.8 Analyse der Contact-for-Contact Strategie

Um die Effektivität der C4C-Strategie zu prüfen, wird in diesem Abschnitt eine formale Analyse durchgeführt. Annahme dabei (und bei den Simulationen in Abschnitt 4.9) ist, dass kooperative Teilnehmer die C4C-Strategie nutzen. Eine Argumentation darüber, warum diese Annahme nicht unrealistisch ist, ist in Abschnitt 4.10 enthalten.

Die Nutzung der C4C-Strategie hat zwei Effekte. Der erste Effekt ist, dass Teilnehmer mit einer DPC-Strategie keinen Nutzen von zusätzlichen Kontakten anderer Teilnehmer haben. Es folgt, dass ihre Anfragen die gleiche Anzahl an Weiterleitungsschritten haben wie in einem Content-Addressable Network ohne zusätzliche Kontakte. Der zweite Effekt ist, dass Anfragen von Teilnehmern mit zusätzlichen Kontakten (im Vergleich) eine kürzere Pfadlänge besitzen (vergleiche Abschnitt 4.6.1). Es sei angemerkt, dass dem so ist, selbst wenn nur der Teilnehmer, der zusätzliche Kontakte besitzt, diese nutzt und sonst keine weiteren Teilnehmer. Die durchschnittliche Pfadlänge für Anfragen von Teilnehmern mit einer DPC-Strategie D_{opt}^{dpc} ist $\frac{\sqrt{n}}{2}$. Daher gilt $D_{opt} = \frac{\sqrt{n}}{2} = D_{opt}^{dpc}$ in Gleichung (4.10). Die Anzahl der Weiterleitungen für Anfragen kooperativer Teilnehmer (D_{worst}^{coop}) ist deutlich geringer (vergleiche Appendix A.1)

Das folgende Ergebnis wird vom Vergleich der zu erwartenden Auszahlungen von Teilnehmern mit einer DPC-Strategie und kooperativen Teilnehmern abgeleitet.

Lemma 4.8.1 *Sei die Wahrscheinlichkeit P_F , dass eine Anfrage weitergeleitet wird kleiner als 100% (d.h. $P_F < 1$). Dann führt Deliberately Poor Connectedness (mit Kooperation auf Anfrageebene) zu höheren Auszahlungen als vollständige Kooperation, wenn die folgende Ungleichung gilt:*

$$a < \frac{\frac{|K^+|}{|K^+| - |K^f|} \cdot (D_{avg} - 1)}{(1 - P_F^{D_{opt}^{dpc} - D_{worst}^{coop}})} \cdot f + \frac{1}{(1 - P_F^{D_{opt}^{dpc} - D_{worst}^{coop}})} \cdot c \cdot \bar{\gamma}$$

Appendix A.3 enthält einen Beweis. Lemma 4.8.1 besitzt eine wichtige Implikation: Je stärker kooperative Teilnehmer in zusätzliche Kontakte investieren, desto geringer wird die Pfadlänge für ihre Anfragen (D_{worst}^{coop}). Gleichzeitig ist dies für die Pfadlänge der Anfragen von Teilnehmern mit einer DPC-Strategie (D_{opt}^{dpc}) nicht der Fall. Es folgt, dass die Differenz ($D_{opt}^{dpc} - D_{worst}^{coop}$) der beiden Pfadlängen groß wird, und $P_F^{D_{opt}^{dpc} - D_{worst}^{coop}}$ damit klein wird. Es folgt, dass der Nenner in Ungleichung 4.8.1 um Größenordnungen größer ist als in der gleichen Situation ohne die Verwendung der C4C-Strategie. (Dort war der Nenner $1 - P_F$.) Insbesondere in großen Netzwerken ist vollständige Kooperation lohnender als Deliberately Poor Connectedness, wenn kooperative Teilnehmer die C4C-Strategie nutzen und viele zusätzliche Kontakte besitzen. In anderen Worten: Kooperative Teilnehmer können mit der C4C-Strategie die Vorteile von Deliberately Poor Connectedness verhindern, wenn sie selbst zusätzliche Kontakte verwenden. Des Weiteren sei angemerkt, dass wenn kooperative Teilnehmer die C4C-Strategie nutzen, eine korrekte Klassifizierung unkooperativer Teilnehmer, d.h. P_F nahe 100%, nicht mehr notwendigerweise dazu führt, dass Deliberately Poor Connectedness lohnender ist als vollständige Kooperation.

Beispiel: Es seien alle Parameter so gewählt wie im vorherigen Beispiel zu Deliberately Poor Connectedness (vergleiche Abschnitt 4.6.4) mit dem Unterschied, dass kooperative Teilnehmer die C4C-Strategie nutzen. Kooperation führt zu höheren Auszahlungen als Deliberately Poor Connectedness, wenn die Auszahlung durch den Erhalt einer Antwort (a) größer als 39 ist. Ohne die C4C-Strategie hätte a größer als 600 sein müssen. ■

Es sei angemerkt, dass die Auszahlung durch den Erhalt einer Antwort (a) größer sein muss als die Weiterleitungskosten der zugehörigen Anfrage plus die Kosten für die Beantwortung der Anfrage, damit die Teilnehmer des Systems überhaupt einen Anreiz haben, im System zu verbleiben. Andernfalls ist die durchschnittliche Auszahlung negativ und ihre Auszahlung maximierende Teilnehmer würden das System verlassen. In einem CAN mit 1.000.000 Teilnehmern bedeutet dies, dass die durchschnittliche Pfadlänge 500 Schritte beträgt. D.h. Deliberately Poor Connectedness ist in einem solchen System nicht lohnender als vollständige Kooperation.

Für die Berechnung von Lemma 4.8.1 wurde der schlechteste Fall für die Anfragen kooperativer Teilnehmer angenommen: Lemma 4.8.1 gilt, auch wenn eine Anfrage eines kooperativen Teilnehmers nur ein einziges Mal (nämlich vom Erzeuger) mit zusätzlichen Kontakten weitergeleitet wird. Falls zusätzliche Kontakte auch in weiteren Weiterleitungsschritten verwendet werden, so sind die Ergebnisse für kooperative Teilnehmer sogar noch besser.

4.9 Validierung

In der formalen Analyse der C4C-Strategie wurden einige Annahmen gemacht:

KAPITEL 4. ANREIZE ZUR VERNETZUNG

1. Alle Zonen besitzen die gleiche Größe.
2. Teilnehmer mit einer DPC-Strategie haben keinen einzigen zusätzlichen Kontakt.
3. Alle Teilnehmer haben Reputationsinformationen über andere Teilnehmer, d.h. können auf ein Reputationssystem zurückgreifen.

Um einen Einblick in die Nützlichkeit der C4C-Strategie zu erhalten, der über die Erkenntnisse aus der formalen Analyse hinausgeht, werden diese Annahmen in diesem Kapitel fallen gelassen.

Das Fallenlassen der Annahmen (1) bis (3) hat die folgenden Vorteile: Das Fallenlassen von Annahme (1), d.h. dass die Teilnehmer gleich große Zonen besitzen, führt zu realistischeren Systemen. Das Fehlen von Annahme (2), d.h. dass Teilnehmer mit einer DPC-Strategie (wenige) zusätzliche Kontakte besitzen, reduziert möglicherweise die Effektivität der C4C-Strategie. Selbst wenn kooperative Teilnehmer die C4C-Strategie nutzen, können Teilnehmer mit einer DPC-Strategie (mit wenigen zusätzlichen Kontakten) einige zusätzliche Kontakte von kooperativen Teilnehmern benutzen. Dennoch ist es ein Entwurfsziel der C4C-Strategie, dass kooperative Teilnehmer eine höhere Auszahlung erhalten als Teilnehmer mit einer DPC-Strategie. Das Fallenlassen von Annahme (3) macht die Simulationen realistischer, da viele Systeme kein Reputationssystem besitzen. Anstelle eines Reputationssystems [RKZF00] nutzen die simulierten Teilnehmer Schwellenwert-Strategien. (Es sei an dieser Stelle erneut darauf hingewiesen, dass nach der üblichen Definition eines Reputationssystems [RKZF00] Schwellenwert-Strategien kein Reputationssystem darstellen, da eine solche Strategie allein nicht zu einer Verteilung der Reputationsinformationen führt. Dies ist jedoch eine zentrale Aufgabe eines Reputationssystems. Dennoch ist die lokale Organisation von Reputationsinformationen durch eine Schwellenwert-Strategie, wie z.B. in [Sch08] gezeigt, meist ausreichend.)

4.9.1 Aufbau der Simulationen

Die Zuordnung der Zonen zu den Teilnehmern erfolgt wie in der originalen Content-Addressable Network Veröffentlichung [RFH⁺01]: Ein Teilnehmer, der ins Content-Addressable Network eintreten will, muss mindestens einen Teilnehmer im System kennen. Dieser Teilnehmer gilt als der Startpunkt des neuen Teilnehmers. Der neue Teilnehmer erzeugt eine zufällige Anfrage, d.h. er erzeugt eine zufällige Koordinate k im Adressraum. Analog zur Anfragebearbeitung findet der neue Teilnehmer das Ziel der Anfrage, also den Teilnehmer, der für k verantwortlich ist. Sei dieser Teilnehmer mit t bezeichnet. Die Zone von Teilnehmer t wird geteilt und der neue Teilnehmer und Teilnehmer t verwalten jeweils eine der zwei neuen Zonen. (Zonen werden abwechselnd in jeder Dimension geteilt.) Es sei angemerkt, dass die Zuordnung der Zonen im Content-Addressable Network nicht zu Zonen mit identischer Größe führt.

Jeder Teilnehmer mit einer DPC-Strategie hat zwischen null und fünf zusätzliche Kontakte, d.h. es gilt: $|K^+| \in [0, 5]$. Die Anzahl an Kontakten ist zufällig und folgt einer Gleichverteilung. Vollständig kooperative Teilnehmer haben zehn zusätzliche Kontakte ($|K^+| = 10$).

Unkooperative Teilnehmer (auf Anfrageebene) lassen zwischen 50% und 100% aller Anfragen fallen, die sie erhalten, d.h. es gilt: $u_a, u_f \in [0.5, 1]$. Der Anteil an Anfragen, die fallengelassen werden, ist zufällig und folgt einer Gleichverteilung.

Um unkooperative Teilnehmer erkennen zu können, benutzen die simulierten Teilnehmer Schwellenwert-Strategien. Dies ist durch [SBSV06] motiviert. In [SBSV06] wird gezeigt, dass menschliche Teilnehmer in Content-Addressable Networks solche Strategien nutzen. In vorbereitenden Experimenten wurde ein geeigneter Schwellenwert ermittelt: Es hat sich herausgestellt, dass ein Schwellenwert von 0,1% geeignet ist, um unkooperative Teilnehmer zu klassifizieren. Es sei angemerkt, dass ein gut gewählter Schwellenwert essentiell ist: Ist er zu niedrig, dann erkennen kooperative Teilnehmer unkooperative Teilnehmer auf Anfrageebene nicht und Freifahren ist lohnender als Kooperation. Wenn hingegen der Schwellenwert zu hoch ist, dann werden kooperative Teilnehmer als unkooperativ klassifiziert.

In den Simulationen werden die folgenden exogenen Parameter variiert:

- Der Anteil an Teilnehmern mit einer Deliberately Poor Connectedness Strategie, die sich kooperativ auf Anfrageebene verhalten.
- Der Anteil an Teilnehmern, die sich unkooperativ auf Anfrageebene verhalten.
- Die Größe des Netzwerkes.

Diese exogenen Parameter haben (möglicherweise) einen Einfluss auf die Effektivität der C4C-Strategie. Um zu verstehen, welche Strategien zu höheren Auszahlungen führen als andere, werden die Auszahlungen und die Einnahmen von vollständig kooperativen Teilnehmern und von Teilnehmern mit einer DPC-Strategie (plus Kooperation auf Anfrageebene) gemessen.

Um die Ergebnisse der formalen Analysen aus den Abschnitten 4.6.3 und 4.8 zu validieren, werden neben den Auszahlungen auch die Vorhersagen für die Auszahlungen aus der formalen Analyse (z_c^+, z_d^+) präsentiert. Für die Vorhersagen werden eine false-positive Rate von 0,025 und eine false-negative Rate von 0,035 angenommen. Es sei angemerkt, dass für die formalen Analysen angenommen wird, dass Teilnehmer mit einer DPC-Strategie keine zusätzlichen Kontakte besitzen. Daher kann hier erwartet werden, dass die vorhergesagte Auszahlung kleiner ist als die gemessene.

Des Weiteren werden die folgenden Parameter in den Simulationen verwendet: Es wird ein zweidimensionales Content-Addressable Network simuliert. Die Auszahlung für den Erhalt einer Antwort a ist 100 Punkte. Weiterleitungskosten f und Unterhaltskosten für zusätzliche Kontakte (pro Runde) c sind beide 1 Punkt. Das Erzeugen einer Anfrage kostet 2 Punkte und das Beantworten einer Anfrage 5 Punkte.

Vor Beginn der Messungen werden initial 500 Runden simuliert. Dies ist wichtig, um störende Starteffekte zu verhindern und um ein eingeschwungenes System zu erhalten. Die simulierten Teilnehmer formen z.B. in diesen initialen Runden das zusätzliche Netzwerk.

4.9.2 Resultate

Der Einfluss von Deliberately Poor Connectedness auf die C4C-Strategie

Wie in Abschnitt 4.8 gezeigt wurde, ist die C4C-Strategie effektiv, wenn *alle* Teilnehmer sie benutzen. Mit der ersten Reihe von Experimenten wird getestet, wie der Anteil an Teilnehmern, der die C4C-Strategie nutzt, die Effektivität der Strategie beeinflusst. Es wird daher der Anteil an vollständig kooperativen Teilnehmern und Teilnehmern mit einer DPC-Strategie variiert. Es sei angemerkt, dass Teilnehmer mit einer DPC-Strategie die C4C-Strategie nicht benutzen. Die Simulation findet in einem System aus 5.000 Teilnehmern mit einem festen Anteil an Freifahrern auf Anfrageebene statt. Dieser Anteil beträgt 5%. Abbildung 4.1 zeigt die Resultate. Die X-Achse zeigt den Anteil an Teilnehmern mit einer DPC-Strategie. Je größer dieser wird, desto weniger Teilnehmer mit einer kooperativen Strategie auf Kontaktebene existieren. Beispielsweise 0% Teilnehmer mit einer DPC-Strategie bedeutet, dass nur noch vollständig kooperative Teilnehmer (95%) und Freifahrer auf Anfrageebene (5%) vorhanden sind. Konsequenterweise werden nur die Messergebnisse vorhandener Strategien gezeigt. Die Y-Achse zeigt die Auszahlung pro Teilnehmer und Runde. Das wichtigste Ergebnis der Simulationen ist, dass vollständig kooperative Teilnehmer immer eine höhere Auszahlung erhalten als Teilnehmer mit einer DPC-Strategie. Teilnehmer mit einer DPC-Strategie haben zwar geringere Unterhaltskosten für zusätzliche Kontakte haben aber auch (wie aus Abbildung 4.1 deutlich wird) weniger Einkommen als kooperative Teilnehmer, wenn diese die C4C-Strategie nutzen.

Der Einfluss der Netzwerkgröße auf die C4C-Strategie Wie bereits erwähnt, hängen die Einnahmen eines Teilnehmers (a) von der durchschnittlichen Pfadlänge seiner Anfragen ab und (b) von dem Anteil an unkooperativen Teilnehmern. Daher wird in einer zweiten Reihe von Simulationen der Einfluss der Netzwerkgröße (und in einer dritten Reihe der Einfluss unkooperativer Teilnehmer) untersucht. Es wurden unterschiedliche Netzwerkgrößen zwischen 100 und 10.000 Teilnehmern simuliert. Die Verteilung der Strategien im System ist dabei fest: 5% Freifahrer auf Anfrageebene, 47,5% vollständig kooperative Teilnehmer und 47,5% Teilnehmer mit einer DPC-Strategie, die kooperativ auf Anfrageebene sind. Abbildung 4.2 zeigt die Auszahlung, die Einnahmen und die durch die formale Analyse vorhergesagten Auszahlungen von vollständig kooperativen Teilnehmern und von Teilnehmern mit einer DPC-Strategie.

Große Netzwerke haben gegenüber kleinen zwei Nachteile: Erstens, die durch-

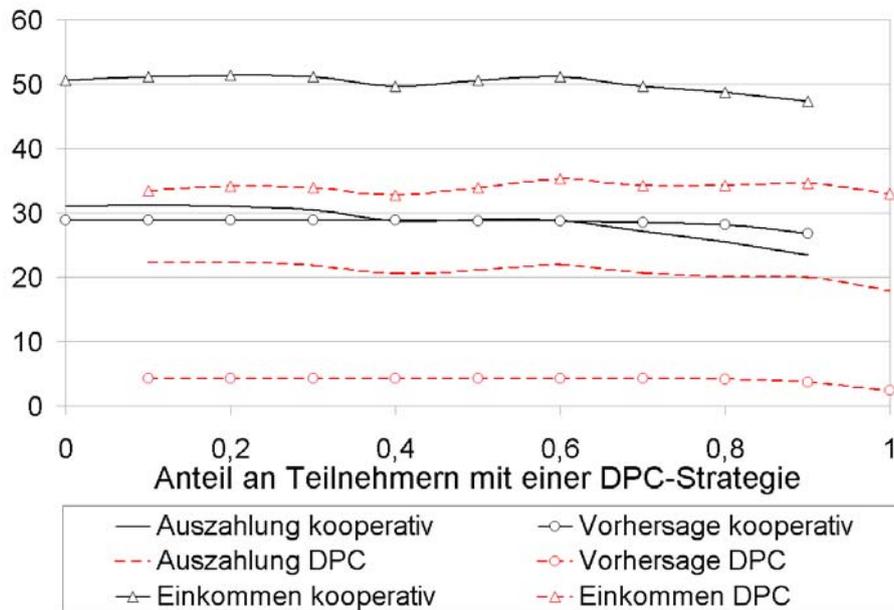


Abbildung 4.1: Einfluss der Teilnehmer mit einer DPC-Strategie auf die C4C-Strategie

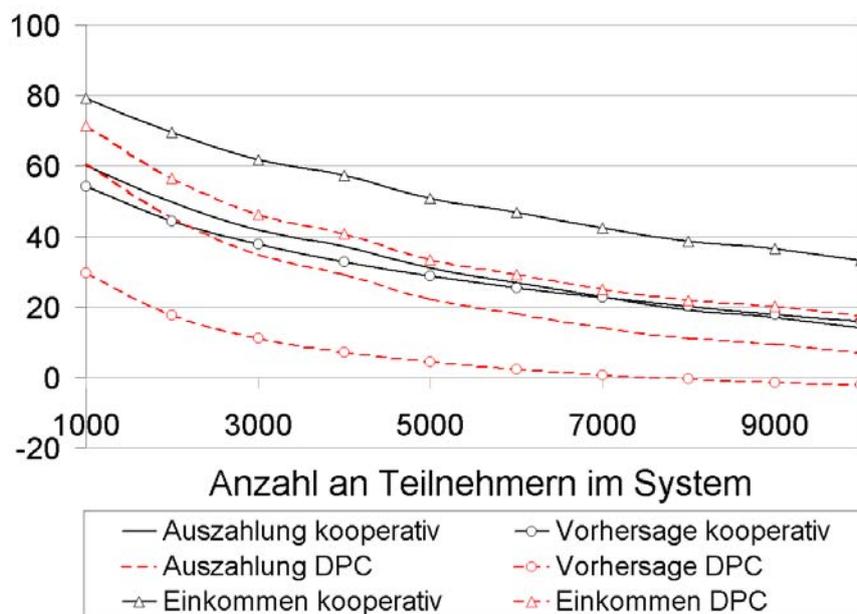


Abbildung 4.2: Einfluss der Netzwerkgröße auf die C4C-Strategie

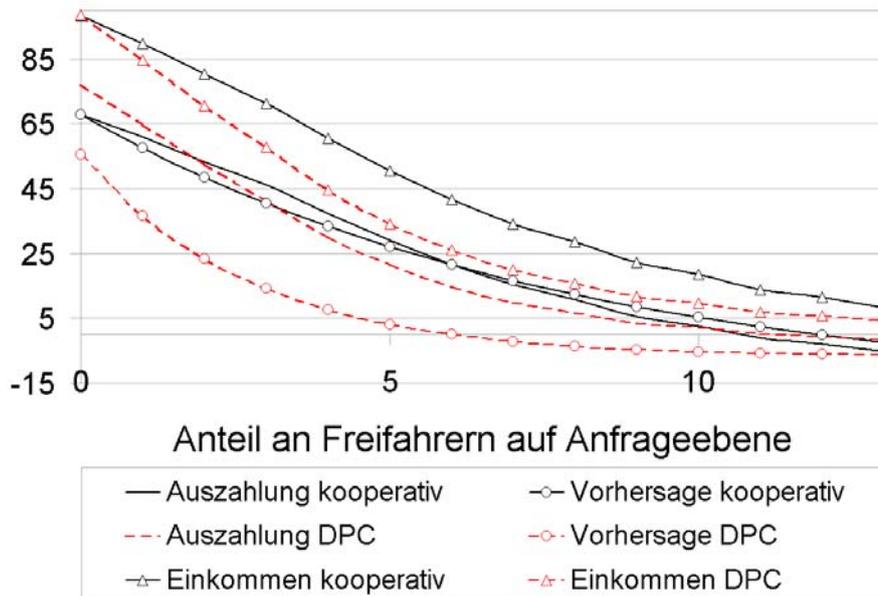


Abbildung 4.3: Einfluss an Freifahren auf die C4C-Strategie

schnittliche Pfadlänge erhöht sich, d.h. die Weiterleitungslast erhöht sich und in der Konsequenz steigen die Kosten für kooperative Teilnehmer. Zweitens, die Einnahmen verringern sich, da mehr Anfragen verlorengehen als in kleinen Netzwerken. Daher haben Teilnehmer großer Netze eine geringere Auszahlung als in kleinen Netzwerken.

Abbildung 4.2 zeigt, dass Teilnehmer mit einer DPC-Strategie weniger Auszahlung erhalten als kooperative Teilnehmer. Das heißt, dass sie weniger Nutzen vom System haben als kooperative Teilnehmer: Vollständige Kooperation ist lohnender als Deliberately Poor Connectedness.

Der Einfluss von Freifahrern auf die C4C-Strategie In einer dritten Reihe von Simulationen wurde der Anteil an Freifahrern auf Anfrageebene in einem System von 5.000 Teilnehmern variiert. Die weiteren Teilnehmer im System sind zu 50% vollständig kooperativ und zu 50% Teilnehmer mit einer DPC-Strategie (plus Kooperation auf Anfrageebene). Abbildung 4.3 zeigt die Ergebnisse: Solange weniger als 1% unkooperative Teilnehmer im System vorhanden sind, ist Deliberately Poor Connectedness lohnender als Kooperation. Beträgt der Anteil an Freifahrern mehr als 1%, so ist ein kooperatives Verhalten lohnender als Deliberately Poor Connectedness. Steigt der Anteil an Freifahrern über 11%, so haben kooperative Teilnehmer und Teilnehmer mit einer DPC-Strategie keinen Nutzen mehr vom System. In diesem Fall würden Teilnehmer, die ihre Auszahlung maximieren, das System verlassen.

Validierung der vorhergesagten Auszahlungen Die gemessenen Auszahlungen und die durch die formale Analyse vorhergesagten Auszahlungen gleichen sich nicht exakt, sind aber in der Regel vergleichbar. Eine Ausnahme entsteht, sobald der Anteil unkooperativer Teilnehmer sehr hoch ist. Die Diskrepanz kann wie folgt erklärt werden: In der formalen Analyse werden feste und für alle Teilnehmer gleiche false-positive und false-negative Werte angenommen. Im Gegensatz dazu bestimmt in den Simulationen jeder Teilnehmer den Kooperationsgrad anderer Teilnehmer für sich. Bei einem hohen Anteil an Freifahrern auf Anfrageebene empfangen die Teilnehmer kaum noch Antworten auf ihre Anfragen. In dieser Situation ist es schwer – wenn nicht unmöglich – sinnvolle statistische Daten zu sammeln. Die beobachteten Unterschiede sind die Folge.

Insgesamt stimmen die Simulationsergebnisse mit den Vorhersagen aus der formalen Analyse überein. Die Effektivität der C4C-Strategie ist robust gegen unterschiedliche Verteilungen an kooperativen Teilnehmern, Freifahrern auf Anfrageebene und Teilnehmern mit einer DPC-Strategie. Ebenfalls funktioniert sie in unterschiedlich großen Systemen. In den Experimenten wird deutlich, dass Deliberately Poor Connectedness nur in wenigen Situationen zu höheren Auszahlungen führt als Kooperation. Im Einzelnen ist dies dann der Fall, wenn der Anteil an Freifahrern auf Anfrageebene kleiner als 1% ist. Wenn kooperative Teilnehmer auch in dieser Situation eine höhere Auszahlung als Teilnehmer mit einer DPC-Strategie erhalten wollen, so müssen sie – nach Lemma 4.8.1 – eine größere Anzahl an zusätzlichen Kontakten wählen.

4.10 Diskussion

In den vorhergehenden Abschnitten wurde gezeigt, dass die C4C-Strategie ein effektives Mittel gegen Deliberately Poor Connectedness ist. Im Folgenden wird diskutiert, warum es wahrscheinlich ist, dass Teilnehmer eines verteilten Systems mit Interaktions- und Kontaktwahl diese Strategie nutzen. Am Ende dieser Diskussion wird erörtert werden, warum die Ergebnisse auch auf andere Systeme, wie strukturierte Peer-to-Peer Systeme oder Soziale Netzwerke, anwendbar sind.

Vergleichbare Kooperationsprobleme Wie-Du-mir-so-ich-Dir (engl. tit-for-tat), d.h. sich so zu verhalten, wie es der Kontakt oder der Spielpartner getan hat, ist eine erfolgreiche Strategie im wiederholten Gefangenendilemma [Axe06]. Die C4C-Strategie nutzt dieses Prinzip: Nur diejenigen Teilnehmer, die selbst zusätzliche Kontakte besitzen, können die zusätzlichen Kontakte anderer Teilnehmer benutzen. Obwohl Wie-Du-mir-so-ich-Dir-Strategien Kosten durch die vorzuhaltende Statistik pro Kontakt verursachen, werden sie im wiederholten Gefangenendilemma genutzt. Da die C4C-Strategie keine Kosten verursacht, d.h. da sie weder zusätzliche Nachrichten noch Statistik oder komplexe Berechnungen erfordert, kann angenommen werden,

KAPITEL 4. ANREIZE ZUR VERNETZUNG

dass Teilnehmer sie nutzen (wenn sie vom Entwickler eines Systems vorgeschlagen wird).

Weißwaschen Weißwaschen, also ein Verhalten, bei dem ein Teilnehmer mit schlechter Reputation das Netzwerk verlässt, um es mit neuer Identität, und daher mit einer neutralen Reputation betritt, ist ein ernstes Problem in Peer-to-Peer Systemen [FPCS06]. Die C4C-Strategie ist robust gegenüber diesem Verhalten, da sie ohne Historie – also zustandslos – ist (d.h. dass die C4C-Strategie keine Statistik benötigt).

Die Grenzen der C4C-Strategie In einigen Situationen ist die C4C-Strategie nicht effizient. Dies ist der Fall, wenn keine Anfragen im System verloren gehen ($P_F = 1$). Dieser Zustand tritt ein, wenn die Einschätzung des Kooperationsgrads anderer Teilnehmer perfekt durchgeführt wird und daher keine unkooperativen Teilnehmer auf Anfrageebene vorhanden sind. Des Weiteren dürfen keine Anfragen aufgrund von technischen Fehlern verloren gehen. In diesem Fall wird die Auszahlung eines Teilnehmers nicht dadurch beeinflusst, über wie viele Weiterleitungsschritte eine Anfrage gesendet wird. Daher hat ein Teilnehmer auch keinen Vorteil von zusätzlichen Kontakten. Im Gegenteil kann er die Kosten für den Unterhalt zusätzlicher Kontakte sparen. In diesem Fall steigt allerdings die Weiterleitungslast an. Dies führt zu ineffizienten Systemen, falls beim Weiterleiten von Anfragen Kosten entstehen. Da es unwahrscheinlich ist, dass die oben genannten Bedingungen gleichzeitig eintreten, ist es unwahrscheinlich, dass dieser Fall eintritt.

Strukturierte Peer-to-Peer Systeme und Soziale Netzwerke Deliberately Poor Connectedness ist nicht auf Content-Addressable Networks beschränkt. Ohne Gegenmaßnahmen ist es auch in anderen Systemen vorteilhaft, nur wenige Kontakte zu unterhalten. Ein Teilnehmer kann über diese wenigen Kontakte das gesamte Netzwerk erreichen.

Beispiel: Chord [SMK⁺01] ist ein strukturiertes Peer-to-Peer System, das als initiale Struktur einen Ring aus Teilnehmern aufweist. Benachbarte Teilnehmer auf dem Ring sind feste Kontakte. Darüber hinaus werden durch das System weitere (entfernte) Kontakte festgelegt, deren Abstand zum jeweiligen Peer exponentiell wächst. Ein Teilnehmer mit einer DPC-Strategie würde diese Kontakte fallenlassen und würde seine Anfragen an andere Teilnehmer über die beiden Nachbarn im Ring leiten. Obwohl die entfernten Kontakte vom System vorgegeben sind, kann ein Teilnehmer sie dadurch ignorieren, dass er ihre Anfragen fallen lässt oder technische Probleme vorgibt, die es ihm unmöglich machen, mit diesen zu kommunizieren. ■

Verallgemeinerung Die Ergebnisse dieses Kapitels sind nicht nur für die Entwickler von Peer-to-Peer Systemen interessant, sondern für ein breiteres Publikum. Netzwerke, in denen ein Distanzmaß anwendbar ist und die aus autonomen Entitäten

bestehen, die Kontakte wählen, können von den Ergebnissen profitieren. Ihre Teilnehmer können die C4C-Strategie nutzen, um anderen Teilnehmern einen Anreiz zu geben, Verbindungen aufzubauen und damit zur Netzwerkstruktur beizutragen.

4.11 Fazit

Eine große Klasse an komplexen Netzwerken, d.h. Netzwerken in denen Teilnehmer Anfragen typischerweise über Zwischenknoten an ihr Ziel weiterleiten, ist durch eine hinreichende Vernetzung zwischen den Teilnehmern effizient [JW96, BG00a]. Da der Unterhalt zusätzlicher Kontakte Kosten verursacht, ist Deliberately Poor Connectedness – also die Strategie kaum vernetzt zu sein – vorteilhaft. In diesem Kapitel wurde gezeigt, dass Deliberately Poor Connectedness in einer großen Menge von Parameterkonfigurationen lohnender ist als Kooperation. Um dieses Problem in den Griff zu bekommen, wurde eine Anreiz-basierte Strategie gegen Deliberately Poor Connectedness vorgeschlagen: die C4C-Strategie. Diese Strategie bedeutet, dass ein Teilnehmer einen Kontakt nur in dem Maß an seinen zusätzlichen Kontakten teilhaben lässt wie der Kontakt seine zusätzlichen Kontakte einbringt. Teilnehmer, die sich reziprok auf Kontaktebene verhalten, geben anderen Teilnehmern einen Anreiz, sich zu vernetzen und zusätzliche Kontakte zu unterhalten. Teilnehmer von DPC-Strategien müssen in der Konsequenz selbst zusätzliche Kontakte unterhalten, um von der (zusätzlichen) Netzwerkstruktur einen Nutzen erhalten zu können. Eine formale Analyse als auch Experimente bestätigen die Effektivität der C4C-Strategie. Die C4C-Strategie ist kostenneutral, skalierbar und unter wechselnden Verteilungen unterschiedlicher Strategien effektiv.

Die C4C-Strategie kann genutzt werden, um die Anzahl an zusätzlichen Kontakten pro Teilnehmer zu beeinflussen. Dennoch ist unklar, wie die Teilnehmer eines Systems zur verteilten Anfrageverarbeitung ihre zusätzlichen Kontakte verteilen. Ebenfalls ist unklar, ob die von ihnen angestrebte Verteilung zu effizienten Netzwerkstrukturen führt. Die nötige Untersuchung dieser Fragestellungen ist Bestandteil des nächsten Kapitels.

KAPITEL 4. ANREIZE ZUR VERNETZUNG

5 Gleichgewichte in Systemen mit Interaktions- und Kontaktwahl

Entwickler von Software, die auf Basis eines Peer-to-Peer Systems oder eines Sozialen Netzwerks geplant wird, müssen bei dem Entwurf ihrer Software beachten, dass die Teilnehmer des Systems sich autonom verhalten. Dies bedeutet, dass im Gegensatz zum klassischen Softwareentwurf, bei dem der Entwickler die Ausführung des Programms durch seine Software kontrolliert, keine Kontrolle über die Teilnehmer des Systems vorhanden ist. Daher ist eine strategische Analyse der Interessen der Teilnehmer nötig. Unter der Annahme, dass die Teilnehmer ihren Nutzen maximieren, kann überprüft werden, ob die Teilnehmer ein Interesse haben, von einer gegebenen Strategie abzuweichen. Ist dem nicht so, kann davon ausgegangen werden, dass ein entwickeltes System stabil ist. Ein solches System wird als Gleichgewicht bezeichnet. Sollte allerdings ein Anreiz bestehen, von dem vorgegebenen System abzuweichen, so ist das System instabil und es muss damit gerechnet werden, dass die Teilnehmer des Systems eine andere Strategie wählen.

Im Gegensatz zum vorherigen Kapitel, in dem überprüft wurde, ob die Teilnehmer verteilter Systeme überhaupt einen Anreiz haben, Kontakte zu wählen, soll dieses Kapitel eine grundsätzliche Frage klären: Haben Teilnehmer eines verteilten Systems einen Anreiz, von einer vorgegebenen effizienten Strategie abzuweichen?

Bisher war unklar, welche Gleichgewichte in Systemen mit Interaktions- und Kontaktwahl existieren. In diesem Kapitel wird untersucht, ob (und wann) ein kooperatives Netzwerk aus Teilnehmern, die eine Netzwerkstruktur besitzen, die es erlaubt, Anfragen in wenigen Schritten weiterzuleiten, ein Gleichgewicht ist. Es wird gezeigt, dass die folgenden Gleichgewichte existieren: Erstens, ein ineffizientes Gleichgewicht, bei welchem die Teilnehmer eine Strategie wählen, bei der sie weder Anfragen bearbeiten noch absenden und darüber hinaus keine Kontakte unterhalten. Zweitens, ein Gleichgewicht, bei dem die Teilnehmer ihre Kontakte gleichmäßig über den Adressraum verteilen. Dieses zweite Gleichgewicht ist ebenfalls nicht effizient. Eine Strategie, die in diesem Kapitel vorgeschlagen wird, führt zu einem Gleichgewicht, das optimal ist: In ihm erzielen die Teilnehmer die höchsten Auszahlungen über alle Gleichgewichte. Wenn die Teilnehmer ihre Entscheidung, einen Kontakt zu wählen, davon abhängig machen, wie schnell die Anfragen, die sie ihm geschickt haben, bearbeitet werden, dann können optimale Systemzustände Gleichgewichte sein.

5.1 Einleitung

Entwickler von verteilten Systemen oder von Anwendungen, die auf verteilten Systemen aufbauen, haben ein Interesse an effizienten Systemen. Unklar ist jedoch, ob ein einzelner Teilnehmer ebenfalls ein Interesse hat, nicht von einer vorgegebenen effizienten Strategie – also von einem globalen Standpunkt aus gesehenen optimalen Zustand – abzuweichen. Ist die effiziente Strategie für den einzelnen Teilnehmer nicht optimal, kann dieser also durch Wählen einer anderen Strategie eine höhere Auszahlung erzielen, so liegt es nahe, dass die vom Systementwickler vorgeschlagene Strategie nicht stabil ist.

Forschungsfragen Dieses Kapitel untersucht die folgenden Forschungsfragen:

- Ist der effiziente Systemzustand ein Gleichgewicht? Oder in anderen Worten: Hat ein Teilnehmer einen Anreiz, einseitig von seiner Strategie abzuweichen, wenn alle Teilnehmer eine Strategie verwenden, die zu einem effizienten System führt? Ist dies der Fall, so ist davon auszugehen, dass ein System, das von einem System-Entwickler vorgeschlagen wird und welches effizient ist, nicht stabil ist.
- Die schnelle Antwort auf eine Anfrage ist in manchen Anwendungen wertvoller als eine späte Antwort. Andersherum ist dieser Sachverhalt für manche Systeme nicht der Fall. Dieser Sachverhalt wird in diesem Kapitel durch die Verwendung eines exogenen Diskontierungsfaktors modelliert. Wie beeinflusst dieser Parameter die Strategien der Teilnehmer?
- Das Verhalten zwischen einzelnen Knoten wird davon beeinflusst, wie gut diese den Kooperationsgrad anderer Teilnehmer bestimmen können. Wie bestimmt die Qualität der Einschätzung das Verhalten der Teilnehmer?

Es sei angemerkt, dass die Untersuchung von Spielen mit mehr als zwei Teilnehmern im Allgemeinen schwierig ist [GGS03] und zusätzlich dadurch erschwert wird, dass die Teilnehmer ihre Kontakte wählen können [VR06]. Beides ist in dieser Arbeit der Fall.

Beiträge In einem ersten Schritt wird ein Strategieraum definiert, der sehr abstrakt ist. Dieser ist in der Lage, Strategien, die aus der Literatur bekannt sind, sowie weitere Strategien zu modellieren. Während in [Kle06] Netzwerkstrukturen identifiziert wurden, die optimal für Geiziges Weiterleiten sind, ist unklar, ob diese Netzwerkstrukturen Gleichgewichte sind. In diesem Kapitel wird eine formale Analyse durchgeführt, die zeigt, dass zwei Gleichgewichte existieren.

1. Das *ineffiziente Netzwerk*: Wenn die Teilnehmer nicht hinreichend kooperativ sind, sinkt der Nutzen des Systems so stark, dass die Teilnehmer Anfragen weder senden noch bearbeiten und dass sie keine Kontakte wählen.
2. Das *kooperative Gleichgewicht*: Unkooperative Teilnehmer – also Freifahrer auf Anfrageebene – werden erkannt und erhalten geringe Auszahlungen. Dadurch wird diese Strategie unattraktiv und Strategien mit hohem Kooperationsgrad sind die Folge.

Wenn die Teilnehmer bei der Kontaktwahl nicht die Geschwindigkeit, mit der ihre Anfragen von einem Teilnehmer verarbeitet werden, berücksichtigen, dann ist das gleichmäßige Verteilen der Kontakte über den Adressraum eine optimale Strategie (aus Sicht des einzelnen Teilnehmers). Die daraus resultierende Netzwerkstruktur ist in großen Netzen nicht effizient [Kle00, BG00a]. Effizientere Systeme sind möglich, wenn die Teilnehmer nicht nur zwischen kooperativen und unkooperativen Teilnehmern unterscheiden, sondern auch die Verarbeitungszeiten ihrer Anfragen berücksichtigen, wenn sie Kontakte wählen. Dazu wird in diesem Kapitel eine Strategie – genannt *Drop-Slow-Contacts (DSC)* (deutsch: lass-langsame-Kontakte-fallen) – vorgeschlagen. Diese Strategie ist wie folgt: Teilnehmer lassen die zusätzlichen Kontakte fallen, bei denen die Bearbeitung ihrer Anfragen zu viel Zeit in Anspruch nimmt. Der Hauptbeitrag dieses Kapitels ist es, dass gezeigt wird, dass die DSC-Strategie in ein Gleichgewicht mit hohem Gemeinwohl und hohen Auszahlungen führt.

Wenn die Teilnehmer die DSC-Strategie nutzen, dann geben sie ihren Kontakten einen Anreiz, ihre Kontakte in optimaler Weise zu verteilen, da sonst keine schnelle Abarbeitung ihrer Anfragen möglich ist. Die Drop-Slow-Contacts Strategie ist kostenneutral, d.h. sie benötigt keine zusätzlichen Anfragen oder komplexen Berechnungen und führt zu hohen Auszahlungen.

Während das vorherige Kapitel mit der C4C-Strategie eine Strategie vorschlägt, die die Anzahl von zusätzlichen Kontakten pro Teilnehmer beeinflusst, beeinflusst die DSC-Strategie die Verteilung der zusätzlichen Kontakte. Beide Strategien haben also eine unterschiedliche Zielrichtung und können orthogonal zueinander angewendet werden. Um die Effekte der Strategien voneinander trennen zu können, werden beide Strategien untersucht, ohne dass die Verwendung der anderen Strategie angenommen wird.

Gliederung des Kapitels In Abschnitt 5.1 werden Definitionen und Annahmen (Abschnitt 5.2) erörtert. Es folgt die Analyse der Gleichgewichte in Abschnitt 5.3 und ein Fazit 5.4.

5.2 Definitionen und Annahmen

Im Folgenden werden die in diesem Kapitel zugrunde gelegte feste Netzwerkstruktur sowie der Strategieraum erläutert.

5.2.1 Netzwerkeigenschaften

Fundamentale Eigenschaften Die fundamentalen Eigenschaften des Netzes seien wie in Abschnitt 3.1 definiert. Um die folgende formale Analyse der Gleichgewichte zu vereinfachen, wird angenommen, dass die feste Netzwerkstruktur ein Ring ist: Jeder Teilnehmer i hat zwei feste Kontakte, den Teilnehmer links und den Teilnehmer rechts im Ring, d.h. Teilnehmer $(i - 1) \bmod n$ und Teilnehmer $(i + 1) \bmod n$.

Kleinberg-Verteilungen Kleinberg untersucht in [Kle00] die Weiterleitungskomplexität von Geizigem Weiterleiten in Zufallsnetzen mit einer initialen Struktur. Nach [Kle00] ist die Anzahl an Weiterleitungen in einem Ring mit einer zufälligen zusätzlichen Netzwerkstruktur minimal, wenn alle Paare von Teilnehmern (u, v) mit einer Wahrscheinlichkeit proportional zu $d(u, v)^{-r_k}$ Kontakte sind. Dabei sei $d(u, v)$ definiert als die Anzahl an Schritten im minimalen Pfad zwischen u und v in der initialen Ringstruktur. Sei r_k die Dimension des Netzwerks, es gilt also $r_k = 1$ im Ring. In dieser Arbeit werden solche Netzwerke als *Netzwerke mit einer Kleinberg-Verteilung* bezeichnet. Kleinberg zeigt in [Kle00], dass Teilnehmer Anfragen mit hoher Wahrscheinlichkeit in $\beta \cdot \log^2(n)$ Schritten weiterleiten können, dass also eine entsprechende Konstante β existiert.

[Kle00, Kle06] nimmt an, dass die Teilnehmer seiner Netze kooperieren und untersucht damit die Interaktionswahl-Strategien der Teilnehmer nicht. Auch wenn die Ergebnisse in [Kle00, Kle06] von einer algorithmischen Perspektive interessant sind, so ist doch unklar, ob die beschriebene Kontaktverteilung im Interesse der Teilnehmer ist und ein Gleichgewicht darstellt.

Einfaches vs. Wiederholtes Spiel In diesem Abschnitt werden die Auszahlungen von Teilnehmern berechnet, die ein one-shot Spiel mit jedem anderen Teilnehmer spielen, d.h. die eine Anfrage für jeden anderen Teilnehmer erzeugen (vergleiche Abschnitt 3.4). Weiterleitungsketten führen dabei zu wiederholten Interaktionen zwischen Kontakten.

Auszahlungsfunktion Die Auszahlungsfunktion sei so wie in Abschnitt 3.5 definiert. Einschränkend wird in diesem Abschnitt davon ausgegangen, dass der Nutzen, den ein Teilnehmer erhält, wenn er eine Antwort nach wenigen Zeiteinheiten empfängt, höher ist, als wenn er sie erst nach vielen Zeiteinheiten erhält. D.h. es gilt $\delta \in [0, 1)$ – eine Diskontierung in der Anfragebearbeitung wird also durchgeführt.

Unsicherheit Teilnehmer verteilter Systeme nutzen oft Reputationssysteme, um sich über den Kooperationsgrad anderer Teilnehmer zu informieren. Um von unterschiedlichen Reputationssystemen zu abstrahieren, wird in diesem Abschnitt davon ausgegangen, dass das von den Teilnehmern genutzte Reputationssystem fehlerhaft ist: In *reportingError* Prozent der Fälle liefert es einen fehlerhaften Kooperationsgrad.

Die im vorherigen Kapitel genutzten false-positive und false-negative Raten erlauben prinzipiell eine genauere Analyse der Qualität eines Reputationssystems. Gleichzeitig erschweren sie jedoch die Analyse der Strategien. Es wird daher in diesem Kapitel auf false-positive und false-negative Raten verzichtet.

Teilnehmer benutzen Schwellenwert-Strategien [SBV07], haben also einen Schwellenwert definiert, ab dem ein Teilnehmer nicht mehr als kooperativ angesehen wird. Dieses Verhalten zusammen mit den fehlerhaften Informationen des Reputationssystems führt zu einem Gesamtverhalten, das durch Algorithmus 6 gegeben ist. Da unklar ist, wie sich der Fehler eines Reputationssystems im Allgemeinen verhält, wird hier vereinfacht angenommen, dass die Entscheidung, ob ein Kontakt kooperativ ist, in *reportingError* Prozent der Fälle falsch getroffen wird. Es sei angemerkt, dass es bei dem in Algorithmus 6 angewendeten Verfahren lediglich eine ja/nein-Entscheidung getroffen wird.

Algorithmus 6: *istKooperativ*(Teilnehmer *i*, Schwellenwert *t*)

Input : Teilnehmer *i*, Schwellenwert *t*
Output : {WAHR, FALSCH}

```
1 Gleitkommazahl wahrerKooperationsgrad = C(i);
2 if wahrerKooperationsgrad ≥ t then
3   | if nimmZufallswertAusIntervall[0, 1] ≥ reportingError then
4   | | return WAHR ;
5   | else
6   | | return FALSCH ;
7 else
8   | if nimmZufallswertAusIntervall[0, 1] < reportingError then
9   | | return WAHR ;
10  | else
11  | | return FALSCH ;
```

Entfernungen in der Ringstruktur Zwei nützliche Hilfsfunktionen sind die *Entfernung zwischen Teilnehmern im Ring* und die *normalisierte Distanz*. Die erste ist die Anzahl an Zwischenknoten in der festen Netzwerkstruktur und die zweite die Normalisierung davon.

KAPITEL 5. GLEICHGEWICHTE IN SYSTEMEN MIT INTERAKTIONS- UND KONTAKTWAHL

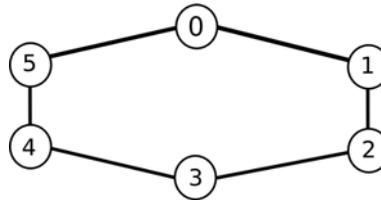


Abbildung 5.1: Ring mit sechs Teilnehmern

Definition 5.1 (Entfernung zwischen Teilnehmern im Ring)

Die Distanz zwischen zwei Teilnehmern i, j ist:

$$d(i, j) = \min(|i - j|, n + \min(i, j) - \max(i, j))$$

Beispiel: Abbildung 5.1 zeigt einen Ring mit sechs Teilnehmern. Der Abstand zwischen Teilnehmer 4 und 0 ist zwei: $d(4, 0) = \min(|4 - 0|, 6 + \min(4, 0) - \max(4, 0)) = \min(|4 - 0|, 6 + 0 - 4) = \min(4, 2) = 2$. Der Abstand zwischen Teilnehmer 1 und 3 ist ebenfalls zwei: $d(1, 3) = \min(|1 - 3|, 6 + 1 - 3) = \min(2, 4) = 2$. ■

Definition 5.2 (Normalisierter Abstand)

Der normalisierte Abstand $\Delta(\cdot, \cdot)$ zwischen Teilnehmer $i \in N$ und $j \in N$ ist:

$$\Delta(i, j) = \frac{4}{n-1} \cdot d(i, j)$$

Faktor $\frac{4}{n-1}$ ist dadurch motiviert, dass der maximale Abstand $\frac{n-1}{2}$ beträgt, d.h. die Hälfte des Rings, und dass Anfragen im Durchschnitt nur den halben Weg zurücklegen müssen.

5.2.2 Strategieraum

Das Verhalten von Teilnehmern kann sehr unterschiedlich ausfallen. Es gibt Aspekte des menschlichen Verhaltens in der Interaktionswahl und Kontaktwahl, die bereits bekannt sind (und sich nicht ändern), und andere, die unbekannt sind. Bekannte Aspekte werden daher hier in Form von Pseudocode wiedergegeben. Unklar ist allerdings, welche genauen Schwellenwerte genutzt werden. Daher werden die Teilnehmer in dem hier verwendeten Modell diese Werte selbstständig wählen, wie im Verlaufe des Abschnitts deutlich werden wird. Abgesehen von Schwellenwert-Strategien sind die meisten Aspekte des Verhaltens von Teilnehmern in der Kombination aus Interaktions- und Kontaktwahl unbekannt. Daher werden diese Aspekte mit Hilfe eines Strategievektors S modelliert. Dieser enthält die Parameter, die für eine Strategie entscheidend sind. Die folgende Beschreibung dieser Parameter beginnt mit der Interaktionswahl und endet mit der Kontaktwahl. Tabelle 5.1 dient als Zusammenfassung.

5.2. DEFINITIONEN UND ANNAHMEN

Parameter	Typ	Beschreibung
<i>sendRequest</i>	binär	falls wahr, dann erzeugt der Teilnehmer Anfragen
<i>contactC</i>	Gleitkommazahl $\in [0, 1]$	minimaler Kooperationsgrad, damit eine Anfrage verarbeitet wird
<i>ownCoop</i>	Gleitkommazahl $\in [0, 1]$	Anteil der verarbeiteten Anfragen von Erzeugern und Weiterleitern, die als kooperativ eingeschätzt worden sind
<i>maxC</i>	natürliche Zahl $\in [0, N - K^f - 1]$	maximale Anzahl zusätzlicher Kontakte
<i>additionalCC</i>	Gleitkommazahl $\in [0, 1]$	minimaler Kooperationsgrad, den Kontakte erfüllen müssen, um nicht fallengelassen zu werden
<i>maxProcessingTime</i>	Gleitkommazahl $\in [0, \infty)$	maximale Verarbeitungszeit für Anfragen
<i>distribution</i>	Verteilungsfunktion	Verteilung der Kontakte im Adressraum

Tabelle 5.1: Parameter des Strategieraums

KAPITEL 5. GLEICHGEWICHTE IN SYSTEMEN MIT INTERAKTIONS- UND KONTAKTWAHL

Abschließend wird beschrieben, wie aus der Literatur bekannte Strategien mit dem hier genutzten Strategieraum modelliert werden können.

Erzeugen von Anfragen Ein Teilnehmer kann Anfragen erzeugen, oder es sein lassen. (Diese Arbeit untersucht keine Strategien, in denen ein Teilnehmer manchmal Anfragen erzeugt und es manchmal sein lässt. Dies ist dadurch begründet, dass angenommen wird, dass es sich entweder lohnt am System zu partizipieren oder eben nicht.) Falls der binäre Parameter *sendRequest* wahr ist, erzeugt und sendet ein Teilnehmer Anfragen.

Bearbeiten von Anfragen Menschen in Koordinator-freien Systemen nutzen Schwellenwert-Strategien beim Bearbeiten von Anfragen [SBV07]. Beim Erzeugen und Senden von Anfragen an einen Kontakt *j* misst ein Teilnehmer *i* den Anteil der zugehörigen Antworten, den er erhält. Wenn der Anteil beantworteter Anfragen geringer ist als ein Schwellenwert, dann lässt Teilnehmer *i* die Anfragen von Kontakt *j* fallen. Das bedeutet, dass menschliche Teilnehmer ihre Kontakte für das Verhalten der gesamten Weiterleitungskette verantwortlich machen, selbst wenn die Kontakte die Anfragen nicht fallengelassen haben. Dies macht Sinn, da die Kontakte wiederum dafür verantwortlich sind, sich kooperative Kontakte zu suchen. Die hier modellierten Teilnehmer verhalten sich in gleicher Art und Weise, d.h. wie in Algorithmus 7 beschrieben. Für jede Anfrage testet der Teilnehmer, ob der Vorgänger in der Weiterleitungskette sich qualifiziert (Zeile 1). „Sich zu qualifizieren“ bedeutet, dass der Vorgänger als kooperativer wahrgenommen wird, als es der Teilnehmer verlangt, d.h. er muss mindestens einen wahrgenommenen Kooperationsgrad von *contactC* (contact-cooperation Schwellenwert) besitzen. „Wahrgenommener Kooperationsgrad“ bedeutet, dass der vom Reputationssystem – möglicherweise falsche – Kooperationsgrad höher sein muss als Schwellenwert *contactC*. Ein Teilnehmer kann in dem hier vorgestellten Modell entscheiden, ob er sich kooperativ verhält oder nicht. Er kann Anfragen fallenlassen, selbst dann, wenn sich der Vorgänger qualifiziert hat. Anfragen von qualifizierten Vorgängern werden in *ownCoop* Prozent der Fälle bearbeitet (Zeile 1).

Algorithmus 7: Bearbeiten von Anfragen

Input : Anfrage *req*, Vorgänger in der Weiterleitungskette *p*

- 1 **if** *istKooperativ*(*p*, *contactC*) **und**
 nimmZufallswertAusIntervall[0, 1] < *ownCoop* **then**
- 2 | bearbeite *req*;
- 3 **else**
- 4 | lasse *req* fallen;

5.2. DEFINITIONEN UND ANNAHMEN

Kontaktwahl Die Kontaktwahl-Strategie eines Teilnehmers spezifiziert, unter welchen Bedingungen er einen Kontakt wählt. Das Modell nutzt vier Parameter beziehungsweise vier Bedingungen, um Kontaktwahl-Strategien zu modellieren. Es sei dabei angemerkt, dass alle vier Bedingungen gelten müssen, damit ein Kontakt ausgewählt wird beziehungsweise weiterhin unterhalten wird.

1. Ein Teilnehmer kann mit Parameter $maxC$ spezifizieren, wie viele zusätzliche Kontakte er maximal unterhalten möchte.
2. Der wahrgenommene Kooperationsgrad eines zusätzlichen Kontakts muss höher sein als der Schwellenwert $additionalCC$ (additional-contact-cooperation Schwellenwert).
3. Die durchschnittliche Verarbeitungszeit der Anfragen, die Teilnehmer i an Kontakt j leitet, d.h. $\bar{h}_{\Delta}(i|j)$ (vergleiche Abschnitt 3.6), muss größer oder gleich Schwellenwert $maxProcessingTime$ sein.
4. Ein Teilnehmer spezifiziert die Lage seiner Kontakte im Adressraum mit Hilfe einer Wahrscheinlichkeitsverteilung (oder mit Hilfe einer Dichtefunktion im Fall eines kontinuierlichen Adressraums). Die Wahrscheinlichkeitsverteilung wird durch den Parameter $distribution$ repräsentiert. Wenn ein Teilnehmer k Bedingungen (1) bis (3) erfüllt, so wird er Kontakt von Teilnehmer i mit der Wahrscheinlichkeit, die seiner Lage in der Wahrscheinlichkeitsfunktion zugeordnet ist (vergleiche [Kle00]).

Übliche Strategien Es soll jetzt kurz wiedergegeben werden, wie mit Hilfe des definierten Strategieraums Strategien, die aus der Literatur bekannt sind, modelliert werden können:

- Ein *Interaktionswahl Altruist* [FF03] nutzt das System nicht, bearbeitet aber die Anfragen von anderen Teilnehmern ($ownCoop = 1$, $sendRequest = falsch$ und $contactC = 0$).
- Ein *Kontaktwahl Altruist* [FF03] akzeptiert Kontakte mit einem beliebigen Kooperationsgrad ($additionalCC = 0$ und $maxC$ ist eine große Konstante).
- Ein *Freifahrer auf Anfrageebene* bearbeitet keine Anfragen von anderen Teilnehmern: $ownCoop = 0$ [FLSC04].
- Ein Teilnehmer mit einer DPC-Strategie (vergleiche Abschnitt 4) unterstützt den Aufbau der Netzwerkstruktur nicht, d.h. $maxC = 0$ [HBKS07].
- Ein *Aussteiger* nutzt weder das Netzwerk noch trägt er zu ihm bei. (Es folgt, dass $sendRequest = falsch$ und $maxC = ownCoop = 0$ gilt).

KAPITEL 5. GLEICHGEWICHTE IN SYSTEMEN MIT INTERAKTIONS- UND KONTAKTWAHL

- *Schwellenwert-Strategien*: Menschen kooperieren miteinander, wenn der Anteil an bearbeiteten Anfragen pro Kontakt einen Schwellenwert überschreitet, d.h. Menschen setzen $contactC$ auf einen bestimmten Wert. Das gleiche Verhalten ist für Kontakte bekannt [HBV08], d.h. $additionalCC$ wird auf einen bestimmten Wert gesetzt.

Es können also aus der Literatur bekannte Strategien sowie weitere Strategie mit dem Strategieraum modelliert werden.

Um die im Folgenden durchgeführte Analyse zu vereinfachen, wird angenommen, dass $maxC$ ein exogener Parameter ist (vergleiche [BG00a, Jac03]). In Kapitel 6 werden Werte für $maxC$ experimentell bestimmt.

Der Strategieraum S eines Teilnehmers i ist wie folgt:

$$S = \{sendRequest, contactC, ownCoop, additionalCC, maxProcessingTime, distribution\} \quad (5.1)$$

Des Weiteren sei $S(i)$ der Strategievektor des Teilnehmers i .

5.3 Effizienz des Netzwerks

Ein Ziel in diesem Kapitel ist es, vorgegebene Systemzustände daraufhin zu überprüfen, ob sie ein Gleichgewicht sind. In der formalen Analyse wird mit Gleichgewichten begonnen, die für zwei Teilnehmer gelten. Später werden diese Resultate genutzt und für den generellen Fall verallgemeinert (also für $n \geq 2$).

Kooperatives Gleichgewicht und ineffizientes Netzwerk In der formalen Analyse dieses Kapitels wird gezeigt werden, dass sich das System in einem Gleichgewicht befindet, wenn alle Teilnehmer die folgende Strategie verfolgen:

$$S_c = \{wahr, 1.0, 1.0, 1.0, \bar{h}_\Delta(i), kleinberg\} \quad (5.2)$$

Dieses Gleichgewicht sei als *kooperatives Gleichgewicht* bezeichnet: Alle Teilnehmer senden Anfragen, kooperieren und haben Kleinberg-verteilte zusätzliche Kontakte, d.h. alle Teilnehmer haben Strategie S_c . Des Weiteren akzeptieren sie keine Teilnehmer als zusätzliche Kontakte, die ihre Anfragen langsamer verarbeiten als der Durchschnitt ihrer zusätzlichen Kontakte. Diese Strategie führt zu hohem Gemeinwohl, wie in Abschnitt 5.3.2 beschrieben werden wird.

Neben dem kooperativen Netzwerk existiert ein weiteres. Dieses sei als *ineffizientes Netzwerk* bezeichnet. Alle Teilnehmer haben dort die Strategie S_i : Die Teilnehmer partizipieren weder an der Interaktionswahl noch an der Kontaktwahl und haben deshalb keine Verteilung an Kontakten oder Schwellenwerte für die Verarbeitungszeit ihrer Anfragen, beides mit „-“ bezeichnet:

$$S_i = \{falsch, 0.0, 0.0, 0.0, -, -\} \quad (5.3)$$

Negative Auszahlungen Es kann ausgeschlossen werden, dass im Gleichgewicht Strategien mit negativer Auszahlung existieren. Strategie „Aussteiger“ (mit Auszahlung null) führt in jedem Fall zu einer höheren Auszahlung.

5.3.1 Verhalten von Kontakten

Die formale Analyse beginnt mit einem einfachen Fall: Es wird das Verhältnis zwischen zwei Teilnehmern, Teilnehmer i und Teilnehmer j untersucht. Dieser Fall erlaubt die Untersuchung des Verhältnisses zwischen zwei Kontakten. Es wird in diesem Abschnitt gezeigt, dass es mindestens zwei Gleichgewichte in dieser Situation gibt: das ineffiziente Netzwerk und das kooperative Gleichgewicht. Während das ineffiziente Netzwerk immer ein Gleichgewicht ist, hängt das kooperative Gleichgewicht von einigen Bedingungen ab, wie im Verlauf des Abschnitts erklärt werden wird.

Im Folgenden wird bei symmetrischen Fällen immer nur die Lage für einen Teilnehmer erörtert, nämlich für Teilnehmer i .

Ineffizientes Netzwerk Beide Teilnehmer seien Aussteiger ($sendRequest = falsch$; vergleiche Abschnitt 5.2.2). Keiner von beiden hat einen Anreiz das System aus zwei Spielern wieder zu betreten, da der andere Teilnehmer die Anfragen ignoriert. Die Strategie Aussteiger mit Auszahlung null ist lohnender als andere Strategien (da diese zu Ausgaben führen, aber nicht zu Einnahmen).

Kooperatives Gleichgewicht Im Folgenden werden Parameterwerte ermittelt, so dass vollständige Kooperation, d.h. $C(i) = 1.0$, ein Gleichgewicht darstellt.

Lemma 5.3.1 *Zwei Kontakte, Teilnehmer i und Teilnehmer j , kooperieren ($C(i) = C(j) = 1.0$), falls die folgenden Bedingungen gelten. (Es sei angemerkt, dass die Bedingungen auch im symmetrischen Fall gelten müssen, in dem i und j vertauscht sind.)*

$$istKooperativ(i, additionalCC_j = 1)$$

und

$$istKooperativ(i, contactC_j = 1)$$

und

$$\frac{ownCoop_i}{|R_{i \rightarrow j}|} < \frac{a \cdot ownCoop_j - q - c \cdot r(i)}{w \cdot |W_{j \rightarrow i}| + f \cdot |F_{j \rightarrow i}|}$$

Appendix B.1 enthält einen Beweis. Die dritte Bedingung beschreibt, welches Verhältnis die Auszahlungsfunktionskoeffizienten und die Anzahl an Anfragen, die Teilnehmer i an j schickt (d.h. $|R_{i \rightarrow j}|$) haben müssen, damit das Gleichgewicht be-

KAPITEL 5. GLEICHGEWICHTE IN SYSTEMEN MIT INTERAKTIONS- UND KONTAKTWAHL

stand hat. Solange der Gewinn beim Erhalt einer Antwort (a) groß ist im Verhältnis zu den Kosten des Erzeugens und Sendens einer Anfrage (q) oder dem Unterhalt von zusätzlichen Kontakten (c) sowie die Bearbeitungskosten für Anfragen ($w \cdot |W_{j \rightarrow i}| + f \cdot |F_{j \rightarrow i}|$), gilt die dritte Bedingung. Die ersten zwei Bedingungen sind bedeutender: Ein Teilnehmer muss vollständig kooperativ sein, da der andere Spieler nur vollständig kooperative Kontakte akzeptiert und sonst seine Anfragen ignoriert.

Es sei angemerkt, dass Lemma 5.3.1 eine wichtige Implikation beinhaltet: Im kooperativen Gleichgewicht, kooperieren Kontakte zu einem Grad, der von ihren Kontakten erwartet wird (und andersherum). Daher sind nur bilaterale Kontakte, d.h. Teilnehmer i vertraut Teilnehmer j und umgekehrt, stabil.

Im effizienten Zustand ist der Kooperationsgrad maximal, d.h. es gilt: $ownCoop_i = contactC_j = additionalCC_j = 1$. Dieser Zustand ist nur ein Gleichgewicht, wenn der Fehler beim Ermitteln des Kooperationsgrads kleiner als 50% ist. Appendix B.2 enthält einen Beweis.

Schossor [Sch08] kommt für Systeme ohne Kontaktwahl zu ähnlichen Ergebnissen.

5.3.2 n -Spieler

Im Gegensatz zum einfachen System aus zwei Teilnehmern, können die Teilnehmer des allgemeinen Falles, d.h. im n -Spieler System, zusätzlich zu der Interaktionswahl auch Kontakte wählen. Es ist aus [GGS03, VR06] bekannt, dass das Analysieren von Spielen mit n Spielern, die die Struktur des Netzes beeinflussen können, im Allgemeinen schwierig ist. Daher wird in diesem Abschnitt eine statistische Herangehensweise an das Problem verwendet: Die Teilnehmer des Spiels wählen ihre Kontakte nach einer Dichtefunktion (beziehungsweise nach einer Wahrscheinlichkeitsverteilung).

Wie in Abschnitt 5.3.1 gezeigt wurde, kooperieren die Teilnehmer mit ihren Kontakten. Daher kann im Folgenden angenommen werden, dass die Teilnehmer Geiziges Weiterleiten verwenden, dass sie also bei einer gegebenen Anfrage den Kontakt als Weiterleiter auswählen, der dem Ziel der Anfrage am nächsten liegt. Diese Annahme wird in Kapitel 6 relaxiert.

In einem n -Spieler System haben die Teilnehmer eine hohe Weiterleitungslast und die Struktur des Netzwerks spielt eine Rolle. Die Kontaktwahl-Strategien der Teilnehmer beeinflussen also die Effizienz des Systems. Im Folgenden wird die Kontaktwahl der Teilnehmer im Gleichgewicht bestimmt.

Im Gegensatz zum 2-Spieler System (vergleiche Abschnitt 5.3.1) kann sich die Netzwerkstruktur im n -Spieler System ändern, und damit auch die Kosten des Weiterleitens von Anfragen. Der zu erwartende Nutzen ändert sich folglich auch, wenn sich die Netzwerkstruktur ändert. Während sich die Definition des Einkommens aus Abschnitt 3.5 nicht ändert, ändert sich der Wert der zu erwartenden Ausgaben, d.h.

$E[Ausgaben(i)]$. Im Einzelnen ändern sich die Kosten für das Weiterleiten von Anfragen und die Kosten für den Unterhalt von zusätzlichen Kontakten:

$$E[Ausgaben(i)] = q \cdot |R_i| + ownCoop_i \cdot (f \cdot |F_i| + w \cdot |W_i|) + c \cdot r(i) \quad (5.4)$$

Wenn gilt, dass die Einnahmen kleiner als die Ausgaben sind, d.h. wenn

$$E[Einkommen(i)] < E[Ausgaben(i)]$$

gilt, dann ist Strategie Aussteiger lohnender als ein kooperatives Verhalten, andernfalls lohnt sich eine kooperative Strategie. Gleichung (5.4) hängt von der Weiterleitungslast $|F_i|$ von Teilnehmer i ab. Die Weiterleitungslast wiederum von der Netzwerkstruktur.

Kontaktwahl – Einfacher Fall

Die Netzwerkstruktur eines Systems hängt von den Kontaktwahl-Strategien der Teilnehmer ab. Im Folgenden werden zwei Fälle untersucht.

1. Im ersten Fall beachten die Teilnehmer bei der Auswahl ihrer zusätzlichen Kontakte den zu erwartenden Gewinn.
2. Im zweiten Fall hingegen wird die Bearbeitungszeit der Anfragen, die an einen Kontakt gesendet werden, berücksichtigt.

Wie wir im Folgenden sehen werden, sind die resultierenden Gleichgewichte für die beiden Fälle unterschiedlich. Es wird mit der Analyse des ersten Falles begonnen.

Weiterleitungslast in der festen Netzwerkstruktur Für die Analyse ist es wichtig, die Weiterleitungslast in der festen Netzwerkstruktur – also in der Ring-Netzwerkstruktur – zu verstehen. In einem solchen Netzwerk ist es wahrscheinlich, dass die Weiterleitungslast hoch ist. Es sei \bar{C} der durchschnittliche Kooperationsgrad im System. Die Wahrscheinlichkeit $p(h)$, dass eine Anfrage über h Schritte weitergeleitet wird, ist wie folgt: $p(h) = \bar{C}^h$.

Beispiel: Abbildung 5.2 zeigt Wahrscheinlichkeit $p(h)$ in einem kleinen System aus 40 Teilnehmern. Jeder Teilnehmer leitet Anfragen mit einer Wahrscheinlichkeit von 80% weiter. Da die initiale Netzwerkstruktur ein Ring ist, beträgt die maximale Anzahl an Weiterleitungen 20 Schritte. Mit jedem Schritt sinkt die Wahrscheinlichkeit, dass die Anfrage erfolgreich bearbeitet wird. ■

Für Systeme mit vielen Teilnehmern konvergiert die Wahrscheinlichkeitsfunktion $p(h)$ gegen eine kontinuierliche Funktion. Da der Fokus dieser Arbeit auf großen Systemen liegt, wird im Folgenden die kontinuierliche Funktion verwendet.

KAPITEL 5. GLEICHGEWICHTE IN SYSTEMEN MIT INTERAKTIONS- UND KONTAKTWAHL

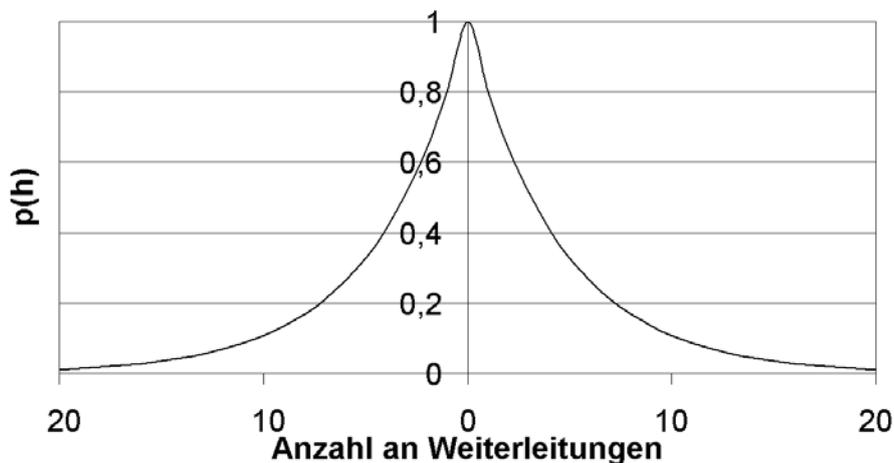


Abbildung 5.2: Weiterleitungswahrscheinlichkeit ohne zusätzliche Kontakte

Erwartete Einnahmen ohne zusätzliche Kontakte Hat ein Teilnehmer keine zusätzlichen Kontakte, so folgt, dass sein zu erwartendes Einkommen $E_h(h)$ beim Erzeugen und Senden einer Anfrage über h Weiterleitungsschritte wie folgt ist:

$$E_h(h) = p(h) \cdot \delta^h \cdot a - q \quad (5.5)$$

Das zu erwartende Einkommen $E(n)$ beim Erzeugen und Senden von einer Anfrage an jeden anderen Teilnehmer, also das zu erwartende Einkommen beim Erzeugen und Senden von $n - 1$ Anfragen ist:

$$E(n) = 2 \cdot \sum_{i=1}^{(n-1)/2} (p(i) \cdot \delta^i \cdot a - q) \quad (5.6)$$

Die Summe wird mit zwei multipliziert, da ein Teilnehmer $\frac{n-1}{2}$ Teilnehmer *rechts* und *links* von sich in der Ringstruktur hat. (Es wird in dieser Arbeit auf Rundungsoperatoren verzichtet, da durch sie die Darstellung verschlechtert wird, der Nutzen jedoch gering ist.) Obwohl Gleichung 5.6 das zu erwartende Einkommen beschreibt, sei an dieser Stelle eine Approximation für Gleichung 5.6 eingeführt, die im Folgenden verwendet wird. Diese erleichtert die mathematischen Operationen, die im Folgenden durchgeführt werden.

$$E_a(n) = 2 \cdot \int_1^{\frac{n-1}{2}} (p(h) \cdot \delta^h \cdot a - q) \, dh \quad (5.7)$$

$E_a(n)$ konvergiert gegen $E(n)$ für eine große Anzahl von Teilnehmern, d.h. es gilt $\lim_{n \rightarrow \infty} |E(n) - E_a(n)| = 0$.

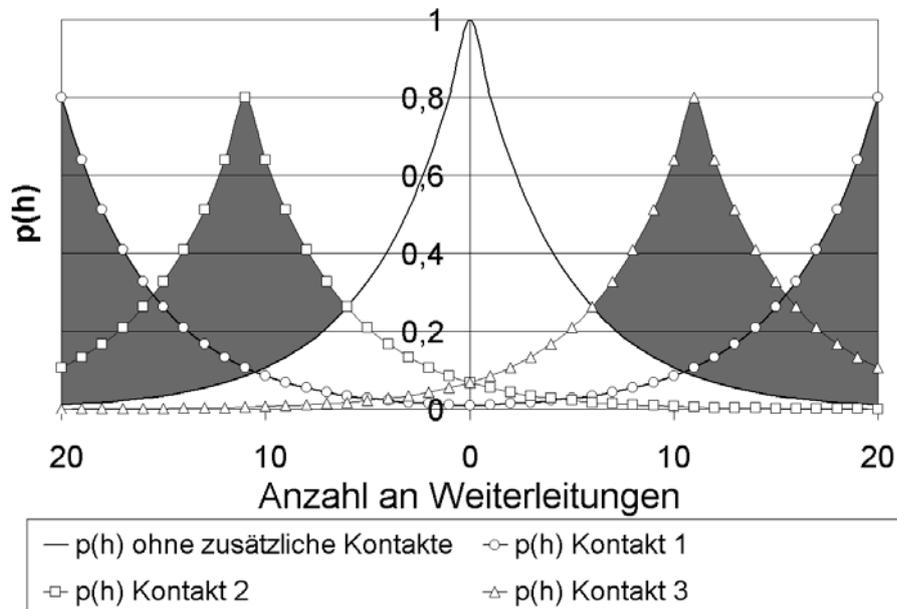


Abbildung 5.3: Weiterleitungswahrscheinlichkeit mit drei zusätzlichen Kontakten

Erwartete Einnahmen mit zusätzlichen Kontakten Mit der Verwendung von zusätzlichen Kontakten verändert sich das zu erwartende Einkommen $E(n)$.

Beispiel: Abbildung 5.3 zeigt die gleiche Situation wie im vorherigen Beispiel. Im Unterschied zu diesem hat der Teilnehmer nun drei zusätzliche Kontakte. Einer ist genau 20 Weiterleitungsschritte entfernt, die anderen zwei 11 Schritte (der eine rechts auf dem Ring, der andere links). Die graue Fläche ist die gewonnene Wahrscheinlichkeitsmasse, also der Bereich, in dem der Teilnehmer vorher Probleme hatte, Anfragen bearbeitet zu bekommen. ■

Das Einkommen, das durch die Verwendung von zusätzlichen Kontakten entsteht, ist maximal, wenn die dazugewonnene Wahrscheinlichkeitsmasse maximal ist. Wenn die Teilnehmer nicht die gesamte Kontaktverteilung der Teilnehmer einer Weiterleitungskette berücksichtigen, dann ist die Verteilung der Kontakte näherungsweise optimal, wenn die Kontakte denselben Abstand zueinander haben. Die exakten Positionen der Kontakte werden durch Lemma 5.3.2 beschrieben. *Optimal* bedeutet an dieser Stelle, dass durch die zusätzlichen Kontakte das maximale Einkommen entsteht.

Optimale Positionen Ein Teilnehmer, der seine Auszahlung maximieren möchte, muss optimale Positionen $p_i, i \in [1, \dots, |K^+|]$ für seine $|K^+|$ zusätzlichen Kontakte finden. (Aus Gründen der Vereinfachung sei angenommen, dass die Anzahl von Kontakten ungerade ist. Für eine gerade Anzahl von Kontakten kann die Analyse analog durchgeführt werden.) Wenn ein Teilnehmer nicht die Verteilung der Teilneh-

KAPITEL 5. GLEICHGEWICHTE IN SYSTEMEN MIT INTERAKTIONS- UND KONTAKTWAHL

mer einer Weiterleitungskette für eine gegebene Anfrage berücksichtigt, dann sind die optimalen Positionen für seine zusätzlichen Kontakte wie folgt:

Lemma 5.3.2 *Wenn Teilnehmer ihre Anfragen unabhängig davon verarbeiten, wie die Kontaktverteilung ihrer Kontakte gestaltet ist, dann sind die optimalen Positionen der Kontakte eines Teilnehmers mit Position $p_0 = 0$ wie folgt: Ein zusätzlicher Kontakt ist auf der anderen Seite der Ringstruktur, d.h. es gilt $p_{|K^+|} = \frac{n-1}{2}$; jeder andere zusätzliche Kontakt mit Position p definiert die Position eines weiteren zusätzlichen Kontaktes auf der anderen Seite des Rings mit Position $n - p$. Für $i = 1 \dots \frac{|K^+|-1}{2}$ sind die optimalen Positionen wie folgt:*

$$\frac{n-1}{2} \cdot \frac{2 \cdot i}{|K^+| + 1} + 2 \cdot \frac{|K^+| - i}{|K^+| + 1} - 4 \cdot \frac{|K^+| - i}{|K^+| + 1} \cdot \frac{\ln(\frac{1+\delta}{2})}{\ln(C \cdot \delta)}$$

Appendix B.3 beinhaltet einen Beweis.

Beispiel: Teilnehmer i mit Position $p_i = 0$ hat drei zusätzliche Kontakte ($|K_i^+| = 3$). Er berechnet die optimalen Positionen für seine Kontakte. Für einen durchschnittlichen Kooperationsgrad von 80% ($\bar{C} = 0.8$), 40 Teilnehmer im System und einem Diskontierungsfaktor von 99% ($\delta = 0.99$) sind die Positionen (11, 20, 29) optimal (vergleiche Abbildung 5.3): $\text{rundeAufInteger} \left([n-] \left(20 \cdot \frac{2}{4} + 2 \cdot \frac{2}{4} - 4 \cdot \frac{\ln(\frac{1+0.99}{2})}{\ln(0.8 \cdot 0.99)} \right) \right)$ ■

Fehlende Effizienz Die Situation in Abbildung 5.3 ist nicht effizient. Das Problem ist, dass zusätzliche Kontakte und die Kontakte von Kontakten ungefähr die gleichen Positionen haben. Das wiederum bedeutet, dass der Nutzen von Kontakten von Kontakten gering ist. Insbesondere sei angemerkt, dass Teilnehmer, die ihre Kontakte nach der oben genannten Strategie verteilen, keine zusätzlichen Kontakte in ihrer direkten Umgebung haben.

Die Implikationen des Lemma 5.3.2 sind überraschend: Ohne eine Strategie oder einen Mechanismus, die das Formen effizienter Netzwerke fördern, werden Teilnehmer, die ihre Auszahlung maximieren, ihre zusätzlichen Kontakte regelmäßig über den Adressraum verteilen. Sie werden also gleichverteilte zusätzliche Kontakte anstreben. Dieser Zustand ist – wie erläutert – nicht effizient.

Kontaktwahl – Allgemeiner Fall

Im Folgenden wird untersucht, wie Teilnehmer Netze formen, wenn sie ihre Kontaktwahl dadurch beeinflussen lassen, wie schnell ihre Kontakte ihre Anfragen verarbeiten (also danach, wie die Verteilung der Kontakte in der Weiterleitungskette gestaltet ist). D.h. es wird im folgenden Fall 2 untersucht. Es wird damit begonnen, zu erörtern,

welche Eigenschaften zusätzliche Kontakte erfüllen müssen, um nicht fallengelassen zu werden.

Das Verhältnis zwischen zwei Kontakten hat zwei Aspekte: Erstens, ein Teilnehmer benötigt zusätzliche Kontakte, um die Wahrscheinlichkeit, Antworten auf seine Anfragen zu erhalten, zu erhöhen. Zweitens, ein Teilnehmer möchte selbst als zusätzlicher Kontakt ausgewählt werden: Dies ist wichtig, da Kontaktverbindungen bilateral sind, und dies daher eine Voraussetzung für den ersten Fall ist. Der erste Aspekt führt zum Folgenden: Teilnehmer versuchen, solche Teilnehmer als zusätzliche Kontakte zu wählen, die wiederum ihre Kontakte nach einer Kleinberg-Verteilung gewählt haben (da in diesem Fall die Bearbeitung am schnellsten verläuft). Wegen Aspekt 2, also wegen der Wechselseitigkeit der Kontakt-Beziehung, hat ein Teilnehmer einen Anreiz, seine Kontakte ebenfalls nach einer Kleinberg-Verteilung zu wählen. Andernfalls ist er nicht attraktiv genug, um von anderen Teilnehmern als zusätzlicher Kontakt ausgewählt zu werden. Während Kleinberg in [Kle06] zeigt, dass eine Kleinbergverteilung optimal für Geiziges Weiterleiten ist, ist bisher unklar, ob Teilnehmer, die ihre Auszahlung maximieren, einen Anreiz haben, solche Netzwerke zu formen. Ein Beitrag dieser Arbeit ist es, zu zeigen, dass dem in der Tat so ist, d.h. dass eine Kleinberg-Verteilung der zusätzlichen Kontakte ein Gleichgewicht ist. Um dies zu zeigen, werden Kleinberg-verteilte zusätzliche Kontakte angenommen und dann überprüft, ob die Teilnehmer einen Anreiz besitzen von dieser Strategie abzuweichen. Es sei erneut angemerkt, dass aus Abschnitt 5.3.1 bereits bekannt ist, dass Kontakte miteinander kooperieren.

Nützliche Kontakte Aufgrund der Tatsache, dass eine schnelle Antwort meist nützlicher ist als eine verspätete, aufgrund der Tatsache also, dass diskontiert wird, verringert sich das Einkommen aus einer Antwort mit jedem Weiterleitungsschritt, der zwischen dem Erzeugen und Senden der Anfrage und dem Erhalt der Anfrage vergeht. Zwei Tatsachen beeinflussen die Anzahl der Weiterleitungen:

1. die Anzahl an Kontakten eines Teilnehmers
2. die Verteilung der Kontakte über den Adressraum

Um die Anzahl an zusätzlichen Kontakten zu erhöhen (Punkt 1), muss ein Teilnehmer als Kontakt attraktiv für andere Teilnehmer sein. Ein Teilnehmer ist dann attraktiv, (wenn er kooperativ ist und) wenn er die Anfragen der Kontakte schnell verarbeiten kann (Punkt 2). Nach [Kle00] leiten Teilnehmer Anfragen in optimaler Art und Weise weiter, wenn sie eine Kleinberg-Verteilung verwenden.

Evaluierung der Kontakte Da ein Teilnehmer die Verteilung der Kontakte seiner Kontakte nicht direkt beobachten kann, muss eine andere Technik verwendet werden. Eine Möglichkeit ist es, die Anzahl an Weiterleitungen (oder die Zeit) zu messen,

KAPITEL 5. GLEICHGEWICHTE IN SYSTEMEN MIT INTERAKTIONS- UND KONTAKTWAHL

die es braucht, um Anfragen zu verarbeiten. Wenn Anfragen, die an einen bestimmten Kontakt geleitet wurden, zu viele Weiterleitungsschritte benötigen (und wenn der Diskontierungsfaktor niedrig ist), dann ist der Kontakt nicht nützlich.

Drop-Slow-Contacts (DSC) Strategie: Ein Teilnehmer i misst die Zeit, die seine Anfragen benötigen, um beantwortet zu werden, d.h. er misst $\bar{h}_\Delta(i)$ (vergleiche Abschnitt 3.6). Er lässt zusätzliche Kontakte fallen, die Anfragen langsamer verarbeiten als der Durchschnitt seiner zusätzlichen Kontakte, d.h. wenn $\bar{h}_\Delta(i) < \bar{h}_\Delta(i|j)$ gilt.

Appendix B.4 zeigt, dass die Verwendung der Drop-Slow-Contacts-Strategie das kooperative Gleichgewicht nicht stört. Der Beweis nutzt die Tatsache, dass der Austausch eines vollständig kooperativen Kontaktes j , der Anfragen langsamer verarbeitet als ein vollständig kooperativer Teilnehmer i , mit Teilnehmer i , zu einer höheren Auszahlung führt.

Die DSC-Strategie gibt Teilnehmern einen Anreiz, Anfragen schnell zu verarbeiten, wenn sie selbst an zusätzlichen Kontakten interessiert sind. Dies ist der Fall, da ein Teilnehmer ständig mit anderen Teilnehmern im Wettbewerb um zusätzliche Kontakte steht. Es sei angemerkt, dass die DSC-Strategie kostenneutral ist und keine zusätzlichen Anfragen oder komplexen Berechnungen benötigt.

Beispiel: Ein Teilnehmer i habe drei zusätzliche Kontakte: Kontakte j , k , l . Angenommen, dass die Anfragen von Teilnehmer i im Durchschnitt in $\bar{h}_\Delta(i) = 2.6$ Schritten verarbeitet werden und dass die Verarbeitungszeit für Anfragen, die i an seine Kontakte sendet, wie folgt ist: $\bar{h}_\Delta(i|j) = 2$, $\bar{h}_\Delta(i|k) = 2$ und $\bar{h}_\Delta(i|l) = 4$. Dann würde Teilnehmer i nach der DSC-Strategie versuchen, Kontakt l gegen einen nützlicheren Kontakt zu tauschen, da $\bar{h}_\Delta(i) < \bar{h}_\Delta(i|l)$ gilt. ■

Im Folgenden werden die Implikationen der DSC-Strategie erörtert. Zuerst wird dabei ein spezieller Fall untersucht: Gibt es einen Anreiz, von Kleinberg-verteilten Kontakten zu einer Gleichverteilung der Kontakte zu wechseln? Danach wird der generelle Fall behandelt.

Es sei angenommen, dass alle Teilnehmer Kleinberg-verteilte zusätzliche Kontakte besitzen, mit Ausnahme von Teilnehmer u . Dieser hat gleichverteilte zusätzliche Kontakte. Des Weiteren ist u zusätzlicher Kontakt von Teilnehmer i . Sollte i nun u gegen einen anderen Teilnehmer ersetzen? Um die Frage zu beantworten, wird im Folgenden die zu erwartende Anzahl an Weiterleitungen von Anfragen von Teilnehmer i in beiden Fällen miteinander verglichen.

Kontakt mit gleichverteilten Kontakten Wenn Teilnehmer i seinen zusätzlichen Kontakt u behält, so ist die erwartete Anzahl an Weiterleitungen gleich $(1 + \beta \cdot \log^2(n))$ (mit einer festen Konstante β ; vergleiche [Kle00]): Die gleichverteilten Kontakte von Kontakt u sind mit hoher Wahrscheinlichkeit nicht nützlich für Teilnehmer i , wenn i seine Anfrage in die Umgebung von u leiten möchte. Die Ursache ist, dass Teilnehmer u keine zusätzlichen Kontakte in seiner direkten Umgebung hat (da er diese möglichst gut auf den Adressraum verteilt). Deshalb kann u die Anfrage nur

an einen festen Kontakt geben. Da dieser Kleinberg-verteilte Kontakte hat, ist nun die Weiterleitungskomplexität $\beta \cdot \log^2(n)$ [Kle00]. Das bedeutet, dass die Anfrage eine zusätzliche Weiterleitung erfährt, da Kontakt u keine Kontakte in seiner Umgebung besitzt.

Kontakt mit Kleinberg-verteilten zusätzlichen Kontakten Wenn Teilnehmer i Kontakt u mit einem Kontakt k austauscht, der – im Gegensatz zu u – Kleinberg-verteilte zusätzliche Kontakte hat, dann ist der Erwartungswert für die Anzahl an Weiterleitungen $\beta \cdot \log^2(n)$ [Kle00] und nicht $1 + \beta \cdot \log^2(n)$.

Lemma 5.3.3 *Die Auszahlung von Teilnehmer i erhöht sich für jede Anfrage, die er zu Teilnehmer k mit Kleinberg-verteilten zusätzlichen Kontakten sendet, verglichen mit dem Fall, in dem i seinen Kontakt u mit gleichverteilten Kontakten nutzt, um den folgenden Betrag:*

$$a \cdot \bar{C}^{\beta \cdot \log^2(n)} \cdot \delta^{\beta \cdot \log^2(n)} \cdot (1 - \bar{C} \cdot \delta)$$

Appendix B.5 enthält einen Beweis. Die Beweisidee ist, die zu erwartenden Auszahlungen von Teilnehmer i mit Kontakt u und mit Kontakt k zu vergleichen. Die Konsequenzen von Lemma 5.3.3 sind wie folgt: In der Gegenwart unkooperativer Teilnehmer oder einer Diskontierung (d.h. $\delta \neq 1$) tauscht Teilnehmer i Kontakt u gegen k aus, da dies zu einer höheren Auszahlung führt. Das bedeutet, dass Teilnehmer Kontakte mit Kleinberg-verteilten Kontakten solchen vorziehen, deren Kontakte gleichverteilt sind. In anderen Worten, ein Teilnehmer mit Kleinberg-verteilten Kontakten kann nicht dadurch seine Auszahlung maximieren, dass er seine Strategie zu gleichverteilten Kontakten abändert: Seine zusätzlichen Kontakte würden ihn gegen andere Teilnehmer austauschen und er hätte damit Probleme, überhaupt zusätzliche Kontakte zu finden. Dies würde zu einer geringeren Auszahlung führen.

Weitere Verteilungen Kontakte, die Kleinberg-verteilte Kontakte besitzen, die nicht optimal verteilt sind, d.h. für die $r_k \neq 1$ gilt (vergleiche Kapitel 2 oder [Kle00]), oder die eine andere Verteilung ihrer Kontakte besitzen, haben ebenfalls keine optimale Strategie. Wenn ein Kontakt seine Kontakte nicht in einer optimalen Art und Weise verteilt hat, dann erhöht sich die Anzahl an Weiterleitungen, die notwendig ist, um eine Anfrage zum Ziel zu leiten. D.h. es existiert ein $\epsilon > 0$, so dass sich die Anzahl an Weiterleitungen um ϵ Schritte erhöht. Die Konsequenz ist, dass die Auszahlung eines Teilnehmers, der einen zusätzlichen Kontakt mit einer suboptimalen Verteilung hat, verringert.

Lemma 5.3.4 *In der Gegenwart unkooperativer Teilnehmer oder einer Diskontierung (d.h. $\delta \neq 1$) erhöht sich die Auszahlung eines Teilnehmers i für jede Anfrage, die er zu Teilnehmer k mit Kleinberg-verteilten Kontakten schickt, verglichen mit dem*

KAPITEL 5. GLEICHGEWICHTE IN SYSTEMEN MIT INTERAKTIONS- UND KONTAKTWAHL

Fall, in dem er die Anfrage zu einem Kontakt u mit einer Verteilung an zusätzlichen Kontakten schickt, die nicht der optimalen Kleinberg-Verteilung entspricht.

Appendix B.6 enthält einen Beweis. Die Beweisidee ist erneut der Vergleich der zu erwartenden Auszahlungen. Lemma 5.3.4 folgend, wird ein seine Auszahlung maximierender Teilnehmer Kontakte mit einer nicht optimalen Verteilung an zusätzlichen Kontakten gegen solche austauschen, die Kleinberg-verteilte zusätzliche Kontakte besitzen. Dies bedeutet, dass Teilnehmer mit einer nicht-optimalen Kontaktverteilung ihre Kontakte verlieren und damit eine geringere Auszahlung besitzen. Sie haben also kein Interesse, vom kooperativen Gleichgewicht abzuweichen.

Zusammenfassung In Lemma 5.3.1 wird gezeigt, unter welchen Bedingungen Kontakte kooperieren. Des Weiteren benutzen die Teilnehmer entweder eine DSC-Strategie oder sie nehmen davon Abstand. Für den zweiten Fall zeigt Lemma 5.3.2 wie Teilnehmer ihre Kontakte wählen sollten. Es sei angemerkt, dass diese Situation nicht effizient ist (vergleiche Abschnitt 5.3.2). Das Gemeinwohl ist im ersten Fall größer, in dem die Teilnehmer die DSC-Strategie nutzten. Lemmata 5.3.3 und 5.3.4 zeigen, dass alle Teilnehmer, die diese Strategie benutzen, Kleinberg-verteilte Kontakte im Gleichgewicht besitzen. Da nach [Kle00] eine Kleinbergverteilung optimal für Geiziges Weiterleiten ist, folgt, dass das Gleichgewicht effizient ist.

5.4 Fazit

Entwickler verteilter Systeme geben ihren Teilnehmern oft ein Verhalten vor. Ob die Teilnehmer von diesem Verhalten abweichen, oder nicht, kann durch das Gleichgewichtskonzept überprüft werden. In einem Gleichgewicht hat kein Teilnehmer ein Interesse, einseitig von seiner Strategie abzuweichen. In diesem Abschnitt wurden zwei Gleichgewichte ermittelt: (1) Teilnehmer kooperieren nicht. In dieser Situation ist es sinnvoll, nicht mehr am System teilzunehmen. (2) Teilnehmer nutzen die DSC-Strategie, die in diesem Kapitel vorgeschlagen wird. Diese Strategie ist wie folgt: Die Teilnehmer messen die Zeit (beziehungsweise die Anzahl der Weiterleitungsschritte), die eine Anfrage benötigt, um beantwortet zu werden. Ist diese Zeit größer als die durchschnittliche Weiterleitungszeit aller Kontakte, so lässt ein Teilnehmer den Kontakt fallen, an den er die Anfrage gesendet hatte. Des Weiteren kooperieren Kontakte miteinander und alle Teilnehmer haben Kleinberg-verteilte zusätzliche Kontakte. Die DSC-Strategie ist kostenneutral und die Teilnehmer können das Gemeinwohl erhöhen, indem sie die DSC-Strategie verwenden.

Während in diesem Kapitel davon ausgegangen wird, dass alle Teilnehmer eines Systems im Gleichgewicht die selbe Strategie haben, werden im nächsten Kapitel Systemzustände untersucht, die auch dann noch stabil sind, wenn einzelne zufällig eingestreute zufällige Strategien das Gleichgewicht stören.

6 Evolution der Interaktions- und Kontaktwahl

Heutige Soziale Netzwerke und Peer-to-Peer Systeme unterliegen einer ständigen Dynamik. Neben der Dynamik des Ein- und Austretens von Teilnehmern oder den dynamischen Veränderungen des Wissens und der Daten im System, unterliegen Soziale Netzwerke und Peer-to-Peer Systeme dynamischen Änderungen der Strategien der Teilnehmer. Auf diese konzentriert sich dieses Kapitel. Menschen mit einer bestimmten Strategie lernen, ob ihre Strategie im Vergleich zu anderen Strategien erfolgreich ist oder nicht. Falls dem nicht so ist, passen sie ihre Strategie gegebenenfalls an. Dies kann durch Veränderung ihrer eigenen Strategie, durch Kopieren anderer Strategien oder durch Rekombination der eigenen Strategie mit einer anderen Strategie passieren. Oft sind die Strategieanpassungen nicht perfekt, d.h. führen nicht dazu, dass die Strategie, die zur maximalen Auszahlung führt, gefunden wird. Dies liegt (wahrscheinlich) daran, dass das zugrunde liegende Problem der Interaktionswahl und Kontaktwahl so komplex ist, dass es schwierig ist, eine optimale Strategie zu finden [GG03, VR06]. Es sei daher in diesem Kapitel angenommen, dass sich die Teilnehmer begrenzt rational verhalten (engl. bounded rationality) [BEG06]. Dass sie also ihre (erfolglosen) Strategien durch Anpassungen, durch Kopieren oder durch Rekombination verändern und jene Strategien behalten, die in Relation zu anderen Strategien erfolgreich sind. Solche Anpassungen werden auch als evolutionäre Verbesserungen bezeichnet. Unklar ist, welche Strategien sich in diesen dynamischen Systemen durchsetzen und unter welchen Voraussetzungen. Es ist ebenfalls unklar, ob die dynamischen Prozesse zu kooperativen Systemen führen, d.h. zu Systemen, in denen Teilnehmern ein positiver Nutzen aus der Zusammenarbeit mit anderen Teilnehmern entsteht.

Evolutionäre Stabilität *Evolutionär stabile Strategien* (ESS) sind eine Untermenge von Nash-Gleichgewichten [MS74, BEG06]. Nutzen alle Teilnehmer eine ESS, so befindet sich das System im Gleichgewicht [BEG06]. Dennoch haben sie eine andere Interpretation.

Eine Strategie S ist eine evolutionär stabile Strategie, wenn die zwei folgenden Bedingungen gelten (vgl. [MS74]):

- Strategie S wird nicht durch eine Strategie S' , die durch Mutationen oder Rekombination aus S hervorgegangen ist, verdrängt. Es sei dabei angemerkt, dass

KAPITEL 6. EVOLUTION DER INTERAKTIONS- UND KONTAKTWAHL

mit einer genügend hohen Anzahl an Mutationen jede Strategie aus jeder anderen Strategie hervorgehen kann. Eine Verdrängung ist wie folgt definiert: Eine Strategie S' verdrängt Strategie S , wenn sie beim Zusammentreffen auf eine Population der Strategie S zu einer höheren Auszahlung führt.

- Falls die Auszahlung durch Strategie S' beim Auftreffen auf eine Population der Strategie S gleich der Auszahlung durch Strategie S ist, so muss gelten, dass Strategie S beim Auftreffen auf eine Population der Strategie S' zu einer höheren Auszahlung führt.

In diesem Kapitel werden mittels Simulationen evolutionär stabile Strategien identifiziert, um die Dynamik von Strategien in der Interaktions- und Kontaktwahl zu untersuchen.

Fragestellung Die wissenschaftlichen Fragen, die in diesem Kapitel beantwortet werden sollen, sind wie folgt:

1. Ist das identifizierte kooperative Gleichgewicht aus Kapitel 5 evolutionär stabil?
2. In welchen Zustand führen evolutionäre Verbesserungen zufälliger Strategien? Findet das System einen Gleichgewichtszustand, also einen Zustand, in dem das System trotz der evolutionären Anpassungen stabil ist? Falls ja, wie sieht dieser Zustand aus?

Die Beantwortung dieser Fragen ist nicht trivial. Simulationen bergen die Gefahr, dass der Experimentator durch den Entwurf und Aufbau der Simulationen die Simulationen so beeinflusst, dass das Simulationsergebnis bereits impliziert wurde. Die Simulation evolutionärer Prozesse versucht, diesen Effekt zu verhindern: Sie bedient sich eines bekannten und wohl erforschten Algorithmus – der Evolution – und behandelt die strategischen Auswahlmöglichkeiten eines Teilnehmers als *Genom*. Es ist daher für den Experimentator nicht nötig, Annahmen über die Strategien oder die Verteilung der Strategien im System vorzunehmen. Es ist lediglich nötig, einen Strategieraum zu definieren. Theoretische Analysen, die ebenfalls genutzt werden können, um dynamische Strategieveränderungen in der Interaktionswahl und Kontaktwahl zu untersuchen, sind exakter, führen jedoch zu umfangreichen und komplexen Formeln, die zum Teil nicht mehr analytisch gelöst werden können beziehungsweise die nicht mehr handhabbar sind.

Beiträge Dieses Kapitel hat zwei wichtige Beiträge:

1. Die theoretischen Ergebnisse aus Kapitel 5 werden bestätigt.

2. Es wird gezeigt, dass Kooperation – wenn auch auf niedrigem Niveau – aus Netzwerken mit Teilnehmern, die eine zufällige Anfangsstrategie besitzen, entstehen kann, wenn diese Teilnehmer ihre Strategien durch evolutionäre Verbesserungen anpassen.

Es sei angemerkt, dass sich der zweite Beitrag vom ersten unterscheidet: Um zu zeigen, dass eine Strategie ein Gleichgewicht ist, wird angenommen, dass alle Teilnehmer diese Strategie nutzen. Es wird dann überprüft, ob sie einen Anreiz haben, die Strategie abzuändern. Der zweite Beitrag ist davon verschieden, da die Teilnehmer mit zufälligen Strategien starten und diese dann durch evolutionäre Anpassungen verändern.

Die Beiträge sind wichtig, da sie eine Erklärung geben, warum kooperatives Verhalten in Systemen mit Interaktions- und Kontaktwahl entstehen kann. Ebenfalls zeigt Beitrag (1), dass einmal entstandene Kooperation stabil ist.

6.1 Versuchsaufbau

Netzwerk-Spiel Die simulierten Teilnehmer spielen das folgende Spiel, das dem Verhalten der Teilnehmer aus Kapitel 5 nachempfunden ist. Das Spiel sei als *Netzwerk-Spiel* bezeichnet. Die Spieler spielen das Spiel in Runden, d.h. Zeit ist diskret:

1. Jeder Spieler kann eine Anfrage pro Runde erzeugen und senden – oder es sein lassen. Er kann sie an einen Kontakt senden.
2. Am Anfang jeder Runde erhält ein Spieler i alle Anfragen, die ihm seine Kontakte in der vorherigen Runde geschickt haben. Spieler i entscheidet für jede Anfrage, ob er sie fallen lässt oder ob er sie verarbeitet (d.h. sie beantwortet oder weiterleitet).
3. Ein Teilnehmer kann zusätzliche Kontakte nach eigenem Ermessen fallenlassen und kann Spieler zu zusätzlichen Kontakten machen, mit denen er bereits einmal interagiert hat. Zum Beispiel kann Spieler i Spieler j zu einem zusätzlichen Kontakt machen, falls Spieler j mindestens eine Anfrage erzeugt hat, die Spieler i erreicht hat. Das Fallenlassen und Hinzufügen von zusätzlichen Kontakten ist in jeder Runde möglich.

Die Spieler haben dabei den aus Kapitel 5 bekannten Strategieraum, der durch den Strategievektor S repräsentiert wird (vgl. Gleichung (5.1)). Dieser wird noch um die maximale Anzahl an zusätzlichen Kontakten $maxC$ und um einen weiteren Parameter, nämlich Parameter *locationWeight*, erweitert. Dieser Parameter beschreibt, welchen Einfluss die Lage eines Kontaktes im Adressraum bei der Anfragebearbeitung hat.

KAPITEL 6. EVOLUTION DER INTERAKTIONS- UND KONTAKTWAHL

locationWeight – Einfluss der Lage eines Kontaktes Beim Weiterleiten von Anfragen mag es eine Abwägung zwischen der Lage eines Kontaktes relativ zum Ziel einer Anfrage und seinem Kooperationsgrad geben. Zum Beispiel kann es sein, dass ein mittelmäßig kooperativer Kontakt, der sehr nah am Ziel einer Anfrage gelegen ist, eine bessere Wahl darstellt als ein vollständig kooperativer Kontakt, der weit weg vom Ziel der Anfrage gelegen ist.

Im Gegensatz zur formalen Analyse aus Kapitel 5 wird in diesem Abschnitt nicht angenommen, dass die Teilnehmer beim Weiterleiten von Anfragen stets den Kontakt auswählen, der dem Ziel der Anfrage am nächsten liegt. Vielmehr wird diese Annahme in diesem Kapitel gelockert: Ein Teilnehmer wählt einen Kontakt, der näher zum Ziel der Anfrage gelegen ist als er selbst.

Der Strategieraum-Parameter *locationWeight* definiert, wie wichtig einem Spieler die Lage seines Kontaktes relativ zu seinem Kooperationsgrad ist. Zum Beispiel bedeutet *locationWeight* = 1, dass der Teilnehmer immer den Kontakt auswählt, der dem Ziel der Anfrage am nächsten ist, unabhängig davon, wie kooperativ der Kontakt ist.

Algorithmus 8: Der Einfluss des Parameters *locationWeight*

Input : Menge der Kontakte K , Ziel der Anfrage t

- 1 $\max(\Delta) = \max_{k \in K} \Delta(k, t)$;
- 2 $\min(\Delta) = \min_{k \in K} \Delta(k, t)$;
- 3 $\max(C) = \max_{k \in K} C(k)$;
- 4 $\min(C) = \min_{k \in K} C(k)$;
- 5 Gleitkommazahl bewertung = -1.0;
- 6 Kontakt kontakt = null;
- 7 **foreach** Kontakt $k \in K$ **do**
- 8 Gleitkommazahl $r =$
 $\frac{\max(\Delta) - \Delta(k, t)}{\max(\Delta) - \min(\Delta)} \cdot \text{locationWeight} + \frac{C(k) - \min(C)}{\max(C) - \min(C)} \cdot (1 - \text{locationWeight})$;
- 9 **if** $r > \text{bewertung}$ **then**
- 10 bewertung = r ;
- 11 kontakt = k ;
- 12 **return** kontakt;

Algorithmus 8 zeigt, wie ein Spieler einen Kontakt beim Weiterleiten einer Anfrage auswählt: In Abhängigkeit von Parameter *locationWeight* berechnet der Algorithmus eine Bewertung für jeden Kontakt. Algorithmus 8 identifiziert den Kontakt mit der größten (kleinsten) Distanz zum Ziel der Anfrage und speichert diese Distanz (Zeile 1 und 2). Ebenfalls wird der höchste (geringste) Kooperationsgrad der Kontakte gespeichert (Zeile 3 und 4). Danach berechnet der Algorithmus eine Bewertung für jeden Kontakt (Zeilen 7-8) und merkt sich dabei den Kontakt, mit der höchsten

Bewertung (Zeilen 9-11). Der Teilnehmer sendet die Anfrage an den Kontakt mit der höchsten Bewertung.

Genom Die Strategie $S(i)$ eines Teilnehmers i wird durch ein Genom $G(i)$ repräsentiert. In diesem Paragraph wird kurz darauf eingegangen, wie dieses Genom gestaltet ist. Das Genom besteht aus den folgenden Teilen:

- Der binäre Parameter *sendRequest* wird durch ein Bit kodiert.
- Die folgenden Parameter werden durch sieben Bits kodiert: *contactC*, *ownCoop*, *additionalCC*, *locationWeight* und *distribution*. Sie können damit auf 128 Zustände quantisiert werden. Da alle Parameter den Wertebereich $[0, 1]$ abdecken, ergibt sich eine Genauigkeit der Quantifizierung von über 99%.
- Parameter *maxC* ist mit acht Bit kodiert und kann Werte im Bereich $[0, 497]$ annehmen. Der Fehler in der Quantifizierung beträgt daher ebenfalls weniger als 1%.

Die Werte werden im GrayCode kodiert. Dadurch unterscheiden sich benachbarte Codewörter in der Hamming-Distanz nur um eins. Dies führt dazu, dass der Fehler, der bei der Quantifizierung auftritt, auf ein Codewort begrenzt ist.

Parameter *maxProcessingTime* ist als Java-Double mit 80 Bit kodiert, da er direkt im DSC-Algorithmus verwendet wird.

Simulationsumgebung Es werden in diesem Kapitel jeweils 500 Teilnehmer simuliert. Ihre Strategien werden dynamisch angepasst. Es wird dabei angenommen, dass Strategien, die zu einer geringeren Auszahlung führen als andere, schneller aussterben als andere Strategien. Des Weiteren wird angenommen, dass neue Strategien mit einer gewissen Wahrscheinlichkeit entstehen. Nach 1.000 Runden im Netzwerk-Spiel wird ein evolutionärer Schritt vollzogen: Strategien mit einer hohen Auszahlung bleiben bestehen, Strategien mit einer geringen Auszahlung sterben aus. Im Einzelnen wird der evolutionäre Prozess wie folgt simuliert:

1. Es wird single-point crossover verwendet und eine crossover-Rate von 20%. Dies bedeutet, dass von 500 Teilnehmern im System 100 ihre Strategien kombinieren. Sie teilen ihren Strategievektor an einer zufälligen Stelle und tauschen je einen Teil miteinander aus.
2. Es wird eine Mutationswahrscheinlichkeit von 10% verwendet. Dies bedeutet, dass 50 von 500 Teilnehmern ihre Strategie zufällig abändern.

Die Idee hinter Maßnahme (1) ist, dass die Teilnehmer ihre Strategien mit anderen Teilnehmern kombinieren, um erfolgreicher zu werden. Maßnahme (2) sorgt dafür,

KAPITEL 6. EVOLUTION DER INTERAKTIONS- UND KONTAKTWAHL

dass neue Strategien entstehen können. Die Intuition dahinter ist, dass einige Teilnehmer neue Ideen darüber haben, wie ihre Strategien aussehen sollten. Es sei dabei angemerkt, dass es äquivalent ist, ob ein Teilnehmer seine Strategie ändert oder ob er das Netzwerk verlässt und ein neuer Teilnehmer mit einer neuen Strategie das Netzwerk betritt: Kein Teilnehmer hat eine Statistik oder eine Historie der Ereignisse aus vorhergehenden evolutionären Schritten. (Innerhalb eines evolutionären Schrittes, also innerhalb der 1.000 Runden des Netzwerk-Spiels kann dies sehr wohl der Fall sein.) Diese Modellierung ist natürlich, da die Statistik oder die Historie sich auf Strategien beziehen würde, die nicht mehr existieren. Die Statistik oder Historie wäre daher semantisch sinnlos.

Eine Strategie kommt mit einer Wahrscheinlichkeit proportional zu ihrer Fitness in den nächsten evolutionären Schritt. Die Fitness-Funktion setzt die Auszahlung eines Teilnehmers ins Verhältnis zur durchschnittlichen Auszahlung.

Definition 6.1 (Fitness-Funktion)

Die Fitness eines Teilnehmers i ist wie folgt:

$$fitness(i) = \frac{Auszahlung(i) - \min_{j \in N} \{Auszahlung(j)\}}{Auszahlung - \min_{j \in N} \{Auszahlung(j)\}} \quad (6.1)$$

Exogene Parameter Neben den dynamischen Strategieveränderungen werden exogene Parameter verändert:

1. Die Unterhaltskosten für zusätzliche Kontakte c .
2. Die Qualität der Reputationsinformationen über den Kooperationsgrad der Teilnehmer, d.h. Parameter *reportingError*.
3. Der Wert einer Antwort in Relation zur Zeit, die nötig ist, die zugehörige Anfrage zu beantworten: Diskontierungsfaktor δ .

Die weiteren vorhandenen exogenen Parameter werden hier nicht variiert, da eine Berücksichtigung ihrer Effekte schon in Kapitel 4 durchgeführt wurde.

Basiswerte Bei den Simulationen wird eine Menge von Basiswerten für die exogenen Parameter verwendet. Je ein exogener Parameter wird in den Simulationen variiert. Die anderen Parameter behalten die Basiswerte. Die Basiswerte sind wie folgt: Für jeden zusätzlichen Kontakt zahlt ein Spieler einen Punkt pro Runde ($c = 1$). Beim Bestimmen des Kooperationsgrads eines anderen Teilnehmers existiert ein Fehler von 5% (*reportingError* = 0,05). Wenn ein Spieler eine Anfrage erhält, so gilt ein Diskontierungsfaktor von 90% ($\delta = 0,9$).

6.2. ERGEBNISSE – KOOPERATIVES GLEICHGEWICHT

Auszahlungsfunktionskoeffizienten Teilnehmer Sozialer Netzwerke im Internet oder von Peer-to-Peer Systemen haben in der Regel geringe Kosten für das Weiterleiten, das Beantworten oder das Erzeugen von Anfragen sowie geringe Kosten für den Unterhalt von Kontakten. Auf der anderen Seite erfährt ein Teilnehmer einen Nutzen, wenn er eine Antwort auf seine Anfrage erhält. Daher sind die Auszahlungsfunktionskoeffizienten in den Simulationen wie folgt: Die Auszahlung, die ein Spieler beim Erhalt einer Anfrage erhält (ohne dass der Diskontierungsfaktor berücksichtigt ist), beträgt 1.000 Punkte. Die Kosten für das Erzeugen und Versenden von Anfragen beträgt 2 Punkte, Weiterleitungskosten 1 Punkt und die Kosten für das Beantworten von Anfragen betragen 5 Punkte.

Bis auf den Gewinn a beim Erhalt einer Antwort entsprechen die Werte den Werten, die in den Simulationen in Abschnitt 4.9 verwendet wurden. Ist der Gewinn beim Erhalt einer Anfrage zu gering im Verhältnis zu den Kosten der Anfragebearbeitung, so ist das Aussteigen aus dem System die Folge. Da die Kosten für die Verarbeitung einer Anfrage in einem System aus 500 Teilnehmern, die in einem Ring angeordnet sind, über 250 Punkte betragen können, wurde hier ein entsprechend hoher Wert für den Gewinn beim Erhalt einer Anfrage gewählt.

Kontaktverteilung Um die Ergebnisse aus dem vorherigen Kapitel zu validieren, wird angenommen, dass die Teilnehmer Kleinberg-verteilte Kontakte besitzen. D.h. ein Spieler i wählt einen Teilnehmer j mit einer Wahrscheinlichkeit, die zu $d(i, j)^{-r}$ proportional ist, als zusätzlichen Kontakt (wenn Parameter $maxC$ und $additionalCC$ dies zulassen). Abbildung 6.1 illustriert den Effekt von Parameter r . Spieler mit einem höheren r -Wert bevorzugen zusätzliche Kontakte, die nahe bei ihnen gelegen sind, im Gegensatz zu zusätzlichen Kontakten, die weit entfernt sind. Es sei angemerkt, dass Parameter r Teil des Strategieraums ist (vergleiche Kapitel 5, Parameter *distribution*).

6.2 Ergebnisse – Kooperatives Gleichgewicht

In diesem Abschnitt wird überprüft, ob das in Kapitel 5 ermittelte Gleichgewicht auch evolutionär stabil ist.

Initiale Strategie Die simulierten Teilnehmer benutzen die folgende Strategie am Anfang der Simulation, d.h. vor dem ersten evolutionären Schritt: Die Teilnehmer nehmen am System teil und kooperieren ($sentRequest = wahr$; $ownCoop = 100\%$); sie nutzen eine Schwellenwert-Strategie und ignorieren Teilnehmer, die einen Kooperationsgrad von weniger als 95% besitzen ($contactC = additionalCC = 95\%$). Beim Weiterleiten von Anfragen oder beim Senden eigener Anfragen legen sie kaum Wert auf den Kooperationsgrad des Kontaktes, an den sie die Anfrage weiterleiten ($locationWeight = 95\%$); die Spieler haben fünf zusätzliche Kontakte und

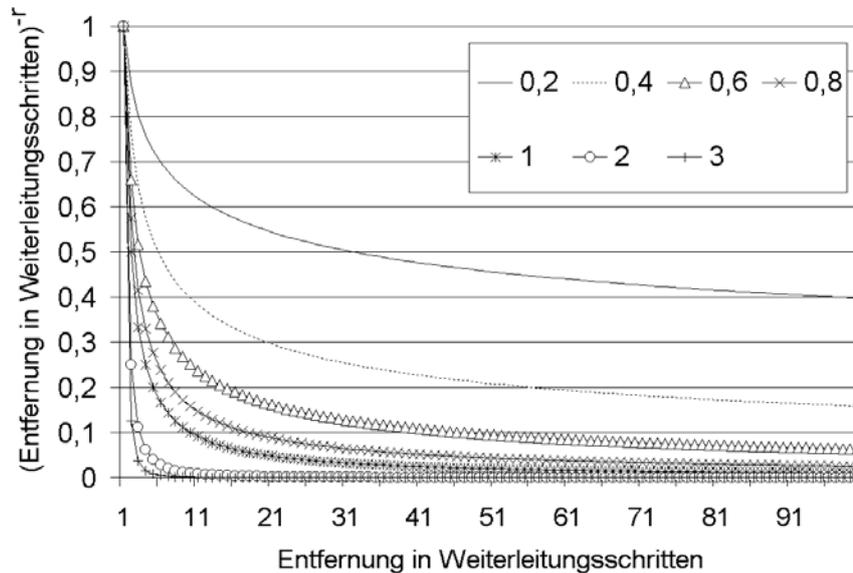


Abbildung 6.1: Einfluss des Parameters r

wählen Kontakte mit einer höheren Wahrscheinlichkeit, wenn diese näher zu ihnen gelegen sind ($maxC = 5$; $r = 1$).

Ergebnisse – Basiswerte Abbildungen 6.2 und 6.3 zeigen die Resultate der Simulation mit den Basiswerten. Abbildung 6.2 zeigt die Evolution in der Interaktionswahl und Abbildung 6.3 die Evolution der zusätzlichen Kontakte. Beide Abbildungen zeigen jeweils den Durchschnitt über alle Strategien. Für Abbildung 6.3 wurden die zusätzlichen Kontakte nach ihrer Distanz zum Teilnehmer in acht äquidistante Klassen gruppiert. Die Klasse, die am nächsten am Teilnehmer liegt (d.h. Klasse B1) wird am unteren Ende von Abbildung 6.3 gezeigt, die nächstgelegene (d.h. B2) darüber, und so fort. Für die beiden Abbildungen wurden insgesamt 100.000 Spieler, die 100.000.000 Anfragen erzeugt haben, simuliert. Es sei angemerkt, dass – obwohl hier nur ein repräsentatives Ergebnis gezeigt wird – alle Simulationen mindestens 25 mal durchgeführt wurden, um sicherzustellen, dass es sich bei den Ergebnissen nicht um Ausreißer handelt.

Die Evolution der Interaktionswahl ist wie folgt (vgl. Abbildung 6.2): Die Spieler nehmen am System teil ($sendRequest = 1$; $Auszahlung(i) > 0$, beides der Übersicht halber nicht dargestellt), haben einen hohen Kooperationsgrad ($ownCoop \approx 1$) und nutzen hohe Schwellenwerte für ihre Kontakte (es gilt $contactC \approx 0,5$ und $additionalCC \approx 0,85$). Es sei angemerkt, dass sich kein signifikanter Wert für den $contactC$ -Schwellenwert einstellt: Über alle Simulationen hat der Wert eine große Varianz (während alle anderen Ergebnisse repräsentativ sind). Eine Erklärung dafür

6.2. ERGEBNISSE – KOOPERATIVES GLEICHGEWICHT

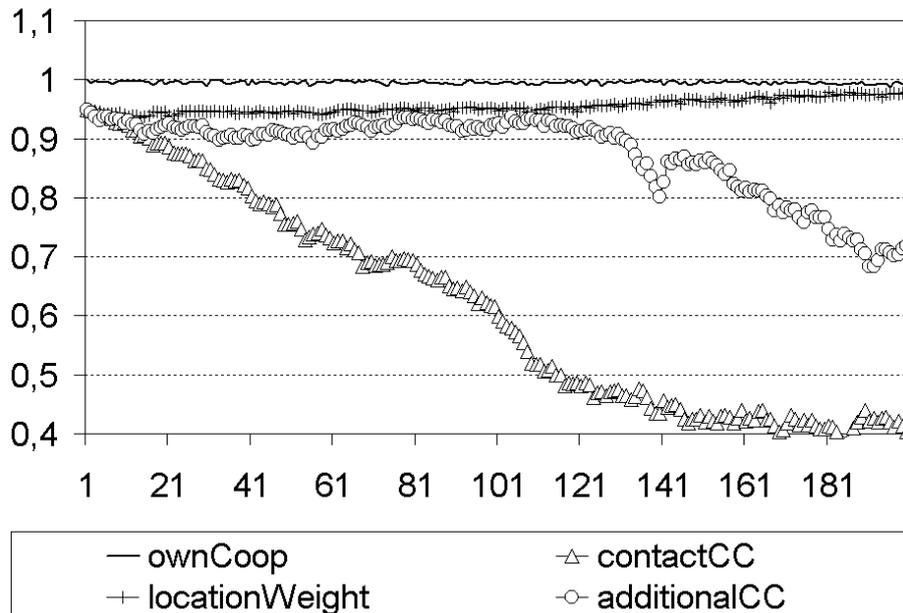


Abbildung 6.2: Interaktionswahl-Strategie im evolutionären Prozess

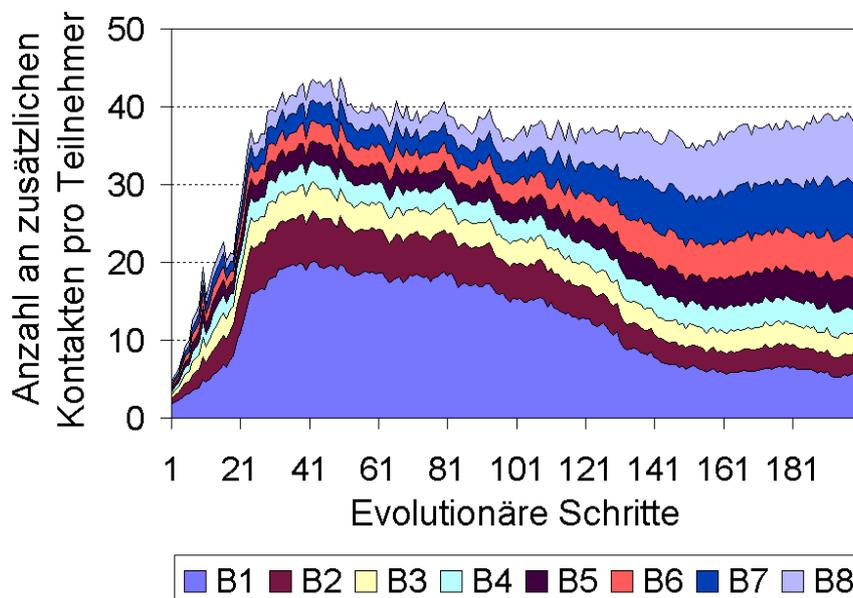


Abbildung 6.3: Kontaktwahl-Strategie im evolutionären Prozess

KAPITEL 6. EVOLUTION DER INTERAKTIONS- UND KONTAKTWAHL

könnte sein, dass Spieler mit zusätzlichen Kontakten nicht von ihren festen Kontakten abhängig sind und dass daher der Wert kaum eine Rolle spielt. Beim Weiterleiten gewichten die Teilnehmer die Lage eines Kontaktes stärker als seinen Kooperationsgrad (*locationWeight* $\approx 0,95$). Dies ist natürlich, da sie bereits kooperative Kontakte gewählt haben.

Bei der Kontaktwahl gibt es die folgenden Ergebnisse (vergleiche Abbildung 6.3): Die Spieler wählen 40 zusätzliche Kontakte. Am Anfang des dynamischen Prozesses bevorzugen die Spieler zusätzliche Kontakte, die ihnen nahe gelegen sind: Im 37. evolutionären Schritt haben die Spieler im Durchschnitt 20 zusätzliche Kontakte nahe an ihrer eigenen Position (*B1*), jedoch nur noch fünf zusätzliche Kontakte im 200. evolutionären Schritt. Am Ende der Messung verteilen die Spieler ihre zusätzlichen Kontakte fast gleichmäßig über den Adressraum: Klassen *B1* bis *B8* haben fast die gleiche Anzahl an Kontakten. Einzige Ausnahme ist, dass die Teilnehmer sehr nahe gelegene Kontakte und sehr weit entfernte Kontakte etwas mehr bevorzugen als andere Kontakte.

Ergebnisse – Variation der Unterhaltskosten für zusätzliche Kontakte

In diesem Abschnitt wird der Einfluss der Unterhaltskosten für zusätzliche Kontakte auf die Dynamik des Systems beziehungsweise auf das kooperative Gleichgewicht untersucht. Die Kosten für den Unterhalt wurden in zehn Schritten von $c = 0$; $c = 0,5$; ... bis $c = 4,5$ variiert. Alle anderen exogenen Parameter entsprechen den oben eingeführten Basiswerten. Die Ergebnisse sind wie folgt:

1. Je höher die Unterhaltskosten für zusätzliche Kontakte sind, desto weniger zusätzliche Kontakte werden ausgewählt. Während ein Spieler viele zusätzliche Kontakte wählt, wenn die Unterhaltskosten gering sind ($c = 0,5 \Rightarrow \max C > 60$), hat ein Spieler nur etwa 25 zusätzliche Kontakte bei Kosten $c = 2$ und nur zehn zusätzliche Kontakte für Kosten $c = 4$.
2. Je weniger zusätzliche Kontakte kosten, desto höherer ist die Erwartung an den Kooperationsgrad der zusätzlichen Kontakte. Schwellenwert *additionalCC* ist um die 83% für $c = 0$ und um die 90% für $c = 0,5$. Für höhere Unterhaltskosten für zusätzliche Kontakte fällt der Schwellenwert auf 81% ($c = 2$), 72% ($c = 3$), und auf 70% für $c = 4$.

Ergebnisse – Einfluss der Qualität des Reputationssystems Die Qualität der Reputationsinformationen über den Kooperationsgrad der Teilnehmer ist, wie in Kapitel 5 gezeigt, bedeutend. In diesem Paragraph wird der Fehler in den Reputationsinformationen *reportingError* variiert. Im Gegensatz zu Kapitel 5, in welchem eine Auskunft über den Kooperationsgrad eines Kontaktes entweder korrekt oder falsch war, wird die Information in diesem Kapitel proportional zum Parameter *reportingError* verfälscht (vergleiche Algorithmus 9). Algorithmus 9 funktio-

6.2. ERGEBNISSE – KOOPERATIVES GLEICHGEWICHT

Algorithmus 9: istKooperativ(Teilnehmer i , Schwellenwert t)

Input : Teilnehmer i , Schwellenwert t

Output : {WAHR, FALSCH}

- 1 Gleitkommazahl wahrerKooperationsgrad = $C(i)$;
 - 2 Gleitkommazahl offset = $((\text{nimmZufallswertAusIntervall}[0,1] - \text{wahrerKooperationsgrad}) \cdot \text{reportingError})$;
 - 3 Gleitkommazahl verfälschterKooperationsgrad = wahrerKooperationsgrad + offset;
 - 4 **return** verfälschterKooperationsgrad $\geq t$;
-

niert wie folgt: Es wird ein Offset, der maximal um *reportingError* Prozent vom originalen Kooperationsgrad abweicht, auf den wahren Kooperationsgrad aufgeschlagen oder abgezogen (Zeilen 1-3).

Der Fehler des Reputationssystems *reportingError* wird in fünf Schritten von 0%, 5%, ... bis 20% variiert. Weitere exogene Parameter entsprechen den oben eingeführten Basiswerten. Die Ergebnisse der Simulationen sind wie folgt:

- Ist der Fehler im Reputationssystem zu groß, dann bricht das System zusammen und alle Teilnehmer steigen aus dem System aus, d.h. sie haben Strategievektor S_i (vgl. Abschnitt 5.3). Dies passiert insbesondere dann, wenn der Fehler größer als 15% ist. (Im Folgenden werden nur noch Ergebnisse aus Systemen präsentiert, die nicht zusammenbrechen.)
- Je höher der Fehler in den Reputationsinformationen, desto mehr zusätzliche Kontakte wählen die Teilnehmer. Dies entspricht der Erwartung, da durch die zusätzlichen Kontakte die Pfadlängen der Anfragen sinken und dadurch die Wahrscheinlichkeit erhöht wird, dass eine Anfrage erfolgreich verarbeitet wird.
- Je höher der Fehler im Reputationssystem, desto mehr bevorzugen die Teilnehmer Kontakte, die ihnen nahe gelegen sind, sowie Kontakte die möglichst weit von ihnen entfernt sind. Der erste Effekt ist stärker als der zweite, d.h. die Spieler wählen mehr Kontakte in ihrer Umgebung als solche, die weit entfernt sind.
- Der Kooperationsgrad sinkt (langsam) umso mehr, je höher der Fehler in den Reputationsinformationen ist, bis das System zusammenbricht.

Ein niedriger Fehler in den Reputationsinformationen über den Kooperationsgrad der Teilnehmer ist essentiell für die Effizienz des Systems. Je geringer der Fehler ist, desto mehr Spieler kooperieren und desto weniger zusätzliche Kontakte sind nötig.

KAPITEL 6. EVOLUTION DER INTERAKTIONS- UND KONTAKTWAHL

Ergebnisse – Einfluss der Diskontierung Der letzte exogene Parameter, der in diesem Kapitel variiert wird, ist der Diskontierungsfaktor δ . Je geringer er ist, desto geringer ist der Nutzen einer Antwort pro Weiterleiter der zugehörigen Anfrage. Der Diskontierungsfaktor wird wie folgt variiert: $\delta = 0,5$; $\delta = 0,55$; $\delta = 0,6$; ...; $\delta = 0,95$ und $\delta = 1$. Alle anderen exogenen Parameter entsprechen den oben eingeführten Basiswerten. Die Ergebnisse sind wie folgt:

- Niedrige Diskontierungsfaktoren (die aber jedoch größer als $\delta = 50\%$ betragen) führen nicht dazu, dass das ineffiziente Netzwerk entsteht. Sie führen also nicht dazu, dass das System zusammenbricht. Dies ist ein überraschendes Ergebnis, da ein Diskontierungsfaktor von z.B. 50% bedeutet, dass eine Anfrage mit jedem Weiterleitungsschritt 50% ihres Wertes verliert.
- Ein geringer Diskontierungsfaktor führt zu mehr zusätzlichen Kontakten. Dies ist natürlich, da mehr zusätzliche Kontakte weniger Weiterleitungsschritte bedeuten und da daher der Nutzen einer Antwort weniger stark sinkt.
- In einem System mit einem Diskontierungsfaktor von 100%, d.h. der Wert einer Anfrage hängt nicht von der Anzahl an Weiterleitungsschritten ab, wählen die Spieler kaum noch zusätzliche Kontakte ($maxC \approx 5$). Dies ist ebenfalls natürlich, da es (in Bezug auf den Diskontierungsfaktor) unwichtig ist, wie viele Weiterleitungsschritte eine Anfrage benötigt, bis sie ihr Ziel erreicht. Es sei jedoch angemerkt, dass durch unkooperative Teilnehmer immer noch Anfragen verloren gehen können und dass es daher immer noch sinnvoll ist, einige zusätzliche Kontakte zu besitzen.

Zusammenfassung Die drei variierten exogenen Parameter beeinflussen das kooperative Gleichgewicht auf unterschiedliche Weise. Die Kosten für den Unterhalt zusätzlicher Kontakte beeinflusst die Anzahl zusätzlicher Kontakte sowie die Schwellenwerte der Teilnehmer. Der Fehler in den Reputationsinformationen beeinflusst den Kooperationsgrad der Teilnehmer und ist essentiell für die Effizienz des Systems. Der Diskontierungsfaktor hat den geringsten Effekt auf die evolutionäre Stabilität des Gleichgewichts. Dennoch beeinflusst er die durchschnittliche Anzahl an zusätzlichen Kontakten pro Teilnehmer. Exogene Parameter können also – wenn möglich – genutzt werden, um die Dynamik der Strategien zu beeinflussen.

6.3 Das Entstehen kooperativer Systeme

Im Gegensatz zum vorherigen Abschnitt werden in diesem keine initialen Strategien angenommen: Jeder Spieler hat ein zufälliges Genom. Lediglich durch den evolutionären Prozess werden erfolgreiche Strategien, also solche, die zu hohen Auszah-

6.3. DAS ENTSTEHEN KOOPERATIVER SYSTEME

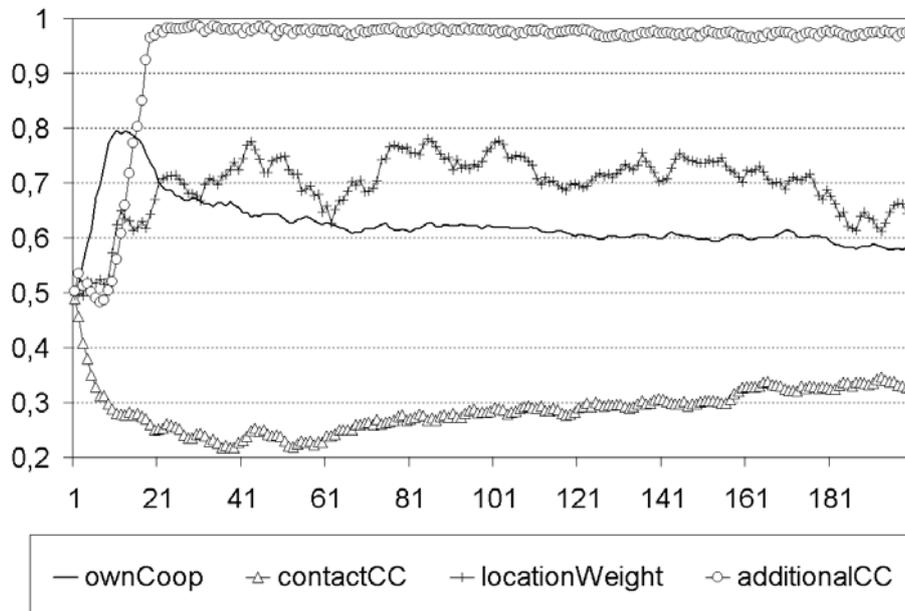


Abbildung 6.4: Erfolgreiche Interaktionswahl-Strategien aus zufälligen Strategien

lungen führen, ausgewählt und stärker verbreitet. Der angewendete Prozess ist dabei derselbe, wie der in Abschnitt 6.1, d.h. im Versuchsaufbau, beschriebene.

Die wissenschaftliche Fragestellung, die es zu klären gilt, ist, ob Kooperation aus zufälligen Strategien entstehen kann. Unter „Kooperation“ ist ein System-Zustand zu verstehen, bei dem die Teilnehmer dadurch positive Auszahlungen erhalten, dass sie mit anderen Teilnehmern zusammenarbeiten. Für wesentlich trivialere Systeme, z.B. dem wiederholten Gefangenendilemma, ist bekannt, dass Kooperation entstehen kann [Axe06]. Im Gegensatz zu [Axe06] werden allerdings hier keine Strategien, die von Menschen geschrieben wurden, verwendet, sondern solche, die aus evolutionären Strategieanpassungen aus (initial) zufälligen Strategien entstehen.

Die Antwort auf die Frage ist bedeutend, da sie eine Erklärung liefert, wie menschliche Kooperation in Systemen der Interaktions- und Kontaktwahl entstehen kann. Es sei angemerkt, dass Altruismus bei der Erklärung dieses Effekts keine Rolle spielt, sondern dass die Spieler sich egoistisch verhalten. Wenn die Antwort auf die Frage „ja“ lautet, dann entsteht Kooperation trotz – oder gerade wegen – den Interessen der einzelnen Spieler. Es sei des Weiteren angemerkt, dass schon eine einzige Simulation, die zu einem kooperativen System führt, ausreichend ist, um zu zeigen, dass Kooperation entstehen kann.

Die kurze Antwort auf die oben gestellte Frage ist „ja“. Dennoch sind einige Unterschiede zu den in Abschnitt 6.2 beobachteten Ergebnissen zu nennen, die im Folgenden beschrieben werden.

KAPITEL 6. EVOLUTION DER INTERAKTIONS- UND KONTAKTWAHL

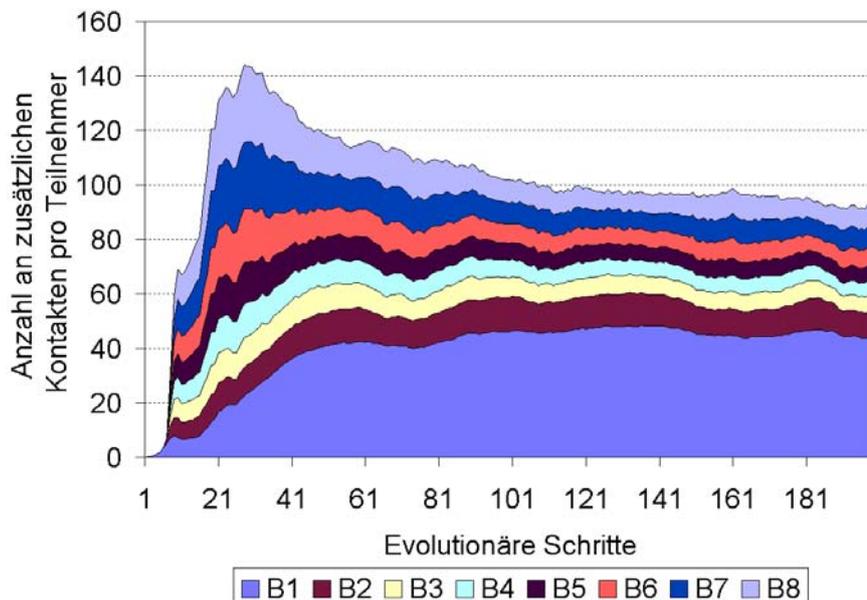


Abbildung 6.5: Erfolgreiche Kontaktwahl-Strategien aus zufälligen Strategien

Ergebnisse – Kooperation Die Spieler kooperieren, zeigen jedoch einen geringeren Kooperationsgrad als im kooperativen Gleichgewicht. Abbildung 6.4 zeigt die Interaktionswahl-Strategie für die Basiswerte. Im Gegensatz zum vorherigen Abschnitt (d.h. Abschnitt 6.2) ist der durchschnittliche Kooperationsgrad geringer, d.h. um die 60%. Dies erscheint ein mäßiger Wert zu sein, es sei jedoch angemerkt, dass das System nicht zusammenbricht, d.h. dass die Teilnehmer einen Nutzen aus der gegenseitigen Zusammenarbeit ziehen können. Die Erwartung der Teilnehmer an den Kooperationsgrad ihrer Kontakte ist sehr hoch, d.h. fast 100% (vergleiche Parameter *additionalCC*). Ein interessantes Ergebnis ist, dass die Stabilität des Systems wesentlich robuster gegen Fehler in den Reputationssystemen ist als vergleichbare Systeme in Abschnitt 6.2: Systeme mit einem Fehler in den Reputationssystemen von bis zu 60% brechen nicht zusammen. Eine Erklärung für die Robustheit ist, dass die Strategien der Teilnehmer wesentlich heterogener sind als im kooperativen Gleichgewicht.

Ergebnisse – Kontaktwahl Verglichen mit dem kooperativen Gleichgewicht wählen die Spieler deutlich mehr zusätzliche Kontakte. Abbildung 6.5 zeigt die durchschnittliche Kontaktwahl für die Basiswerte: Die Spieler wählen um die 100 zusätzlichen Kontakte. Dies macht Sinn, da der Kooperationsgrad der Teilnehmer geringer ist als im effizienten System. Die Verteilung der Kontakte ist wie folgt: Die Spieler bevorzugen Teilnehmer als zusätzliche Kontakte, die ihnen nahe gelegen sind. Sie erreichen fast eine Kleinbergverteilung ihrer zusätzlichen Kontakte.

Ergebnisse – Diskontierung Der Diskontierungsfaktor spielt im Gegensatz zu den Ergebnissen in Abschnitt 6.2 eine wesentlich größere Rolle: Systeme mit einem Diskontierungsfaktor, der kleiner als 85% ist, brechen zusammen. Es scheint für die Entstehung von Kooperation wichtig zu sein, dass der Gewinn beim Erhalt einer Anfrage hinreichend groß ist. Wenn der Diskontierungsfaktor zu gering ist, dann ist dies nicht der Fall. Während der Diskontierungsfaktor die Interaktionswahl-Strategien beeinflusst, scheint er überraschenderweise keinen Einfluss auf die Kontaktwahl zu haben. Die Untersuchung möglicher Ursachen mag ein interessanter Fortsetzungspunkt dieser Arbeit sein.

6.4 Fazit

Die Simulation evolutionärer Strategien zeigt, dass die in Kapitel 5 identifizierten Gleichgewichte stabil sind. Des Weiteren kann aus der Simulation evolutionärer Verbesserungen zufälliger Strategien geschlossen werden, dass Kooperation aus den Eigeninteressen der Teilnehmer in der Interaktions- und Kontaktwahl entstehen kann. Dies ist ein wichtiges Ergebnis, da es eine Erklärung für Kooperation in Systemen mit Interaktions- und Kontaktwahl liefert.

KAPITEL 6. EVOLUTION DER INTERAKTIONS- UND KONTAKTWAHL

7 Das menschliche Verhalten in der Interaktions- und Kontaktwahl

Teilnehmer verteilter Systeme mit Interaktions- und Kontaktwahl sind entweder Menschen oder Software-Artefakte, die von Menschen kontrolliert werden. Menschen haben also einen großen Einfluss auf die Strategien, die das Verhalten solcher Systeme bestimmen. Es ist daher wichtig zu untersuchen, wie menschliche Teilnehmer Netze formen und wie sie sich in ihren Interaktionen verhalten. Da unklar ist, ob theoretische Vorhersagen in allen Situationen das Systemverhalten vorhersagen können, ist die Beobachtung von menschlichem Verhalten unumgänglich. Dieses Kapitel beantwortet die Frage, ob menschliche Teilnehmer, die in einem Netzwerk interagieren und Kontakte wählen, Strategien finden und nutzen, die zu effizienten Systemen führen. Des Weiteren soll geklärt werden, welche Charakteristiken die entstehenden Netzwerke haben und wie sie durch exogene Parameter beeinflusst werden. Um diese Fragen zu beantworten, wird menschliches Verhalten in einem Laborexperiment mit insgesamt 120 menschlichen Teilnehmern untersucht, bei dem diese die Netzwerkstruktur formen und mit ihren Kontakten interagieren. Das wichtigste Ergebnis der Untersuchung ist, dass Kontaktwahl den Kooperationsgrad erhöht und dass die Teilnehmer Netzwerke erzeugen, die – wie im Verlauf des Kapitels erörtert wird – relativ gesehen optimal sind. Es kann also aus den Ergebnissen abgeleitet werden, dass Kontaktwahl das Gemeinwohl im System fördert. Dies ist ein wichtiges Ergebnis, da Systementwickler den Nutzen ihrer Systeme allein dadurch anheben können, dass sie ihre Teilnehmer selbstständig Kontakte wählen lassen.

7.1 Einleitung

Knoten verteilter Systeme sind autonom und werden durch Menschen kontrolliert, sei es, weil sie Menschen sind, oder weil sie von Menschen ausgewählte beziehungsweise manipulierte Programme sind. Da unklar ist, ob (a) Menschen in der Interaktions- und Kontaktwahl effiziente Strategien finden und ob sie (b) diese Strategien nutzen, ist es nötig, menschliches Verhalten in der Kontakt- und Interaktionswahl zu untersuchen. So zeigt z.B. [CT94], dass menschliche Teilnehmer sich oft nicht so verhalten, wie sich ein Agent verhalten würde, der seine Auszahlung maximiert. Des Weiteren

KAPITEL 7. DAS MENSCHLICHE VERHALTEN IN DER INTERAKTIONS- UND KONTAKTWAHL

existieren Arbeiten, die nicht das menschliche Verhalten berücksichtigen und damit zu nicht effizienten Systemen führen, die hohe Folgekosten nach sich ziehen. So zeigt z.B. [ZA07], dass die Mechanismen, die das Gemeinwohl der Peer-to-Peer Systeme EMule und BitTorrent erhöhen sollen, nicht den erwarteten Nutzen bringen: Der Kooperationsgrad in den Systemen ist niedriger als es die Theorie vorhersagen würde.

Dieses Kapitel untersucht die folgenden wissenschaftlichen Fragen:

- Welche Interaktionswahl- und Kontaktwahl-Strategien nutzen Menschen? Im Detail soll untersucht werden, wie das vergangene Verhalten anderer Teilnehmer das Verhalten eines Teilnehmers beeinflusst. Von einfachen Systemen ist bekannt, dass das Kontaktwahl-Verhalten einen Einfluss auf die Strategien der Teilnehmer hat [BV04, GVR05]. Der Einfluss auf ein realistischeres Szenario, d.h. auf die Interaktions- und Kontaktwahl, soll hier untersucht werden.
- Wie beeinflussen exogene Parameter (z.B. die Unterhaltskosten für Kontakte) das Verhalten der Teilnehmer und damit die Netzwerkstruktur?
- Zu welchen Netzwerkstrukturen führt das Verhalten der Teilnehmer? Sind diese Strukturen effizient? Wie optimal sind sie an exogene Parameter angepasst?

Die letzte Fragestellung ist insbesondere aus der Perspektive eines Informatikers interessant: Wenn effiziente Netzwerke entstehen, d.h. wenn die Summe der Auszahlungen maximal ist, oder anschaulicher, wenn die Teilnehmer kooperieren und Anfragen in wenigen Weiterleitungsschritten bearbeitet werden, dann sind Anreiz-basierte Mechanismen oder Strategien zur Steigerung der Systemeffizienz nicht nötig. Andernfalls müssen solche Mechanismen oder Strategien entwickelt werden. Im Gegensatz zu Studien, die das Verhalten in existierenden Systemen beobachten, sollen die hier durchgeführten Experimente es möglich machen, menschliches Verhalten vorherzusagen, bevor entsprechende Systeme entworfen werden, um fehlgeleitete Entwicklungen zu umgehen.

Das Finden von Antworten auf die obigen Fragen ist nicht trivial. Es kann nicht durch das Studium existierender Systeme erreicht werden. Die Argumentation ist wie folgt: Das Verhalten von menschlichen Teilnehmern ist in der Regel abhängig von äußeren Beeinflussungen, z.B. dem Einfluss von Kosten für bestimmte Aktionen. Diese sind oft unbekannt und jenseits der Kontrolle des Experimentators. Des Weiteren wird in den in diesem Kapitel beschriebenen Experimenten der Einfluss von störendem Kontext verhindert. Z.B. wird darauf verzichtet, Anfragen einer bestimmten Applikation oder aus einer bestimmten Domäne zu verwenden. Ein anderes Problem existierender Systeme ist, dass in diesen bereits eine Netzwerkstruktur existiert und dass die Knoten in diesen Netzwerken eine unterschiedliche Bedeutung beziehungsweise ein unterschiedliches Gewicht besitzen. Um diesem Einfluss auf das menschliche Verhalten auszuschließen, werden die Experimente dieses Kapitels nicht auf einer Struktur, die den Knoten ein unterschiedliches Gewicht verleiht, durchgeführt, sondern auf einer initialen Struktur, in der alle Teilnehmer gleichberechtigt sind.

Vorgehen Da das menschliche Verhalten der wesentliche Faktor in den Strategien eines Systems ist, werden in diesem Kapitel Experimente mit menschlichen Teilnehmern – kurz *Subjekte*¹ genannt – durchgeführt. Es werden 20 Experimente mit zehn Gruppen à 12 Subjekten vorgenommen. Insgesamt nehmen 120 Subjekte an den Experimenten teil. Die Subjekte können ihre Interaktionswahl-Strategien und ihre Kontaktwahl-Strategien frei wählen. Dabei erzeugt verrichtete Arbeit, wie das Erzeugen, Senden, Weiterleiten und Beantworten von Anfragen sowie der Unterhalt von Kontakten Kosten, während der Erhalt einer Antwort zu Einnahmen führt.

Beiträge Die durchgeführten Experimente führen zu den folgenden Erkenntnissen:

- Zwischen der Interaktionswahl und der Kontaktwahl bestehen Abhängigkeiten: Die Kontaktwahl erhöht den Kooperationsgrad im System verglichen mit Systemen, die keine Kontaktwahl der Teilnehmer zulassen.
- Exogene Parameter beeinflussen die Kontaktwahl und damit die Bildung von Netzwerken: Subjekte wählen umso mehr Kontakte, je weniger sie dafür bezahlen müssen.
- Kontakte, die als unkooperativ (auf Anfrageebene) eingeschätzt werden, werden von den Subjekten fallengelassen.
- Das Verhalten der Teilnehmer in der Kontaktwahl führt zu Netzwerken, die relativ gesehen optimal sind (was im Verlauf des Kapitels genauer erörtert werden wird). Kontaktwahl ist also ein wichtiger Schritt, um das Gemeinwohl im System zu erhöhen.

Die Ergebnisse decken sich mit den Ergebnissen der Untersuchungen in Kapitel 6 und sind darüber hinaus bedeutend: Viele technische verteilte Systeme erlauben es ihren Teilnehmern nicht, Kontakte nach eigenen Interessen zu wählen [LCP⁺05]. Z.B. ist die Topologie vieler strukturierter Peer-to-Peer Systeme fest vorgegeben. Dennoch scheint das Erlauben der Kontaktwahl ein einfacher (also kostengünstiger) Weg zu sein, um das Gemeinwohl zu steigern, d.h. die Auszahlung aller Teilnehmer und damit den Nutzen zu erhöhen. Da die Akzeptanz eines Systems, das nicht effizient ist, geringer sein sollte als die Akzeptanz eines Systems mit hohem Gemeinwohl beziehungsweise die Akzeptanz eines effizienten Systems, wird die Bedeutung der Ergebnisse deutlich.

Gliederung des Kapitels Abschnitt 7.2 erläutert Anpassungen des Modells aus Kapitel 3 auf die Experimente. Abschnitt 7.3 motiviert und erläutert Hypothesen, die

¹Subjekte, Spieler und Teilnehmer werden synonym verwendet.

KAPITEL 7. DAS MENSCHLICHE VERHALTEN IN DER INTERAKTIONS- UND KONTAKTWAHL

in den Experimenten getestet werden. Es folgen eine Beschreibung des Experimentaufbaus (Abschnitt 7.4) und die Ergebnisse der Experimentreihen (Abschnitt 7.5). Abschnitt 7.6 schließt das Kapitel mit einem Fazit ab.

7.2 Definitionen und Anpassungen

Bei der Untersuchung von menschlichem Verhalten und der Analyse menschlicher Strategien ist es wichtig, darauf zu achten, dass die Teilnehmer das zugrunde liegende Problem beziehungsweise die Auszahlungsfunktion verstehen. Andernfalls kann nicht erwartet werden, dass die Teilnehmer sinnvolle Strategien hervorbringen. Es werden daher in diesem Kapitel einige Vereinfachungen des in Kapitel 3 vorgestellten Modells genutzt und im Folgenden besprochen.

Auszahlung Die Auszahlungsfunktion entspricht der in Kapitel 3 vorgestellten. Zur Vereinfachung und um der durch die Subjekte vorgenommenen Diskontierung nicht vorwegzugreifen, wurde auf eine Diskontierung verzichtet, d.h. in den Experimenten verringert sich der Wert einer Antwort nicht mit der Anzahl an Weiterleitungsschritten, die die zugehörige Anfrage benötigt. Es gilt also $\delta = 1$. Die Auszahlungsfunktion ist demnach wie folgt:

$$\text{Auszahlung}(i) = a \cdot |A_i| - q \cdot |R_i| - f \cdot |F_i| - w \cdot |W_i| - c \cdot r(i)$$

In den Experimenten sind die Auszahlungskoeffizienten wie folgt:

- Das Erzeugen und Senden einer Anfrage kostet zwei Punkte ($q = 2$).
- Das Weiterleiten einer Anfrage kostet einen Punkt ($f = 1$).
- Fünf Punkte werden einem Subjekt beim Beantworten einer Anfrage abgezogen ($w = 5$).
- Der Erhalt einer Antwort auf eine Anfrage führt zu einem Einkommen von 20 Punkten ($a = 20$).
- Zur Klärung der in Abschnitt 7.1 genannten Fragestellungen werden vier Experimentreihen durchgeführt. Diese unterscheiden sich in den Unterhaltskosten für zusätzliche Kontakte. Je nach Experimentreihe kostet der Unterhalt eines zusätzlichen Kontakts pro Runde null, ein, drei oder fünf Punkte ($c = 0, 1, 3, 5$).

Tabelle 7.1 fasst die Werte zusammen.

Aktion/Ereignis	Auszahlungs- koeffizient	Wert im Experiment
Erzeugen einer Anfrage	q	2
Weiterleiten einer Anfrage	f	1
Beantworten einer Anfrage	w	5
Erhalt einer Antwort	a	20
Unterhalt von zusätzlichen Kontakten	c	0, 1, 3, 5

Tabelle 7.1: Auszahlungskoeffizienten in den Experimenten

7.3 Hypothesen

Simulationen und formale Modelle können auf Annahmen basieren, die das menschliche Verhalten nicht immer treffend wiedergeben.

Beispiel: Die Spieltheorie sagt für das (nicht wiederholte) Gefangenendilemma voraus, dass die Gefangenen sich gegenseitig verraten also dass sie defektieren. Dennoch ist bekannt, dass menschliche Teilnehmer nicht immer dieses Verhalten zeigen (vergleiche [CDR96]). ■

In diesem Kapitel werden unter anderem theoretische Modelle genutzt, um Erwartungen an das Verhalten menschlicher Teilnehmer zu formulieren, d.h. um Hypothesen zu formulieren. Diese werden im Verlauf des Kapitels auf ihre Übereinstimmung mit dem menschlichen Verhalten überprüft.

Kontaktwahl-Strategie Es ist unklar, wie die Tatsache, dass Teilnehmer Kontakte wählen können, das Interaktionsverhalten beeinflusst. Insbesondere ist unklar, wie der Kooperationsgrad durch die Kontaktwahl der Teilnehmer beeinflusst wird. Grundsätzlich existieren zwei sich gegenseitig beeinflussende und entgegengesetzt wirkende Effekte.

Auf der einen Seite ist zu erwarten, dass die Kontaktwahl den Kooperationsgrad senkt: In einer festen Netzwerkstruktur kann ein Teilnehmer lernen, wie sich seine Kontakte verhalten und entsprechend ihrem Kooperationsgrad reagieren. In einer zusätzlichen Kontaktstruktur mit sich eventuell schnell und oft wechselnden zusätzlichen Kontakten mag es schwerer sein, den Kooperationsgrad zu bestimmen. Kontaktwahl könnte zu einer besonderen Form des Weißwaschens [FCSS05] führen. Diese würde wie folgt aussehen: Ein Teilnehmer bearbeitet die Anfragen seiner zusätzlichen Kontakte nicht und wechselt diese so schnell, dass sie nicht auf sein Verhalten reagieren können (da sie seine Strategie erst durchschauen, wenn er sie gegen andere Teilnehmer austauscht.)

Auf der anderen Seite ist es denkbar, dass Kontaktwahl die Kooperation im System erhöht. In einer festen Netzwerkstruktur muss ein Teilnehmer unabhängig davon

KAPITEL 7. DAS MENSCHLICHE VERHALTEN IN DER INTERAKTIONS- UND KONTAKTWAHL

mit seinen Kontakten leben, ob sie kooperieren oder nicht. Intuitiv sind Netzwerke, in denen Kontakte gewählt werden können, aus zwei Gründen kooperativer als solche, in denen das nicht möglich ist: Erstens, ein Teilnehmer hat einen natürlichen Anreiz, kooperative zusätzliche Kontakte auszuwählen, da sonst der Nutzen des Systems für ihn sinken würde. Zweitens, Freifahrer auf Anfrageebene antizipieren dieses Verhalten und fürchten, dass sie Schwierigkeiten haben, sich zu vernetzen, wenn kooperative Teilnehmer ausschließlich kooperative zusätzliche Kontakte wählen. Dies könnte die Strategie des Freifahrens unwahrscheinlich werden lassen. Es folgt die Hypothese:

Hypothese 7.3.1 *Im Durchschnitt wird der Kooperationsgrad durch die Kontaktwahl erhöht.*

Position vs. Kooperationsgrad Jeder Kontakt hat eine bestimmte Position im Adressraum und einen bestimmten Kooperationsgrad. Ein Teilnehmer kann einen Kontakt wählen, (a) da er eine vielversprechende Position im Adressraum hat, oder (b) weil der Kontakt einen hohen Kooperationsgrad hat. Beide Punkte haben unterschiedliche Vorteile. Punkt (a) verringert die Anzahl an Weiterleitungsschritten und erhöht damit die Wahrscheinlichkeit, dass Anfragen beantwortet werden. Punkt (b) ist aussichtsreich, da ein Teilnehmer annehmen kann, dass ein Kontakt, der in der Vergangenheit kooperativ war, dies auch in der Zukunft sein wird. Es ist unklar, welches Motiv stärker ist. Simulationen von Agenten, die ihre Auszahlung maximieren, haben gezeigt, dass Agenten Kontakte mit dem höchsten Kooperationsgrad wählen [ZL06]. Diese Arbeit testet daher die folgende Hypothese:

Hypothese 7.3.2 *Ein Subjekt wählt in der Anfragebearbeitung Kontakte nach dem Kooperationsgrad aus.*

Fallenlassen von Freifahrern Zwei Kontakte, die miteinander kooperieren oder nicht kooperieren, sind in der gleichen Situation wie die Gefangenen im Gefangenendilemma. Kooperieren beide, so haben sie beide Ausgaben und Einnahmen dadurch. Kooperiert nur einer, so hat dieser Einnahmen und sein Kontakt Ausgaben. Sind beide unkooperativ, so haben beide nur Kosten, bzw. keine Auszahlung, wenn sie den anderen Kontakt fallenlassen. Insgesamt entspricht dies dem Gefangenendilemma.

Axelrod [Axe06] hat die Dynamik von Strategien im wiederholten Gefangenendilemma untersucht. Dabei sterben die Strategien mit der geringsten Auszahlung aus, beziehungsweise imitieren erfolgreichere Strategien. Das Ergebnis der mittels Simulationen durchgeführten Analyse ist, dass die sogenannte Tit-for-Tat Strategie (Wie-Du-mir-so-ich-Dir Strategie) am erfolgreichsten ist. Erfolgreich bedeutet in diesem Kontext, dass die Strategie die anderen Strategien überlebte. Ein Teilnehmer, der eine Tit-for-Tat Strategie nutzt, verhält sich so, wie es sein Spielpartner in der vor-

hergehenden Runde getan hat. Aus diesem Verhalten wird die folgende Hypothese abgeleitet:

Hypothese 7.3.3 *Teilnehmer lassen zusätzliche Kontakte fallen, die sich in der Vergangenheit unkooperativ verhalten haben.*

Kosten für zusätzliche Kontakte Goyal und Vega-Redondo [GVR05] zeigen, dass die Spieler eines Koordinationsspiels, die ihre Kontakte wählen, ihr Verhalten an die Kosten für zusätzliche Kontakte anpassen. Ein Koordinationsspiel ist nicht mit der Definition der Interaktionswahl in dieser Arbeit zu vergleichen. Es bezeichnet in [GVR05] ein Spiel, bei dem die Spieler, die die gleiche Aktion wählen wie ihr Spielpartner, eine hohe Auszahlung erhalten und Spieler, die gegensätzliche Aktionen wählen, geringe Auszahlungen. Dabei führen geringe Kosten für zusätzliche Kontakte in [GVR05] nicht zu effizienten Spielzuständen, während hohe Kosten zu effizienten Systemen führen. Eine Erklärung ist, dass die Spieler ihre Kontakte sorgfältiger wählen, wenn die Kosten hoch sind. Obwohl Koordinationsspiele und die Interaktionswahl verschieden sind, kann aus den Ergebnissen in [GVR05] abgeleitet werden, dass die Kontaktwahl sorgfältiger durchgeführt wird, wenn die Kosten für den Unterhalt zusätzlicher Kontakte hoch sind. Eine Erwartung an die menschlichen Teilnehmer ist, dass sie unkooperative zusätzliche Kontakte eher fallenlassen, wenn diese hohe Unterhaltskosten verursachen, als wenn dies nicht der Fall ist. Da die Kontakte dieses Verhalten wahrscheinlich voraussehen, ist anzunehmen, dass sie sich kooperativ verhalten, wenn sie auf zusätzliche Kontakte angewiesen sind. Die folgende Hypothese kann abgeleitet werden:

Hypothese 7.3.4 *Subjekte verhalten sich kooperativer gegenüber zusätzlichen Kontakten, wenn die Unterhaltskosten für zusätzliche Kontakte hoch sind.*

Anzahl an Kontakten Der Unterhalt von Kontakten verursacht Kosten. Subjekte, die bereits kooperative Kontakte besitzen und die den zu erwarteten Nutzen eines neuen zusätzlichen Kontakts zu den Unterhaltskosten ins Verhältnis setzen, werden die Anzahl ihrer zusätzlichen Kontakte limitieren. Dies ist damit begründet, dass sobald der Nutzen eines neuen zusätzlichen Kontakts geringer ist als die Kosten, es nach der Maximierung der Auszahlung keinen Sinn mehr macht, diesen Kontakt zu wählen (ohne einen anderen fallenzulassen). Es kann daher erwartet werden, dass Subjekte umso weniger Kontakte wählen, desto höher die Unterhaltskosten für zusätzliche Kontakte sind.

Hypothese 7.3.5 *Die Anzahl zusätzlicher Kontakte pro Subjekt und die Kosten für den Unterhalt zusätzlicher Kontakte sind negativ korreliert.*

Falls diese Hypothese Bestand hat, dann hat sie die folgende Konsequenz: Da die Kosten für den Unterhalt zusätzlicher Kontakte von der jeweiligen Anwendung beziehungsweise dem jeweiligen Sozialen Netzwerk abhängen, ist es schwierig, für die

KAPITEL 7. DAS MENSCHLICHE VERHALTEN IN DER INTERAKTIONS- UND KONTAKTWAHL

Entwickler dieser Systeme vorherzusehen, ob sie effizient sind, beziehungsweise wie hoch das jeweilige Gemeinwohl ist. Falls das menschliche Verhalten zu effizienten oder fast effizienten Systemen führt, so ist es nicht nötig, Mechanismen zur Steigerung des Gemeinwohls, z.B. durch Manipulation der Unterhaltskosten für zusätzliche Kontakte, zu entwickeln und zu nutzen.

7.4 Experimentaufbau

Experimentreihen 120 Subjekte nehmen in zehn Gruppen mit je 12 Spielern an 20 Einzelexperimenten teil. Eine *Experimentreihe* (engl. treatment) besteht aus mehreren sich in Bezug auf exogene Parameter gleichenden Einzelexperimenten. Insgesamt gibt es vier Experimentreihen. Diese unterscheiden sich in den Kosten für den Unterhalt von zusätzlichen Kontakten: Ein Subjekt verliert 0, 1, 3 oder 5 Punkte pro zusätzlichem Kontakt pro Runde. Die vier entsprechenden Experimentreihen seien mit T_0 , T_1 , T_3 und T_5 bezeichnet. Jedes Subjekt nimmt an zwei Einzelexperimenten in zwei Experimentreihen teil. Von den 20 Einzelexperimenten wurden vier in der Experimentreihe T_0 durchgeführt. (Eine Experimentreihe schlug aufgrund von technischen Fehlern fehl.) Sechs Einzelexperimente wurden in der Experimentreihe T_1 durchgeführt, sieben in T_3 und drei in T_5 .

Es wurden Standardmethoden der experimentellen Forschung [Chr06] genutzt, um die Experimente durchzuführen und sicherzustellen, dass keine signifikanten Verfälschungen durch den Experimentaufbau induziert werden, wie beispielsweise Verfälschungen durch Experimentator-Effekte oder Verfälschungen durch unerwünschte Kommunikation der Teilnehmer. Während die softwareseitige Experimentumgebung durch Infrastruktur der Universität Karlsruhe (TH) (heute Karlsruher Institut für Technologie (KIT)) bereitgestellt wurde, wurden die Experimente am Lehrstuhl für Empirische Wirtschaftsforschung an der Otto-von-Guericke-Universität Magdeburg durchgeführt. So konnte der Experimentator Effekt reduziert werden. Des Weiteren wurden die Spieler voneinander isoliert, um Kommunikation zwischen ihnen zu verhindern.

Auszahlung Um ein natürliches Verhalten der Subjekte zu erreichen, werden ihnen finanzielle Anreize gegeben, die denen der Anfragebearbeitung in verteilten Systemen entsprechen. Die Spieler verdienen Geld relativ zu der Anzahl an Punkten, die sie im Experiment erreichen. Sie wurden dabei vor der Durchführung des Experiments informiert, wie die Auszahlungsfunktion gestaltet ist, also darüber, welche Kosten und welche Einnahmen durch Aktionen oder Ereignisse entstehen. Zum Vergleich siehe Tabelle 7.1.

Die Wahl der Höhe der Punkte soll approximativ die zu verrichtende Arbeit in einem verteilten System zur Anfragebearbeitung widerspiegeln. Einfache Aktionen wie das Weiterleiten von Anfragen oder das Erzeugen (und Senden) einer Anfrage

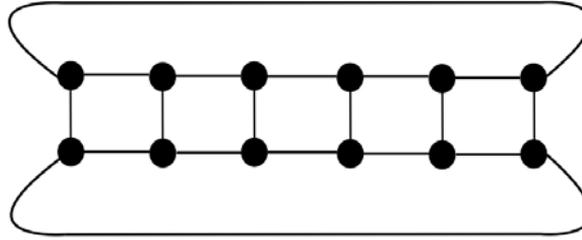


Abbildung 7.1: Initiale Netzwerkstruktur

sowie der Unterhalt eines zusätzlichen Kontaktes pro Runde kosten dabei je einen Punkt. Das Erzeugen einer Antwort ist in der Regel teurer und kostet daher fünf Punkte. Der Nutzen beim Erhalt einer Antwort auf eine Anfrage sollte höher sein als die Summe der Kosten der Anfragebearbeitung, andernfalls ist es im Sinne der Maximierung der Auszahlung optimal das System zu verlassen.

Des Weiteren hat jedes Subjekt einen initialen Punktestand von 100 Punkten. Es wurden 2 € pro 100 Punkte gezahlt. Für das Erscheinen beim Experiment wurden 3 € ausgezahlt.

Netzwerkstruktur Jedem Subjekt wurde zufällig eine Position im Netzwerk zugeweiht. Die initiale Netzwerkstruktur ist wie folgt: 12 Spieler formen einen 6×2 Torus, vergleiche Abbildung 7.1. Mit dieser initialen Struktur sollen zwei Dinge sichergestellt werden: Erstens, dass alle Spieler die gleichen Startbedingungen erhalten, und zweitens, dass die Struktur nicht durch den Ausfall oder das unkooperative Verhalten eines Spielers fragmentiert wird beziehungsweise stark eingeschränkt ist. Das erste Ziel ist dadurch begründet, dass in [RW90] gezeigt wird, dass die Position eines Knotens in einem Netzwerk seine Strategie beeinflusst. Es ist daher wichtig, eine einfache initiale Netzwerkstruktur zu wählen, die gleiche Bedingungen für alle Subjekte sicherstellt. Dabei muss die Struktur gewährleisten, dass sich alle Spieler kennenlernen können. Da eine Listen-Struktur oder eine Kreis-Struktur schnell fragmentiert – also in kleinere Teile zerfällt – wenn ein oder zwei Teilnehmer sich unkooperativ verhalten, werden in dieser Arbeit zwei Kreise verwendet, die miteinander verbunden sind (d.h. ein 6×2 Torus). Diese Struktur ist redundant und toleriert damit bis zu einem gewissen Grad unkooperative Teilnehmer. Die Subjekte kennen vom Netzwerk nur ihre Kontakte und potentielle Kontakte, also solche Spieler, mit denen sie bereits interagiert haben. D.h. die Spieler kennen das Netzwerk im Allgemeinen nicht.

Obwohl die Subjekte ihre Kontaktwahl-Strategie frei gestalten können, wird in dieser Arbeit ein besonderer Fokus auf bidirektionale Verbindungen gelegt. Unidirektionale Verbindungen, bei denen ein Subjekt (der Sender) Anfragen an seinen Kontakt (den Empfänger) senden kann, aber der Empfänger keine Anfragen an den Sender schicken kann, sind problematisch: Der Empfänger erhält keinen erkennbaren Nut-

KAPITEL 7. DAS MENSCHLICHE VERHALTEN IN DER INTERAKTIONS- UND KONTAKTWAHL

zen vom Sender und kann darüber hinaus nicht selbst überprüfen, ob der Sender grundsätzlich kooperativ ist, da er sein Verhalten nicht beobachten kann. Wenn der Empfänger sich reziprok gegenüber dem Sender verhalten will, so kann er dies nur, wenn er mit Dritten über den Kooperationsgrad des Senders kommuniziert. Auch wenn dies ein interessantes wissenschaftliches Thema ist, ist in diesem Experiment keine Kommunikation über den Kooperationsgrad von Dritten möglich. Vielmehr fokussiert dieses Kapitel auf die Analyse von Strategien, die auf eigenen Beobachtungen der Subjekte beruhen. Da unidirektionale Verbindungen unter den im Experiment gegebenen Voraussetzungen nicht stabil sind (da sie dem Empfänger nur Kosten verursachen), konzentriert sich die im Folgenden durchgeführte Analyse auf *stabile Verbindungen*. Eine stabile Verbindung ist eine bidirektionale und dauerhafte Verbindung, d.h. es gilt, dass ab einer beliebigen Runde x für alle Runden größer x aus $g(i, j)$ mit $i, j \in N$ auch $g(j, i)$ folgt. Des Weiteren können kaum strategisch sinnvolle Strategien, wie z.B. reziprokes Verhalten, erwartet werden, wenn die Subjekte nicht das Verhalten ihrer direkten Interaktionspartner beobachten können. Stabile bilaterale Verbindungen ermöglichen es beiden beteiligten Subjekten, das Verhalten des anderen Teilnehmers zu beobachten.

In [BV04] wird festgestellt, dass Subjekte eine geringere Anpassungsgeschwindigkeit in der Kontaktwahl zeigen als in der Interaktionswahl. Daher können die Subjekte während des Netzwerk-Spiels nur alle fünf Runden Kontakte wählen oder zusätzliche Kontakte fallenlassen, d.h. es gilt $k = 5$.

Terminierung der Einzelexperimente Ist die Anzahl an Runden eines Spiels bekannt, so kann durch Rückwärtsinduktion gezeigt werden, dass unkooperatives Verhalten der Auszahlungsmaximierung entspricht. Die Subjekte spielen daher das Netzwerk-Spiel als ein quasi-unendliches Spiel, d.h., kein Spieler weiß, wann das Spiel enden wird. Ab der 20. Runde wird durch die Experimentatoren ein zehneitiger Würfel gewürfelt (d.h. das Spiel bricht mit einer Wahrscheinlichkeit von 10% ab). Zeigt dieser eine Eins, so endet das Spiel, andernfalls startet eine weitere Runde. (In den Experimenten gab es eine Ausnahme: Ein Spiel endete aufgrund von technischen Schwierigkeiten nach der zehnten Runde. Dennoch gilt auch für dieses Spiel, dass den Subjekten nicht bekannt war, wann das Spiel endet. Daher wird auch dieses Spiel in der Evaluation herangezogen.)

Informationen Genau wie in einem realen System haben die Teilnehmer im Experiment nur lokal verfügbare Informationen – also Informationen über Aktionen und Ereignisse, die sie selbst beobachten können. So gibt es im Experiment z.B. kein Reputationssystem, bei welchem sich die Spieler über andere Spieler informieren können. Insbesondere weiß ein Spieler nicht, ob eine Anfrage, die er abgeschickt hat und für die er noch keine Antwort erhalten hat, fallengelassen wurde (und von wem) oder ob sie immer noch bearbeitet wird.

Ihr Name:		Finanzstatus																																																						
2		Kontostand:	144																																																					
		Ausgaben:	16																																																					
		Einnahmen:	60																																																					
Fremde Anfragen. (Beliebig viele pro Runde möglich.)		Kontakte																																																						
<table border="1"> <thead> <tr> <th>Von</th> <th>Beschreibung</th> </tr> </thead> <tbody> <tr><td>11</td><td>Weiterleiten</td></tr> <tr><td>23</td><td>Weiterleiten</td></tr> <tr><td>20</td><td>Weiterleiten</td></tr> <tr><td>9</td><td>Weiterleiten</td></tr> <tr><td>21</td><td>Weiterleiten</td></tr> <tr><td>18</td><td>Beantworten</td></tr> <tr><td>3</td><td>Weiterleiten</td></tr> <tr><td>3</td><td>Weiterleiten</td></tr> </tbody> </table>	Von	Beschreibung	11	Weiterleiten	23	Weiterleiten	20	Weiterleiten	9	Weiterleiten	21	Weiterleiten	18	Beantworten	3	Weiterleiten	3	Weiterleiten	<table border="1"> <thead> <tr> <th colspan="5">Anfragen (von Ihnen ausgegangen) (bei Ihnen eingegangen)</th> </tr> <tr> <th>Name</th> <th>gesendet</th> <th>bearbeitet</th> <th>empfangen</th> <th>bearbeitet</th> </tr> </thead> <tbody> <tr><td>3</td><td>2</td><td>50.00 %</td><td>3</td><td>0.00 %</td></tr> <tr><td>20</td><td>2</td><td>50.00 %</td><td>1</td><td>0.00 %</td></tr> <tr><td>22</td><td>0</td><td>0.00 %</td><td>0</td><td>0.00 %</td></tr> <tr><td>13</td><td>0</td><td>0.00 %</td><td>0</td><td>0.00 %</td></tr> <tr><td>14</td><td>1</td><td>100.00 %</td><td>1</td><td>100.00 %</td></tr> </tbody> </table>			Anfragen (von Ihnen ausgegangen) (bei Ihnen eingegangen)					Name	gesendet	bearbeitet	empfangen	bearbeitet	3	2	50.00 %	3	0.00 %	20	2	50.00 %	1	0.00 %	22	0	0.00 %	0	0.00 %	13	0	0.00 %	0	0.00 %	14	1	100.00 %	1	100.00 %
Von	Beschreibung																																																							
11	Weiterleiten																																																							
23	Weiterleiten																																																							
20	Weiterleiten																																																							
9	Weiterleiten																																																							
21	Weiterleiten																																																							
18	Beantworten																																																							
3	Weiterleiten																																																							
3	Weiterleiten																																																							
Anfragen (von Ihnen ausgegangen) (bei Ihnen eingegangen)																																																								
Name	gesendet	bearbeitet	empfangen	bearbeitet																																																				
3	2	50.00 %	3	0.00 %																																																				
20	2	50.00 %	1	0.00 %																																																				
22	0	0.00 %	0	0.00 %																																																				
13	0	0.00 %	0	0.00 %																																																				
14	1	100.00 %	1	100.00 %																																																				
		Abgesetzte Anfragen																																																						
		<table border="1"> <thead> <tr> <th>An</th> <th>Antwort</th> </tr> </thead> <tbody> <tr><td>3</td><td>(erhalten)</td></tr> <tr><td>14</td><td>(erhalten)</td></tr> <tr><td>20</td><td>(erhalten)</td></tr> <tr><td>20</td><td>Warte seit: 1 Runden</td></tr> <tr><td>3</td><td>Warte seit: 3 Runden</td></tr> </tbody> </table>	An	Antwort	3	(erhalten)	14	(erhalten)	20	(erhalten)	20	Warte seit: 1 Runden	3	Warte seit: 3 Runden																																										
An	Antwort																																																							
3	(erhalten)																																																							
14	(erhalten)																																																							
20	(erhalten)																																																							
20	Warte seit: 1 Runden																																																							
3	Warte seit: 3 Runden																																																							
Aktionen (Ausführung beendet Runde)																																																								
Ignorieren	Verarbeiten / Weiterleiten	Anfrage absetzen	Nichts tun																																																					

Abbildung 7.2: Software-Client

7.4.1 Experiment-Rahmenwerk

Ein entscheidender Punkt im Entwurf eines verhaltensökonomischen Experiments ist, dass die menschlichen Teilnehmer sich auf ihre strategischen Entscheidungen konzentrieren können. Es soll daher vermieden werden, dass die Teilnehmer sich um nicht-strategische Entscheidungen, wie den Einsatz der Hash-Funktion zur Bestimmung des Ziels einer Anfrage oder das Sammeln von statistischen Informationen, kümmern müssen. Um sicherstellen zu können, dass die Spieler sich auf ihre Strategien konzentrieren können, interagieren sie mit Hilfe eines Experiment-Rahmenwerks² (engl. experiment framework). Dieses sammelt lokal vorhandene Informationen und aggregiert diese. Des Weiteren führt das Rahmenwerk nicht-strategische Handlungen durch, wie das Berechnen des Ziels einer Anfrage mittels der öffentlich bekannten Hash-Funktion.

Die Interaktion zwischen den Spielern erfolgt mit Hilfe eines Software-Clients. Dieser ist in Abbildung 7.2 zu sehen. Er besteht aus sechs Sektionen: dem Namensfeld (links oben), dem Finanzstatus (rechts oben), der „fremde Anfragen“-Sektion (links Mitte), der Kontakte-Sektion (rechts Mitte), der „abgesetzte Anfragen“-Sektion und der Aktionen-Sektion (rechts unten).

²Teile des Experiment-Rahmenwerks werden bereits in [Sch08] verwendet.

KAPITEL 7. DAS MENSCHLICHE VERHALTEN IN DER INTERAKTIONS- UND KONTAKTWAHL

Namensfeld und Finanzstatus Die zwei oberen Sektionen sind das Namensfeld (links) mit der Identifikationsnummer des Spielers und die Finanzstatus-Sektion.

Beispiel: Die Identifikationsnummer des Spieler in Abbildung 7.2 ist die 2. ■

Die Finanzstatus-Sektion zeigt die Auszahlung des Spielers. Diese ist aufgeschlüsselt in Einnahmen und Ausgaben.

Beispiel: Spieler 2 aus Abbildung 7.2 hat eine Auszahlung von 144 Punkten: Zusätzlich zu seinen 100 initial vorhandenen Punkten hat er Einnahmen von 60 Punkten und Ausgaben von 16 Punkten. ■

Fremde Anfragen Die Sektion „Fremde Anfragen“ zeigt alle Anfragen, die ein Spieler von anderen Spielern erhält und die er noch nicht bearbeitet oder fallengelassen hat.

Beispiel: Spieler 2 in Abbildung 7.2 hat im Moment acht Anfragen, die Spieler 2 bearbeiten kann. Von Spieler 11 hat Spieler 2 eine Anfrage erhalten, die er weiterleiten oder fallenlassen kann. Von Spieler 18 hat er eine Anfrage erhalten, die er beantworten oder fallenlassen kann. ■

Wenn ein Spieler sich entscheidet, eine Anfrage zu bearbeiten oder fallenzulassen, dann wählt er die Anfrage mit der Maus aus. Danach drückt er dann entweder den „Verarbeiten/Weiterleiten“-Knopf oder den „Ignorieren“-Knopf. Im Fall des Weiterleitens öffnet sich ein Menü, mit dessen Hilfe der Spieler einen Kontakt auswählen kann, an den die Anfrage weitergeleitet wird. In diesem Menü sind die Kontakte nach ihrer Distanz zum Ziel der Anfrage sortiert. (Kontakte, die weiter vom Ziel der Anfrage entfernt sind als der Spieler selbst, sind in dem Menü nicht enthalten.) Ein Spieler hat also die Möglichkeit, Kontakte nach ihrer Distanz zum Ziel der Anfrage zu wählen.

Kontakte Der Spieler findet Informationen über seine Kontakte in der Kontakte-Sektion. Diese besitzt fünf Spalten. Die erste nennt die Identifikationsnummer des Kontakts. Die zweite und dritte Spalte beziehen sich auf Anfragen, die der Spieler an den Kontakt gesendet hat: Die zweite Spalte zeigt die Anzahl der Anfragen, die an den Kontakt gesendet wurden, und die dritte Spalte den Anteil an diesen Anfragen, der bis zu diesem Zeitpunkt erfolgreich bearbeitet wurde. Die vierte und fünfte Spalte betreffen Anfragen, die der Kontakt dem Spieler gesendet hat: Die vierte Spalte zeigt die Anzahl der Anfragen und die fünften Spalte den Anteil, den der Spieler bearbeitet hat.

Die ersten drei Zeilen der Kontakte-Sektion zeigen Informationen über feste Kontakte, die restlichen Zeilen zeigen Informationen über zusätzliche Kontakte an.

Beispiel: Nach Abbildung 7.2 hat Teilnehmer 3 zwei Anfragen von Spieler 2 erhalten. Von diesen hat er eine – also 50% – bearbeitet. Spieler 2 hat von seinem festen Kontakt mit Identifikationsnummer 3 bisher drei Anfragen erhalten, von denen er (bisher) keine bearbeitet hat. ■

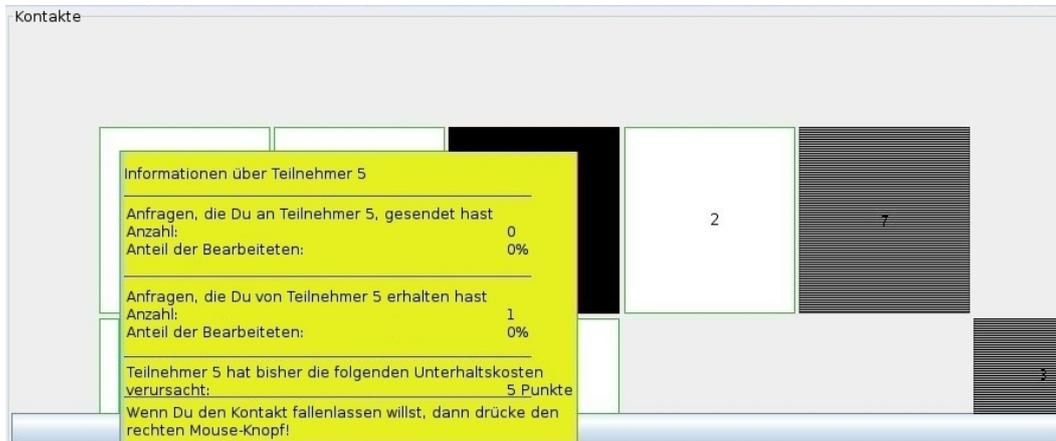


Abbildung 7.3: Kontaktwahl Fenster

Abgesetzte Anfragen Die „Abgesetzte Anfragen“-Sektion listet alle Anfragen auf, die der Spieler erzeugt und gesendet hat. Die linke Spalte zeigt die Identifikationsnummer des Kontakts, an den der Spieler die Anfrage gesendet hat. Die rechte Spalte zeigt an, ob die Anfrage beantwortet wurde oder seit wie vielen Runden der Spieler auf die Anfrage wartet.

Aktionen Unterhalb der „Abgesetzte Anfragen“-Sektion sind zwei Knöpfe: „Anfrage absetzen“ und „Nichts tun“. Der Spieler kann pro Runde eine Anfrage erzeugen und absetzen – oder es lassen. Wenn der Spieler eine Anfrage erzeugt, dann öffnet sich dasselbe Menü wie für weiterzuleitende Anfragen. Sobald er eine Anfrage erzeugt und abgesetzt hat, ist die Runde für ihn zu Ende.

Kontaktwahl Ein Spieler kann zusätzliche Kontakte auswählen oder fallenlassen. Abbildung 7.3 zeigt das Kontaktwahl-Fenster, mit dem zusätzliche Kontakte gewählt und fallengelassen werden können. Rechtecke repräsentieren die Zonen der Teilnehmer. Das Rechteck des Spielers ist schwarz. Seine festen Kontakte sind diejenigen Teilnehmer, die Zonen besitzen, die mit der Zone des Spielers eine Kante gemeinsam haben. Kontakte haben weiße Zonen. Gestreifte Zonen repräsentieren *potentielle Kontakte*. Also solche Teilnehmer, die der Spieler in der Interaktionswahl kennengelernt hat, die er aber nicht als zusätzliche Kontakte ausgewählt hat. Wenn Spieler i Spieler j zu einem zusätzlichen Kontakt macht, dann wird Spieler i bei Spieler j ein potentieller Kontakt. Das Kontaktwahl-Fenster zeigt Informationen zu allen Kontakten des Spielers an. Darunter sind diejenigen Informationen, die auch in der Kontakte-Sektion angezeigt werden, sowie die Kosten, die der Unterhalt eines zusätzlichen Kontaktes verursacht haben.

Beispiel: In Abbildung 7.3 hat der zusätzliche Kontakt mit Identifikationsnum-

KAPITEL 7. DAS MENSCHLICHE VERHALTEN IN DER INTERAKTIONS- UND KONTAKTWAHL

mer 5 Unterhaltskosten von fünf Punkten verursacht. ■

Es sei angemerkt, dass ein Spieler die Netzwerkstruktur nicht komplett sieht. Ihm werden nur seine Kontakte sowie diejenigen Teilnehmer, mit denen er (direkt) interagiert hat, und ihre Zonen angezeigt. In Abbildung 7.3 bleiben beispielsweise einige Rechtecke frei. Der Spieler, der das Kontaktwahl-Fenster kontrolliert, hat mit diesen Teilnehmern noch nicht interagiert.

7.4.2 Strategie-Spiel

Nach jedem Einzelexperiment wird mit jedem einzelnen Subjekt ein Strategie-Spiel veranstaltet. Per Umfrage wird jeder Spieler gebeten, seine Strategie anhand der Beantwortung von Fragen zu beschreiben. (Die Ergebnisse aus den Fragebögen können auf http://dbis.ipd.kit.edu/Hartmann_Experimentdaten.php eingesehen werden.) Jedes Subjekt schreibt seine Strategie in natürlicher Sprache auf. Vor dem Strategie-Spiel wird den Teilnehmern erklärt, dass ihre Strategien später implementiert werden und dass die implementierten Strategien in einem simulierten Wettstreit gegeneinander antreten und die Spieler relativ zum Erfolg ihrer Strategie ausbezahlt werden. Dies ist als Anreizmechanismus gedacht, um sinnvolle Strategien zu erhalten. Des Weiteren soll durch das Strategie-Spiel überprüft werden, ob das Spiel verstanden wurde.

Die Durchführung eines Strategie-Spiels ist eine anerkannte Methode in der experimentellen Wirtschaftsforschung (vergleiche [SMU97]). Mit Hilfe eines Strategie-Spiels ist es möglich, Strategien zu gewinnen, die von Spielern stammen, die bereits Erfahrungen mit ihren Strategien im Experiment gemacht haben. Des Weiteren hat ein Subjekt die Möglichkeit, die eigene Strategie in voller Länge wiederzugeben und seine Gedankengänge zu motivieren. In den Experimenten kann lediglich das Verhalten eines Spielers beobachtet werden. Die zugrunde liegende Strategie kann daraus nicht beobachtet, sondern lediglich abgeleitet werden.

7.5 Ergebnisse

Im Folgenden werden die Ergebnisse der Untersuchung der Hypothesen durch die Experimente wiedergegeben und im Anschluss Beobachtungen bezüglich der Kontaktwahl der Teilnehmer erläutert.

Eine Auflistung der wichtigsten Daten aus den Experimenten finden sich im Appendix C.

7.5.1 Kontaktwahl

Um herauszufinden, ob zusätzliche Kontakte das Gemeinwohl eines Systems verbessern, wird der durchschnittliche Kooperationsgrad \bar{C} der einzelnen Experimentreihen

Maß	T_0	T_1	T_3	T_5
\bar{C}	0,72	0,69	0,64	0,53
\bar{C}_c	0,67	0,72	0,61	0,8

Tabelle 7.2: Kooperationsgrad in den Experimentreihen

„ \bar{C} in T_i ist größer als \bar{C} in T_j .“	p-Wert
$i = 0, j = 1: T_0, T_1$	0,135
$i = 0, j = 3: T_0, T_3$	< 0,01
$i = 0, j = 5: T_0, T_5$	< 0,01
$i = 1, j = 3: T_1, T_3$	0,10
$i = 1, j = 5: T_1, T_5$	< 0,01
$i = 3, j = 5: T_3, T_5$	< 0,01

Tabelle 7.3: Vergleich des Kooperationsgrads zwischen den Experimentreihen

verglichen. Tabelle 7.2 gibt eine Übersicht. Der durchschnittliche Kooperationsgrad sinkt von Experimentreihe T_0 zu T_5 . Wie in Abschnitt 7.5.5 gezeigt werden wird, gilt dies auch für die Anzahl an Kontakten pro Subjekt. Es gilt also, dass Experimentreihen mit mehr zusätzlichen Kontakten auch einen höheren durchschnittlichen Kooperationsgrad haben. Um herauszufinden, ob Hypothese „Im Durchschnitt wird der Kooperationsgrad durch die Kontaktwahl erhöht“ gilt, wird der Kooperationsgrad aller Knoten aller Experimentreihen mit einem Wilcoxon Rangsummentest [Wil45] untersucht. Die Ergebnisse finden sich in Tabelle 7.3 und bestätigen die Hypothese.

Des Weiteren wurden die Subjekte im Strategie-Spiel gefragt, (a) wie sie die Anfragen von Kontakten im Allgemeinen bearbeiten und (b) wie sie die Anfragen, die ihnen von zusätzlichen Kontakten zugesendet wurden, bearbeitet haben. In beiden Fällen zeigt die Mehrheit ein kooperatives Verhalten – 83% für (a) und 93% für (b). Die Beobachtung, dass die Subjekte sich kooperativ verhalten, ist signifikant. Es sei angemerkt, dass hier und im Folgenden immer dann von Signifikanz gesprochen wird, wenn die Ereignisse auf einem 1%-Niveau signifikant sind, d.h., wenn der p-Wert kleiner als 1% ist. Dabei wird die Annahme gemacht, dass die Spieler unabhängig voneinander agieren. Das kooperative Verhalten kann in zwei Unterklassen kategorisiert werden:

1. Die Subjekte bearbeiten die Mehrheit der Anfragen unabhängig vom Verhalten der Kontakte, von denen sie die Anfragen erhalten haben.
2. Die Spieler verarbeiten Anfragen, wenn der Kontakt, von dem sie die Anfrage erhalten haben, kooperativ ist.

Für Anfragen von zusätzlichen Kontakten zeigen 44% der Subjekte Verhalten (1) und 49% Verhalten (2). Im Allgemeinen, d.h. unabhängig davon, ob ein Kontakt ein fester

KAPITEL 7. DAS MENSCHLICHE VERHALTEN IN DER INTERAKTIONS- UND KONTAKTWAHL

Kontakt oder ein zusätzlicher ist, zeigen 12% Verhalten (1) und 72% Verhalten (2).

Eine Erklärung für das beobachtete Verhalten ist, dass die Subjekte unkooperative Teilnehmer fallenlassen und damit nur kooperative zusätzliche Kontakte besitzen. Bei nur kooperativen zusätzlichen Kontakten ist es nicht mehr nötig, den Kooperationsgrad zu beachten, da die darauf aufbauende Unterscheidung bereits in der Kontaktwahl getroffen wurde. Bei festen Kontakten hingegen ist dies nicht der Fall. Daher liegt es nahe, dass die Subjekte den Kooperationsgrad eines festen Kontakts in der Anfragebearbeitung berücksichtigen. Es folgt der Unterschied in den Beobachtungen.

7.5.2 Position vs. Kooperationsgrad

Wie leiten Subjekte Anfragen durch das Netzwerk? Bei der Beantwortung dieser Frage stellen sich weitere Fragen. So ist unklar, ob die Subjekte kooperative Kontakte beim Weiterleiten auswählen oder ob sie Kontakte mit einer vielversprechenden Position bezüglich des Ziels der Anfrage auswählen. Um diese Fragen beantworten zu können, wird im Folgenden das Strategie-Spiel analysiert. Darin wurden alle Subjekte gefragt, an welche Kontakte sie (a) ihre eigenen Anfragen, d.h. die von ihnen erzeugten Anfragen, senden und (b) an wen sie Anfragen von anderen Teilnehmern weiterleiten. Es sei angemerkt, dass die Ergebnisse der Untersuchung nur von begrenzter Aussagekraft sind, da die Experimente in kleinen Netzwerken von nur 12 Teilnehmern stattfanden. In diesen spielt naturgemäß die Position eines Kontakts eine geringere Rolle als in einem großen Netzwerk. (Dieser Umstand wird in Kapitel 8 erneut betrachtet und gelöst.) Es sei jedoch ebenfalls angemerkt, dass durch die Beobachtung von insgesamt 120 Strategien ein breites Spektrum an Strategien beleuchtet wird. Während sich die genauen Parameterwerte einer Strategie, wie z.B. Schwellenwerte, natürlicherweise mit der Größe eines Netzwerks verändern, wird oft angenommen, dass dies nicht für die genutzten Strategien gilt [Sel73]. Z.B. sollte die Tatsache, dass Subjekte Anfragen von unkooperativen Teilnehmern fallenlassen, nicht von der Größe des Netzwerks abhängen.

Die Ergebnisse der Auswertung des Strategie-Spiels finden sich in Tabelle 7.4. Die Werte wurden durch manuelle Klassifikation der Ergebnisse des Strategie-Spiels erlangt. Subjekte leiten ihre Anfragen an kooperative Teilnehmer weiter: In der Anfragebearbeitung ist der Kooperationsgrad (Kategorie B) wichtiger als die Position eines Kontaktes (Kategorie C). Die Verwendung eines Binomialtests zeigt, dass diese Beobachtung signifikant ist. Dennoch können Unterschiede in den Strategien bezüglich eigener Anfragen und Anfragen von anderen Teilnehmern beobachtet werden: Die Position eines Kontakts scheint beim Weiterleiten eigener Anfragen eine Rolle zu spielen. Für Anfragen anderer Teilnehmer hingegen spielt die Position des Kontakts keine Rolle: Kein Subjekt leitet Anfragen von anderen Teilnehmern an einen Kontakt weiter, weil dieser eine besondere Position im Adressraum inne hatte. „Unabhängige Strategien“, also solche, die nicht vom Verhalten des Kontakts abhängen (Katego-

Auswahl-Kriterium	Eigene Anfragen	Anfragen von anderen Teilnehmern	Kategorie
„Unabhängige Strategie“	1	29	A
„Kooperationsgrad“	66	67	B
„Kooperationsgrad plus Position“	30	1	B
„Position“	11	0	C
„Andere Strategie“	8	23	D

Tabelle 7.4: Verhalten beim Weiterleiten von Anfragen

rie A) und nicht klassifizierbare Strategien (Kategorie D) spielen keine signifikante Rolle. Daher kann Hypothese „Ein Subjekt wählt in der Anfragebearbeitung Kontakte nach dem Kooperationsgrad aus“ – wenigstens für Netze der im Experiment genutzten Größe – angenommen werden.

7.5.3 Fallenlassen von Freifahrern

Im Strategie-Spiel nennen die Subjekte unterschiedliche Motive, warum sie zusätzliche Kontakte wählen. So geben zum Beispiel einige Subjekte an, dass sie sich davon einen positiven Nutzen erwarten. Keins der Motive ist signifikant oder besonders häufig. Im Gegensatz zum Hinzufügen von zusätzlichen Kontakten hat das Fallenlassen von zusätzlichen Kontakten ein eindeutiges Muster: Subjekte lassen zusätzliche Kontakte fallen, die sich ihnen gegenüber in der Vergangenheit nicht kooperativ verhalten haben. Das Strategie-Spiel zeigt, dass 95 von 111 Subjekten diese Strategie nutzen. Mit Hilfe eines Binomialtests bestätigt dies Hypothese „Teilnehmer lassen zusätzliche Kontakte fallen, die sich in der Vergangenheit unkooperativ verhalten haben“ auf einem 1%-Niveau.

7.5.4 Kosten für zusätzliche Kontakte

Die Ergebnisse der Experimente sind konsistent mit Hypothese „Subjekte verhalten sich kooperativer gegenüber zusätzlichen Kontakten, wenn die Unterhaltskosten für zusätzliche Kontakte hoch sind“: Der durchschnittliche Kooperationsgrad bezüglich zusätzlicher Kontakte \bar{C}_c steigt mit steigenden Kosten für den Unterhalt für zusätzliche Kontakte. Tabelle 7.2 bietet eine Übersicht.

KAPITEL 7. DAS MENSCHLICHE VERHALTEN IN DER INTERAKTIONS- UND KONTAKTWAHL

Effizienz Es bleibt zu klären, welchen Einfluss die Kosten für den Unterhalt von zusätzlichen Kontakten auf die Effizienz des Systems haben. Tabelle 7.5 zeigt die durchschnittliche Auszahlung $Auszahlung(i)$ sowie die durchschnittliche Auszahlung ohne die Kosten für den Unterhalt zusätzlicher Kontakte. Wie zu erwarten gewesen ist, haben die Spieler in der Experimentreihe T_0 die höchsten Auszahlungen, da sie keine Unterhaltskosten für zusätzliche Kontakte zahlen. Ebenso ist zu erwarten gewesen, dass die Spieler von Experimentreihe T_5 die geringsten Auszahlungen erhalten, da sie die höchsten Kosten für zusätzliche Kontakte haben. Interessanterweise sind negative Auszahlungen möglich (und sind in diesem Experiment auch vorgekommen), da die Subjekte bis zu fünf Runden benötigen, um das System zu verlassen: Zusätzliche Kontakte können nur alle fünf Runden gewählt oder fallengelassen werden.

Interessanter als der Vergleich der durchschnittlichen Auszahlung, der aufgrund der unterschiedlichen Kosten für den Unterhalt von zusätzlichen Kontakten wenig aussagekräftig ist, ist der Vergleich der Auszahlungen ohne die Kosten für den Unterhalt von Kontakten (vergleiche rechte Spalte von Tabelle 7.5). Es stellt sich heraus, dass die Subjekte in Experimentreihe T_1 die höchsten Auszahlungen (ohne Kontaktkosten) haben. Mit anderen Worten, moderate Kosten für den Unterhalt von zusätzlichen Kontakten haben einen positiven Effekt auf das Gemeinwohl im System.

Relative Effizienz Des Weiteren wurde untersucht, welche Experimentreihe die höchsten relativen Auszahlungen besitzt. Es werden dazu die echten Kosten für den Unterhalt von zusätzlichen Kontakten gegen virtuelle Kosten ersetzt: Dazu werden die echten Kosten herausgerechnet und durch die Kosten der anderen Experimentreihen ersetzt. So entstehen virtuelle Auszahlungen.

Beispiel: Subjekt i ist in Experimentreihe T_0 und hat keine Kosten für den Unterhalt von zusätzlichen Kontakten. i hat eine Auszahlung von 200 Punkten in T_0 . Des Weiteren hat Subjekt i fünf Kontakte über 20 Runden. In dem Vergleich der relativen Auszahlungen werden Subjekt i nun 1, 3 und 5 Punkte Kosten für seine zusätzlichen Kontakte berechnet. Die (virtuellen) Auszahlungen von i wären also: $T_1 = 100$ Punkte, $T_3 = -100$ Punkte und $T_5 = -400$ Punkte. In diesem Fall ist die echte Auszahlung von Subjekt i höher als die virtuellen Auszahlungen. ■

Die Ergebnisse sind wie folgt: Die Experimentreihen T_0 , T_1 und T_5 führen zu höheren Auszahlungen verglichen mit den entsprechenden virtuellen Auszahlungen. Die Experimentreihe T_3 weicht als einzige Experimentreihe von dieser Beobachtung ab: Subjekte aus der Experimentreihe T_5 mit virtuellen Kosten für den Unterhalt von zusätzlichen Kontakten von 3 Punkten sind erfolgreicher als Subjekte aus Experimentreihe T_3 . Es folgt, dass die Subjekte Auszahlungen erhalten, die optimal (d.h. relativ gesehen maximal) sind, in dem Sinn, dass das Verhalten der Subjekte auf die entspre-

Experimentreihe	Durchschnittliche Auszahlung	$Auszahlung(i)$ ohne Kosten für den Unterhalt zusätzlicher Kontakte
T_0	5.04	5.04
T_1	3.39	6.79
T_3	-0.02	5.53
T_5	-1.23	1.97

Tabelle 7.5: Auszahlung pro Runde

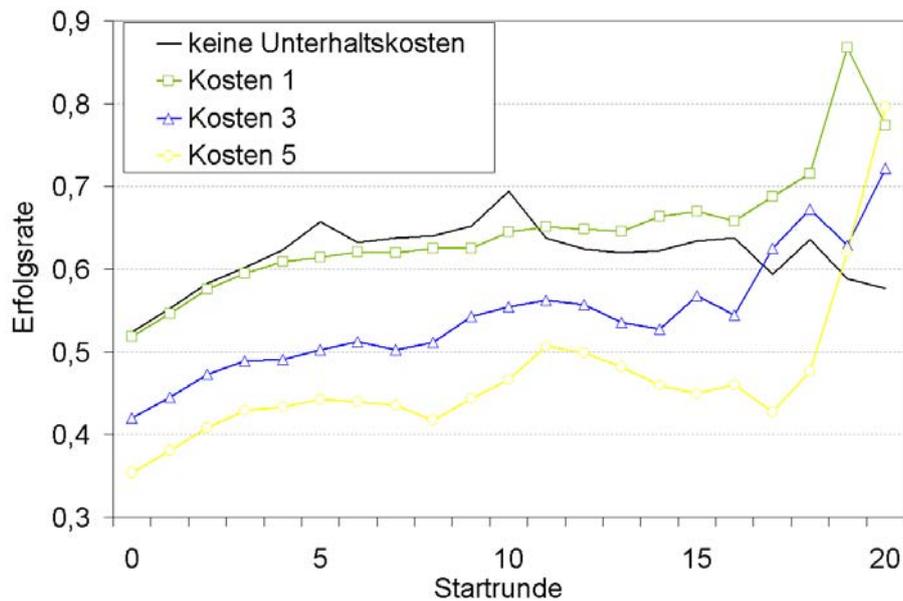


Abbildung 7.4: Erfolgsrate

chenden Kosten optimal abgestimmt ist. Es sei allerdings angemerkt, dass aufgrund der Tatsache, dass nur vier Datenpunkte (d.h. die vier Experimentreihen) vorhanden sind, dieses Ergebnis lediglich eine Beobachtung beziehungsweise einen Trend darstellt und kein signifikantes Ergebnis. Eine weitergehende Untersuchung erscheint sinnvoll.

Erfolgsrate Ein weiterer Effekt der Kosten für den Unterhalt von zusätzlichen Kontakten ist, dass sie einen positiven Effekt auf die Erfolgsrate haben. Die Erfolgsrate ist der Anteil beantworteter Anfragen an den Anfragen, die ein Subjekt erzeugt und gesendet hat (vergleiche Kapitel 3.6). Abbildung 7.4 zeigt die Erfolgsrate. Die X-Achse zeigt die Runde, ab der die Erfolgsrate für alle folgenden Runden berechnet wird. Man kann sehen, dass die Erfolgsrate für die Experimentreihen T_1 , T_3 und T_5

KAPITEL 7. DAS MENSCHLICHE VERHALTEN IN DER INTERAKTIONS- UND KONTAKTWAHL

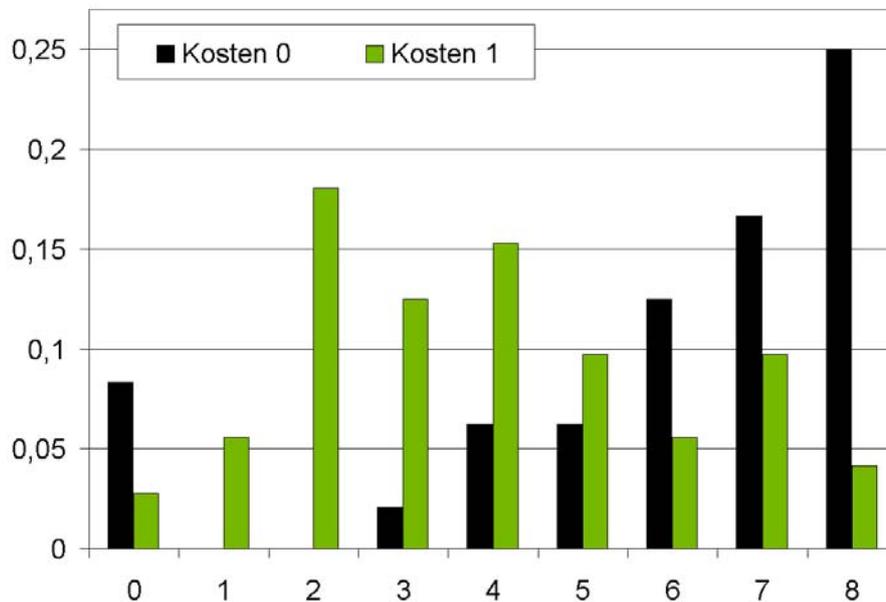


Abbildung 7.5: Anzahl an Kontakten für die Experimentreihe T_0 und T_1

– also für die Experimentreihen mit Kosten für zusätzliche Kontakte – mit der Zeit ansteigt. Dies ist für Experimentreihe T_0 (ohne Kosten für den Unterhalt zusätzlicher Kontakte) nicht der Fall. Es kann daraus geschlossen werden, dass Kosten für den Unterhalt zusätzlicher Kontakte das Einkommen eines Subjekts erhöhen.

7.5.5 Anzahl an Kontakten

Anzahl an Kontakten Die Anzahl an zusätzlichen Kontakten pro Subjekt verringert sich mit steigenden Kosten für ihren Unterhalt: Subjekte in den Experimentreihen T_0 (T_1 , T_3 , T_5) haben im Durchschnitt 5,17 (3,40; 1,85; 0,64) zusätzliche Kontakte. Diese Beobachtungen sind signifikant, wenn man die Anzahl zusätzlicher Kontakte pro Subjekt mit Hilfe eines Wilcoxon Rangsummentests vergleicht. Daher kann Hypothese „Die Anzahl zusätzlicher Kontakte pro Subjekt und die Kosten für den Unterhalt zusätzlicher Kontakte sind negativ korreliert“ bestätigt werden.

Strukturelle Eigenschaften Neben der Anzahl zusätzlicher Kontakte wurden auch die strukturellen Eigenschaften der entstehenden Netzwerke untersucht. Dafür wird als Vergleichsmaß das Connections Model (vergleiche Abschnitt 2.2) herangezogen, obwohl sich das in dieser Arbeit genutzte Modell vom Connections Model unterscheidet (da z.B. Interaktionswahl im Connections Model nicht berücksichtigt wird). Es wird also untersucht, ob die Teilnehmer in Abhängigkeit von den Kosten

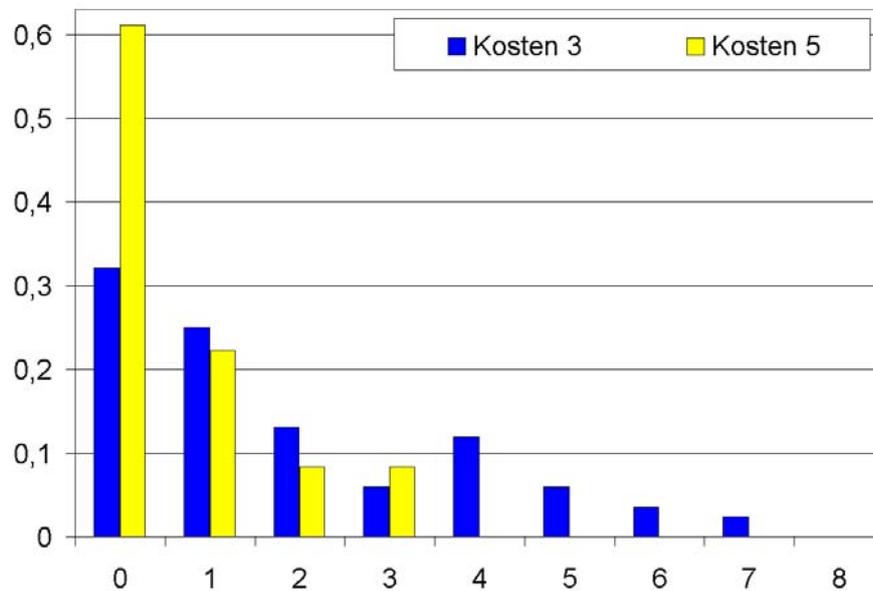


Abbildung 7.6: Anzahl an Kontakten für die Experimentreihe T_3 und T_5

für den Unterhalt zusätzlicher Kontakte Stern-Netzwerke, komplette Netzwerke oder leere Netzwerke bilden. Dazu wurden vier Histogramme erzeugt, die für die vier Experimentreihen den Anteil der Subjekte zeigen, der eine bestimmte Anzahl an zusätzlichen Kontakten besitzt. Dabei sei angemerkt, dass ein Subjekt in einem Netzwerk aus 12 Subjekten mit je drei festen Kontakten maximal acht zusätzliche Kontakte haben kann. Abbildung 7.5 und Abbildung 7.6 zeigen die Histogramme. Die X-Achse zeigt die Anzahl an zusätzlichen Kontakten und die Y-Achse den Anteil an Subjekten, der die entsprechende Anzahl an zusätzlichen Kontakten hat. Die Ergebnisse sind wie folgt: In Experimentreihe T_0 sind die Subjekte fast vollständig miteinander vernetzt. In Experimentreihe T_5 haben die Subjekte wenige Kontakte und in den Experimentreihen T_1 und T_3 liegen die Netzwerke zwischen kompletten und leeren Netzwerken. Es können keine Stern-Netzwerke beobachtet werden. Bei diesen müsste das Histogramm zeigen, dass elf Zwölftel der Teilnehmer einen zusätzlichen Kontakt haben und ein Teilnehmer acht zusätzliche Kontakte (was elf Kontakten entspricht). Obwohl die Subjekte in Experimentreihe T_3 Netzwerke formen, die Stern-Netzwerken ähnlich sind, d.h. sie formen Netzwerke, in denen wenige Knoten stark vernetzt sind, ist diese Beobachtung nicht signifikant. Dennoch – wie in Abschnitt 7.5.4 gezeigt – verhalten sich die Subjekte relativ gesehen effizient.

7.6 Fazit

Autonome Teilnehmer verteilter Systeme wählen ihre Interaktionswahl-Strategie und ihre Kontakte. Das menschliche Verhalten wurde in solchen Systemen bisher kaum erforscht. Da die Maximierung der Auszahlung nicht immer menschliches Verhalten vorhersagen kann, wurden in diesem Kapitel Laborexperimente mit menschlichen Teilnehmern genutzt, um gleichzeitige Interaktions- und Kontaktwahl zu untersuchen.

Das wichtigste Ergebnis der Untersuchung ist, dass Kontaktwahl – also die Tatsache, dass die Teilnehmer ihre Kontakte selbstständig aus- und abwählen – den Kooperationsgrad und das Gemeinwohl im System positiv beeinflusst. Je geringer die Kosten für den Unterhalt zusätzlicher Kontakte, desto mehr Kontakte wählen die Teilnehmer aus. Dabei zeigt sich, dass die Teilnehmer umso stärker kooperieren, je mehr Kontakte ausgewählt werden. Das bedeutet, dass Systementwickler keine teuren Mechanismen oder Strategien entwickeln müssen, um den Kooperationsgrad im System zu erhöhen. Kosten für den Unterhalt zusätzlicher Kontakte haben keinen positiven Effekt auf den Kooperationsgrad. Die Teilnehmer verteilter Systeme bilden Netzwerkstrukturen, die auf exogene Parameter abgestimmt sind und die relativ gesehen optimal sind. Aus den Ergebnissen kann geschlossen werden, dass Systementwickler keine Netzwerkstrukturen vorschreiben müssen, wenn die Teilnehmer ihre Kontakte selbstständig wählen können.

Da die Laborexperimente mit kleinen Gruppen durchgeführt wurden, ist es sinnvoll, die Ergebnisse in einem größeren System zu validieren und zusätzliche Beobachtungen zu machen. Dies wird im folgenden Kapitel durchgeführt.

8 PostGame – Eine Studie

Das menschliche Verhalten in der Interaktions- und Kontaktwahl kann auf unterschiedliche Arten untersucht werden. Zwei übliche Methoden sind formale Analysen und verhaltensökonomische Experimente. Während in formalen Analysen davon ausgegangen wird, dass die Teilnehmer ihre Auszahlung maximieren, werden verhaltensökonomische Experimente oft mit kleinen Experiment- und Kontrollgruppen durchgeführt. Da unklar ist, ob Menschen stets eine Strategie wählen, die ihre Auszahlung maximiert, ist die erste Methode (allein) nicht ausreichend, um menschliches Verhalten zu analysieren [HBB⁺01]. Die zweite Methode – d.h. verhaltensökonomische Experimente – haben den Vorteil, dass mit ihrer Hilfe nicht nur Verhalten beobachtet werden kann, sondern, dass auch die Signifikanz einer Beobachtung bestimmbar ist. Sie haben allerdings den folgenden Nachteil: Da Soziale Netzwerke und Peer-to-Peer Systeme oft wesentlich größer sind als die in den Experimenten untersuchten Gruppen, ist unklar, ob alle Effekte, die in großen verteilten Systemen auftreten auch in Labor-Experimenten beobachtet werden können. Es liegt daher nahe, weitergehende Untersuchungen über das menschliche Verhalten in der Interaktionswahl und Kontaktwahl durchzuführen. In diesem Kapitel werden die Ergebnisse einer Studie über das menschliche Verhalten in der Interaktions- und Kontaktwahl motiviert und dargestellt. Den Teilnehmern der Studie wurde das Problem der Strategiewahl in verteilten Systemen als Online-Spiel präsentiert: Die Subjekte spielen Briefträger in einem föderalen Postsystem in einem erdachten Königreich. Der Aufbau der Studie wurde so gewählt, dass Störeffekte wie das Rauschen, das durch unerwünschte Gruppenstrategien hervorgerufen wird, verhindert wird oder wie mangelndes Problemverständnis der Teilnehmer reduziert wird. Das Ergebnis der Studie ist, dass die Strategien der Teilnehmer zu einem System führen, das beinahe effizient ist, das also ein hohes Gemeinwohl hat. Überraschenderweise überschätzen die Teilnehmer den Nutzen von kooperativen Kontakten und behalten kooperative aber für sie kostspielige zusätzliche Kontakte. Dieses Ergebnis ist wichtig, da Strategien oder Mechanismen, die helfen, dieses Verhalten zu korrigieren, das Gemeinwohl weiter anheben könnten.

8.1 Einleitung

Soziale Netzwerke können eine enorme Größe besitzen. So haben Beispiele von Sozialen Netzwerken wie E-Mail Graphen [DMW03] oder Soziale-Netzwerke-Seiten

KAPITEL 8. POSTGAME – EINE STUDIE

wie Facebook, Xing, Orkut oder LinkedIn Millionen von Teilnehmern¹. Da unklar ist, ob alle Effekte großer Systeme auch in kleinen Netzen beobachtet werden können, ist eine Validierung der im vorangegangenen Kapitel erzielten Ergebnisse in einem größeren Netzwerk nötig.

Ziele und Herausforderungen Das Ziel dieses Kapitels ist es, die Ergebnisse des vorherigen Kapitels mit Hilfe einer Studie zu validieren und – falls vorhanden – neue Erkenntnisse zu gewinnen. Die Studie soll dabei in der Lage sein, die üblichen Beschränkungen kontrollierter Experimente zu überwinden, d.h. sie sollte in der Lage sein, eine größere Gruppe an Subjekten über einen längeren Zeitraum zu beobachten.

Klar ist, dass Studien – im Gegensatz zu kontrollierten Experimenten – nicht zu Ergebnissen mit einer bestimmbarer Signifikanz führen. Dennoch sind sie eine in der Wissenschaft anerkannte Methode, um Beobachtungen aus der realen Welt zu sammeln. Trotz der Tatsache, dass eine Studie nicht die Qualität eines kontrollierten Experiments erreichen kann, soll die in dieser Arbeit vorgestellte Studie die Ergebnisse des vorherigen Kapitels ergänzen. Darüber hinaus soll – wenn möglich – verhindert werden, dass störende Effekte die Beobachtungen verfälschen.

Bei der Umsetzung der Studie gibt es einige Herausforderungen. Es muss sichergestellt werden, (a) dass die Subjekte motiviert sind, dass aber keine falschen Anreize gesetzt werden, die das Verhalten beeinflussen. (b) Verfälschungen der Ergebnisse müssen verhindert werden. (c) Auf technischer Seite muss sichergestellt werden, dass die Teilnehmer nicht unsynchronisiert auf das System zugreifen und dass sie das Netzwerk jederzeit verlassen und ihm jederzeit beitreten können.

(a) Teilnehmer verteilter Systeme versprechen sich Nutzen von der Teilnahme am System (vergleiche Abschnitt 3.5). Dies gilt ebenfalls für die Teilnehmer der Studie. Auf der einen Seite ist das Geld, welches die Teilnehmer durch ihre Strategien erhalten können, ein Anreiz. Dennoch sind weitere Anreize nötig, um genügend Subjekte motivieren zu können. So wird eine Rangliste genutzt, um die Subjekte zur Teilnahme zu motivieren. Dies ist eine übliche Methode in Online-Spielen. Wahrscheinlich ziehen die Teilnehmer einen Nutzen aus der Tatsache, dass sie in einem Spiel erfolgreich sind. Dennoch muss beachtet werden, dass diese zusätzlichen Anreize nicht Effekte hervorrufen, die gegenläufig zu den Effekten, die durch die Auszahlungsfunktion hervorgerufen werden, verlaufen. So gibt z.B. eine Rangliste einen Anreiz, andere Spieler zu diskriminieren, da ein Spieler, der dafür sorgt, dass andere Spieler weniger Auszahlung erhalten, relativ gesehen besser dasteht.

(b) Die Subjekte müssen das Problem beziehungsweise das Spiel, das sie spielen, verstehen. Andernfalls kann nicht erwartet werden, dass sinnvolle Strategien beobachtet werden können. Während der Beobachtung des Verhaltens der Subjekte muss verhindert werden, dass Störeffekte, wie ungewollte Gruppenstrategien, die Messungen verrauschen. So soll z.B. die Realwelt-Freundschaft von zwei Subjekten nicht ihr

¹vgl. www.facebook.com, www.xing.com, www.orkut.com, www.linkedin.com

gegenseitiges Verhalten beeinflussen.

(c) Eine technische Herausforderung entsteht aus der Tatsache, dass die Teilnehmer eines Online-Spiels das System nach eigenem Ermessen betreten und verlassen. In der Konsequenz muss sichergestellt werden, dass das System dynamisch wachsen und schrumpfen können muss, ohne dass Datenstrukturen durcheinander geraten. Des Weiteren darf der gleichzeitige Zugriff der Spieler auf das System nicht zu Inkonsistenzen führen.

Beiträge Um die Gefahr zu verringern, dass das Problemverständnis der Teilnehmer nicht ausreichend ist, wurde das Modell eines verteilten Systems zur Interaktions- und Kontaktwahl (vergleiche Kapitel 3) als Online-Spiel präsentiert. In diesem Spiel – genannt *PostGame* – spielen die Subjekte einen Postboten in einem erdachten Königreich. In diesem Königreich ist die Verwaltung der Postzustellung föderal organisiert, d.h. es gibt keine zentrale Kontrolle. 193 Subjekte nahmen an dem Spiel, das über 67 Tage online war, teil. Durch das Online-Spiel wurde die Beobachtung von menschlichen Strategien in der Interaktions- und Kontaktwahl in einem größeren System möglich. Einige Ergebnisse beziehen sich auf die Interaktionswahl, andere auf die Kontaktwahl und einige auf beide Strategiedimensionen. Die Ergebnisse sind wie folgt:

- Die Subjekte nutzen in der Interaktionswahl einige Kriterien, die in [Axe06] (und in diesem Kapitel) beschrieben werden. Strategien, die diese Kriterien nutzen, scheinen erfolgreicher als andere Strategien zu sein. Insbesondere sollte eine Strategie freundlich, zurückschlagend, vergebend und nicht neidisch sein (vergleiche Abschnitt 8.3).
- Das gebildete Netzwerk ist performant, aber nicht 100% effizient. Möglicherweise liegt dies daran, dass die Subjekte den Wert jener kooperativer Kontakte überschätzen, die ihnen trotz ihres hohen Kooperationsgrads hohe Kosten verursachen. Die Subjekte maximieren in dieser Hinsicht nicht ihre Auszahlung, sondern kooperieren mit kooperativen zusätzlichen Kontakten und ignorieren dabei eventuell für sie auftretende Nachteile.
- Wesentlich deutlicher als es in den Experimenten in kleinen Gruppen (vergleiche Kapitel 7) zu beobachten gewesen ist, beachten die Subjekte die Position zusätzlicher Kontakte in der Anfragebearbeitung ihrer eigenen Anfragen. Es ist anzunehmen, dass dieses Ergebnis nur möglich ist, da es der Entwurf des Experiments erlaubt, größere Gruppen von Subjekten zu untersuchen.

Durch die Studie kann das menschliche Verhalten in der Interaktions- und Kontaktwahl besser verstanden werden, insbesondere die Kontaktwahl in großen Systemen. Systementwickler können die Ergebnisse nutzen, um Mechanismen oder Strategien zu entwickeln, die das bereits hohe Gemeinwohl (bis zur Effizienz) erhöhen.

Gliederung des Kapitels Die für die Studie nötigen Anpassungen an das in Kapitel 3 vorgestellte Modell finden sich in Abschnitt 8.2. Hypothesen, die in der Studie untersucht werden, finden sich in Abschnitt 8.3. Es folgen die Beschreibung der Anforderungen an die Experimentumgebung (Abschnitt 8.4) und eine Beschreibung des Experimentaufbaus (Abschnitt 8.5). Die Ergebnisse der Studie (Abschnitt 8.8) werden von einem Fazit abgeschlossen (Abschnitt 8.9).

8.2 Definitionen und Anpassungen

Strategien menschlicher Teilnehmer sind oft nur dann sinnvoll, wenn die Teilnehmer das zugrunde liegende Problem verstanden haben. So ist es zum Beispiel eine Voraussetzung für die Maximierung der eigenen Auszahlung, dass man die Auszahlungsfunktion versteht. Andernfalls kann keine gezielte Maximierung derselben erwartet werden.

8.2.1 Anpassung der Auszahlungsfunktion

Die in Abschnitt 3.5 eingeführte Auszahlungsfunktion wird für die in diesem Kapitel beschriebene Studie vereinfacht und angepasst.

Ein Ziel beim Experimentaufbau ist es, die Spieler zur Teilnahme am Spiel zu motivieren beziehungsweise sicherzustellen, dass sie sich häufig am Spiel beteiligen. Dies ist notwendig, da eine einzelne Anfrage von mehreren Teilnehmern bearbeitet werden muss, um sie erfolgreich zu beantworten. Beteiligen sich die Spieler selten am Spiel, so verzögert sich die Anfragebearbeitung. Um einen Anreiz zur häufigen Beteiligung am Spiel zu schaffen, nimmt der Wert einer Anfrage pro verstrichener Zeiteinheit linear ab. Die Einnahmen eines Subjektes i sind daher wie folgt:

Definition 8.1 (Einkommen)

Das Einkommen von Subjekt i ist:

$$\text{Einkommen}(i) = \sum_{req \in A_i} \left(\max \left\{ a - \frac{t(req)}{t_a}, 0 \right\} \right)$$

Dabei ist $t(req)$ die Anzahl an Sekunden (statt Runden), die eine Anfrage req benötigt, um beantwortet zu werden. Alle t_a Sekunden ist eine Antwort einen Punkt weniger Wert.

Für die Evaluation und die Durchführung der Studie ist es für die Subjekte neben ihrer eigenen Auszahlung wichtig zu wissen, welchen Nutzen (d.h. welche Auszahlung) sie von einem zusätzlichen Kontakt erhalten.

Definition 8.2 (Einkommen durch einen Kontakt)

Das Einkommen von Subjekt i , das durch ein Subjekt j verursacht wurde, ist das

Zeitfenster	Divisor	Standard-Wert
Interaktionswahl	t_a	17.280 s (4.8 h)
Kontaktwahl	t_c	7.200 s (2 h)

Tabelle 8.1: Intervalle im PostGame

Einkommen jener Anfragen von i , die Subjekt i zu Subjekt j gesendet hat:

$$Einkommen(i|j) = \sum_{req \in A_{i \rightarrow j}} \left(\max \left\{ a - \frac{t(req)}{t_a}, 0 \right\} \right)$$

Die Kosten für den Unterhalt zusätzlicher Kontakte wurden ebenfalls angepasst. Die genutzte Zeiteinheit sind Sekunden (statt Runden).

$$nr(i, j) = \text{Anzahl der **Sekunden** an denen } g(i, j) \text{ gilt} \quad (8.1)$$

Die Kosten für den Unterhalt von zusätzlichen Kontakten sind c Punkte pro t_c Sekunden. $r(i)$ ist die Anzahl dieser Ereignisse:

$$r(i) = \sum_{n \in N} \frac{nr(i, n)}{t_c}$$

Tabelle 8.1 gibt eine Übersicht über die neuen Abkürzungen (vergleiche Tabelle 3.1).

Definition 8.3 (Ausgaben)

Die Ausgaben von Subjekt i , die durch ein Subjekt j verursacht werden, sind wie folgt:

$$Ausgaben(i|j) = q \cdot |R_{i \rightarrow j}| + f \cdot |F_{j \rightarrow i}| + w \cdot |W_{j \rightarrow i}| + c \cdot \frac{nr(i, j)}{t_c}$$

Die Auszahlung eines Subjektes i , die durch ein Subjekt j verursacht werden, sind die (jeweiligen) Einnahmen minus die Ausgaben.

8.3 Hypothesen

Anschaulich führt ein hoher Kooperationsgrad sowie eine Netzwerkstruktur, die es erlaubt, Anfragen in wenigen Weiterleitungsschritten zu verarbeiten, zu effizienten Systemen. Im Folgenden werden Hypothesen formuliert, die Aspekte menschlicher

Strategien überprüfen sollen, die das Gemeinwohl im System beeinflussen. Die Hypothesen sind dafür in drei Kategorien unterteilt: Hypothesen, die sich auf die Kontaktwahl der Subjekte beziehen (Abschnitt 8.3.1), Hypothesen, die Aspekte der Interaktionswahl der Teilnehmer testen (Abschnitt 8.3.2), und Hypothesen, die sich sowohl auf die Kontaktwahl als auch auf die Interaktionswahl der Subjekte beziehen (Abschnitt 8.3.3).

8.3.1 Hypothesen zur Kontaktwahl

Eine fundamentale Frage der Netzwerkbildung ist: Wie wählen Teilnehmer ihre Kontakte? Um diese Frage zu beantworten, ist zu klären, welche Kriterien menschliche Teilnehmer in der Kontaktwahl als wichtig erachten. Wählen die Subjekte kooperative Kontakte oder wählen sie Kontakte mit einer vielversprechenden Position im Netzwerk? Während die erste Frage ein einzelnes Subjekt betrifft, ist die folgende Frage aus der Sicht des gesamten Netzwerks wichtig: Führt die Kontaktwahl der Teilnehmer zu effizienten Netzwerkstrukturen? Die Fragen führen zu den folgenden Hypothesen:

Hypothese 8.3.1 *Der Kooperationsgrad wird in den Kontaktwahl-Strategien häufiger genutzt als andere Kriterien.*

Hypothese 8.3.2 *Die Strategien der Subjekte führen zu effizienten Netzwerken.*

8.3.2 Hypothesen zur Interaktionswahl

Axelrod [Axe06] hat vier Kriterien identifiziert, die zu erfolgreichen Strategien im wiederholten Gefangenendilemma führen.

1. „Sei freundlich“: Eine Strategie sollte in der ersten Interaktion kooperieren und nicht defektieren, um ein kooperatives Verhalten zu signalisieren.
2. „Schlage zurück“: Defektiert ein Kontakt, so soll gegenüber diesem Kontakt ebenfalls nicht kooperiert werden.
3. „Sei nachsichtig“: Kooperiert ein Kontakt, nachdem er zuvor defektiert hat, so soll ebenfalls wieder kooperiert werden. Dies ermöglicht prinzipiell einen Wechsel von gegenseitigem unkooperativen Verhalten hin zu gegenseitiger Kooperation.
4. „Sei nicht neidisch“: Hat ein Spieler eine deutlich höhere Auszahlung, so ist dies allein kein Grund, ihm gegenüber zu defektieren.

Ein Ziel der in diesem Kapitel beschriebenen Studie ist es, zu klären, (a) ob die Subjekte die vier genannten Strategieaspekte nutzen und (b) ob dies zu höheren Auszahlungen führt.

Es sei angemerkt, dass das in dieser Studie genutzte Modell sich vom wiederholten Gefangenendilemma unterscheidet, aber auch einige Gemeinsamkeiten mit diesem hat: Feldman et al. [FLSC04] modelliert die Interaktionen der Endpunkte einer Weiterleitungskette in einem verteilten Peer-to-Peer System mittels des Gefangenendilemmas. Weiterleitungsketten und Kontaktwahl werden dabei nicht berücksichtigt. In diesem Kapitel soll untersucht werden, ob die Ergebnisse aus [Axe06] auch im Netzwerkspiel (vergleiche Abschnitt 7.4) gelten. Die zugehörige Hypothese ist wie folgt:

Hypothese 8.3.3 *Subjekte, die die von Axelrod identifizierten Kriterien in ihrer Strategie – insbesondere in der Interaktionswahl – nutzen, sind erfolgreicher als andere Teilnehmer.*

Basierend auf ihrer eingeschränkten Kenntnis der Netzwerkstruktur, minimieren die Teilnehmer eines verteilten Systems die Anzahl der Weiterleitungen ihrer Anfragen, wenn sie diese an Kontakte senden, die nahe dem Ziel der Anfrage gelegen sind [Kle06]. Dieses Verhalten ist eine Voraussetzung für Effizienz. Die verhaltensökonomischen Experimente des vorherigen Kapitels konnten nicht mit Signifikanz zeigen, dass menschliche Teilnehmer sich in der Tat so verhalten. Eine Ursache mag gewesen sein, dass die untersuchten Systeme zu klein waren und dass damit der gesuchte Effekt ebenfalls zu klein ausfiel, um mit einer hinreichenden Signifikanz gemessen werden zu können.

Hypothese 8.3.4 *Subjekte nutzen die Position eines Kontaktes als Kriterium, wenn sie entscheiden, welchem Kontakt sie eine Anfrage zusenden.*

8.3.3 Hypothesen zur Kombination von Interaktions- und Kontaktwahl

Kontaktwahl erlaubt eine strategische Auswahlmöglichkeit. Menschliche Teilnehmer können unkooperative feste Kontakte dadurch kompensieren, dass sie zusätzliche Kontakte wählen, die diese überbrücken. Ohne zusätzliche Kontakte steigt die Wahrscheinlichkeit, eine Anfrage zu verlieren, mit der Anzahl der Weiterleitungen an (da sich die Wahrscheinlichkeit erhöht, dass die Anfrage auf einen unkooperativen Teilnehmer trifft). Die Anzahl der Weiterleitungen hingegen wächst mit der Distanz im Adressraum, die eine Anfrage überwinden muss, um ihr Ziel zu erreichen. Subjekte können zusätzliche Kontakte nutzen, um große Distanzen im Adressraum zu überbrücken und um damit die Anzahl der Weiterleitungen zu vermindern. Eine Folge

der Nutzung zusätzlicher Kontakte sollte also folglich sein, dass die Wahrscheinlichkeit, eine Anfrage zu verlieren, nicht mehr von der Länge, die diese Anfrage im Adressraum zu überwinden hat, um ihr Ziel zu erreichen, abhängt. Eine weitere Folge ist, dass das Netzwerk – aufgrund geringerer Weiterleitungslast und aufgrund weniger verlorener Anfragen – als Ganzes performanter wird. Daraus folgt die folgende Hypothese:

Hypothese 8.3.5 *Die Erfolgsrate hängt nicht von der Distanz ab, die eine Anfrage vom Erzeuger zum Ziel der Anfrage überbrücken muss, um beantwortet zu werden.*

Subjekte, die mehr zusätzliche Kontakte besitzen als andere, sollten erfolgreicher damit sein, unkooperative Teilnehmer zu überbrücken:

Hypothese 8.3.6 *Subjekte mit mehr zusätzlichen Kontakten als andere Teilnehmer haben eine größere Erfolgsrate als diese.*

8.4 Anforderungen an die Umgebung der Studie

Ein Rahmenwerk, das es erlaubt, die Hypothesen des vorherigen Abschnitts zu testen, muss bestimmte Anforderungen (engl. requirements) erfüllen. Diese werden im Folgenden beschrieben.

Reduzierung störender Einflüsse $R_{\text{Rauschreduzierung}}$ Subjekte, die sich aus dem realen Leben kennen, möchten möglicherweise als Gruppe agieren, um so ihre Auszahlungen zu erhöhen. Obwohl die Analyse von Gruppenstrategien ein interessantes Wissenschaftsfeld ist, ist diese Studie nicht an solchen Strategien interessiert, sondern an dem grundsätzlichen Verhalten einzelner Teilnehmer. Daher ist es wichtig, Gruppenstrategien zu vermeiden oder zu minimieren, da deren Effekte unerwünschtes Rauschen in den Messergebnissen erzeugen kann.

Vergleichbarkeit $R_{\text{Vergleichbarkeit}}$ Der Entwurf des der Studie zugrunde liegenden Spiels muss im Einklang mit den Charakteristika von Sozialen Netzwerken und Peer-to-Peer Systemen stehen. Andernfalls kann es nicht als Modell für solche Systeme genutzt werden. Ein weiterer Aspekt, der die Vergleichbarkeit mit existierenden Netzwerken verfälschen kann, ist das Wissen, dass die Teilnehmer benötigen, um erfolgreiche Strategien verfolgen zu können. Da anzunehmen ist, dass die Subjekte keine Experten für Soziale Netzwerke oder Peer-to-Peer Systeme sind, sollte das Wissen, das nötig ist, um das Spiel zu verstehen, minimal sein.

8.4. ANFORDERUNGEN AN DIE UMGEBUNG DER STUDIE

Motivation der Subjekte $R_{Motivation}$ Zur Beobachtung der Effekte eines großen Netzwerkes ist es nötig, eine möglichst große Anzahl an Teilnehmern für die Studie gewinnen zu können. Daher sollte die Studie so geringe Einstiegskosten wie möglich verursachen. Andernfalls können die Einstiegskosten potentielle Teilnehmer davon abhalten, an der Studie teilzunehmen.

Eine Möglichkeit, um Subjekte zu motivieren, ist es, Anreize zur Teilnahme zu setzen. Neben der Auszahlung, die ein Subjekt durch seine Strategie am Ende der Studie erhält, sind Ranglisten eine übliche Form eines solchen Anreizes. Dennoch muss darauf geachtet werden, dass die gesetzten Anreize sich nicht widersprechen. So sollten die Anreize der Rangliste nicht gegenläufig zur Auszahlungsfunktion laufen. Es muss z.B. verhindert werden, dass Teilnehmer durch die Anreize einer Rangliste von anderen diskriminiert werden.

Beobachtungszeitraum $R_{Beobachtungszeitraum}$ Da es eins der Ziele dieses Kapitels ist, Subjekte über einen langen Zeitraum zu beobachten, sollten die Kosten, die ein Subjekt hat, um das System zu nutzen, so gering wie möglich sein. Die Subjekte sollten in der Lage sein (a) das System unabhängig von ihrem Aufenthaltsort und (b) unabhängig von der Tageszeit zu nutzen.

Gleichzeitigkeit des Zugriffs $R_{Gleichzeitigkeit}$ Unterschiedliche Subjekte nutzen das Rahmenwerk der Studie gleichzeitig. Dennoch darf gleichzeitiger Zugriff nicht zu inkonsistenten Systemzuständen führen.

Dynamisches Netzwerk $R_{Dynamik}$ Die Subjekte entscheiden, wann und ob sie das System verlassen beziehungsweise wann und ob sie dem System beitreten. Daher ist es eine natürliche Anforderung an das Rahmenwerk der Studie, dass es dynamisch wachsen und schrumpfen können muss. Dies gilt insbesondere für das Netzwerk, das die Teilnehmer bilden.

Sicherheit $R_{Sicherheit}$ Es ist bekannt, dass einige Nutzer von Online-Spielen versuchen, das Spiel durch Betrug zu gewinnen beziehungsweise ihren Ranglistenplatz durch Betrug zu verbessern. Daher muss das Rahmenwerk verhindern, (a) dass Nutzer Aktionen durchführen, die sie nach der Definition des Strategieraums nicht durchführen können, und (b) dass sie sich nicht als andere Spieler ausgeben können.

Beispiel: Nutzer *Heinz* hat Sitzungsnummer „123“. Er möchte sich als Nutzer *Otto* ausgeben. Er sendet eine Systemnachricht, die impliziert, dass *Otto Irmgard* als zusätzlichen Kontakt hinzugefügt hat. Die Nutzer-Kontrolleinheit des Rahmenwerks prüft nun, ob Nutzer *Otto* die Sitzungsnummer „123“ hat. Da dies nicht der Fall ist, ignoriert das Rahmenwerk die Systemnachricht. ■

8.4.1 Korrektheit

Um zu überprüfen, ob das Rahmenwerk die Anforderungen Motivation der Subjekte $R_{Motivation}$, Reduzierung störender Effekte $R_{Rauschreduzierung}$ und Vergleichbarkeit $R_{Vergleichbarkeit}$ erfüllt, werden den Hypothesen die drei folgenden Hypothesen hinzugefügt:

1. Motivation der Subjekte C_M : Die Subjekte sind motiviert an der Studie teilzunehmen.
2. Pseudonymität C_P : Die Subjekte können sich nicht gegenseitig identifizieren.
3. Verständnis C_U : Die Subjekte verstehen die Spielregeln und das Ziel des Spieles.

Diese Korrektheitskriterien werden in Abschnitt 8.8 überprüft.

8.5 Rahmenwerk des PostGames

Dieser Abschnitt beschreibt das Rahmenwerk, das für die Studie genutzt wird.

8.5.1 Dienstmerkmale

Webbasiert Um Anforderung $R_{Beobachtungszeitraum}$ (d.h. geringe Nutzungskosten zur Erreichung einer lang anhaltenden Nutzung) gerecht zu werden, nutzt das PostGame eine Client-Server Architektur. Der Client ist webbasiert und ist mit Hilfe des Google Web Toolkits in JavaScript geschrieben. Das bedeutet, dass ein Subjekt, das an der Studie teilnehmen möchte, dafür lediglich einen aktuellen Browser benötigt, was im Allgemeinen als gegeben vorausgesetzt werden kann. Die Subjekte können also ohne große Einstiegskosten an der Studie teilnehmen. Der Zugriff auf das PostGame erfolgt über das Internet und ist damit von (praktisch) jedem Ort aus und zu jeder Zeit gesichert.

Leicht verständlich Das PostGame verbirgt das Problem des Findens sinnvoller Strategien in der Anfragebearbeitung verteilter Systeme hinter einem Spiel. Durch diese Maßnahme wird vermieden, dass die Subjekte Wissen über Soziale Netzwerke oder Peer-to-Peer Systeme benötigen (vergleiche Anforderung $R_{Vergleichbarkeit}$). Des Weiteren muss jeder Teilnehmer der Studie eine online-Übung durchlaufen, in der der Strategieraum und die Auszahlungsfunktion mittels Aufgaben erklärt werden, die das Subjekt lösen muss. Durch die online-Übung wird sichergestellt, dass die Teilnehmer den Strategieraum verstanden haben.

Dynamisches Netzwerk Die Teilnehmer der Studie können das Netzwerk jederzeit nach eigenem Ermessen verlassen oder betreten. Dieses Dienstmerkmal entspricht Anforderung $R_{Dynamik}$. Einem Subjekt i , das dem Netzwerk beiträgt, wird durch das Rahmenwerk eine zufällige Position im Adressraum zugeteilt. Das dynamische Anwachsen und Schrumpfen der Netzwerkstruktur folgt den Algorithmen, die in [RFH⁺01] beschrieben werden. Das Rahmenwerk teilt eine zufällig ausgewählte Adressraum-Zone z und weist i eine Hälfte, d.h. z_2 , zu. Die andere Hälfte, d.h. z_1 , bleibt unter der Kontrolle des Subjektes, das ursprünglich für die gesamte Zone verantwortlich war. Es sei angemerkt, dass ein Subjekt zu jedem Zeitpunkt nur eine Zone verwaltet und dass jede Zone nur von einem Subjekt verwaltet wird. Sollte Subjekt i das System verlassen, so werden alle Anfragen von i durch das System gelöscht. Des Weiteren wird die Zone von Subjekt i , d.h. z_2 , mit Zone z_1 verschmolzen. (Sollte z_1 in der Zwischenzeit in weitere Zonen, z.B. in z_{11} und z_{12} , geteilt worden sein, so übernimmt ein Subjekt, z.B. das, welches vorher z_{11} verwaltet hat, z_2 und das andere z_1 .)

Die hier verwendeten Algorithmen sorgen dafür, dass die Größe der Zonen nur um einen kleinen Faktor variieren [RFH⁺01]. Die daraus folgenden Ungleichheiten bei der Anzahl der zu beantwortenden Anfragen sind – im Gegensatz zu den Experimenten in Kapitel 7 – hier erwünscht, um eine möglichst realistische Spielumgebung zu erhalten.

Pseudonymität Um Gruppen-Strategien zu verhindern, ordnet das PostGame-Rahmenwerk allen Teilnehmern so Pseudonyme zu, dass sie dadurch gegenüber anderen Teilnehmern anonym werden. Im Einzelnen: Für jeden Kontakt c eines Subjekts i , wird für Subjekt i ein künstlicher Name, d.h. ein Pseudonym, erzeugt, d.h. es wird folgende Abbildung durchgeführt $i \rightarrow i'$, so dass $i' \notin N$ gilt und dass kein Pseudonym zweimal auftritt. Das Rahmenwerk garantiert also, dass zwei Subjekte, die den selben Kontakt haben, ihn mit zwei unterschiedlichen Pseudonymen wahrnehmen und daher nicht durch Absprache klären können, ob sie den selben Kontakt haben. Durch diesen Mechanismus wird sichergestellt, dass die Subjekte kein Wissen dadurch erlangen können, dass sie die Namen ihrer Kontakte austauschen. Dieses Dienstmerkmal folgt der Anforderung $R_{Rauschreduzierung}$.

Rangliste Das Rahmenwerk des PostGames nutzt eine Rangliste, um die Subjekte zu motivieren (vergleiche Anforderung $R_{Motivation}$). Obwohl viele Online-Spiele Ranglisten verwenden, um ihre Spieler zu motivieren, muss dies mit Vorsicht getan werden. Ranglisten führen zu zwei Problemen:

1. Eine Rangliste gibt Teilnehmern einen Anreiz, andere Spieler zu diskriminieren, da ein Subjekt durch die Tatsache, dass andere Subjekte eine geringere Auszahlung erhalten, relativ gesehen – also im Ranglistenplatz – besser dasteht. Um diesen Effekt zu verringern, werden einem Subjekt nur die Punkte-

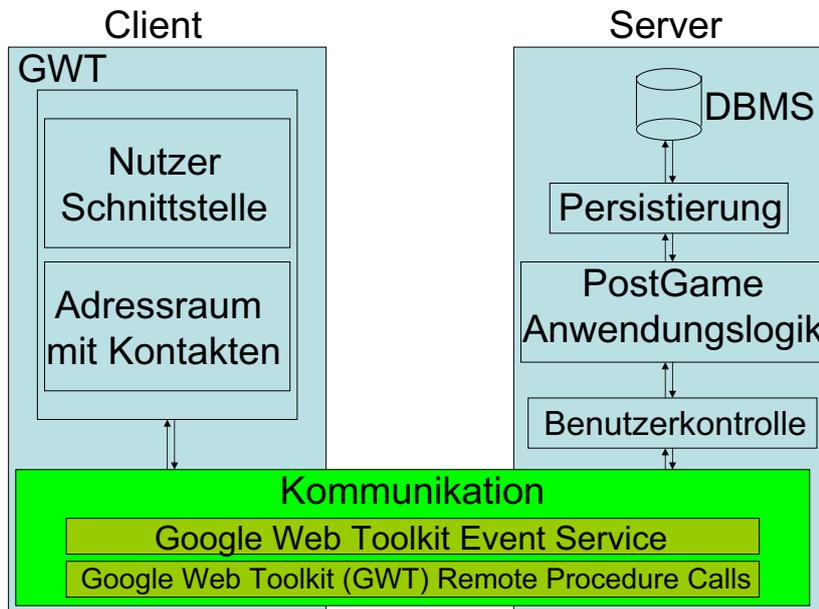


Abbildung 8.1: Architektur des PostGame-Rahmenwerks

stände der Ranglistenplätze gezeigt, die höchstens zwei Ranglistenplätze entfernt sind. Zum Beispiel wird dem Subjekt mit Ranglistenplatz 18 angezeigt, welche Auszahlung die Subjekte auf den Ranglistenplätzen 16, 17, 19 und 20 besitzen. Die Namen der Subjekte sind in der Rangliste durch Fragezeichen „???“ unkenntlich gemacht. (Außer dem Punktestand werden keine weiteren Informationen angezeigt).

2. Die Subjekte, die am Anfang oder am Ende der Rangliste stehen, ändern eventuell ihr Verhalten, da sie denken könnten, dass sie sich nicht mehr besser (oder schlechter) verhalten können. Um diesem Problem zu begegnen, werden am Ende und am Anfang je zwei erfundene Subjekte eingefügt.

Die beiden Maßnahmen verhindern oder reduzieren die negativen Effekte einer Rangliste und erlauben es, diese als Motivation für die Spieler einzusetzen.

8.5.2 Architektur

Das PostGame nutzt eine Client-Server Architektur (vergleiche Abbildung 8.1).

Der Client (auch als Dienstnehmer bezeichnet) visualisiert die Ausgaben und verarbeitet die Eingaben des Nutzers. Der Client ist mit Hilfe des Google-Web Toolkit (GWT) geschrieben und kann dadurch mit einem aktuellen Browser genutzt werden, ohne dass dafür die Installation von zusätzlicher Software nötig ist. Dieser Entwurf stimmt mit der Anforderung $R_{\text{Beobachtungszeitraum}}$ überein.

8.5. RAHMENWERK DES POSTGAMES

Der Server (auch als Dienstgeber bezeichnet) ist in Java geschrieben und verwaltet die Applikationslogik. Die Applikationslogik hat drei Hauptbestandteile: die *Nutzerkontrolle*, die *Anwendungslogik* und die *Persistierung der Daten*.

Nutzerkontrolle Es ist davon auszugehen, dass einige Nutzer des PostGames ihre Auszahlung dadurch verbessern wollen, dass sie betrügen. Daher validiert der Server die gesamten Nutzereingaben. Zum Beispiel werden alle Zeichenketten überprüft, um SQL Injektion Angriffe sowie Cross-Site-Scripting Angriffe zu verhindern. Des Weiteren überprüft die Nutzerkontroll-Einheit weitere Eigenschaften, um Betrug durch die Nutzer zu verhindern. Im Einzelnen:

- Ein Subjekt kann nur diejenigen Anfragen verarbeiten oder fallenlassen, die das Subjekt erhalten oder erzeugt hat und die es nicht bereits verarbeitet oder fallengelassen hat.
- Es wird überprüft, ob das Subjekt, an das eine Anfrage weitergeleitet werden soll, wirklich ein Kontakt des Subjektes ist.
- Ein Subjekt kann nur solche Subjekte als zusätzliche Kontakte hinzufügen, die potentielle Kontakte sind.
- Die Nutzerkontroll-Einheit überprüft, ob nur solche Subjekte als Kontakt fallengelassen werden, die (vorher) zusätzliche Kontakte sind.
- Durch die Verwendung langer, zufälliger Sitzungsnummern (und deren Abgleich) wird verhindert, dass Teilnehmer die Identität anderer Teilnehmer annehmen.

Die genannten Maßnahmen erfüllen Anforderung $R_{Sicherheit}$.

Anwendungslogik Die PostGame-Anwendungslogik-Einheit kapselt den Programmcode, der für das Netzwerkspiel selbst nötig ist. Er führt die Aktionen durch, die durch die Subjekte angestoßen werden und sammelt statistische Daten über das Verhalten der Nutzer.

Persistierung Der Server persistiert – also speichert dauerhaft und robust – alle Daten mit Hilfe eines Datenbank-Verwaltungssystems (engl. data-base management system) (DBMS). Der Zugriff auf die genutzte Datenbank erfolgt ausschließlich über Transaktionen um sicherzustellen, dass gleichzeitiger Zugriff der Subjekte nicht zu Inkonsistenzen führt. Dadurch wird Anforderung $R_{Gleichzeitigkeit}$ erfüllt.

Kommunikation Die Kommunikation zwischen den Client-Programmen und dem Server wird mittels GWT remote procedure calls und dem GWT event service realisiert.

KAPITEL 8. POSTGAME – EINE STUDIE

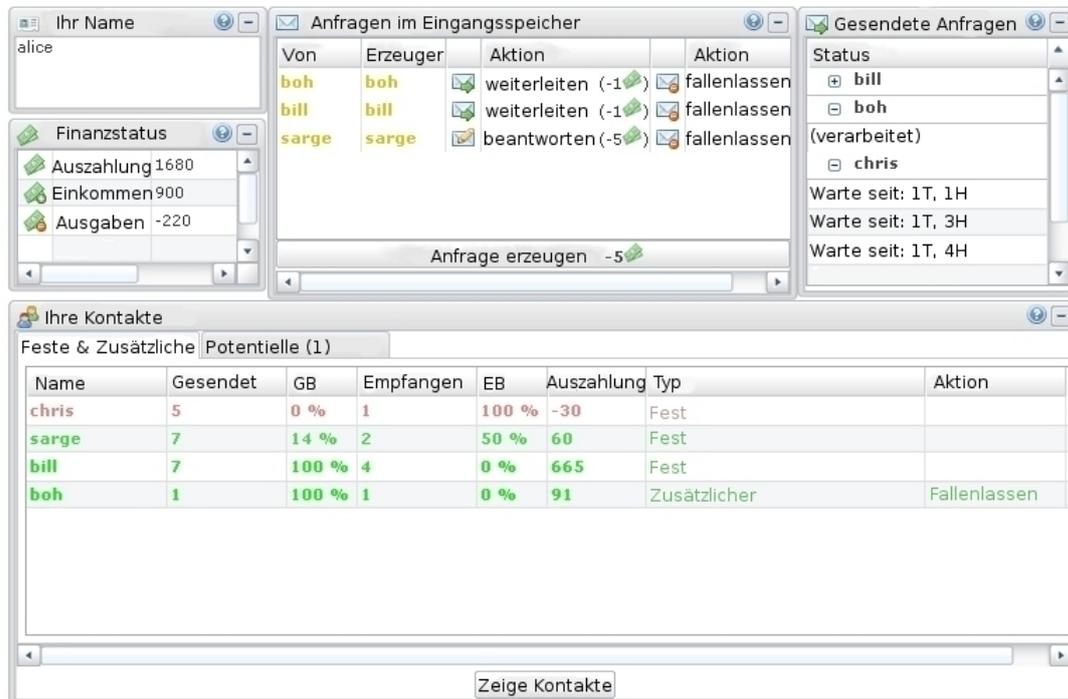


Abbildung 8.2: PostGame Client

8.5.3 Nutzerschnittstelle

Der Unterschied zwischen der Benutzerschnittstelle des PostGames und der in Abschnitt 7.4.1 vorgestellten Benutzerschnittstelle ist, dass für die zweite Benutzerschnittstelle die Einstiegskosten höher sind: Durch den Nutzer müsste Software installiert werden (z.B. eine Java-Laufzeitumgebung). Ansonsten unterscheiden sich die Benutzerschnittstelle im PostGame in ihrer Funktionalität im Wesentlichen nicht von der Benutzerschnittstelle in Abschnitt 7.4.1. Der Vollständigkeit halber wird die Benutzerschnittstelle des PostGames an dieser Stelle dennoch beschrieben.

Die Subjekte interagieren miteinander mittels des PostGame Clients. Dieser ist in den Abbildungen 8.2 und 8.3 abgebildet. Abbildung 8.2 zeigt die Benutzerschnittstelle und Abbildung 8.3 zeigt eine Karte der Kontakte im Adressraum. Beide Abbildungen zeigen die Situation in einem kleinen System bestehend aus sechs Subjekten.

Der Client besteht aus fünf Sektionen: der Namens-Sektion, dem Finanzstatus, der Eingangsspeicher-Sektion, der Gesendete-Anfragen-Sektion und der Kontakt-Sektion.

Namens-Sektion Oben links ist der Name des Subjekts zu sehen. In diesem Fall ist es der Name „alice“.

8.5. RAHMENWERK DES POSTGAMES

Finanzstatus Unter der Namens-Sektion befindet sich der Finanzstatus: Er zeigt die Auszahlung sowie die Einnahmen und Ausgaben des Subjekts.

Eingangsspeicher Rechts vom Finanzstatus ist die Eingangsspeicher-Sektion. Sie zeigt die Anfragen, die ein Subjekt verarbeiten kann. Im Beispiel in Abbildung 8.2 hat das Subjekt *alice* drei Anfragen. Je eine von „*boh*“ und „*bill*“, welche *alice* weiterleiten oder fallenlassen kann. Eine weitere Anfrage ist von „*sarge*“. Diese kann *alice* entweder beantworten oder fallenlassen, da *alice* selbst das Ziel der Anfrage ist. Wenn ein Subjekt eine Anfrage erzeugen möchte, so muss lediglich der Knopf am unteren Ende der Eingangsspeicher-Sektion gedrückt werden. Dann erscheint die entsprechende Anfrage in der Eingangsspeicher-Sektion. Das Erzeugen einer Anfrage kostet 5 Punkte, das Versenden derselben weitere 5 Punkte. Es sei angemerkt, dass ein Subjekt die Möglichkeit hat, eigene Anfragen nach der Erzeugung fallenzulassen. Der Sinn dahinter ist, dass es so möglich wird, erst zu überprüfen, ob das Ziel der Anfrage kooperativ ist oder nicht. Sollte ein Subjekt wissen, dass das Ziel einer erzeugten Anfrage nicht kooperativ ist, so macht es keinen Sinn, die Anfrage zu senden.

Gesendete-Anfragen Die Sektion rechts oben, d.h. die Gesendete-Anfragen-Sektion, zeigt den Status bereits versendeter Anfragen. Der Status kann „verarbeitet“ sein, wenn das Subjekt bereits eine Antwort zu der entsprechenden Anfrage erhalten hat, oder „Warte seit: x T, y H“, wenn das Subjekt bereits seit x Tagen und y Stunden auf die Antwort wartet. Es sei angemerkt, dass ein Subjekt nicht weiß, ob eine Anfrage bearbeitet wird oder ob sie fallengelassen wurde.

Kontakt-Sektion Am unteren Ende des Clients ist die Kontakt-Sektion. Diese zeigt Statistik über die Kontakte (erster Reiter) und über potentielle Kontakte (zweiter Reiter) an. Die erste Spalte zeigt die (pseudonymisierten) Namen der Kontakte, die zweite die Anzahl an eigenen Anfragen, die diesem Kontakt gesendet wurden und die dritte Spalte zeigt den Anteil dieser Anfragen, der bereits beantwortet wurde. Die vierte und fünfte Spalte zeigen die Anzahl an Anfragen, die das Subjekt, d.h. *alice*, von dem Kontakt empfangen hat, und den Anteil, den das Subjekt bisher bearbeitet hat. Es folgt in der sechsten Spalte die Auszahlung, die durch den Kontakt verursacht wurde, und in der siebten der Typ des Kontaktes. Mit Hilfe der letzten Spalte können zusätzliche Kontakte fallengelassen werden. Der zweite Reiter zeigt die gleiche Statistik für potentielle Kontakte, sofern sie vorhanden sind.

Es sei angemerkt, dass die Spalte mit der Auszahlung, die durch einen Kontakt erzeugt wird, aus einer ökonomischen Perspektive heraus wichtig ist: Ein Kontakt ist nur dann nützlich, wenn er eine positive Auszahlung erzeugt.

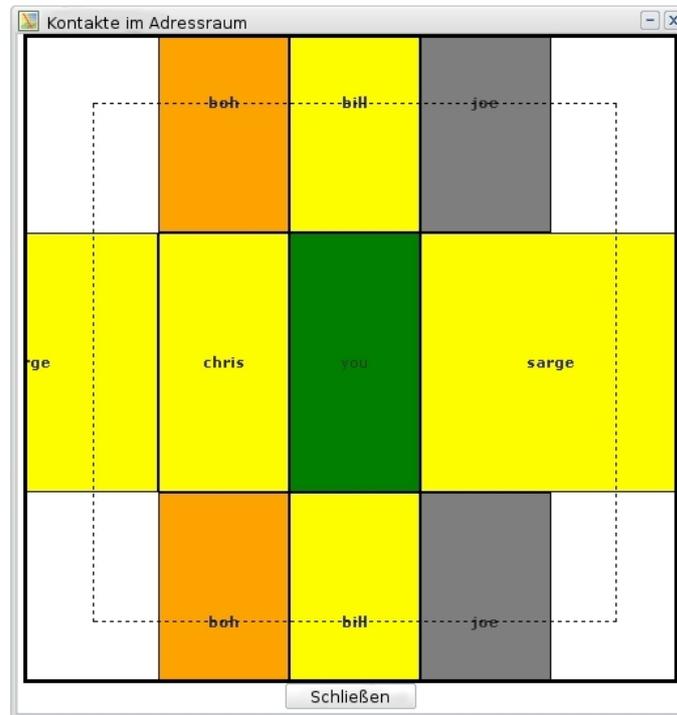


Abbildung 8.3: Adressraum mit Kontakten

Kontakte im Adressraum Abbildung 8.3 zeigt das System aus der Sicht des Subjekts. Das Fenster zeigt Kontakte, zusätzliche Kontakte und die letzten zehn potentiellen Kontakte, mit denen das Subjekt interagiert hat, im Adressraum. Es sei angemerkt, dass die Darstellung aller potentiellen Kontakte zu einer Situation führt, in der die Subjekte eine globale Sicht auf das System haben – allerdings, ohne dafür zu zahlen, was im Fall zusätzlicher Kontakte zwingend erforderlich ist. In anderen Worten, die kostenlose globale Sicht auf das Netzwerk wird dadurch verhindert, dass nur die letzten zehn potentiellen Kontakte angezeigt werden.

In Abbildung 8.3 repräsentieren Rechtecke die Zonen der Subjekte. Das grüne Rechteck repräsentiert das Subjekt, also *alice* in diesem Fall. Die Kontakte von *alice* sind gelb, der zusätzliche Kontakt „boh“ wird durch ein oranges Rechteck repräsentiert. Potentielle Kontakte sind grau. Wenn das Subjekt seinen Mauszeiger über ein Rechteck bewegt, dann wird die selbe Statistik wie in der Kontakt-Sektion gezeigt. Durch Anklicken eines Rechtecks, das einen zusätzlichen Kontakt oder einen potentiellen Kontakt repräsentiert, öffnet sich ein Dialogfenster. Mit diesem Dialogfenster kann ein zusätzlicher Kontakt gewählt oder fallengelassen werden.

8.6 Messwerte

Während der Studie werden die folgenden Messwerte durch direkte Beobachtungen des Verhaltens erhoben.

Anzahl an Subjekten Eine Anforderung ist, die Subjekte zu motivieren, an der Studie teilzunehmen (vergleiche Anforderung $R_{Motivation}$). Um Rückschlüsse auf die Motivation der Teilnehmer zu erhalten, wird die Anzahl der Teilnehmer an der Studie über die Zeit gemessen. Wenn die Subjekte ihre Motivation verlieren, verlassen sie das System. Wenn sie das System nicht verlassen, dann ist ihre Motivation mindestens ausreichend. Es sei angemerkt, dass ein Teilnehmer, der sich über drei Tage nicht in das System einloggt, ebenfalls aus dem System ausscheidet.

Auszahlung Der Erfolg der Strategie eines Subjektes spiegelt sich in der Auszahlung wieder. Daher wird diese gemessen. Des Weiteren wird der Verlauf der Erfolgsrate über die Zeit gespeichert, um einen fein-granularen Einblick darin zu erhalten, wie erfolgreich die Subjekte sind, Antworten auf ihre Anfragen zu erhalten. Wie in Abschnitt 8.8 gezeigt werden wird, existieren Situationen, in denen die Subjekte eine hohe Erfolgsrate haben, aber nur eine durchschnittliche Auszahlung.

Erfolgsrate Die Erfolgsrate wird über die Zeit gemessen. Zusätzlich wird sie (a) in der Relation zur Entfernung, die eine Anfrage zu ihrem Ziel zurücklegen muss, gemessen und (b) in Abhängigkeit zur Anzahl an zusätzlichen Kontakten des Teilnehmers. Punkt (a) ist wichtig, um Hypothese 8.3.5, d.h. die Nützlichkeit zusätzlicher Kontakte, zu beantworten. Punkt (b) ist notwendig, um zu überprüfen, ob zusätzliche Kontakte einen positiven Effekt auf die Erfolgsrate haben.

Anzahl an Kontakten Das Rahmenwerk misst die Anzahl an Kontakten pro Subjekt. Dies ist Voraussetzung, um Hypothese 8.3.6 zu überprüfen. Hypothese 8.3.6 testet, ob die Anzahl an Kontakten die Erfolgsrate beeinflusst.

Anzahl an Weiterleitungen Die durchschnittliche Anzahl an Weiterleitungen ist ein wichtiges Maß, um die Performanz des Netzwerks zu verstehen. Es wird genutzt, um zu zeigen, ob durch die Strategien der Subjekte effektive Netzwerkstrukturen entstehen (vergleiche Hypothese 8.3.2).

8.7 Experimentaufbau

Subjekte Die Subjekte wurden mit Hilfe der Travian GmbH rekrutiert. Die Travian GmbH ist auf Online-Spiele spezialisiert und hat Kontakt zu einer großen Gruppe

KAPITEL 8. POSTGAME – EINE STUDIE

von online-Spielern. Die Studie wurde in einem der Foren der Travian GmbH bekanntgemacht und 193 Subjekte konnten rekrutiert werden. 53% der Subjekte sind weiblich und 47% männlich. Im Durchschnitt sind die Teilnehmer 27 Jahre alt.

Motivation Es wurden zwei Techniken genutzt, um die Subjekte zu motivieren.

1. Jedes Subjekt konnte seine Auszahlung in einer Rangliste mit anderen Subjekten vergleichen.
2. Alle Subjekte erhielten eine (reale) Auszahlung relativ zu ihrem Punktestand (d.h. relativ zu $Auszahlung(i)$). Die Subjekte wurden in Travian Nuggets bezahlt, einer geldwerten Travian GmbH Währung. Im Durchschnitt verdienten die Subjekte 3 €.

(Die Probleme, die durch Methode (1) hervorgerufen werden, wurden bereits in Abschnitt 8.5.1 diskutiert.)

Messtechniken In der Studie wurden zwei Messtechniken verwendet: (a) direkte Beobachtungen und (b) eine Umfrage am Ende der Studie.

(a) Das PostGame Rahmenwerk beobachtet und speichert alle Aktionen jedes Subjektes für die Analyse der Strategien. Des Weiteren wurde eine Umfrage während der Studie durchgeführt. Dafür wurden zufällig 50% der Teilnehmer einer Umfrage-Gruppe zugelost. (Es sei angemerkt, dass nur 50% der Subjekte zugelost wurden, um zu überprüfen, ob die Umfrage während der Studie das Verhalten der Teilnehmer im Verhältnis zu den anderen 50% signifikant beeinflusst. Dies ist nicht der Fall.) Die Subjekte der Umfrage-Gruppe wurden nach jeder ihrer Aktionen mit einer 10%-tigen Wahrscheinlichkeit gefragt, warum sie diese Aktion durchgeführt haben. Die Subjekte konnten dabei unter den folgenden Antworten wählen: „Kooperationsgrad“, „Position“, „Kooperationsgrad und Position“, „zufällig“, „anderer Grund“ und „möchte nicht antworten“. Es sei angemerkt, dass das Rahmenwerk die Antworten stets in einer zufälligen Reihenfolge angezeigt hat, damit keine Antwort durch ihre Position bevorzugt wird.

(b) Während man die Effekte einer Strategie direkt beobachten kann, ist es nicht möglich, eine Strategie zu beobachten. Daher wurden die Subjekte nach dem Ende der Studie mit Hilfe einer online-Umfrage über ihre Strategien befragt. Von den 193 Subjekten, die am PostGame teilgenommen haben, haben 82 die Umfrage begonnen. 72 Subjekte beendeten die Umfrage. Um zu verstehen, wie wichtig (oder unwichtig) diese 72 Subjekte für das Gesamtsystem sind, wurde analysiert, wie erfolgreich und wie aktiv sie sind. Die Subjekte, die an der Umfrage teilgenommen haben, haben eine höhere Auszahlung als jene, die nicht an der Umfrage teilgenommen haben. Dies gilt auch für die Auszahlung pro erzeugter Anfrage. Die Analyse der Subjekte zeigt außerdem, dass sie für 85% der Interaktionen und für 74% der Kontaktwahl-Aktionen verantwortlich sind. Das zeigt, dass die Subjekte, die an der Umfrage am

Strategie	Kontakte wählen	Kontakte fallenlassen
Unabhängig	42	9
Kooperationsgrad (K.)	14	36
K. + Auszahlung		6
K. + Position	2	1
Auszahlung	3	10
Position	6	1
Anderweitige Strategie	4	2

Tabelle 8.2: Umfrage nach Ende der Studie – Kontaktwahl Strategien

Ende der Studie teilgenommen haben, sowohl erfolgreiche Strategien repräsentieren als auch für die Mehrzahl der Aktivitäten im System verantwortlich sind.

8.8 Evaluierung

Die Evaluierung zeigt, dass sich die Subjekte kooperativ verhalten. Sie verarbeiten 78% der Anfragen und erzeugen ein Netzwerk, das es erlaubt, Anfragen in nur wenigen Schritten zu verarbeiten.

8.8.1 Ergebnisse – Hypothesen

Die Studie wurde in einem Zeitraum von 67 Tagen durchgeführt. Insgesamt nahmen 193 Subjekte an der Studie teil, wobei im Durchschnitt 52 Teilnehmer pro Tag online waren.

Hypothese 8.3.1 Tabellen 8.2 und 8.3 zeigen die Ergebnisse der Umfrage am Ende der Studie und der Umfrage während der Studie bezüglich der Kontaktwahl-Strategien der Teilnehmer. Während die Zeilen zeigen, wie viele Subjekte eine bestimmte Strategie wählen, zeigen die Spalten (1) welche Strategie die Teilnehmer beim Wählen zusätzlicher Kontakte nutzen und (2) welche Strategie beim Fallenlassen zusätzlicher Kontakte gewählt wird.

Tabelle 8.2 zeigt, dass kein besonderes Motiv in der Kontaktwahl existiert, d.h. die Subjekte wählen Kontakte aus unterschiedlichen Gründen. Die Mehrheit von 42 (aus 72) Subjekten gibt an, dass sie Kontakte unabhängig von ihrem Verhalten oder ihrer Position gewählt hat. Allerdings prägt sich eine Strategie beim Fallenlassen von zusätzlichen Kontakten heraus: Die Subjekte lassen unkooperative zusätzliche Kontakte fallen. Dies bestätigt Hypothese 8.3.1.

Die Daten, die während der Studie erhoben wurden, d.h. die Daten in Tabelle 8.3, zeigen, dass eine Tendenz existiert, die Position eines zusätzlichen Kontaktes bei der

KAPITEL 8. POSTGAME – EINE STUDIE

Strategie	Kontakte wählen	Kontakte fallenlassen
Zufällig	46	4
Kooperationsgrad (K.)	39	35
K. + Position	21	1
Position	84	3
Anderweitige Strategie	47	13

Tabelle 8.3: Umfrage während der Studie – Kontaktwahl Strategien

Kontaktwahl zu berücksichtigen. In 44,3% der Kontaktwahl-Strategien ist die Position des Kontakts bei der Kontaktwahl zumindestens teilweise von Bedeutung. Es ist überraschend, dass die Subjekte der Position eines zusätzlichen Kontaktes während der Studie Bedeutung beimessen, während sie dies nach der Studie nicht mehr tun. Eine Erklärung ist, dass sie bei der Umfrage während der Studie stets zu einer konkreten Aktion befragt worden sind, d.h. zu einem konkreten zusätzlichen Kontakt. In der Umfrage am Ende der Studie haben sie hingegen ihre generelle Strategie beschrieben.

Die Ergebnisse der Umfragen beinhalten eine weitere überraschende Beobachtung: Die Auszahlung, die durch einen zusätzlichen Kontakt hervorgerufen wird, scheint keine bedeutende Rolle in der Kontaktwahl-Strategie zu spielen. Nur 25% der Subjekte berücksichtigen dieses Maß, wenn sie zusätzliche Kontakte fallenlassen. Beim Hinzufügen von zusätzlichen Kontakten ist die Anzahl noch geringer: 4%. Diese Beobachtung ist überraschend, da die Nützlichkeit eines – eventuell wenig kooperativen – zusätzlichen Kontaktes, der eine große Auszahlung hervorruft, höher sein sollte als die Nützlichkeit eines vollständig kooperativen zusätzlichen Kontaktes, der wenig oder gar keine Auszahlung erzeugt. Dieses Verhalten kann die Effizienz des Systems beeinflussen.

Hypothese 8.3.2 Die Strategien der Subjekte führen zu einer Netzwerkstruktur mit einer hohen Performanz: Die Messungen der Anzahl an Weiterleitungen pro Anfrage ergaben einen Wert von 1,4. Zum Vergleich siehe Abbildung 8.4. Diese zeigt die Anzahl an Weiterleitungen in Relation zur Entfernung im Adressraum, die eine Anfrage zurücklegen muss. Die gestrichelte waagerechte Linie zeigt den Durchschnitt, d.h. 1,4 Weiterleitungen. Dieser Wert ist nahe des Wertes, der im Stern-Netzwerk auftreten würde. Im Stern-Netzwerk müssen alle Anfragen (außer die vom Zentrum des Sterns ein- und ausgehenden) einmal – nämlich vom Zentrum des Sterns – weitergeleitet werden. Dennoch bilden die Subjekte keinen Stern: Hypothese 8.3.2 kann nicht bestätigt werden.

Hypothese 8.3.3 Abbildung 8.5 zeigt die Ergebnisse der Umfrage am Ende der Studie. Die X-Achse zeigt Kombinationen aus den Kriterien von Axelrod, die Y-

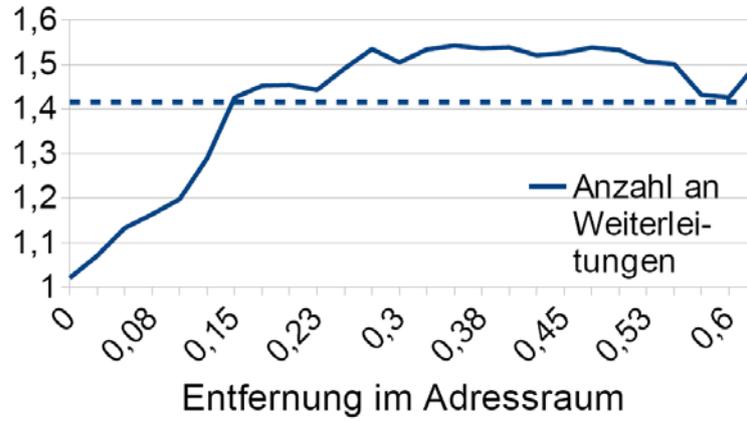


Abbildung 8.4: Anzahl an Weiterleitungen relativ zur Adressraum-Distanz der Anfrage

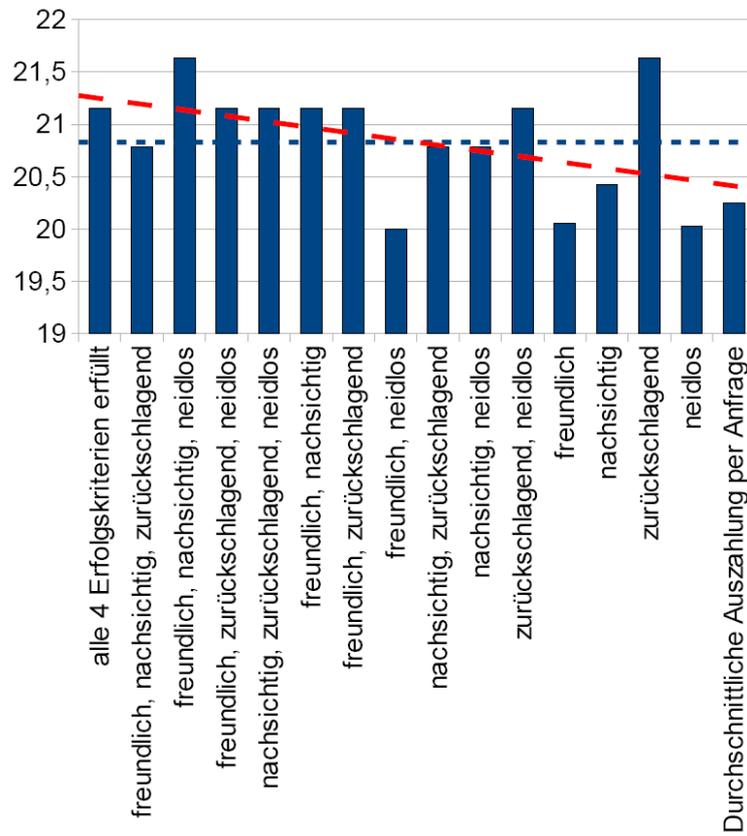


Abbildung 8.5: Auszahlung pro Anfrage relativ zur Anzahl an verwendeten Axelrod-Kriterien

KAPITEL 8. POSTGAME – EINE STUDIE

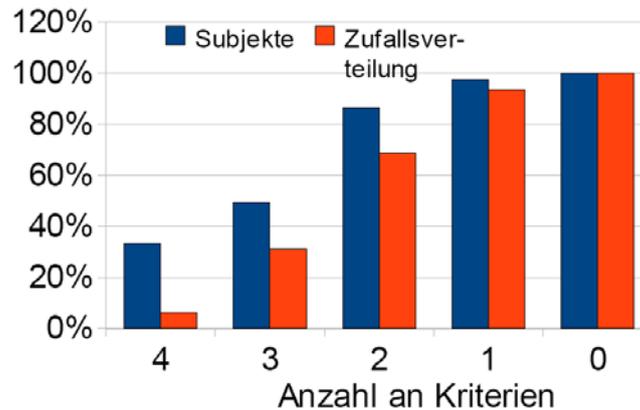


Abbildung 8.6: Kumulativer Anteil der genutzten Axelrod-Kriterien

Achse zeigt die durchschnittliche Auszahlung (pro Anfrage), die Subjekte mit der entsprechenden Strategie-Kombination hatten. Die fein-gestrichelte blaue Linie zeigt den Durchschnitt der Balken an (also den Durchschnitt über die Durchschnitte der Auszahlungen der entsprechenden Strategie-Kombinationen pro Anfrage). Die grob-gestrichelte rote Linie zeigt eine lineare Regression der Werte: Strategien von Subjekten, die mehrere der Kriterien von Axelrod benutzt haben, sind erfolgreicher als Teilnehmer, die weniger Kriterien verwendet haben. Dies bedeutet, dass ein Trend, der Hypothese 8.3.3 unterstützt, beobachtet werden kann. Dennoch sei angemerkt, dass diese Beobachtung aufgrund der Natur der Studie nicht signifikant ist und dass sie daher nur eine begrenzte Aussagekraft hat. So hängt z.B. die Neigung der linearen Regression von der Reihenfolge der Kriterien ab und ist daher mehr als Beobachtung zu werten und weniger als signifikantes Ergebnis.

Abbildung 8.6 zeigt (den kumulativen Anteil) der Subjekte, die x oder mehr Kriterien von Axelrod genutzt haben (rote Balken). Des Weiteren zeigt die Abbildung zum Vergleich den Anteil, den die Subjekte gewählt hätten, wenn sie die Kriterien zufällig (und mit gleicher Wahrscheinlichkeit) gewählt hätten. Da die Teilnehmer Axelrods Kriterien häufiger nutzen als es mit einem zufälligen Verhalten erklärt werden kann, ist es evident, dass die Teilnehmer die Kriterien von Axelrod als nützlich erachten.

Hypothese 8.3.4 Tabelle 8.4 zeigt die Umfrageergebnisse am Ende der Studie in Bezug auf die Interaktionswahl-Strategien der Teilnehmer. Sie zeigt, dass die Subjekte die Position eines Kontaktes berücksichtigen, wenn sie ihre eigenen – also von ihnen erzeugte – Anfragen weiterleiten. Dieses Ergebnis unterscheidet sich von den Ergebnissen in Kapitel 7. Dort spielt der Kooperationsgrad eine stärkere Rolle. Eine Erklärung dafür ist, dass die Netzwerke, die in Kapitel 7 untersucht worden sind, kleiner waren und dass daher die Position eine geringere Rolle gespielt hat. Es sei an-

Strategie	Erzeugte Anfragen	Weiterleiten	Antworten
Unabhängig	3	24	31
Kooperationsgrad (K.)	6	21	18
K. + Position	21		1
K. + Position + Auszahlung	1		
Position	34		3
Auszahlung	2		5
Kontakt		1	2
Kontakt + K.		3	3

Tabelle 8.4: Interaktionswahl-Strategien

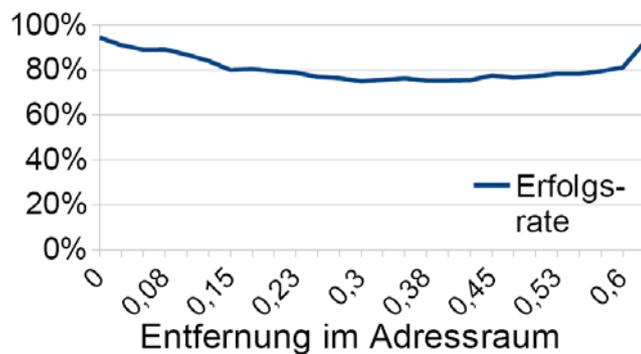


Abbildung 8.7: Erfolgsrate relativ zur Adressraum-Distanz der Anfrage

gemerkt, dass die Subjekte bei der Interaktionswahl nicht die Auszahlung, die durch einen Kontakt erzeugt wurde, berücksichtigen: Während 49% der Subjekte beim Weiterleiten den Kooperationsgrad ihrer Kontakte berücksichtigen, berücksichtigen 0% die Auszahlung durch den Kontakt. Beim Beantworten von Anfragen sind die Ergebnisse wie folgt: 35% berücksichtigen den Kooperationsgrad und nur 8% die Auszahlung. Beides ist überraschend, da die Maximierung der Auszahlung vermuten lassen würde, dass die Auszahlung, die ein Kontakt erzeugt, ein wichtiges Kriterium ist. Dabei gilt zu berücksichtigen, dass die Subjekte – wie bereits in Abschnitt 8.5.3 erwähnt – durch den Client angezeigt bekommen, wie viel Auszahlung ein Kontakt erzeugt. Das bedeutet, dass die Subjekte sich über dieses Maß im Klaren sein sollten, und es kann ausgeschlossen werden, dass die Subjekte diese Strategie nicht nutzen, da sie nicht wissen, wie viel Auszahlung ein Kontakt bei ihnen erzeugt hat.

Hypothese 8.3.5 Abbildung 8.7 zeigt die Erfolgsrate in Relation zur Distanz im Adressraum, die eine Anfrage vom Erzeuger zu ihrem Ziel zurücklegen muss, um beantwortet zu werden. Für geringe Distanzen ist die Erfolgsrate höher als für größere Distanzen. Dies macht Sinn, da in der Regel für geringe Distanzen im Adressraum

KAPITEL 8. POSTGAME – EINE STUDIE

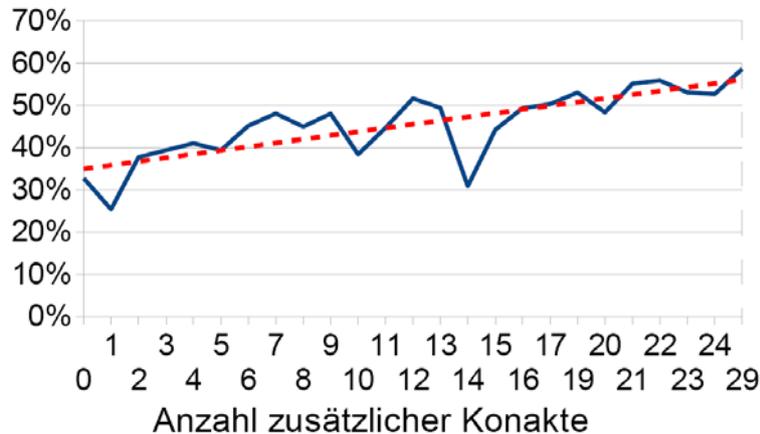


Abbildung 8.8: Erfolgsrate in Abhängigkeit zur Anzahl an zusätzlichen Kontakten

nur wenige Weiterleitungen (im in dieser Arbeit betrachteten Overlay-Netzwerk) benötigt werden (vergleiche Abbildung 8.4). Beim leichten Anstieg der Kurve für die maximale Distanz im Torus kann es sich um einen Ausreißer handeln, da nur wenige Ziele einer Anfrage die maximale Distanz von $\frac{\sqrt{2}}{2}$ haben. Die Erfolgsrate ist immer größer als 75%, unabhängig davon wie groß die Distanz ist, die die Anfrage vom Erzeuger bis zum Ziel zurücklegen muss. Dies unterstützt Hypothese 8.3.5: Die Erfolgsrate ist nicht abhängig von der Distanz im Adressraum.

Hypothese 8.3.6 Abbildung 8.8 zeigt die Nützlichkeit zusätzlicher Kontakte. Die X-Achse ist die Anzahl an zusätzlichen Kontakten und die Y-Achse die Erfolgsrate. Die rote Linie zeigt eine lineare Regression. Wie aus Abbildung 8.8 deutlich wird, haben Subjekte mit einer größeren Anzahl an zusätzlichen Kontakten eine höhere Erfolgsrate. Es sei jedoch angemerkt, dass eine hohe Erfolgsrate nicht zwangsläufig auch zu einer hohen Auszahlung führt. Mit dem in der Studie verwendeten Kosten-Nutzen-Verhältnis (beziehungsweise mit den in der Studie verwendeten Auszahlungsfunktionskoeffizienten) wurde die Auszahlung der Teilnehmer nicht durch die Anzahl der zusätzlichen Kontakte pro Subjekt beeinflusst. Es ist allerdings anzunehmen, dass dies bei einem anderen Kosten-Nutzen-Verhältnis durchaus der Fall ist: Wäre der Gewinn für den Erhalt einer Anfrage größer als in der Studie (und alle anderen Parameter blieben gleich), so wäre auch eine positive Korrelation zwischen der Auszahlung und der Anzahl an zusätzlichen Kontakten zu erwarten, d.h. eine größere Anzahl an zusätzlichen Kontakten würde auch zu einer größeren Auszahlung führen.

8.8.2 Ergebnisse – Korrektheitskriterien

Motivation 193 Subjekte nahmen an der Studie teil. Dies ist eine große Anzahl an Subjekten verglichen mit der Gruppengröße üblicher Laborexperimente.

Am Ende des PostGames wurden die Subjekte in einer Umfrage befragt, warum sie an der Studie teilgenommen haben. Das Ergebnis ist wie folgt: 16% der Subjekte sind am Spiel selbst interessiert. 44% sagen, dass sie Teil eines wissenschaftlichen Experiments sein wollen und 40% geben an, dass sie durch ihre Teilnahme Geld verdienen wollen. Da 60% der Subjekte nach ihren Angaben nicht an der monetären Belohnung interessiert sind, wurde überprüft, ob sie diese nach dem Experiment dennoch eingefordert haben. 67% der Subjekte fordern ihre monetäre Belohnung ein. Diese 67% scheinen also (mindestens partiell) durch die finanziellen Anreize motiviert worden zu sein.

Pseudonymität In der Umfrage wurden die Teilnehmer gefragt, ob sie in der Lage gewesen sind, Personen, die sie aus der Realwelt kennen, identifizieren zu können. Des Weiteren sollten sie angeben, wie sicher sie sich dabei sind. 80% der Teilnehmer geben an, dass sie keine Personen, die sie kennen, identifizieren konnten oder dass sie nicht versucht haben, jemanden zu identifizieren. 15% geben an, dass sie glauben, eine Person identifiziert zu haben. Allerdings waren 45% der Subjekte aus dieser Gruppe nicht sicher, ob dem wirklich so ist. Insgesamt konnte die große Mehrheit der Beteiligten kein Wissen über ihre Kontakte erlangen und war daher nicht in der Lage, Gruppenstrategien zu verwenden.

Problemverständnis Es liegt nahe, dass nur dann sinnvolle Strategien beobachtet werden können, wenn die Subjekte das Spiel verstehen, das sie spielen. Die Analyse der online-Umfrage am Ende der Studie legt den Schluss nahe, dass die Subjekte das Spiel verstanden haben. Darüber hinaus wurden sie explizit befragt, ob und wie gut sie das Spiel verstanden haben. 71% der Subjekte glauben, dass sie das Spiel vollständig oder gut verstanden haben. Die direkten Beobachtungen der Strategien der Teilnehmer während des Experiments stützen diese Aussage. Weitere 17% geben an, dass sie mindestens ein ausreichendes Verständnis haben. Lediglich 6% der Teilnehmer geben in der Umfrage an, dass sie das Spiel nicht verstanden haben. Insgesamt hat die große Mehrheit von 88% das Spiel mindestens ausreichend verstanden. Es ist also davon auszugehen, dass ein hinreichendes Spielverständnis vorhanden ist.

8.9 Fazit

Viele Realwelt-Systeme werden maßgeblich durch den Einfluss menschlichen Handels beziehungsweise menschlicher Strategien beeinflusst. Systeme, die es erlauben, Kontakte zu wählen und bei denen die Interaktionen der Teilnehmer sich auf ein Kooperationsproblem zurückführen lassen, sind eine wichtige Klasse dieser Realwelt-Systeme. Bisher war es unklar, ob das Verhalten der Teilnehmer großer Systeme mit solchen Eigenschaften, zu effizienten Netzwerken führt und wie die Strategien der

KAPITEL 8. POSTGAME – EINE STUDIE

Teilnehmer aussehen. Die Antwort auf diese Frage ist schwierig, da theoretische Ansätze Annahmen treffen, die nicht in allen Fällen menschliches Verhalten korrekt vorhersagen und da kontrollierte Experimente mit steigender Gruppengröße an ihre (organisatorischen) Grenzen stoßen.

In diesem Kapitel werden der Aufbau, die Durchführung und die Ergebnisse einer Studie über das Verhalten menschlicher Teilnehmer in der Interaktions- und Kontaktwahl beschrieben.

Eines der Ziele dieses Kapitels ist es, Ergebnisse mit einer hohen Qualität zu erreichen. In erster Linie kann die Studie die Ergebnisse der kontrollierten Experimente aus Kapitel 7 reproduzieren und bestätigen. Sollte ein systematischer Fehler vorhanden sein, der die Ergebnisse verfälscht, so muss dieser sowohl in der Studie als auch in den verhaltensökonomischen Laborexperimenten in Kapitel 7 vorkommen. Da es sich um verschiedene Techniken und (zum großen Teil) um unterschiedliche Experimentatoren handelt, ist dies unwahrscheinlich und es kann von einer hohen Qualität der Ergebnisse ausgegangen werden.

Die Ergebnisse im Einzelnen: Die Analyse der Strategien zeigt, dass die Teilnehmer vielversprechende Strategien nutzen, die zu fast effizienten Systemen führen. Überraschenderweise fällt auf, dass die Subjekte den Nutzen kooperativer Teilnehmer überschätzen: Sie halten kooperative aber teure Kontakte, selbst wenn diese ihnen mehr Kosten als Einnahmen bringen. Während 55% der Teilnehmer angeben, dass sie zusätzliche Kontakte wegen ihres niedrigen Kooperationsgrads fallenlassen, geben nur 25% der Subjekte an, dass sie Teilnehmer fallenlassen, weil diese nicht genug Nutzen erzeugt haben. Dies ist ein wichtiges Ergebnis: Systementwickler können aufbauend auf den Ergebnissen der Studie Mechanismen oder Strategien entwickeln, die das Gemeinwohl im System steigern. Z.B. könnten Hilfsprogramme bei der Kontaktwahl unterstützend eingreifen: Wie in der Studie deutlich geworden ist, beachten die Subjekte kaum, wie viel Nutzen ihnen ein Kontakt einbringt – ein einfaches Hilfsprogramm könnte den Nutzer darauf hinweisen, dass ein Kontakt keinen Nutzen bringt (selbst wenn er kooperativ ist). Es ist anzunehmen, dass ein solch einfaches Hilfsprogramm in der Lage ist, die Auszahlung der Teilnehmer zu erhöhen. Ähnliche Hilfsprogramme sind vorstellbar, die das Gemeinwohl im System dadurch bis zur Effizienz steigern, dass sie die Teilnehmer bei der Koordination der bereits fast effizienten Netzwerkstruktur unterstützen. Dies kann zu einer (noch) größeren Akzeptanz verteilter Systeme führen.

9 Fazit und Ausblick

9.1 Zusammenfassung

Bedeutung verteilter Systeme Soziale Netzwerke sind durch die Popularität von Plattformen wie Facebook, Xing, Orkut oder LinkedIn einem großen Teil der Öffentlichkeit bekannt geworden.¹ Dennoch sind diese Plattformen, die es erlauben, soziale und kommerzielle Interaktionen mit beliebigen Kontakten zu unterstützen, nur ein kleiner Teil der existierenden Sozialen Netzwerke. Ein Beispiel für ein Soziales Netzwerk mit enormer Bedeutung ist das Netz bestehend aus Freunden und Familienangehörigen. Die Knoten sind Menschen und die Verbindungen zwischen ihnen stellen familiäre oder freundschaftliche Verbindungen dar. Ein Beispiel für weitere Soziale Netzwerke sind die Verbindungen zwischen Firmen oder Beziehungen zwischen Staaten.

Neben Sozialen Netzwerken sind Peer-to-Peer Systeme eine weitere Klasse von verteilten Systemen. Bei ihnen handelt es sich um technische Netzwerke, die im Allgemeinen dafür genutzt werden, (Schlüssel, Wert)-Paare zu verwalten. Anwendungen wie Internet-Telefonie (Skype) oder das gemeinsame Lösen wissenschaftlicher Probleme (SETI@home) sind weitere Beispiele für die Verwendung von Peer-to-Peer Systemen.²

Eigenschaften verteilter Systeme Soziale Netzwerke und Peer-to-Peer Systeme haben die Eigenschaft, dass ihre Teilnehmer (a) entscheiden können, wie sie mit anderen Teilnehmern interagieren und (b) welche Kontakte sie wählen. Während dies für Soziale Netzwerke, in denen meist Menschen oder Organisationen handeln, einleuchtend ist, muss diese These für Peer-to-Peer Systeme, die über Software gesteuert werden, erst noch begründet werden: Obwohl von den Entwicklern eines Peer-to-Peer Systems Software vorgegeben wird, die das Verhalten im System vorgibt, so kann dennoch Software von Dritten manipuliert oder erstellt werden, die nicht zu dem durch die Entwickler des Systems vorgegebenen Verhalten konform ist. Es ist also auch in Peer-to-Peer Systemen davon auszugehen, dass die Teilnehmer sich nach ihren eigenen Interessen verhalten.

¹www.xing.com, www.orkut.com, www.linkedin.com

²www.skype.com, setiathome.berkeley.edu

Problemstellung Während die Tatsache, dass Teilnehmer eines Sozialen Netzwerkes oder eines Peer-to-Peer Systems ihre Interaktionswahl-Strategie wählen, dazu führt, dass Freifahrer-Verhalten möglich wird, erlaubt die Kontaktwahl, kooperative Kontakte zu wählen und so unkooperativem Verhalten entgegenzuwirken. Im Einzelnen: Teilnehmer mit einer Interaktionswahl-Strategie haben die Möglichkeit, Anfragen anderer Teilnehmer zu ignorieren, um so Kosten zu sparen. Da die Endpunkte einer Weiterleitungskette in großen Systemen mit hoher Wahrscheinlichkeit nur selten miteinander kommunizieren, besteht ein ökonomischer Anreiz, die Anfragen von anderen Teilnehmern nicht zu bearbeiten, d.h. eine Freifahrer-Strategie zu nutzen [FLSC04]: Der Freifahrer spart sich die Kosten der Bearbeitung der Anfrage und muss nicht damit rechnen, dass der Erzeuger der Anfrage sich reziprok verhalten kann, da es unwahrscheinlich ist, dass der Freifahrer dem Erzeuger ebenfalls eine Anfrage sendet. Die Kommunikation zwischen Kontakten hingegen findet wiederholt statt. Bei dieser gegenseitigen und vielfach wiederholten Interaktion ist kooperatives Verhalten zu erwarten, da ein Kontakt auf unkooperatives Verhalten ebenfalls unkooperativ reagieren wird [SBV07]. Die Auswahl von kooperativen Kontakten ermöglicht es also, die Zusammenarbeit zwischen kooperativen Teilnehmern zu erhöhen und gleichzeitig unkooperative Teilnehmer vom System auszuschließen.

Da der Effekt des Freifahrens und der Effekt der Kontaktwahl das System gegenläufig beeinflussen, ist unklar, ob Systeme mit Interaktions- und Kontaktwahl effizient sind. Anschaulich bedeutet dies, dass unklar ist, ob die Teilnehmer solcher Systeme kooperieren, d.h. die Anfragen anderer Teilnehmer bearbeiten, und ob sie Netzwerke bilden, die es erlauben, Anfragen in nur wenigen Weiterleitungsschritten zu verarbeiten.

Die Klärung dieser Frage ist aus mehrerer Hinsicht schwierig. Aus der Literatur ist bekannt, dass die ökonomische Analyse von Spielen, an denen n Spieler beteiligt sind, im Allgemeinen algorithmisch komplex und schwierig ist [GGS03]. Ebenfalls ist bekannt, dass die Lösung solcher Spiele dadurch erschwert werden kann, dass die Teilnehmer ihre Kontakte wählen können [BDH⁺08]. Des Weiteren sind einige Nebenbedingungen zu beachten. So beeinflusst z.B. die Qualität eines eventuell genutzten Reputationssystems – also eines Systems zur Bestimmung des Kooperationsgrads von anderen Teilnehmern – die Strategien der Teilnehmer. Ist die Qualität gering, so kommt es zu vielen Fehleinschätzungen und unkooperative Teilnehmer werden nicht oder nur selten erkannt. Daneben führen unterschiedliche exogene Parameter, wie unterschiedliche Kosten für Aktionen oder die Größe des Systems zu unterschiedlichen Ergebnissen.

Ziel der Arbeit Soziale Netzwerke als auch Peer-to-Peer Systeme werden zum einen von Software und zum anderen von menschlichen Teilnehmern gesteuert und kontrolliert. Daher ist es (a) wichtig, zu analysieren, ob effiziente Systemzustände Gleichgewichte darstellen und (b) wichtig, das Verhalten menschlicher Teilnehmer

zu verstehen. Erst wenn beide Problemfelder verstanden sind, können Systementwickler Systeme entwerfen, die sowohl konform zum menschlichen Verhalten sind als auch Gleichgewichtszustände darstellen. Andernfalls ist davon auszugehen, dass die Teilnehmer der Systeme Strategien wählen, die in ihrem eigenen Interesse liegen (und nicht im Interesse der Systementwickler).

Um das Ziel zu erreichen, wurden in dieser Arbeit unterschiedliche Techniken verwendet. Zum einen wurden Simulationen genutzt, um die Nützlichkeit unterschiedlicher Strategien zu bewerten oder um Gleichgewichte von Strategien mit evolutionärer Anpassung zu bestimmen. Des Weiteren wurden Konzepte aus der Spieltheorie angewendet, um vorgegebene Strategien darauf zu untersuchen, ob sie Gleichgewichte darstellen (wenn alle Teilnehmer sie nutzen). Durch verhaltensökonomische Experimente und eine Studie eines Online-Spiels konnte das Interaktionswahl- und Kontaktwahl-Verhalten menschlicher Teilnehmer in Kombination untersucht werden.

Beiträge der Arbeit Aus dieser Arbeit sind mehrere Beiträge hervorgegangen:

(1) Kontaktwahl eröffnet die Möglichkeit, Kontakte strategisch zu wählen. Da Kontakte aber im Allgemeinen Kosten verursachen, ist unklar, in welchen Situationen es sich lohnt, Kontakte zu unterhalten, beziehungsweise in welchen Situationen es lohnend ist, keine Kontakte zu unterhalten. Mittels einer formalen Analyse konnte gezeigt werden, dass es für einen einzelnen Teilnehmer in vielen Situationen sinnvoll ist, keine Kontakte zu unterhalten. Dieses Verhalten, das als *Deliberately Poor Connectedness* bezeichnet wird, tritt insbesondere dann auf, wenn Mechanismen oder Strategien gegen Freifahrer auf Anfrageebene greifen und nur wenige Anfragen verloren gehen. Die Tatsache, dass ein einzelner Teilnehmer kein Interesse daran hat, Kontakte zu unterhalten, ist allerdings nicht für die Gesamtheit der Teilnehmer optimal: Eine Netzwerkstruktur mit wenigen Verbindungen führt zu einer höheren Weiterleitungslast. Um diesem Problem entgegenzuwirken, wurde eine Strategie entwickelt – genannt „*Contact-for-Contact (C4C)*“ – die dafür sorgt, dass ein Teilnehmer (nur) im gleichen Maß von den zusätzlichen Kontakten anderer Teilnehmer profitiert, wie er selbst in zusätzliche Kontakte investiert hat. Durch eine formale Analyse konnte gezeigt werden, dass durch die C4C-Strategie *Deliberately Poor Connectedness* zu einer geringeren Auszahlung führt als kooperatives Verhalten. Dies ist ein wichtiges Ergebnis, da die Vernetzung durch die C4C-Strategie dahingehend beeinflusst werden kann, dass die Teilnehmer zusätzliche Kontakte wählen und damit die Weiterleitungslast im System verringern. Dies erhöht das Gemeinwohl im System und führt zu einem größeren Nutzen für die Gesamtheit der Teilnehmer.

(2) Systementwickler geben den Teilnehmern ihrer Systeme ein Verhalten – in Form von Software – vor. Um zu klären, ob Teilnehmer, die ihre Auszahlung maximieren, von dieser Strategie abweichen, wurden in dieser Arbeit Gleichgewichte für Systeme mit Interaktions- und Kontaktwahl identifiziert. Zum einen existiert ein unkooperatives Gleichgewicht, in dem es sich nicht lohnt, am System teilzunehmen.

KAPITEL 9. FAZIT UND AUSBLICK

Zum anderen ein kooperatives Gleichgewicht, in dem die Teilnehmer die Anfragen anderer Teilnehmer verarbeiten. In diesem zweiten Gleichgewicht wählen die Teilnehmer ihre Kontakte gleichverteilt über den Adressraum, wenn sie ausschließlich den direkten Nutzen ihrer Kontakte als Kriterium bei der Kontaktwahl heranziehen. Dies führt nicht zu einer effizienten Netzwerkstruktur. Durch Beachtung der Verarbeitungszeiten eigener Anfragen, die ein Teilnehmer einem bestimmten Kontakt gesendet hat, kann eine effizientere Netzwerkstruktur erreicht werden. Wenn die Teilnehmer Kontakte, die Anfragen nur langsam bearbeiten (da ihre Kontakte und deren Kontakte nicht optimal über den Adressraum verteilt sind) gegen solche austauschen, die ihre Anfragen schnell bearbeiten, so entsteht eine Netzwerkstruktur, die optimal für die verteilte Anfrageverarbeitung ist. Diese Strategie ist kostengünstig und erreicht auf relativ einfache Art eine Verbesserung des Gemeinwohls im System.

(3) Neben der Maximierung der Auszahlung wird oft angenommen, dass menschliche Teilnehmer komplexe Situationen nicht sofort vollständig erfassen können sondern, dass sie ihre Strategien dynamisch so anpassen, dass sie für die gegebene Situation besser geeignet ist als im vorherigen Zustand. Solche evolutionären Strategieverbesserungen sind eine Möglichkeit, um das Verhalten von Menschen zu analysieren. Mittels Simulationen von evolutionären Strategieanpassungen wird in dieser Arbeit gezeigt, dass kooperatives Verhalten – also die Fähigkeit Anfragen gemeinsam zu bearbeiten – aus zufälligen Strategien entstehen kann. Dieses Ergebnis ist bedeutend, da es eine Erklärung für Kooperation in Systemen der Interaktions- und Kontaktwahl liefert.

(4) Menschen beeinflussen das Verhalten Sozialer Netzwerke und von Peer-to-Peer Systemen maßgeblich. Daher ist es entscheidend, menschliches Verhalten in der Interaktions- und Kontaktwahl in Kombination zu analysieren und zu verstehen. Zu diesem Zweck wurden in dieser Arbeit verhaltensökonomische Experimente durchgeführt, die zu den folgenden Ergebnissen geführt haben:

- Kontaktwahl erhöht den Kooperationsgrad im System: Wird zugelassen, dass Teilnehmer ihre Kontakte selbstständig wählen, so hat dies einen positiven Einfluss. Je mehr Kontakte die Teilnehmer wählen, desto kooperativer verhalten sie sich. Durch das einfache Erlauben der Kontaktwahl kann also das Gemeinwohl im System erhöht werden.
- Teilnehmer lassen Kontakte, die sich ihnen gegenüber unkooperativ verhalten haben, fallen. Dies erklärt, warum die Kontaktwahl einen positiven Einfluss auf das Verhalten der Teilnehmer hat: Unkooperative Teilnehmer haben Probleme, sich zu vernetzen und müssen sich kooperativ verhalten, falls sie Verbindungen zu Kontakten erhalten wollen.
- Menschliche Teilnehmer passen ihr Verhalten selbstständig an exogene Parameter an. Sie bilden Netzwerkstrukturen, die relativ gesehen zu den Strukturen aus Experimenten mit anderen Parameterbelegungen optimal sind. Dies zeigt,

dass die Teilnehmer zum großen Teil selbst dazu in der Lage sind, sinnvolle Netzwerkstrukturen zu bilden. Die Netzwerkstrukturen sind performant, aber nicht zu 100% effizient.

Mittels einer Studie an Online-Spielern konnten die Ergebnisse der verhaltensökonomischen Experimente bestätigt werden. Überraschenderweise konnte dabei beobachtet werden, dass die Teilnehmer des Online-Spiels den Nutzen kooperativer Kontakte überschätzen: Kooperative Kontakte, die aber zu negativen Auszahlungen führen, werden nicht fallengelassen, obwohl dies im Sinne der Maximierung der Auszahlung optimal wäre. Diese Beobachtung ist wichtig, da Mechanismen oder Strategien, die die Teilnehmer in der Kontaktwahl unterstützen, das bereits hohe Gemeinwohl weiter bis zur Effizienz erhöhen können.

9.2 Ausblick

Das Verständnis des Verhaltens menschlicher Teilnehmer sowie das Verständnis der Strategien von Agenten, die ihre Auszahlung maximieren, haben einen Einfluss auf Anwendungen, die auf der Anfragebearbeitung in Sozialen Netzwerken oder Peer-to-Peer Systemen basieren. So haben z.B. populäre Social Search Plattformen wie Aardvark oder Yahoo answers mehr und mehr Anhänger und es existieren z.B. für die Plattform facebook bereits 550.000 Anwendungen.^{3 4} Es ist anzunehmen, dass die Anzahl und die Komplexität von Anwendungen, die auf Sozialen Netzwerken oder Peer-to-Peer Systemen beruhen, in der Zukunft noch weiter steigen wird. Entscheidend für die Akzeptanz einer Anwendung sind neben der Bedienbarkeit und ihrer Popularität natürlich das Verständnis des Verhaltens der Teilnehmer sowie der Nutzen, den ein Teilnehmer an dieser Applikation hat. Diese Arbeit beschreibt neben dem Verhalten der Teilnehmer mehrere Strategien, die Systementwickler nutzen können, um das Gemeinwohl im System in Richtung Effizienz zu erhöhen. Dies ist ein wichtiger Beitrag für die Steigerung des Gemeinwohls verteilter Systeme, die auf der Anfragebearbeitung sozialer Netzwerke oder Peer-to-Peer Systemen aufbauen.

Neben Anwendungen – also Programmen – die auf Sozialen Netzwerken basieren, gibt es menschliche Handlungen oder Aktivitäten von Organisationen, die auf Sozialen Netzwerken aufbauen. So sind Firmen meist mit anderen Firmen vernetzt, sei es, weil sie in einem Besitzverhältnis zueinander stehen, oder sei es, weil ihre Produkte voneinander abhängig sind (z.B. weil eine Firma Zulieferer einer anderen Firma ist). Obwohl viele Firmen sich über Verträge absichern, ist es bei internationalen Beziehungen nicht immer sinnvoll, lediglich auf Verträge zu setzen, da die vereinbarte Leistung nur schwer beziehungsweise nur mit hohen Kosten eingeklagt werden kann [Lor99]. Vielmehr spielt das Vertrauen und der Kooperationsgrad der Firmen

³vark.com, answers.yahoo.com

⁴vgl. <http://www.facebook.com/press/info.php?statistics>, August 2010

KAPITEL 9. FAZIT UND AUSBLICK

eine Rolle [Lor99]. Da bei dem Netzwerk bestehend aus Firmen davon auszugehen ist, dass diese sowohl ihre Interaktionswahl-Strategie als auch ihre Kontakte/Partner wählen, können Firmen von dieser Arbeit profitieren. Durch die Ergebnisse dieser Arbeit können Firmen beispielsweise bestimmen, ob eine weitere Vernetzung sinnvoll ist oder nicht. Des Weiteren können die Ergebnisse der Arbeit genutzt werden, um andere Firmen in ihrem Verhalten zu beeinflussen. Z.B. kann ein Anreiz für eine stärkere oder geringe Vernetzung von Firmen-Kontakten gegeben werden. So kann eine Firma auf die C4C-Strategie zurückgreifen (oder dies bewusst nicht tun). Durch die C4C-Strategie kann erreicht werden, dass andere Teilnehmer nur insoweit von den eigenen Kontakten profitieren, wie sie selbst in ihre zusätzlichen Kontakte investieren. Gleiches ist für die Art der Vernetzung und die Verarbeitungsgeschwindigkeit durch die DSC-Strategie möglich. Diese sorgt dafür, dass Teilnehmer, die Anfragen schnell bearbeiten, besser vernetzt sind (beziehungsweise sein können, wenn sie es denn wollen) als andere Teilnehmer.

Durch diese Arbeit kann das wichtige Ziel der Steigerung des Gemeinwohls in Systemen mit Interaktionswahl und Kontaktwahl verwirklicht werden.

Appendix

A Deliberately Poor Connectedness

A.1 Pfadlänge einer Anfrage

Eine notwendige Voraussetzung um die Effektivität der C4C-Strategie zu bestimmen ist es, die durchschnittliche Pfadlänge von Anfragen zu kennen, die von einem Teilnehmer stammen, der zusätzliche Kontakte verwendet. Im schlechtesten Fall wird die Anfrage nur im ersten Schritt über zusätzliche Kontakte geleitet. Dieser Fall tritt ein, wenn alle Teilnehmer einer Weiterleitungskette (außer dem Erzeuger der Anfrage) nur feste Kontakte verwenden. Im Folgenden wird eine Abschätzung dieses Falles durchgeführt. Abbildung A.1 dient als Illustration. Was ist die zu erwartende *kürzeste Distanz*, d.h. die minimale Anzahl an Schritten vom Erzeuger der Anfrage zum Beantworter der Anfrage? Die *Entfernung* zwischen zwei Teilnehmern sei dabei die Anzahl an Weiterleitungsschritten zwischen ihnen. In Abbildung A.1 repräsentieren schwarze Zonen zusätzliche Kontakte des Teilnehmers t_i . Hellgraue Zonen repräsentieren Teilnehmer, die die gleiche Distanz zum Ziel der Anfrage besitzen als der zusätzliche Kontakt von Teilnehmer t_i , der dem Ziel der Anfrage am nächsten gelegen ist, d.h. Kontakt 2. Diese zum Ziel der Anfrage äquidistanten Teilnehmer formen einen Rhombus (und keinen Kreis). Dies liegt daran, dass ein Teilnehmer in einem zweidimensionalen Content-Addressable Network vier feste Kontakte besitzt (zwei pro Dimension). Dies sind die Teilnehmer, deren Zone eine gemeinsame Kante mit der Zone von t_i haben.

Das hier untersuchte Problem ist äquivalent zum „Law of Distribution of the Nearest Neighbor in a Random Distribution of Particles“ [Cha43]. Der Unterschied liegt darin, dass zwei unterschiedliche Metriken verwendet werden und der hier angenommene Raum ein Torus ist. Im Gegensatz zu den Partikeln in [Cha43] sind der Erzeuger einer Anfrage und seine zusätzlichen Kontakte Gegenstand der Untersuchung. Im Folgenden stehe K'_i für diese Menge an $|K'_i| + 1$ Teilnehmern.

Die Schätzung besteht aus zwei Schritten: Zuerst wird eine Näherung für die Wahrscheinlichkeit $\omega(r)$ berechnet. Diese Wahrscheinlichkeit beschreibt, dass ein oder mehr Teilnehmer aus der Menge K'_i den Abstand r zum Ziel der Anfrage besitzen und dass kein Teilnehmer aus K'_i einen geringeren Abstand besitzt. Danach wird der gewichtete Durchschnitt berechnet (wobei $\omega(r)$ das Gewicht ist).

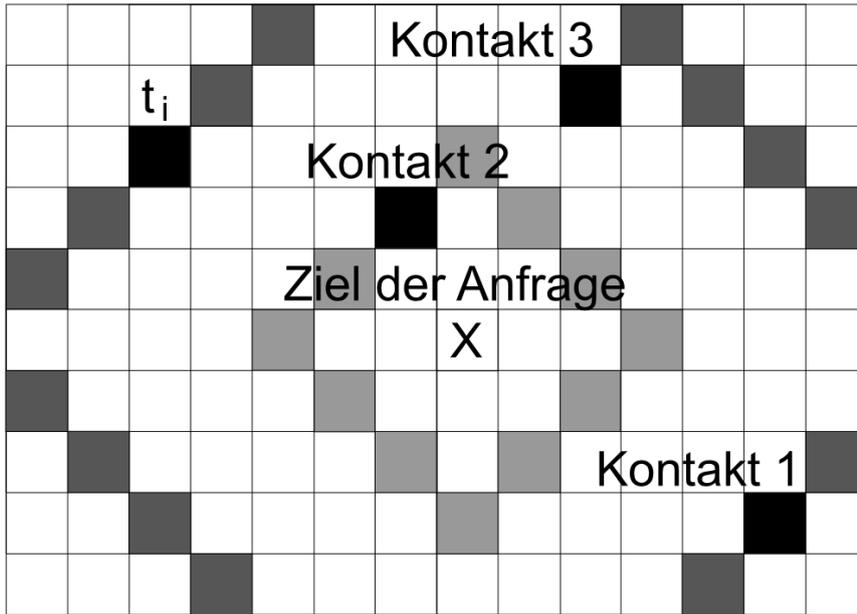


Abbildung A.1: Teilnehmer, die äquidistant zum Ziel der Anfrage sind

Im Folgenden gelte:

$$\omega(r) := \begin{cases} \omega_0(r) & \text{falls } r \in [0, \frac{\sqrt{n}}{2}] \\ \omega_1(r - \frac{\sqrt{n}}{2}) & \text{andernfalls} \end{cases} \quad (\text{A.1})$$

Die beiden Fälle sind notwendig, da für Radien r , die größer sind als $\frac{\sqrt{n}}{2}$, die Torus Eigenschaft eine Rolle spielt: In diesem Fall formen äquidistante Punkte keinen Rhombus mehr, sondern einen „Rhombus ohne Ecken“. Abbildung A.1 illustriert diesen Fall.

Beispiel: Die dunkelgrauen Teilnehmer haben einen Abstand von acht Schritten zum Ziel der Anfrage. Sie formen einen Rhombus ohne Ecken. ■

Der Beweis beginnt mit der Berechnung von $\omega_0(r)$. Die Wahrscheinlichkeit p , dass eine gegebene Zone zu einem Teilnehmer $k \in K'_i$ gehört, ist äquivalent zum Verhältnis der Anzahl der Teilnehmer in K'_i zur Gesamtanzahl der Teilnehmer, d.h. es gilt $p = (|K^+| + 1)/n$. Mit gegebenen p können wir die Wahrscheinlichkeit berechnen, dass mindestens ein Teilnehmer aus K'_i den Abstand r zum Ziel der Anfrage besitzt. Zu diesem Zweck muss die Wahrscheinlichkeit berechnet werden, dass die Menge der Teilnehmer, die Abstand r zum Ziel der Anfrage besitzen, mindestens einen Teilnehmer aus der Menge K'_i enthält. Diese Wahrscheinlichkeit ist die Wahrscheinlichkeit p multipliziert mit dem Umfang der Sphäre mit Radius r . Der Umfang der Sphäre beträgt $4 \cdot r$. Radius r wird in Anzahl von Weiterleitungsschritten gemessen. Z.B. haben in Abbildung A.1 $12 = 4 \cdot 3$ Teilnehmer (repräsentiert durch graue Zonen)

A.1. PFADLÄNGE EINER ANFRAGE

eine Distanz von drei Weiterleitungsschritten zum Ziel der Anfrage.

Es gelte $r \leq \frac{\sqrt{n}}{2}$ und es sei mit $\omega_0(r) \cdot dr$ die Wahrscheinlichkeit bezeichnet, dass die kürzeste Distanz von einem Teilnehmer in K'_i zum Ziel der Anfrage zwischen r und $r + dr$ liegt. Diese Wahrscheinlichkeit ist gleich der Wahrscheinlichkeit, dass mindestens ein Teilnehmer Distanz r besitzt, was wiederum gleich $4 \cdot r \cdot p$ ist, und dass gleichzeitig kein Teilnehmer in K'_i eine kürzere Distanz zum Ziel besitzt ($1 - \int_0^r \omega_0(r) dr$). Es gilt die folgende Gleichung:

$$\omega_0(r) = [1 - \int_0^r \omega_0(r) dr] \cdot 4 \cdot r \cdot p \quad (\text{A.2})$$

Von Gleichung (A.2) leiten wir das Folgende ab:

$$\begin{aligned} \frac{\omega_0(r)}{4 \cdot r \cdot p} &= [1 - \int_0^r \omega_0(r) dr] \\ \left(\frac{\omega_0(r)}{4 \cdot r \cdot p} \right)' &= -4 \cdot r \cdot p \frac{\omega_0(r)}{4 \cdot r \cdot p} \\ \frac{\left(\frac{\omega_0(r)}{4 \cdot r \cdot p} \right)'}{\frac{\omega_0(r)}{4 \cdot r \cdot p}} &= -4 \cdot r \cdot p \end{aligned} \quad (\text{A.3})$$

Die Formel für den logarithmischen Differentialquotienten ($(\ln(f))' = \frac{f'}{f}$) erlaubt es $\omega_0(r)$ von Gleichung (A.3) abzuleiten:

$$\begin{aligned} \left(\ln \left(\frac{\omega_0(r)}{4 \cdot r \cdot p} \right) \right)' &= -4 \cdot r \cdot p \\ \ln \left(\frac{\omega_0(r)}{4 \cdot r \cdot p} \right) &= -2 \cdot r^2 \cdot p \\ \frac{\omega_0(r)}{4 \cdot r \cdot p} &= e^{-2 \cdot r^2 \cdot p} \\ \omega_0(r) &= 4 \cdot r \cdot p \cdot e^{-2 \cdot r^2 \cdot p} \end{aligned} \quad (\text{A.4})$$

Um die durchschnittliche Distanz abzuleiten, muss die Summe über alle möglichen Distanzen multipliziert mit der Wahrscheinlichkeit ihres Auftretens gebildet werden:

$$D_{max}^{H_1} = \int_0^{\frac{\sqrt{n}}{2}} \omega_0(r) \cdot r \cdot dr \quad (\text{A.5})$$

Obwohl Gleichung (A.5) nicht analytisch gelöst werden kann, kann eine beliebig präzise Annäherung berechnet werden.

Es sei angemerkt, dass Gleichung (A.5) nur gilt, wenn $r \leq \frac{\sqrt{n}}{2}$ gilt. Im Folgenden wird eine Näherung für die Wahrscheinlichkeit ω_1 berechnet, die beschreibt, dass die

ANHANG A. DELIBERATELY POOR CONNECTEDNESS

Distanz des am nächsten zum Ziel der Anfrage gelegenen Teilnehmer aus K'_i größer ist als $\frac{\sqrt{n}}{2}$. Sei der Radius von r nun größer als $\frac{\sqrt{n}}{2}$, und sei $t := r - \frac{\sqrt{n}}{2}$. Dann gilt in Analogie zu den Gleichungen (A.2) und (A.4), dass die Wahrscheinlichkeit, dass die kürzeste Distanz eines Teilnehmers in K'_i zum Ziel der Anfrage $\omega_1(r)$ ist, wie folgt ist:

$$\omega_1(t) = [1 - \omega_0(\frac{\sqrt{n}}{2}) - \int_0^t \omega_1(t)dt](-4 \cdot t + 2\sqrt{n})p \quad (\text{A.6})$$

Mit $\varsigma := 1 - \omega_0(\frac{\sqrt{n}}{2})$ können wir Gleichung (A.6) in eine Formel transformieren, die auf die gleiche Art zu lösen ist wie Gleichung (A.2).

$$\begin{aligned} \frac{\omega_1(t)}{\varsigma} &= [1 - \frac{1}{\varsigma} \int_0^t \omega_1(t)dt](-4t + 2\sqrt{n})p \\ \frac{\omega_1(t)}{(-4t + 2\sqrt{n})p\varsigma} &= [1 - \frac{1}{\varsigma} \int_0^t \omega_1(t)dt] \\ \left(\frac{\omega_1(t)}{(-4t + 2\sqrt{n})p\varsigma} \right)' &= -\frac{1}{\varsigma}\omega_1(t) \\ \left(\frac{\omega_1(t)}{(-4t + 2\sqrt{n})p\varsigma} \right)' &= -(-4t + 2\sqrt{n})p \frac{\omega_1(t)}{(-4t + 2\sqrt{n})p\varsigma} \\ \ln \left(\frac{\omega_1(t)}{(-4t + 2\sqrt{n})p\varsigma} \right)' &= (4t - 2\sqrt{n})p \\ \ln \left(\frac{\omega_1(t)}{(-4t + 2\sqrt{n})p\varsigma} \right) &= (2t^2 - 2\sqrt{nt})p \\ \omega_1(t) &= (-4t + 2\sqrt{n})p\varsigma e^{(2t^2 - 2\sqrt{nt})p} \end{aligned} \quad (\text{A.7})$$

Von den Gleichungen (A.5) und (A.7) können wir die kürzeste durchschnittlich Pfadlänge D_{worst}^{coop} eines Teilnehmers aus K'_i zum Ziel der Anfrage ableiten.

$$D_{worst}^{coop} = \int_0^{\frac{\sqrt{n}}{2}} \left(\omega_0(r) \cdot r + \omega_1(r) \cdot \left(r + \frac{\sqrt{n}}{2} \right) \right) dr \quad (\text{A.8})$$

Erneut kann Gleichung (A.8) nicht analytisch gelöst werden, aber es kann eine beliebig genaue Approximation berechnet werden.

Die durchschnittliche Pfadlänge aller Anfragen, die ihr Ziel erreichen, kann aus Gleichung (A.8) abgeleitet werden. Dazu wird zuerst die durchschnittliche Pfadlänge einschließlich fallengelassener Anfragen D_{avg} berechnet. Wir können diese berechnen, indem wir die Wahrscheinlichkeit berechnen, dass die Anfrage nicht während des Weiterleitens fallengelassen wird. Wenn wir die Wahrscheinlichkeit P_F^r , dass eine Anfrage über r Weiterleitungsschritte geleitet wird, zur Gleichung (A.8) hinzufügen, so erhalten wir die durchschnittliche Pfadlänge von Anfragen kooperativer

Teilnehmer inklusive der Anfragen, die fallengelassen worden sind, D_{avg}^{coop} :

$$D_{avg}^{coop} = \int_0^{\frac{\sqrt{n}}{2}} \left(\omega_0(r) \cdot P_F^r \cdot r + \omega_1(r) \cdot P_F^{r+\frac{\sqrt{n}}{2}} \cdot \left(r + \frac{\sqrt{n}}{2} \right) \right) dr \quad (\text{A.9})$$

Die durchschnittliche Pfadlänge von Anfragen von Teilnehmern, die eine Deliberately Poor Connectedness nutzen, inklusive der Anfragen, die fallengelassen worden sind, D_{avg}^{dpc} ist:

$$D_{avg}^{dpc} \leq \int_0^{\sqrt{n}} r \cdot P_F^r dr \quad (\text{A.10})$$

Wenn wir die Pfadlänge aus den Gleichungen (A.9) und (A.10) mit der Wahrscheinlichkeit, dass Teilnehmer mit der entsprechenden Strategie vorhanden sind, gewichten, so erhalten wir die durchschnittliche Pfadlänge inklusive fallengelassener Anfragen für alle Teilnehmer:

$$D_{avg} = (p^{cw} + p^{uw}) \cdot D_{avg}^{coop} + (p^{cd} + p^{ud}) \cdot D_{avg}^{dpc} \quad (\text{A.11})$$

A.2 Beweis – Lemma 4.6.2

Um zu verstehen, wann Deliberately Poor Connectedness zu höheren Auszahlungen führt als vollständige Kooperation, werden im Folgenden die entsprechenden zu erwartenden Auszahlungen verglichen: die zu erwartende Auszahlung vollständig kooperativer Teilnehmer z_c und die zu erwartende Auszahlung von Teilnehmern mit einer DPC-Strategie z_d , die sich kooperativ auf Anfrageebene verhalten (vergleiche Gleichungen (4.9) und (4.11)).

Immer dann, wenn die Differenz der zu erwartenden Auszahlungen von vollständig kooperativen Teilnehmern und Teilnehmern mit einer DPC-Strategie (mit Kooperation auf Anfrageebene) negativ ist, ist Deliberately Poor Connectedness lohnender als vollständige Kooperation. Die Differenz ist wie folgt:

$$\begin{aligned} Diff &= z_c - z_d \\ &= s(D_{opt}, P_F, P_A) \cdot (1 - P_F) \cdot a \cdot \overline{|R|} - \frac{\overline{|K^+|}}{\overline{|K^+|} - |K^f|} \cdot \\ &\quad s(D_{opt}, P_F, P_A) \cdot (D_{avg} - 1) \cdot f \cdot \overline{|R|} - c \cdot \overline{|K^+|} \cdot r \end{aligned} \quad (\text{A.12})$$

Es folgt aus Gleichung (A.12), dass Deliberately Poor Connectedness lohnender ist als Kooperation, wenn die folgende Ungleichung gilt:

$$\begin{aligned} s(D_{opt}, P_F, P_A) \cdot (1 - P_F) \cdot a \cdot \overline{|R|} &< \frac{\overline{|K^+|}}{\overline{|K^+|} - |K^f|} \cdot \\ s(D_{opt}, P_F, P_A) \cdot (D_{avg} - 1) \cdot f \cdot \overline{|R|} &+ c \cdot \overline{|K^+|} \cdot r \end{aligned} \quad (\text{A.13})$$

Dividieren durch $s(D_{opt}, P_F, P_A) \cdot (1 - P_F) \cdot \overline{|R|}$, in Ungleichung A.13, zusammen mit dem Faktum, dass $r = \overline{|R|}$ gilt (vergleiche Abschnitt 4.4), liefert Lemma 4.6.2:

Lemma A.2.1 *Es sei die Wahrscheinlichkeit, dass eine Anfrage weitergeleitet wird kleiner als 100%, d.h. $P_F < 1$. Ein Teilnehmer, der eine Deliberately Poor Connectedness Strategie bei gleichzeitiger Kooperation auf der Anfrageebene verfolgt, hat eine höhere Auszahlung als ein Teilnehmer, der vollständig kooperativ ist, immer dann, wenn die folgende Ungleichung gilt:*

$$a < \frac{\frac{\overline{|K^+|}}{|\overline{|K^+|} - |K^f|} \cdot (D_{avg} - 1)}{(1 - P_F)} \cdot f + \frac{1}{(1 - P_F)} \cdot c \cdot \overline{|K^+|} \quad (\text{A.14})$$

A.3 Beweis – Lemma 4.8.1

Die C4C-Strategie gibt Teilnehmern mit einer DPC-Strategie einen Anreiz zu kooperieren, wenn ihre Auszahlung (durch die C4C-Strategie) geringer ist als die vollständig kooperativer Teilnehmer. Es sei angenommen, dass alle vollständig kooperativen Teilnehmer die C4C-Strategie nutzen. Um die zu erwartenden Auszahlungen von vollständig kooperativen Teilnehmern z_c^+ und die zu erwartende Auszahlung von Teilnehmern mit einer DPC-Strategie (die sich kooperativ auf Anfrageebene verhalten) z_d^+ zu vergleichen zu können, werden in diesem Abschnitt die Pfadlängen in den entsprechenden Formeln ersetzt. Genauer, die durchschnittliche Pfadlänge ohne fallengelassene Anfragen (D_{opt}) in den Funktionen z_c und z_d wird gegen die Pfadlänge von Anfragen vollständig kooperativer Teilnehmer D_{worst}^{coop} in z_c und mit der Pfadlänge von Anfragen von Teilnehmern mit einer DPC-Strategie (die kooperativ auf Anfrageebene sind) D_{opt}^{dpc} in z_d ersetzt:

$$\begin{aligned} z_c^+ &= s(D_{worst}^{coop}, P_F, P_A) \cdot a \cdot \overline{|R|} - (q \cdot \overline{|R|} + s(D_{avg}, P_F, P_A) \\ &\quad \cdot (D_{avg} - 1) \cdot f \cdot \overline{|R|} + w \cdot \phi + c \cdot \overline{|K^+|} \cdot r) \\ z_d^+ &= s_d(D_{opt}^{dpc}, P_F, P_A) \cdot a \cdot \overline{|R|} - (q \cdot \overline{|R|} + \\ &\quad \frac{|K^f|}{|\overline{|K^+|} + |K^f|} \cdot s(D_{opt}, P_F, P_A) \cdot (D_{avg} - 1) \cdot f \cdot \overline{|R|} + w \cdot \bar{\phi}) \end{aligned} \quad (\text{A.15})$$

Analog zu dem Vorgehen in Appendix A.2, folgt das Lemma.

Lemma A.3.1 *Sei die Wahrscheinlichkeit P_F , dass eine Anfrage weitergeleitet wird kleiner als 100% (d.h. $P_F < 1$). Dann führt Deliberately Poor Connectedness (mit Kooperation auf Anfrageebene) zu höheren Auszahlungen als vollständige Kooperation, wenn die folgende Ungleichung gilt:*

$$a < \frac{\frac{\overline{|K^+|}}{|\overline{|K^+|} - |K^f|} \cdot (D_{avg} - 1)}{(1 - P_F^{D_{opt}^{dpc} - D_{worst}^{coop}})} \cdot f + \frac{1}{(1 - P_F^{D_{opt}^{dpc} - D_{worst}^{coop}})} \cdot c \cdot \bar{\gamma} \quad (\text{A.16})$$

B Gleichgewichte

B.1 Beweis – Lemma 5.3.1

In diesem Abschnitt wird das effiziente Gleichgewicht für Kontakte (d.h. Gruppen aus zwei Teilnehmern) untersucht.

Im ineffizienten Netzwerk senden die Teilnehmer keine Anfragen. Hier wird angenommen, dass sie Anfragen senden ($sendRequest = wahr$) und keine Aussteiger sind. Dennoch ist unklar, ob sie kooperieren.

Nach der Definition eines Gleichgewichts schauen wir uns die Auszahlung der Teilnehmer i und j im Gleichgewicht an. Kann keiner der beiden seine Auszahlung durch einseitiges Abweichen von seiner Strategie verbessern, so befinden sie sich im Gleichgewicht. Es sei dabei angenommen, dass die Teilnehmer verstehen, dass Strategien, die ins ineffiziente Netzwerk führen, die Auszahlung auf null begrenzen.

Es existieren drei Fälle.

1. Teilnehmer i und j erfüllen die Erwartungen des anderen in den Kooperationsgrad.
2. Nur ein Teilnehmer erfüllt die Erwartung des anderen.
3. Kein Spieler erfüllt die Erwartung des anderen.

Fall 1 Der Kooperationsgrad von Teilnehmer i ($ownCoop_i$) ist größer als die Erwartung von Teilnehmer j (und anders herum). Dann gelten die folgenden Formeln:

$$istKooperativ(i, additionalCC_j) \quad (B.1)$$

$$istKooperativ(i, contactC_j) \quad (B.2)$$

Des Weiteren müssen die Einnahmen, die Teilnehmer i durch seinen Kontakte j erhält, größer sein als die Ausgaben:

$$a \cdot |A_{i \rightarrow j}| > q \cdot |R_{i \rightarrow j}| + c \cdot r(i) + (w \cdot |W_{j \rightarrow i}| + f \cdot |F_{j \rightarrow i}|) \cdot ownCoop_i \quad (B.3)$$

Es werden zwei Abschätzungen genutzt, um Gleichung B.3 zu vereinfachen. Erstens, Teilnehmer i empfängt nicht mehr Antworten von j als Teilnehmer j verarbeitet hat, d.h. es gilt: $|A_{i \rightarrow j}| \leq |R_{i \rightarrow j}| \cdot ownCoop_j$. Zweitens, wir nehmen an, dass eine bilaterale Verbindung nur dann stattfindet, wenn beide Teilnehmer sich mindestens

ANHANG B. GLEICHGEWICHTE

eine (eigene) Anfrage zusenden. Es gilt daher: $c \leq c \cdot |R_{i \rightarrow j}|$. Dann gilt die folgende Ungleichung:

$$a \cdot |R_{i \rightarrow j}| \cdot \text{ownCoop}_j > q \cdot |R_{i \rightarrow j}| + c \cdot |R_{i \rightarrow j}| \cdot r(i) + (w \cdot |W_{j \rightarrow i}| + f \cdot |F_{j \rightarrow i}|) \cdot \text{ownCoop}_i \quad (\text{B.4})$$

Lemma 5.3.1 folgt direkt aus den Formeln (B.1), (B.2) und (B.4).

Fall 2 In Fall 2 gilt, dass der Kooperationsgrad von Teilnehmer i geringer ist als die Erwartungen von Teilnehmer j (sowohl $\text{istKooperativ}(i, \text{contact}C_j)$ und/oder $\text{istKooperativ}(i, \text{additional}CC_j)$ gelten nicht), aber es gilt ebenfalls, dass der Kooperationsgrad von Teilnehmer j höher ist als die Erwartung von Spieler i (es gilt also $\text{istKooperativ}(j, \text{contact}C_i)$ und $\text{istKooperativ}(j, \text{additional}CC_i)$). Daher lässt Teilnehmer j die Anfragen von Teilnehmer i fallen. In der Konsequenz hat Teilnehmer i keine Einnahmen mehr. Vielmehr hat er Kosten, wenn er seine Strategie nicht auf die Strategie Aussteiger abändert. Es folgt, dass Strategie Aussteiger das Gleichgewicht bestimmt.

Fall 3 Der Kooperationsgrad beider Teilnehmer ist geringer als die Erwartungen des anderen. Es folgt, dass die Teilnehmer nicht kooperieren und dass sie nur Ausgaben aber keine Einnahmen haben, da beide die Anfragen des anderen fallen lassen. Die Konsequenz ist, dass Strategie Aussteiger für beide lohnender ist und das Gleichgewicht bestimmt.

Wie gezeigt wurde, beschreibt nur Fall 1 das kooperative Gleichgewicht.

B.2 Fehlerhafte Reputationssysteme

Moderne Netzwerke nutzen Reputationssysteme, um den Kooperationsgrad und damit das Gemeinwohl im System zu erhöhen. In dieser Arbeit wird angenommen, dass ein Reputationssystem im untersuchten System existiert und genutzt wird. D.h. es ist in der Lage, Auskunft darüber zu geben, ob der Kooperationsgrad eines Teilnehmers i $C(i)$ größer als ein Schwellenwert ist oder nicht. Des Weiteren wird angenommen, dass das Reputationssystem fehlerhaft ist, d.h. dass in reportingError Prozent der Fälle diese Auskunft falsch ist.

Es wird nun untersucht, wann das kooperative Gleichgewicht in Abhängigkeit von Parameter reportingError stabil ist. Um die Frage zu klären, werden vier Beobachtungen wiedergegeben (es sei angemerkt, dass die Teilnehmer im kooperativen Gleichgewicht kooperieren):

1. Kooperative Teilnehmer verlieren reportingError Prozent ihrer zusätzlichen Kontakte.

B.3. OPTIMALE POSITIONEN ZUSÄTZLICHER KONTAKTE

2. Nicht kooperative Teilnehmer verlieren $1 - reportingError$ Prozent ihrer zusätzlichen Kontakte.
3. Kooperative Teilnehmer verlieren um die $\left(1 - (1 - reportingError)^{\bar{h}}\right)$ Prozent ihrer Anfragen.
4. Nicht kooperative Teilnehmer verlieren um die $(1 - reportingError)^{\bar{h}}$ Prozent ihrer Anfragen.

Wenn es überhaupt nützlich ist, zusätzliche Kontakte zu haben, dann führt Beobachtung (1) zur einer höheren Auszahlung als Beobachtung (2), wenn $reportingError < 0.5$ gilt. Das gleiche Argument gilt für Beobachtungen (3) und (4). Daher ist das kooperative Gleichgewicht nur dann stabil, d.h. es ist nur dann ein Gleichgewicht, wenn $reportingError < 0.5$ gilt.

Offensichtlich ist das System nur dann effizient, wenn Parameter $reportingError$ null ist, d.h. wenn kein Fehler in den Reputationsinformationen vorhanden ist. Andernfalls lassen die Teilnehmer einige Anfragen kooperativer Teilnehmer fallen, da sie sie für unkooperativ halten.

B.3 Optimale Positionen zusätzlicher Kontakte

Im Folgenden werden optimale Positionen p_1, \dots, p_{n_c} für n_c zusätzliche Kontakte formal berechnet. Ohne Beschränkung der Allgemeinheit kann angenommen werden, dass die Positionen der zusätzlichen Kontakte links vom Teilnehmer mit $p_1, \dots, p_{n_c/2}$ bezeichnet werden und dass die Positionen der zusätzlichen Kontakte rechts vom Teilnehmer mit $p_{n_c/2+1}, \dots, p_{n_c}$ bezeichnet werden. Des Weiteren sei angenommen, dass $p_1 > p_2 > \dots > p_{n_c/2} > 0 < p_{n_c/2+1} < \dots < p_{n_c}$ gilt. Gelten die Annahmen, dann haben die Schnittpunkte der Funktionen, die die Wahrscheinlichkeit des Erfolgs bei der Weiterleitung von Anfragen definieren, die folgenden Positionen:

1. Die Schnittpunkte des Teilnehmers mit seinen am nächsten gelegenen zusätzlichen Kontakten sind: $\left\{ \frac{p_{n_c/2}-1}{2}, \frac{1+p_{n_c/2+1}}{2} \right\}$
2. Schnittpunkte für alle Teilnehmer i, j , für die $p_i < p_j$ gilt, sind: $\frac{p_i+p_j}{2}$
3. Schnittpunkte für p_1 und p_{n_c} sind: $\frac{p_{n_c}+p_1}{2}$

ANHANG B. GLEICHGEWICHTE

Daher ist das zu erwartende Einkommen für alle Anfragen wie folgt:

$$\begin{aligned}
 & E(p_1, \dots, p_{n_c}) \\
 &= \int_{\frac{p_{n_c}+p_1}{2}}^{\frac{p_1+p_2}{2}} p(|h - p_1| + 1) \cdot \delta^{|h-p_1|} \cdot a \, dh \\
 &+ \dots \\
 &+ \int_{\frac{p_{n_c/2-2}+p_{n_c/2-1}}{2}}^{\frac{p_{n_c/2-1}+p_{n_c/2}}{2}} p(|h - p_{n_c/2-1}| + 1) \cdot \delta^{|h-p_{n_c/2-1}|} \cdot a \, dh \\
 &+ \int_{\frac{p_{n_c/2-1}+p_{n_c/2}}{2}}^{\frac{p_{n_c/2}-1}{2}} p(|h - p_{n_c/2}| + 1) \cdot \delta^{|h-p_{n_c/2}|} \cdot a \, dh \\
 &+ \int_{\frac{p_{n_c/2}-1}{2}}^{\frac{1+p_{n_c/2+1}}{2}} p(h) \cdot \delta^h \cdot a \, dh \tag{B.5} \\
 &+ \int_{\frac{1+p_{n_c/2+1}}{2}}^{\frac{p_{n_c/2+1}+p_{n_c/2+2}}{2}} p(|h - p_{n_c/2+1}| + 1) \cdot \delta^{|h-p_{n_c/2+1}|} \cdot a \, dh \\
 &+ \int_{\frac{p_{n_c/2+1}+p_{n_c/2+2}}{2}}^{\frac{p_{n_c/2+2}+p_{n_c/2+3}}{2}} p(|h - p_{n_c/2+2}| + 1) \cdot \delta^{|h-p_{n_c/2+2}|} \cdot a \, dh \\
 &+ \dots \\
 &+ \int_{\frac{p_{n_c-1}+p_{n_c}}{2}}^{\frac{p_{n_c}+p_1}{2}} p(|h - p_{n_c}| + 1) \cdot \delta^{|h-p_{n_c}|} \cdot a \, dh
 \end{aligned}$$

Um das Maximum von $E(p_1, \dots, p_{n_c})$ zu bestimmen, wird nun die Ableitung, d.h. der Gradient, gebildet. Aus Gründen der Darstellung wird nur die partielle Ableitung $\frac{\partial E(p_1, \dots, p_{n_c})}{\partial p_1}$ für den ersten Parameter, d.h. für Parameter p_1 , gezeigt. Alle anderen partiellen Ableitungen können analog berechnet werden. Die partielle Ableitung für

B.3. OPTIMALE POSITIONEN ZUSÄTZLICHER KONTAKTE

Parameter p_1 ist wie folgt:

$$\begin{aligned}
 \frac{\partial E(p_1, \dots, p_{n_c})}{\partial p_1} &= \underbrace{\frac{\partial}{\partial p_1} \int_{\frac{p_{n_c}+p_1}{2}}^{\frac{p_1+p_2}{2}} p(|h-p_1|+1) \cdot \delta^{|h-p_1|} \cdot a \, dh}_{\text{Fall (A)}} \\
 &+ \underbrace{\frac{\partial}{\partial p_1} \int_{\frac{p_1+p_2}{2}}^{\frac{p_2+p_3}{2}} p(|h-p_2|+1) \cdot \delta^{|h-p_2|} \cdot a \, dh}_{\text{Fall (B)}} \\
 &+ \underbrace{\frac{\partial}{\partial p_1} \int_{\frac{p_{n_c}-1+p_{n_c}}{2}}^{\frac{p_{n_c}+p_1}{2}} p(|h-p_{n_c}|+1) \cdot \delta^{|h-p_{n_c}|} \cdot a \, dh}_{\text{Fall (C)}} \\
 &= a \cdot \left(\int_{\frac{p_{n_c}+p_1}{2}}^{\frac{p_1+p_2}{2}} \frac{\partial}{\partial p_1} p(|h-p_1|+1) \cdot \delta^{|h-p_1|} \, dh + \frac{1}{2} \cdot p\left(\left|\frac{p_1+p_2}{2} - p_1\right| + 1\right) \cdot \delta^{\left|\frac{p_1+p_2}{2} - p_1\right|} \right. \\
 &\quad - \frac{1}{2} \cdot p\left(\left|\frac{p_{n_c}+p_1}{2} - p_1\right| + 1\right) \cdot \delta^{\left|\frac{p_{n_c}+p_1}{2} - p_1\right|} - \frac{1}{2} \cdot p\left(\left|\frac{p_1+p_2}{2} - p_2\right| + 1\right) \cdot \delta^{\left|\frac{p_1+p_2}{2} - p_2\right|} \\
 &\quad \left. + \frac{1}{2} \cdot p\left(\left|\frac{p_{n_c}+p_1}{2} - p_{n_c}\right| + 1\right) \cdot \delta^{\left|\frac{p_{n_c}+p_1}{2} - p_{n_c}\right|} \right) \\
 &= a \cdot \left(\frac{1}{2} \cdot p\left(\left|\frac{p_1+p_2}{2} - p_2\right| + 1\right) \cdot \delta^{\left|\frac{p_1+p_2}{2} - p_2\right|} - \frac{1}{2} \cdot p\left(\left|\frac{p_{n_c}+p_1}{2} - p_{n_c}\right| + 1\right) \cdot \delta^{\left|\frac{p_{n_c}+p_1}{2} - p_{n_c}\right|} \right)
 \end{aligned} \tag{B.6}$$

Fall A kann wie folgt abgeleitet werden:

$$\frac{\partial}{\partial p_1} \int_{p_1 \cdot d+e}^{p_1 \cdot b+c} f(h, p_1) \, dh = \underbrace{\frac{\partial}{\partial p_1} \int_0^{p_1 \cdot b+c} f(h, p_1) \, dh}_{\text{Es wird nur dieser Fall gezeigt.}} - \frac{\partial}{\partial p_1} \int_0^{p_1 \cdot d+e} f(h, p_1) \, dh$$

(B.7)

ANHANG B. GLEICHGEWICHTE

Der erste Teil der Gleichung (B.7) kann wie folgt abgeleitet werden:

$$\begin{aligned}
 & \frac{\partial}{\partial p_1} \int_0^{p_1 \cdot b + c} f(h, p_1) \, dh \\
 &= \lim_{\Delta x \rightarrow 0} \frac{\int_0^{(x_0 + \Delta x) \cdot b + c} f(h, x_0 + \Delta x) \, dh - \int_0^{x_0 \cdot b + c} f(h, x_0) \, dh}{\Delta x} \\
 &= \lim_{\Delta x \rightarrow 0} \left(\frac{\int_0^{(x_0 + \Delta x) \cdot b + c} f(h, x_0 + \Delta x) \, dh - \int_0^{(x_0 + \Delta x) \cdot b + c} f(h, x_0) \, dh}{\Delta x} \right. \\
 &\quad \left. + \frac{\int_0^{x_0 \cdot b + c} f(h, x_0) \, dh + \int_0^{(x_0 + \Delta x) \cdot b + c} f(h, x_0) \, dh}{\Delta x} \right) \\
 &= \lim_{\Delta x \rightarrow 0} \frac{\int_0^{(x_0 + \Delta x) \cdot b + c} f(h, x_0 + \Delta x) \, dh - \int_0^{(x_0 + \Delta x) \cdot b + c} f(h, x_0) \, dh}{\Delta x} \\
 &\quad + \lim_{\Delta x \rightarrow 0} \frac{-\int_0^{x_0 \cdot b + c} f(h, x_0) \, dh + \int_0^{(x_0 + \Delta x) \cdot b + c} f(h, x_0) \, dh}{\Delta x} \\
 &= \underbrace{\lim_{\Delta x \rightarrow 0} \frac{\int_0^{(x_0 + \Delta x) \cdot b + c} (f(h, x_0 + \Delta x) - f(h, x_0)) \, dh}{\Delta x}}_{\text{Es wird nur dieser Fall gezeigt.}} + \underbrace{\lim_{\Delta x \rightarrow 0} \frac{\int_{x_0 \cdot b + c}^{(x_0 + \Delta x) \cdot b + c} f(h, x_0) \, dh}{\Delta x}}_{b \cdot f(p_1 \cdot b + c, p_1)} \\
 & \tag{B.8}
 \end{aligned}$$

Es gelte $\epsilon \in [x_0, x_0 + \Delta x]$. Dann ist der erste Summand von Gleichung (B.8) wie folgt:

$$\begin{aligned}
 & \lim_{\Delta x \rightarrow 0} \frac{\int_0^{(x_0 + \Delta x) \cdot b + c} (f(h, x_0 + \Delta x) - f(h, x_0)) \, dh}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\int_0^{(x_0 + \Delta x) \cdot b + c} f'(h, \epsilon) \Delta x \, dh}{\Delta x} \\
 &= \lim_{\Delta x \rightarrow 0} \int_0^{(x_0 + \Delta x) \cdot b + c} \frac{\partial f}{\partial x}(h, \epsilon) \, dh = \int_0^{p_1 \cdot b + c} \left(\frac{\partial}{\partial p_1} f(h, p_1) \right) \, dh \\
 & \tag{B.9}
 \end{aligned}$$

Wenn wir die Teilergebnisse zusammenfassen, dann ist der erste Summand wie folgt:

$$\begin{aligned}
 & \frac{\partial}{\partial p_1} \int_{p_1 \cdot d + e}^{p_1 \cdot b + c} f(h, p_1) \, dh \\
 &= \int_{p_1 \cdot d + e}^{p_1 \cdot b + c} \frac{\partial}{\partial p_1} f(h, p_1) \, dh + b \cdot f(p_1 \cdot b + c, p_1) - d \cdot f(p_1 \cdot d + e, p_1) \\
 & \tag{B.10}
 \end{aligned}$$

Fall B aus Gleichung (B.6) kann wie folgt hergeleitet werden:

$$\begin{aligned}
 & \frac{\partial}{\partial p_1} \int_{p_1 \cdot b + c}^a f(h) \, dh = \frac{\partial}{\partial p_1} (F(a) - F(p_1 \cdot b + c)) \\
 &= \underbrace{\frac{\partial}{\partial p_1} F(a)}_{=0} - \frac{\partial}{\partial p_1} F(p_1 \cdot b + c) = -b \cdot f(p_1 \cdot b + c) \\
 & \tag{B.11}
 \end{aligned}$$

B.3. OPTIMALE POSITIONEN ZUSÄTZLICHER KONTAKTE

Fall C aus Gleichung (B.6) kann wie folgt berechnet werden:

$$\begin{aligned} \frac{\partial}{\partial p_1} \int_a^{p_1 \cdot b + c} f(h) \, dh &= \frac{\partial}{\partial p_1} (F(p_1 \cdot b + c) - F(a)) \\ &= \frac{\partial}{\partial p_1} F(p_1 \cdot b + c) - \underbrace{\frac{\partial}{\partial p_1} F(a)}_{=0} = b \cdot f(p_1 \cdot b + c) \end{aligned} \quad (\text{B.12})$$

(Bitte umblättern!)

ANHANG B. GLEICHGEWICHTE

Analog kann der Gradient, d.h. $\frac{\partial E(p_1, \dots, p_{n_c})}{\partial p_1, \dots, \partial p_{n_c}}$, berechnet werden. Er ist wie folgt:

$$\begin{aligned}
 & \left(\begin{aligned}
 & \frac{1}{2} a C^{1+} \left| \frac{p_1+p_2}{2} - p_1 \right|_d \left| \frac{p_1+p_2}{2} - p_1 \right| - \frac{1}{2} a C^{1+} \left| \frac{p_1+p_2}{2} - p_2 \right|_d \left| \frac{p_1+p_2}{2} - p_2 \right| \\
 & - \frac{1}{2} a C^{1+} \left| -p_1 + \frac{p_1+p_{n_c}}{2} \right|_d \left| -p_1 + \frac{p_1+p_{n_c}}{2} \right| + \frac{1}{2} a C^{1+} \left| -p_{n_c} + \frac{p_1+p_{n_c}}{2} \right|_d \left| -p_{n_c} + \frac{p_1+p_{n_c}}{2} \right| \\
 & + \int \frac{p_1+p_2}{2} \left(- \frac{a C^{1+|h-p_1|} d^{|h-p_1|} (h-p_1) \text{Log}[C]}{|h-p_1|} - \frac{a C^{1+|h-p_1|} d^{|h-p_1|} (h-p_1) \text{Log}[d]}{|h-p_1|} \right) dh, \\
 & \dots \\
 & \frac{1}{2} a C^{1+} \left| -p_1 + \frac{p_1+p_{n_c/2-1}}{2} \right|_d \left| -p_1 + \frac{p_1+p_{n_c/2-1}}{2} \right| - \frac{1}{2} a C^{1+} \left| -p_{n_c/2-1} + \frac{p_1+p_{n_c/2-1}}{2} \right|_d \left| -p_{n_c/2-1} + \frac{p_1+p_{n_c/2-1}}{2} \right| \\
 & + \frac{1}{2} a C^{1+} \left| -p_{n_c/2-1} + \frac{p_{n_c/2-1}+p_{n_c/2}}{2} \right|_d \left| -p_{n_c/2-1} + \frac{p_{n_c/2-1}+p_{n_c/2}}{2} \right| - \frac{1}{2} a C^{1+} \left| -p_{n_c/2} + \frac{p_{n_c/2-1}+p_{n_c/2}}{2} \right|_d \left| -p_{n_c/2} + \frac{p_{n_c/2-1}+p_{n_c/2}}{2} \right| \\
 & + \int \frac{p_{n_c/2-1}+p_{n_c/2}}{2} \left(- \frac{a C^{1+|h-p_{n_c/2-1}|} d^{|h-p_{n_c/2-1}|} (h-p_{n_c/2-1}) \text{Log}[C]}{|h-p_{n_c/2-1}|} - \frac{a C^{1+|h-p_{n_c/2-1}|} d^{|h-p_{n_c/2-1}|} (h-p_{n_c/2-1}) \text{Log}[d]}{|h-p_{n_c/2-1}|} \right) dh, \\
 & \frac{1}{2} a C^{1+} \left| \frac{1}{2} (-1+p_{n_c/2}) - p_{n_c/2} \right|_d \left| \frac{1}{2} (-1+p_{n_c/2}) - p_{n_c/2} \right| + \frac{1}{2} a C^{1+} \left| -p_{n_c/2-1} + \frac{p_{n_c/2-1}+p_{n_c/2}}{2} \right|_d \left| -p_{n_c/2-1} + \frac{p_{n_c/2-1}+p_{n_c/2}}{2} \right| \\
 & - \frac{1}{2} a C^{1+} \left| -p_{n_c/2} + \frac{p_{n_c/2-1}+p_{n_c/2}}{2} \right|_d \left| -p_{n_c/2} + \frac{p_{n_c/2-1}+p_{n_c/2}}{2} \right| \\
 & + \int \frac{p_{n_c/2-1}+p_{n_c/2}}{2} \left(- \frac{a C^{1+|h-p_{n_c/2}|} d^{|h-p_{n_c/2}|} (h-p_{n_c/2}) \text{Log}[C]}{|h-p_{n_c/2}|} - \frac{a C^{1+|h-p_{n_c/2}|} d^{|h-p_{n_c/2}|} (h-p_{n_c/2}) \text{Log}[d]}{|h-p_{n_c/2}|} \right) dh \\
 & + \frac{a \left(-\frac{1}{2} C^{p_{n_c/2}/2} d^{p_{n_c/2}/2} \text{Log}[C] - \frac{1}{2} C^{p_{n_c/2}/2} d^{p_{n_c/2}/2} \text{Log}[d] \right)}{\sqrt{C} \sqrt{d} (\text{Log}[C] + \text{Log}[d])}, \\
 & - \frac{1}{2} a C^{1+} \left| -p_{n_c/2+1} + \frac{1+p_{n_c/2+1}}{2} \right|_d \left| -p_{n_c/2+1} + \frac{1+p_{n_c/2+1}}{2} \right| + \frac{1}{2} a C^{1+} \left| -p_{n_c/2+1} + \frac{p_{n_c/2+1}+p_{n_c/2+2}}{2} \right|_d \left| -p_{n_c/2+1} + \frac{p_{n_c/2+1}+p_{n_c/2+2}}{2} \right| \\
 & - \frac{1}{2} a C^{1+} \left| -p_{n_c/2+2} + \frac{p_{n_c/2+1}+p_{n_c/2+2}}{2} \right|_d \left| -p_{n_c/2+2} + \frac{p_{n_c/2+1}+p_{n_c/2+2}}{2} \right| \\
 & + \int \frac{p_{n_c/2+1}+p_{n_c/2+2}}{2} \left(- \frac{a C^{1+|h-p_{n_c/2+1}|} d^{|h-p_{n_c/2+1}|} (h-p_{n_c/2+1}) \text{Log}[C]}{|h-p_{n_c/2+1}|} - \frac{a C^{1+|h-p_{n_c/2+1}|} d^{|h-p_{n_c/2+1}|} (h-p_{n_c/2+1}) \text{Log}[d]}{|h-p_{n_c/2+1}|} \right) dh \\
 & + \frac{a \left(\frac{1}{2} C^{1+\frac{p_{n_c/2+1}}{2}} d^{1+\frac{p_{n_c/2+1}}{2}} \text{Log}[C] + \frac{1}{2} C^{1+\frac{p_{n_c/2+1}}{2}} d^{1+\frac{p_{n_c/2+1}}{2}} \text{Log}[d] \right)}{\sqrt{C} \sqrt{d} (\text{Log}[C] + \text{Log}[d])}, \\
 & \frac{1}{2} a C^{1+} \left| -p_{n_c/2+1} + \frac{p_{n_c/2+1}+p_{n_c/2+2}}{2} \right|_d \left| -p_{n_c/2+1} + \frac{p_{n_c/2+1}+p_{n_c/2+2}}{2} \right| \\
 & - \frac{1}{2} a C^{1+} \left| -p_{n_c/2+2} + \frac{p_{n_c/2+1}+p_{n_c/2+2}}{2} \right|_d \left| -p_{n_c/2+2} + \frac{p_{n_c/2+1}+p_{n_c/2+2}}{2} \right| \\
 & + \frac{1}{2} a C^{1+} \left| -p_{n_c/2+2} + \frac{p_{n_c/2+2}+p_{n_c}}{2} \right|_d \left| -p_{n_c/2+2} + \frac{p_{n_c/2+2}+p_{n_c}}{2} \right| - \frac{1}{2} a C^{1+} \left| -p_{n_c} + \frac{p_{n_c/2+2}+p_{n_c}}{2} \right|_d \left| -p_{n_c} + \frac{p_{n_c/2+2}+p_{n_c}}{2} \right| \\
 & + \int \frac{p_{n_c/2+2}+p_{n_c}}{2} \left(- \frac{a C^{1+|h-p_{n_c/2+2}|} d^{|h-p_{n_c/2+2}|} (h-p_{n_c/2+2}) \text{Log}[C]}{|h-p_{n_c/2+2}|} - \frac{a C^{1+|h-p_{n_c/2+2}|} d^{|h-p_{n_c/2+2}|} (h-p_{n_c/2+2}) \text{Log}[d]}{|h-p_{n_c/2+2}|} \right) dh, \\
 & \dots \\
 & - \frac{1}{2} a C^{1+} \left| -p_1 + \frac{p_1+p_{n_c}}{2} \right|_d \left| -p_1 + \frac{p_1+p_{n_c}}{2} \right| + \frac{1}{2} a C^{1+} \left| -p_{n_c} + \frac{p_1+p_{n_c}}{2} \right|_d \left| -p_{n_c} + \frac{p_1+p_{n_c}}{2} \right| \\
 & + \frac{1}{2} a C^{1+} \left| -p_{n_c/2+2} + \frac{p_{n_c/2+2}+p_{n_c}}{2} \right|_d \left| -p_{n_c/2+2} + \frac{p_{n_c/2+2}+p_{n_c}}{2} \right| - \frac{1}{2} a C^{1+} \left| -p_{n_c} + \frac{p_{n_c/2+2}+p_{n_c}}{2} \right|_d \left| -p_{n_c} + \frac{p_{n_c/2+2}+p_{n_c}}{2} \right| \\
 & + \int \frac{p_1+p_{n_c}}{2} \left(- \frac{a C^{1+|h-p_{n_c}|} d^{|h-p_{n_c}|} (h-p_{n_c}) \text{Log}[C]}{|h-p_{n_c}|} - \frac{a C^{1+|h-p_{n_c}|} d^{|h-p_{n_c}|} (h-p_{n_c}) \text{Log}[d]}{|h-p_{n_c}|} \right) dh
 \end{aligned} \right)
 \end{aligned}$$

(B.13)

B.4. NÜTZLICHKEIT DER DSC-STRATEGIE

Bei optimalen Positionen hat der Gradient den Wert null. Zur Vereinfachung sei angenommen, dass die Anzahl an zusätzlichen Kontakten ungerade ist. Die Position eines zusätzlichen Kontaktes ist auf der anderen Seite des Ringes, es gilt also $p_{n_c} = \frac{n-1}{2}$. Jeder andere zusätzliche Kontakt mit Position p hat einen Zwilling auf der anderen Seite des Ringes mit Position $n-p$. Für $i = 1 \dots \frac{|K^+|-1}{2}$ und $j = \frac{|K^+|-1}{2} \dots 1$ seien die optimalen Positionen wie folgt:

- Für p_i gilt:

$$\frac{n-1}{2} \cdot \frac{2 \cdot i}{|K^+|+1} + \frac{2 \cdot j}{|K^+|+1} - \frac{4 \cdot j}{|K^+|+1} \cdot \frac{\ln(\frac{1+\delta}{2})}{\ln(C \cdot \delta)}$$

- Für $p_{i+\frac{|K^+|-1}{2}}$ gilt:

$$n - \left(\frac{n-1}{2} \cdot \frac{2 \cdot i}{|K^+|+1} + \frac{2 \cdot j}{|K^+|+1} - \frac{4 \cdot j}{|K^+|+1} \cdot \frac{\ln(\frac{1+\delta}{2})}{\ln(C \cdot \delta)} \right)$$

Es sei angemerkt, dass alle Berechnungen durch Hilfsprogramme zur symbolischen Berechnung überprüft worden sind.

B.4 Nützlichkeit der DSC-Strategie

In dieser Arbeit wird vorgeschlagen, dass die Teilnehmer die DSC-Strategie nutzen. Dieser Anhang zeigt, dass es für einen Teilnehmer nützlich ist, dies zu tun.

Es sei angenommen, dass sich die Teilnehmer im kooperativen Gleichgewicht befinden und Kleinberg-verteilte zusätzliche Kontakte besitzen, d.h. die Verarbeitungszeit für Anfragen ist für alle Teilnehmer gleich.

Hat ein Teilnehmer i die Möglichkeit, einen Kontakt s gegen einen Teilnehmer f auszutauschen, der Anfragen schneller bearbeitet als s , wenn also $\bar{h}_\Delta(i|f) < \bar{h}_\Delta(i|s)$ gilt, dann wird er dies nach der DSC-Strategie tun. Während die Kosten für den Unterhalt von s oder f gleich sind, sind die Einnahmen, die durch Teilnehmer f erzeugt werden, höher als die durch s erzeugten. Dies folgt aus der Verwendung des Diskontierungsfaktors: Da die durchschnittliche Anzahl an Weiterleitungen für eine Anfrage $r \in A_{i \rightarrow f}$, die an Teilnehmer f weitergeleitet wurde, geringer ist als die durchschnittliche Anzahl an Weiterleitungen für Anfragen, die an Teilnehmer s weitergeleitet wurden, ist das Einkommen im ersten Fall höher. D.h. es gilt die folgende Ungleichung:

$$a \cdot \sum_{f_r \in A_{i \rightarrow f}} \delta^{h(f_r)} < a \cdot \sum_{s_r \in A_{i \rightarrow s}} \delta^{h(s_r)} \quad (\text{B.14})$$

Es folgt die Behauptung für $\delta \in (0, 1)$. Analog kann bei vorhandenen unkooperativen Teilnehmern gezeigt werden, dass die DSC-Strategie nützlich ist, wenn $\bar{C} \in (0, 1)$ gilt.

B.5 Beweis – Lemma 5.3.3

(I) Das zu erwartende Einkommen einer Anfrage, die über $\beta \cdot \log^2(n)$ Schritte geleitet wird, ist:

$$p(\beta \cdot \log^2(n)) \cdot \delta^{\beta \cdot \log^2(n)} \cdot a - q$$

(II) Das zu erwartende Einkommen einer Anfrage, die über $1 + \beta \cdot \log^2(n)$ Schritte – also einen Schritt mehr als in Fall (I) – geleitet wird, ist:

$$p(1 + \beta \cdot \log^2(n)) \cdot \delta^{1 + \beta \cdot \log^2(n)} \cdot a - q$$

Die Differenz aus (I) und (II) ist:

$$a \cdot \bar{C}^{\beta \cdot \log^2(n)} \cdot \delta^{\beta \cdot \log^2(n)} \cdot (1 - \bar{C} \cdot \delta)$$

Der Zuwachs in der Auszahlung folgt aus der Tatsache, dass ein Weiterleiter weniger die Anfrage fallenlassen kann und daraus dass das Weiterleiten mit weniger Schritten durchgeführt wird und damit schneller ist (δ). Lemma 5.3.3 folgt. \square

Es sei angemerkt, dass unkooperative Teilnehmer oder die Diskontierung dazu führen, dass der oben genannte Term positiv ist, da $(1 - \bar{C} \cdot \delta) \in (0, 1]$ gilt.

B.6 Beweis – Lemma 5.3.4

(I) Das zu erwartende Einkommen einer Anfrage, die über $\beta \cdot \log^2(n)$ Schritte geleitet wird, ist:

$$p(\beta \cdot \log^2(n)) \cdot \delta^{\beta \cdot \log^2(n)} \cdot a - q$$

(II) Das zu erwartenden Einkommen einer Anfrage, die $\epsilon > 0$ Schritte weiter geleitet werden muss ist:

$$p(\epsilon + \beta \cdot \log^2(n)) \cdot \delta^{\epsilon + \beta \cdot \log^2(n)} \cdot a - q$$

Die Differenz aus (I) und (II) ist:

$$a \cdot \bar{C}^{\beta \cdot \log^2(n)} \cdot \delta^{\beta \cdot \log^2(n)} \cdot (1 - \bar{C}^\epsilon \cdot \delta^\epsilon)$$

Es folgt Lemma 5.3.4. \square

Es sei angemerkt, dass unkooperative Teilnehmer oder die Diskontierung dazu führen, dass der oben genannte Term positiv ist, da $(1 - \bar{C}^\epsilon \cdot \delta^\epsilon) \in (0, 1]$ gilt. Daher entspricht es der Auszahlungsmaximierung, wenn ein Teilnehmer einen Kontakt mit einer nicht-optimalen Kontaktverteilung gegen einen Kontakt austauscht, der Kleinberg-verteilte zusätzliche Kontakte besitzt.

C Experimentdaten

C.1 Auszahlung

Es folgend die Auszahlungen in den Experimentreihen T_0 , T_1 , T_3 und T_5 . Jede Tabelle repräsentiert dabei eine Experimentreihe: Tabelle C.1 Experimentreihe T_0 , Tabelle C.2 Experimentreihe T_1 , Tabelle C.3 Experimentreihe T_3 und Tabelle C.4 Experimentreihe T_5 . Die erste Spalte zeigt die ID des Knotens, d.h. eines Teilnehmers am System. (Die IDs werden in jedem Experiment neu zugeordnet. D.h. das Subjekt mit der ID 3 in Experiment 1 ist nicht dasselbe Subjekt, wie das mit der ID 3 in Experiment 2.) Die folgenden Spalten zeigen an, um welches konkrete Experiment es sich gehandelt hat. Die Werte geben die Auszahlung des Subjekts, das den entsprechenden Knoten gesteuert hat, wieder.

Knoten ID	Exp. 2	Exp. 4	Exp. 15	Exp. 17
0	-13	141	98	136
1	40	198	123	61
2	50	209	73	163
3	22	233	200	102
4	23	229	188	30
5	27	223	129	37
6	55	216	193	174
7	60	78	9	172
8	5	134	124	392
9	48	125	208	113
10	-31	157	91	141
11	50	190	170	176

Tabelle C.1: Auszahlung z in der Experimentreihe T_0

ANHANG C. EXPERIMENTDATEN

Knoten ID	Exp. 1	Exp. 3	Exp. 5	Exp. 7	Exp. 9	Exp. 19
0	20	89	91	42	97	43
1	72	183	39	-3	-2	102
2	65	49	108	105	64	15
3	59	119	-48	42	72	130
4	73	242	-12	74	-50	114
5	133	3	119	116	88	126
6	92	132	-8	147	-14	151
7	242	111	13	85	-24	123
8	27	90	31	66	37	9
9	61	108	38	50	45	172
10	46	144	-24	50	13	153
11	186	77	49	82	-8	69

Tabelle C.2: Auszahlung z in der Experimentreihe T_1

Knoten ID	Exp. 6	Exp. 10	Exp. 11	Exp. 13	Exp. 16	Exp. 18	Exp. 20
0	-65	88	-32	12	-31	-31	-82
1	62	-5	-3	-13	-12	-1	-61
2	-48	26	-18	84	-138	22	89
3	5	14	-6	61	169	-146	-82
4	1	-28	-2	54	46	-31	60
5	-11	-18	-56	71	90	70	74
6	-130	-60	174	12	-50	-174	-74
7	34	-29	-131	-27	62	23	63
8	48	-98	-163	54	-2	50	121
9	69	80	-71	-161	3	-86	-28
10	-19	0	62	20	-58	-104	142
11	31	-68	-58	-32	157	-113	37

Tabelle C.3: Auszahlung z in der Experimentreihe T_3

C.2. ZUSÄTZLICHE KONTAKTE

Knoten ID	Exp. 8	Exp. 12	Exp. 14
0	-2	114	-46
1	5	51	-60
2	-117	-8	-5
3	-35	-79	-171
4	24	98	-89
5	-105	-14	-41
6	-33	14	-48
7	-103	-41	4
8	-130	24	-49
9	-191	-123	95
10	-114	10	11
11	-254	7	89

Tabelle C.4: Auszahlung z in der Experimentreihe T_5

C.2 Zusätzliche Kontakte

Es folgen die Daten über die Kontakte in den Experimenten.

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	8	6	2
1	6	4	2
2	4	0	4
3	8	2	6
4	3	0	3
5	4	0	4
6	7	0	7
7	4	0	4
8	6	1	5
9	7	0	7
10	4	0	4
11	5	1	4

Tabelle C.5: Kontakte in Experiment 1 (Experimentreihe T_1)

ANHANG C. EXPERIMENTDATEN

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	5	0	5
1	1	0	1
2	1	0	1
3	4	0	4
4	1	0	1
5	2	0	2
6	1	0	1
7	2	0	2
8	3	0	3
9	4	0	4
10	3	0	3
11	2	0	2

Tabelle C.6: Kontakte in Experiment 2 (Experimentreihe T_0)

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	2	0	2
1	2	0	2
2	7	6	1
3	5	0	5
4	6	4	2
5	7	0	7
6	7	5	2
7	8	0	8
8	5	2	3
9	4	0	4
10	5	0	5
11	8	4	4

Tabelle C.7: Kontakte in Experiment 3 (Experimentreihe T_1)

C.2. ZUSÄTZLICHE KONTAKTE

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	7	0	7
1	7	0	7
2	8	0	8
3	8	0	8
4	5	0	5
5	8	0	8
6	8	0	8
7	8	0	8
8	6	0	6
9	6	0	6
10	5	0	5
11	7	0	7

Tabelle C.8: Kontakte in Experiment 4 (Experimentreihe T_0)

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	3	2	1
1	1	0	1
2	2	1	1
3	2	0	2
4	1	1	0
5	0	0	0
6	0	0	0
7	2	1	1
8	5	2	3
9	1	0	1
10	3	2	1
11	1	0	1

Tabelle C.9: Kontakte in Experiment 5 (Experimentreihe T_1)

ANHANG C. EXPERIMENTDATEN

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	2	2	0
1	1	1	0
2	1	1	0
3	1	1	0
4	0	0	0
5	3	3	0
6	3	3	0
7	2	0	2
8	3	1	2
9	4	1	3
10	1	1	0
11	3	2	1

Tabelle C.10: Kontakte in Experiment 6 (Experimentreihe T_3)

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	5	0	5
1	8	0	8
2	7	0	7
3	3	3	0
4	4	1	3
5	4	2	2
6	5	0	5
7	2	0	2
8	6	0	6
9	8	4	4
10	5	4	1
11	2	0	2

Tabelle C.11: Kontakte in Experiment 7 (Experimentreihe T_1)

C.2. ZUSÄTZLICHE KONTAKTE

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	3	3	0
1	3	3	0
2	5	4	1
3	2	2	0
4	1	1	0
5	3	3	0
6	1	1	0
7	4	3	1
8	5	4	1
9	3	3	0
10	3	0	3
11	5	2	3

Tabelle C.12: Kontakte in Experiment 8 (Experimentreihe T_5)

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	3	0	3
1	4	0	4
2	2	0	2
3	0	0	0
4	3	2	1
5	6	0	6
6	4	1	3
7	3	0	3
8	4	1	3
9	4	3	1
10	4	2	2
11	2	0	2

Tabelle C.13: Kontakte in Experiment 9 (Experimentreihe T_1)

ANHANG C. EXPERIMENTDATEN

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	0	0	0
1	4	0	4
2	3	2	1
3	1	1	0
4	4	3	1
5	1	0	1
6	0	0	0
7	2	1	1
8	4	0	4
9	1	0	1
10	2	1	1
11	2	0	2

Tabelle C.14: Kontakte in Experiment 10 (Experimentreihe T_3)

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	4	2	2
1	4	1	3
2	5	4	1
3	4	2	2
4	4	0	4
5	5	4	1
6	0	0	0
7	4	0	4
8	6	5	1
9	6	2	4
10	1	1	0
11	2	2	0

Tabelle C.15: Kontakte in Experiment 11 (Experimentreihe T_3)

C.2. ZUSÄTZLICHE KONTAKTE

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	0	0	0
1	0	0	0
2	0	0	0
3	2	0	2
4	0	0	0
5	2	2	0
6	1	1	0
7	2	2	0
8	0	0	0
9	2	2	0
10	2	2	0
11	0	0	0

Tabelle C.16: Kontakte in Experiment 12 (Experimentreihe T_5)

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	6	0	6
1	6	2	4
2	2	0	2
3	7	0	7
4	3	2	1
5	3	0	3
6	5	0	5
7	5	0	5
8	3	3	0
9	7	5	2
10	5	5	0
11	2	0	2

Tabelle C.17: Kontakte in Experiment 13 (Experimentreihe T_3)

ANHANG C. EXPERIMENTDATEN

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	5	4	1
1	2	0	2
2	1	1	0
3	3	0	3
4	3	2	1
5	3	2	1
6	2	0	2
7	2	2	0
8	3	2	1
9	0	0	0
10	1	0	1
11	0	0	0

Tabelle C.18: Kontakte in Experiment 14 (Experimentreihe T_5)

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	8	0	8
1	4	0	4
2	5	2	3
3	6	6	0
4	6	0	6
5	8	0	8
6	7	0	7
7	5	0	5
8	4	0	4
9	8	0	8
10	6	0	6
11	6	0	6

Tabelle C.19: Kontakte in Experiment 15 (Experimentreihe T_0)

C.2. ZUSÄTZLICHE KONTAKTE

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	1	0	1
1	4	3	1
2	7	3	4
3	0	0	0
4	2	1	1
5	5	0	5
6	3	3	0
7	1	1	0
8	3	1	2
9	7	4	3
10	2	1	1
11	0	0	0

Tabelle C.20: Kontakte in Experiment 16 (Experimentreihe T_3)

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	7	0	7
1	6	0	6
2	8	0	8
3	0	0	0
4	5	1	4
5	8	1	7
6	8	0	8
7	0	0	0
8	7	0	7
9	8	0	8
10	7	0	7
11	10	1	9

Tabelle C.21: Kontakte in Experiment 17 (Experimentreihe T_0)

ANHANG C. EXPERIMENTDATEN

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	0	0	0
1	0	0	0
2	0	0	0
3	6	0	6
4	1	0	1
5	0	0	0
6	6	0	6
7	0	0	0
8	1	0	1
9	4	0	4
10	7	3	4
11	3	3	0

Tabelle C.22: Kontakte in Experiment 18 (Experimentreihe T_3)

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	7	0	7
1	8	4	4
2	5	0	5
3	7	3	4
4	7	0	7
5	8	0	8
6	8	1	7
7	5	0	5
8	6	0	6
9	7	4	3
10	4	2	2
11	4	1	3

Tabelle C.23: Kontakte in Experiment 19 (Experimentreihe T_1)

C.3. KOOPERATIONSGRAD

Knoten ID	Kontakte hinzugefügt	Kontakte entfernt	Zusätzliche Kontakte
0	7	0	7
1	4	3	1
2	5	3	2
3	5	0	5
4	1	0	1
5	6	4	2
6	4	0	4
7	6	6	0
8	2	1	1
9	5	0	5
10	3	2	1
11	4	1	3

Tabelle C.24: Kontakte in Experiment 20 (Experimentreihe T_3)

C.3 Kooperationsgrad

Im Folgenden wird der Kooperationsgrad der Teilnehmer in den Experimenten tabellarisch aufgelistet.

Knoten ID	Exp. 2	Exp. 4	Exp. 15	Exp. 17
0	0,8333	0,8214	0,7586	1,0
1	0,5385	0,8919	0,5152	0,9038
2	0,7692	0,7391	0,8056	0,9516
3	0,9286	0,8478	0,8667	0,4839
4	0,5	0,587	0,9286	0,2545
5	0,6316	0,8788	0,5385	0,9872
6	0,8667	0,6905	0,8529	0,9565
7	0,6154	0,7838	0,9688	0,5946
8	0,4615	0,5625	0,72	0,3736
9	0,7	0,7667	0,9714	0,1667
10	0,8667	0,7632	0,7667	0,5686
11	0,2353	0,9459	0,8182	0,5185

Tabelle C.25: Kooperationsgrad C in der Experimentreihe T_0

ANHANG C. EXPERIMENTDATEN

Knoten ID	Exp. 1	Exp. 3	Exp. 5	Exp. 7	Exp. 9	Exp. 19
0	0,6071	0,55	0,6818	0,68	0,7273	0,6667
1	0,6875	0,7179	0,8571	0,7333	0,5238	0,7778
2	0,6977	0,8	0,75	0,8333	0,3333	1,0
3	0,7955	0,9	0,2632	0,5862	0,2222	0,8438
4	0,5833	0,6154	0,1053	0,7368	0,5	0,8421
5	0,6757	0,75	0,9375	0,8636	0,7273	0,9667
6	0,5962	0,6667	0,6667	0,6667	0,7917	0,9583
7	0,5833	0,7442	0,7407	0,9032	0,2632	0,9286
8	0,4865	0,7857	0,5556	0,8824	0,8	0,7297
9	0,75	0,8197	0,5714	0,8462	0,64	0,7273
10	0,7857	0,9643	0,3182	0,8	0,5	0,7188
11	0,5652	0,8049	0,5	0,7407	0,6429	0,7714

Tabelle C.26: Kooperationsgrad C in der Experimentreihe T_1

Knoten ID	Exp. 6	Exp. 10	Exp. 11	Exp. 13	Exp. 16	Exp. 18	Exp. 20
0	0,6667	0,0385	0,8095	0,9355	0,7586	0,7727	0,9
1	0,625	0,5429	0,7879	0,6444	0,7059	0,4783	0,5667
2	0,5556	0,8	0,55	0,6774	0,9375	0,5385	0,8
3	0,75	0,375	0,6364	0,75	0,5758	0,2414	0,75
4	0,5625	0,5556	0,8947	0,8276	0,64	0,5926	0,8333
5	0,4286	0,4706	0,7692	0,7857	0,9655	0,6667	0,7027
6	0,25	0,8	0	1,0	0,7576	0,6071	0,65
7	0	0,3158	0,68	0,9737	0,5278	0,35	0,7429
8	0,7333	0,0952	0,5517	0,8261	0,3	0,4516	0,7742
9	0,5909	0,6765	0,907	0,4667	0,6154	0,7692	0,9063
10	0,6667	0,6842	0,6667	0,8947	0,9118	0,9524	0,9583
11	0,5758	0,7083	0,7273	0,4688	0,3793	0,2188	0,75

Tabelle C.27: Kooperationsgrad C in der Experimentreihe T_3

C.3. KOOPERATIONSGRAD

Knoten ID	Exp. 8	Exp. 12	Exp. 14
0	0,3636	0,5217	0,9
1	0,5185	0,6522	0,5
2	0,6667	0,7083	0,5217
3	0,4722	0,619	0,5
4	0,4	0,25	0,72
5	0,3333	0,6875	0,5833
6	0,4194	0,7037	0,6667
7	0,4286	0,2381	0,3
8	0,9231	0	0,6176
9	0,3929	0,6087	0,4
10	0,5313	0,5926	0,8
11	0,4516	0,7619	0,3226

Tabelle C.28: Kooperationsgrad C in der Experimentreihe T_5

ANHANG C. EXPERIMENTDATEN

Literaturverzeichnis

- [AH00] Eytan Adar and Bernardo A. Huberman. Free Riding on Gnutella. *First Monday*, 5(10), 2000.
- [AK01] Guillermo Abramson and Marcelo Kuperman. Social Games in a Social Network. *Physical Review E*, 63(3), 2001.
- [ATS04] Stephanos Androutsellis-Theotokis and Diomidis Spinellis. A Survey of Peer-to-Peer Content Distribution Technologies. *ACM Computing Surveys*, 36(4):335–371, 2004.
- [Axe06] Robert M. Axelrod. *The Evolution of Cooperation*. Perseus Books Group, 2006.
- [BB04] Erik Buchmann and Klemens Böhm. FairNet – How to Counter Free Riding in Peer-to-Peer Data Structures. In *Proceedings of Cooperative Information Systems*, 2004.
- [BC88] Béla Bollobás and Fan R. K. Chung. The Diameter of a Cycle Plus a Random Matching. *SIAM Journal on Discrete Mathematics*, 1(3):328–333, 1988.
- [BDH⁺08] Richard Baron, Jacques Durieu, Hans Haller, Rahul Savani, and Philippe Solal. Good neighbors are hard to find: computational complexity of network formation. *Review of Economic Design*, 12(1):1–19, 2008.
- [BEG06] Siegfried K Berninghaus, Karl-Martin Ehrhart, and Werner Güth. *Strategische Spiele*. Springer, 2006.
- [BEO10] Siegfried K Berninghaus, Karl-Martin Ehrhart, and Marion Ott. Cooperation and Forward-Looking Behavior in Hawk-Dove Games in Endogenous Networks: Experimental Evidence. Working paper, Karlsruher Institut für Technologie (KIT), 2010.
- [BEOV07] Siegfried Berninghaus, Karl-Martin Ehrhart, Marion Ott, and Bodo Vogt. Evolution of Networks – an Experimental Analysis. *Journal of Evolutionary Economics*, 17(3):317–347, 2007.

Literaturverzeichnis

- [BG00a] Venkatesh Bala and Sanjeev Goyal. A Strategic Analysis of Network Reliability. *Review of Economic Design*, 5(3):205–228, 2000.
- [BG00b] Venkatesh Bala and Sanjeev Goyal. A noncooperative model of network formation. *Econometrica*, 68(5):1181–1230, September 2000.
- [BV04] Siegfried Berninghaus and Bodo Vogt. Network Formation and Coordination Games. *Advances in Understanding Strategic Behaviour: Game Theory, Experiments, and Bounded Rationality*, 2004.
- [CB10] Rense Corten and Vincent Buskens. Co-Evolution of Conventions and Networks: An Experimental Study. *Social Networks*, 32:4–15, 2010.
- [CC05] N. Christin and J. Chuang. A Cost-Based Analysis of Overlay Routing Geometries. In *Proceedings of the IEEE Conference on Computer Communications*, 2005.
- [CD08] Dean Corbae and John Duffy. Experiments with Network Formation. *Games and Economic Behavior*, 64(1):81–120, 2008.
- [CDR96] Russell Cooper, Douglas W. DeJong, and Thomas W. Ross. Cooperation without Reputation: Experimental Evidence from Prisoner’s Dilemma Games. *Games and Economic Behavior*, 12(2):187–218, 1996.
- [CFSK04] Byung-Gon Chun, Rodrigo Fonseca, Ion Stoica, and John Kubiatiowicz. Characterizing Selfishly Constructed Overlay Routing Networks. In *Proceedings of the IEEE Conference on Computer Communications*, 2004.
- [Cha43] Subrahmanyan Chandrasekhar. Stochastic Problems in Physics and Astronomy. *Reviews of Modern Physics*, 15(1):1–89, 1943.
- [Chr06] Larry B. Christensen. *Experimental Methodology*. Boston: Allyn and Bacon, 2006.
- [CT94] Leda Cosmides and John Tooby. Better than Rational: Evolutionary Psychology and the Invisible Hand. *American Economic Review*, 84:327–332, 1994.
- [CW79] J. Lawrence Carter and Mark N. Wegman. Universal Classes of Hash Functions. *Journal of Computer and System Sciences*, 18(2):143–154, 1979.
- [DMW03] Peter S. Dodds, Roby Muhamad, and Duncan J. Watts. An Experimental Study of Search in Global Social Networks. *Science*, 301(5634):827–829, 2003.

- [EFM07] Amir Epstein, Michal Feldman, and Yishay Mansour. Strong Equilibrium in Cost Sharing Connection Games. In *Proceedings of the ACM Conference on Electronic Commerce*, 2007.
- [FCSS05] Michal Feldman, John Chuang, Ion Stoica, and Scott Shenker. Hidden-Action in Multi-Hop Routing. In *Proceedings of the ACM Conference on Electronic Commerce*, 2005.
- [FF03] Ernst Fehr and Urs Fischbacher. The nature of human altruism. *Nature*, 425(6960):785–791, 2003.
- [FK03] Armin Falk and Michael Kosfeld. It’s all about Connections: Evidence on Network Formation. *IZA DP 777*, 2003.
- [FLSC04] Michal Feldman, Kevin Lai, Ion Stoica, and John Chuang. Robust Incentive Techniques for Peer-to-Peer Networks. In *Proceedings of the ACM Conference on Electronic Commerce*, 2004.
- [FM86] Drew Fudenberg and Eric Maskin. The Folk Theorem in Repeated Games with Discounting or with Incomplete Information. *Econometrica*, 54(3):533–554, 1986.
- [FPCS06] Michal Feldman, Christos Papadimitriou, John Chuang, and Ion Stoica. Free-Riding and Whitewashing in Peer-to-Peer Systems. *IEEE Journal on Selected Areas in Communications*, 24(5):1010–1019, 2006.
- [GFJ⁺03] Zihui Ge, Daniel R. Figueiredo, Sharad Jaiswal, Jim Kurose, and Don Towsley. Modeling Peer-Peer File Sharing Systems. In *Proceedings of the IEEE International Conference on Computer Communications*, 2003.
- [GGS03] Georg Gottlob, Gianluigi Greco, and Francesco Scarcello. Pure Nash Equilibria: Hard and Easy Games. In *Proceedings of the ACM Conference on Theoretical Aspects of Rationality and Knowledge*, 2003.
- [GM09] Veronika Grimm and Friederike Mengel. Cooperation in Viscous Populations—Experimental Evidence. *Games and Economic Behavior*, 66(1):202–220, 2009.
- [GRU09] Jacob K. Goeree, Arno Riedl, and Aljaž Ule. In Search of Stars: Network Formation among Heterogeneous Agents. *Games and Economic Behavior*, 67(2):445–466, 2009.
- [GVR05] Sanjeev Goyal and Fernando Vega-Redondo. Learning, Network Formation and Coordination. *Games and Economic Behavior*, 50, 2005.

Literaturverzeichnis

- [HB10a] Björn-Oliver Hartmann and Klemens Böhm. Incentivizing connectivity in structured Peer-to-Peer systems. *Web Intelligence and Agent Systems An International Journal*, 8(2):123–147, 2010.
- [HB10b] Björn-Oliver Hartmann and Klemens Böhm. Towards Efficient Equilibria of Combinations of Network-Formation and Interaction Strategies. In *Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology*, 2010.
- [HBB⁺01] Joseph Henrich, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath. In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies. *American Economic Review*, 91(2):73–78, 2001.
- [HBKS07] Björn-Oliver Hartmann, Klemens Böhm, Andranik Khachatryan, and Stephan Schosser. The Dangers of Poorly Connected Peers in Structured P2P Networks and a Solution Based on Incentives. In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence*, 2007.
- [HBV08] Björn-Oliver Hartmann, Klemens Böhm, and Bodo Vogt. The Network Game: Analyzing Network-Formation and Interaction Strategies in Tandem. In *Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology*, 2008.
- [Jac03] Matthew O. Jackson. A Survey of Models of Network Formation: Stability and Efficiency. In *Group Formation in Economics: Networks, Clubs, and Coalitions*. Cambridge University Press, 2003.
- [JF04] Radu Jurca and Boi Faltings. “CONFESS”. An Incentive Compatible Reputation Mechanism for the Online Hotel Booking Industry. In *Proceedings of the IEEE Conference on E-Commerce*, 2004.
- [JIB07] Audun Jøsang, Roslan Ismail, and Colin Boyd. A Survey of Trust and Reputation Systems for Online Service Provision. *Decision Support Systems*, 43(2):618–644, 2007.
- [JK08] J. Stephen Judd and Michael Kearns. Behavioral Experiments in Networked Trade. In *Proceedings of the ACM Conference on Electronic Commerce*, 2008.
- [JW96] Matthew O. Jackson and Asher Wolinsky. A Strategic Model of Social and Economic Networks. *Journal of Economic Theory*, 71(1):44–74, 1996.

- [KB78] Peter Killworth and H. Russel Bernard. Reverse Small World Experiment. *Social Networks*, 1(2):159–192, 1978.
- [Kle00] Jon Kleinberg. The Small-World Phenomenon: An Algorithmic Perspective. In *ACM Symposium on Theory of Computing*, 2000.
- [Kle06] Jon Kleinberg. Complex Networks and Decentralized Search Algorithms. In *International Congress of Mathematicians (ICM)*. ACM Press, 2006.
- [KvW00] Claudia Keser and Frans van Winden. Conditional Cooperation and Voluntary Contributions to Public Goods. *Scandinavian Journal of Economics*, 102(1):23–39, 2000.
- [LCP⁺05] Eng K. Lua, Jon Crowcroft, Marcelo Pias, Ravi Sharma, and Steven Lim. A Survey and Comparison of Peer-to-Peer Overlay Network Schemes. *IEEE Communications Surveys & Tutorials*, 7(2):72–93, 2005.
- [LJ06] Haydée Lugo and Raúl Jiménez. Incentives to Cooperate in Network Formation. *Computational Economics*, 28(1):15–27, 2006.
- [Lor99] Edward Lorenz. Trust, Contract and Economic Cooperation. *Cambridge Journal of Economics*, 23(3):301–15, 1999.
- [Mil67] Stanley Milgram. The Small World Problem. *Psychology Today*, 1967.
- [Mil09] Holger Miller. Design, Implementation and Evaluation of a Web-Based Platform to Study Online Networks Equilibria. Diplomarbeit, Karlsruhe Institute of Technology (KIT), 2009.
- [MMG⁺07] Alan Mislove, Massimiliano Marcon, Krishna P. Gummadi, Peter Druschel, and Bobby Bhattacharjee. Measurement and Analysis of Online Social Networks. In *Proceedings of the ACM SIGCOMM Conference on Internet Measurement*, 2007.
- [MS74] J. Maynard Smith. The Theory of Games and the Evolution of Animal Conflicts. *Journal of Theoretical Biology*, 47(1):209–221, 1974.
- [Nas54] John F. Nash. Non-Cooperative Games. *Annals of Mathematics*, 54(2):286–295, 1954.
- [Now06] Martin Nowak. Five Rules for the Evolution of Cooperation. *Science*, 314(5805):1560–1563, 2006.
- [NS05] Martin Nowak and Karl Sigmund. Evolution of Indirect Reciprocity. *Nature*, 437(7063):1291–1298, 2005.

Literaturverzeichnis

- [Qui93] John Quiggin. *Generalized Expected Utility Theory: The Rank-dependent Model*. Springer, 1993.
- [RFH⁺01] Sylvia Ratnasamy, Paul Francis, Mark Handley, Richard Karp, and Scott Shenker. A Scalable Content-Addressable Network. In *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, 2001.
- [RKZF00] Paul Resnick, Ko Kuwabara, Richard Zeckhauser, and Eric Friedman. Reputation Systems. *Communications of the ACM*, 43(12):45–48, 2000.
- [RLS⁺03] Ananth Rao, Karthik Lakshminarayanan, Sonesh Surana, Richard Karp, and Ion Stoica. Load Balancing in Structured P2P Systems. *Lecture Notes in Computer Science*, 2735:68–79, 2003.
- [RS07] Paul Resnick and Rahul Sami. The Influence Limiter: Provably Manipulation-Resistant Recommender Systems. In *Proceedings of the ACM Conference on Recommender Systems*, 2007.
- [RT02] Tim Roughgarden and Éva Tardos. How Bad Is Selfish Routing? *Journal of the ACM*, 49(2):236–259, 2002.
- [RU02] Arno Riedl and Aljaž Ule. Exclusion and Cooperation in Social Network Experiments. Working paper, University of Amsterdam, 2002.
- [RW90] Werner Raub and Jeroen Weesie. Reputation and Efficiency in Social Interactions: An Example of Network Effects. *American Journal of Sociology*, 96(3):626–654, 1990.
- [SBSV06] Stephan Schosser, Klemens Böhm, Rainer Schmidt, and Bodo Vogt. Incentives Engineering for Structured P2P Systems – a Feasibility Demonstration using Economic Experiments. In *Proceedings of the ACM Conference on Electronic Commerce*, 2006.
- [SBV07] Stephan Schosser, Klemens Böhm, and Bodo Vogt. Indirect Partner Interaction in Peer-to-Peer Networks: Stimulating Cooperation by Means of Structure. In *Proceedings of the ACM Conference on Electronic Commerce*, 2007.
- [Sch08] Stephan Schosser. Strategische Analyse von Anreizmechanismen in strukturierten Peer-to-Peer Systemen. Dissertation, Fakultät für Informatik, Universität Karlsruhe (TH), 2008.
- [Sel73] Reinhard Selten. A simple model of imperfect competition, where 4 are few and 6 are many. *International Journal of Game Theory*, 2(1):141–201, 1973.

- [SGG02] Stefan Saroiu, P. Krishna Gummadi, and Steven D. Gribble. A Measurement Study of Peer-to-Peer File Sharing Systems. In *Proceedings of the Conference of Multimedia Computing and Networking*, 2002.
- [SMK⁺01] Ion Stoica, Robert Morris, David Karger, M. Frans Kaashoek, and Hari Balakrishnan. Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications. In *Proceedings of the ACM SIGCOMM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, 2001.
- [SMU97] Reinhard Selten, Michael Mitzkewitz, and Gerald R. Uhlich. Duopoly Strategies Programmed by Experienced Players. *Econometrica*, 65(3):517–556, 1997.
- [SW02] Subhabrata Sen and Jia Wang. Analyzing Peer-to-Peer Traffic Across Large Networks. In *Proceedings of the ACM SIGCOMM Workshop on Internet Measurement*, 2002.
- [Ule05] Aljaž Ule. *Exclusion and Cooperation in Networks*. PhD thesis, Tinbergen Institute, Research Series no. 354, 2005.
- [USRC09] Aljaž Ule, Arthur Schram, Arno Riedl, and Timothy N. Cason. Indirect Punishment and Generosity Towards Strangers. *Science*, 326:1701–1704, 2009.
- [Van02] Paolo Vanin. *Network Formation in the Lab: A Pilot Experiment*. Universitat Pompeu Fabra, 2002.
- [VR06] Greg Valiant and Tim Roughgarden. Braess’s Paradox in Large Random Graphs. In *Proceedings of the ACM Conference on Electronic Commerce*, 2006.
- [Wil45] Frank Wilcoxon. Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, 1(6):80–83, 1945.
- [ZA07] Manaf Zghaibeh and Kostas G. Anagnostakis. On the Impact of P2P Incentive Mechanisms on User Behavior. In *ACM Joint Workshop on The Economics of Networked Systems and Incentive-Based Computing*, 2007.
- [ZL06] Shiwu Zhang and Jiming Liu. From Local Behaviors to the Dynamics in an Agent Network. In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence*, 2006.