



TU Clausthal
Institut für Informatik

Masterarbeit

**Empfehlungssysteme für
kleine Online-Communities
mit regionaler Bindung**

Sven Strickroth

23. Oktober 2011

Masterarbeit

Empfehlungssysteme für kleine Online-Communities mit regionaler Bindung

eingereicht bei

Betreuender Prüfer: Prof. Dr. Niels Pinkwart

Zweitgutachter: Prof. Dr. Jörg P. Müller

Institut für Informatik

Abteilung für Wirtschaftsinformatik

Fakultät für Mathematik/Informatik und Maschinenbau

Technische Universität Clausthal

von

Sven Strickroth, B. Sc.

Studienrichtung: Informatik, M. Sc.

Matrikelnummer: 352752

Datum: 23. Oktober 2011

Inhaltsverzeichnis

Abbildungsverzeichnis	iii
Tabellenverzeichnis	iv
Abkürzungsverzeichnis	v
Symbolverzeichnis	vi
1 Einleitung	1
2 Online-Communities	3
2.1 Definition „Community“	3
2.2 Definition „Online-Community“	5
2.3 Vergleich von Online-Communities	7
2.3.1 Beispiel einer großen Online-Community	7
2.3.2 Beispiel einer kleinen Online-Community	11
2.3.3 Unterschiede und Gemeinsamkeiten	14
2.4 Zusammenfassung	15
3 Empfehlungssysteme	17
3.1 Definition	17
3.2 Definition „Bewertung“	20
3.3 Ähnlichkeitsmaße	22
3.3.1 Definition	22
3.3.2 Euklidischer Abstand	22
3.3.3 Jaccard-Ähnlichkeitsmaß	23
3.3.4 Pearson-Korrelation	23
3.3.5 Cosinus-Korrelation	24
3.3.6 Angepasste Cosinus-Korrelation	24
3.4 Manuelles Empfehlen	24
3.5 Durchschnittsbasiertes Filtern	25
3.6 Inhaltsbasiertes Filtern	25
3.7 Wissensbasiertes Filtern	26
3.8 Demographisches Filtern	27
3.9 Kollaboratives Filtern	28
3.9.1 Nutzerbasiertes Filtern	28
3.9.2 Elementbasiertes Filtern	32
3.10 Allgemeine Probleme und Lösungsansätze	33
3.10.1 Sparsity	34
3.10.2 Cold-Start-Probleme	35
3.10.3 Transparenz-, Blackbox- bzw. Vertrauens-Problem	36

3.10.4	Datenschutz und Privatsphäre	37
3.10.5	Fehlende Diversifikation	37
3.11	Zusammenfassung	38
4	Hybrider Empfehlungssystem-Prototyp für Mobile2Learn	41
4.1	Anforderungen und Entwurfsentscheidungen	41
4.2	Architektur- und Algorithmusbeschreibung	44
4.2.1	Vorbereitende Schritte und Modellbildung	45
4.2.2	Berechnung von Empfehlungen	50
4.3	Implementation	57
4.4	Präsentation der Empfehlungen	58
4.5	Parameterwahl	61
4.5.1	Evaluationsmetrik	61
4.5.2	Parameteroptimierung	63
4.6	Übertragbarkeit des Ansatzes	69
5	Evaluation	71
5.1	Offline-Evaluation	72
5.2	Expertenbefragung	73
5.3	Feldstudie	76
5.4	Zusammenfassung	79
6	Ausblick	81
7	Zusammenfassung	83
	Literaturverzeichnis	85
	Anhang	95

Abbildungsverzeichnis

2.1	Screenshot der Startseite von eltern.de	8
2.2	Screenshot der eltern.de Foren	10
2.3	Screenshot der Startseite von Mobile2Learn.de	12
3.1	Einordnung Empfehlungssysteme	19
3.2	Bewertungsskalen großer Webportale	20
3.3	Übersicht der Datenquellen aller vorgestellten Filterverfahren	38
4.1	Artikel-Leseverhalten der Mobile2Learn-Benutzer	43
4.2	Gesamtarchitektur des hybriden Empfehlungssystems	45
4.3	Vorgehen bei der Modellbildung	46
4.4	Aktivität: Verweildauer-Analyse durchführen	46
4.5	Wahrscheinlichkeitsdichte mit Bewertungssegmenten und relativer Segmentgröße	48
4.6	Generalisierter (Artikel-)Baum	49
4.7	Architektur des hybriden Empfehlungssystems	51
4.8	Aktivität: Regelbasierte Empfehlungen berechnen	52
4.9	Aktivität: α -Communities berechnen	54
4.10	Übersicht der Besucher-Interessen je Veranstaltungsort	56
4.11	Screenshot der Mobile2Learn.de-Startseite mit Empfehlungen	59
4.12	Ausschnitt einer dynamischen Veranstaltungsankündigungs-E-Mail	60
4.13	Ausschnitt einer allgemeinen Veranstaltungsankündigungs-E-Mail	60
4.14	Übersicht der Parameter-Evaluation	64
4.15	Hit-Ratio der α -Communities	65
4.16	Hit-Ratio in Abhängigkeit der minimalen Anzahl von Bewertungen beim durchschnittsbasierten Filter	65
4.17	Prozentualer Anteil der Artefakte sortiert nach Anzahl von Bewertungen	66
4.18	Hit-Ratio in Abhängigkeit der Signifikanzgewichtung im kollaborativen Filter	66
4.19	Hit-Ratio in Abhängigkeit der Variable $t_{\text{agreement}}$ der α -Communities	67
4.20	Hit-Ratio in Abhängigkeit unterschiedlicher Ähnlichkeitsmaße und Signifikanzgewichtungen	67
4.21	Hit-Ratio in Abhängigkeit unterschiedlicher Filterbot-Implementationen	69
4.22	Hit-Ratios der einzelnen Verfahren und Kombinationen	69
5.1	Hit-Ratio-Vergleich zwischen Prototyp und Zufallsempfehlungen	72
5.2	Hit-Ratios für verschiedene Benutzerklassen	73
5.3	Expertenbefragung: Klassifikation der Benutzer	74
5.4	Auswertung der Expertenbefragung	74
5.5	Normalisierte Auswertung der Expertenbefragung	75
5.6	Screenshot einer abgerufenen Empfehlung mit Feedbackmöglichkeit	77

Tabellenverzeichnis

3.1	Mögliche Abbildungen von Benutzermeinungen auf numerische Werte .	20
3.2	Mögliche Abbildungen von Benutzeraktionen auf numerische Werte . .	21
3.3	Bewertungsmatrix für Beispiel des kollaborativen Filters	31
3.4	Bewertungsmittelwerte der Benutzer für Beispiel des kollaborativen Filters	31
3.5	Ähnlichkeitsvektor für Beispiel des nutzerbasierten kollaborativen Filters	31
3.6	Ähnlichkeitsmatrix für Beispiel des elementbasierten kollaborativen Filters	33
3.7	Übersicht über Probleme und Stärken aller vorgestellten Filterverfahren	39
4.1	Transformationsvorschrift der normalisierten Verweildauern auf die Bewertungsskala	47
4.2	Filterbotbewertungen: Bewertungsvektoren der Artefakte „a“ bis „j“ . .	49
4.3	Klassifizierungsmatrix zur Berechnung von Precision und Recall	62
5.1	Ergebnisse der Expertenbefragung	75
5.2	Feldstudie: Anzahl der Abrufe empfohlener Artefakte und Rückmeldungen in Abhängigkeit vom Lesetyp	78

Abkürzungsverzeichnis

AC	α -Community-Räume
AJAX	Asynchronous JavaScript and XML
API	Application Programming Interface
ARPANet	Advanced Research Projects Agency Network
AVG	durchschnittsbasiertes Filtern
CB	inhaltsbasierte bzw. eigenschaftsbasierte Filter (engl. content-based bzw. feature-based)
CF	kollaboratives Filtern
CGI	Common Gateway Interface
CMS	Content Management System
CSS	Cascading Style Sheets
DAO	Data Access Object
ES	Empfehlungssystem
FTP	File Transfer Protocol
HTML	HyperText Markup Language
HTTP	HyperText Transfer Protokoll
IB	elementbasiertes kollaboratives Filtern
IRS	Information Retrieval System
MMS	Multimedia Messaging Service
MVC	Model View Controller
NLP	Natural Language Processing
ORM	Object Relational Mapping
RFC	Requests for Comments
SMS	Short Message Service, meist Synonym für Textnachricht
SMTP	Simple Mail Transfer Protocol
SVD	Singular Value Decomposition
UB	nutzerbasiertes kollaboratives Filtern
URL	Uniform Resource Locator

Symbolverzeichnis

\mathcal{I}	Menge der Artefakte
\mathcal{I}_u	Menge der Artefakte, die der Benutzer u bewertet hat
k	Maximale Größe einer Nachbarschaft von ähnlichen Artefakten bzw. Benutzern
m	Anzahl der Artefakte ($m = \mathcal{I} $)
N	Maximale Anzahl der empfohlenen Artefakte im TOP- N -Ranking
n	Anzahl der Benutzer ($n = \mathcal{U} $)
$p_{u,i}$	Vorhersage einer Bewertung für Benutzer u und Artefakt i
\mathcal{R}	$n \times m$ Bewertungsmatrix
r_i	Bewertungen des Artefakts i , Spaltenvektor der Bewertungsmatrix \mathcal{R}
r_u	Bewertungen des Benutzers u , Zeilenvektor der Bewertungsmatrix \mathcal{R}
\bar{r}_u	Durchschnittsbewertung des Benutzers u
$r_{u,i}$	Bewertung von Benutzer u für Artefakt i , Eintrag der Bewertungsmatrix \mathcal{R}
\mathcal{S}	$m \times m$ Ähnlichkeitsmatrix
$s(\cdot, \cdot)$	Ähnlichkeitsmaß, vgl. Abschnitt 3.3.1
\mathcal{U}	Menge der Benutzer

1 Einleitung

In der heutigen Zeit wird man mit einer wachsenden Zahl von Informationen konfrontiert: Es gibt unzählige Webseiten, Blogs, News, Foren usw. – täglich werden es mehr. Aus diesem unübersichtlichen Angebot interessante Inhalte gezielt herauszusuchen, ist sehr aufwändig bzw. fast unmöglich. Dieses Problem ist allgegenwärtig: Beim Online-Shopping, beim Ausschuchen von Büchern, Filmen oder Urlaubszielen oder beim Lesen von Online-Zeitungen, Foren und Webseiten, um nur eine kleine Auswahl zu nennen.

Der Einsatz von reinen Informationssystemen ist hier nicht immer hilfreich, wenn relevante Inhalte zwar vorhanden sind, aber von den Benutzern nicht gefunden werden. Eine einfache Suche funktioniert nur bedingt, da Benutzer deren Existenz kennen und zudem noch die „richtigen“ Suchbegriffe verwenden müssen.

Herkömmliche Werbung weist zwar auf neue Produkte hin, spricht dabei aber immer wieder die gleichen unpersonalisierten „Empfehlungen“ aus und versucht so, Menschen zum Kauf zu motivieren.

Häufig entdeckt man wirklich interessante Inhalte nur durch Zufall oder „Mundpropaganda“ von Freunden und Bekannten. Gerade Hinweisen aus dem Bekanntenkreis bringt man ein verstärktes Vertrauen entgegen, da Freunde und Bekannte die eigenen Vorlieben und Interessen kennen. Menschen sind förmlich „hungrig“ nach Empfehlungen (vgl. Konstan u. a. 1998, S. 61).

Genau an diesem Punkt setzen (automatisierte) Empfehlungssysteme an: Sie versuchen die „Mundpropaganda“ nachzubilden bzw. zu verbessern (Shardanand u. Maes 1995): Sie „lernen“ die Interessen und Vorlieben der einzelnen Nutzer und geben basierend darauf personalisierte Ratschläge.

Inzwischen sind Empfehlungssysteme aus der heutigen Welt nicht mehr wegzudenken: Viele Online-Shopping-Webseiten wie Amazon.com, Link-Empfehlungsseiten, Dating-Seiten u. v. a. setzen sie ein. E-Commerce gewinnt an Popularität. Daher ist es von enormer Bedeutung, den Kunden die für sie interessantesten Produkte personalisiert anzubieten, damit ihre Zufriedenheit gesteigert wird und sie eher gewillt sind, die Webseite erneut zu besuchen, wodurch der Umsatz des Anbieters steigt (Sarwar u. a. 2000a; Sollenborn u. Funk 2002). Die Anbieter mit der besten Einstellung auf den Kunden setzen sich durch.

Vorteile durch Personalisierung sind natürlich nicht nur für große Shopping-Portale, sondern auch für kleinere Online-Communities wünschenswert. Bereits das Ausschuchen von interessanten Artikeln aus einer Menge von wenigen hundert ist erfahrungsgemäß nicht einfach und sehr zeitintensiv. Abhängig von der Zielgruppe kann es entscheidend sein, dass die Benutzer schnell die für sie relevanten Informationen finden.

In Literatur und Praxis wurden bzw. werden automatische Empfehlungssysteme nur auf (vergleichsweise) riesige Datenmengen angewendet, bei denen auf jeden Fall eine

kritische Masse von aktiven Benutzern vorhanden ist. Kleine Online-Communities hingegen besitzen eine deutlich kleinere Daten- und Nutzerbasis. Daher soll im Rahmen dieser Arbeit untersucht werden, ob Empfehlungssysteme in einem solchen Kontext einsetzbar und welche Probleme bei Entwicklung sowie Einsatz zu erwarten sind.

Daneben besitzen kleine Online-Communities mit regionalem Bezug spezielle Eigenschaften, die besondere Anforderungen an Empfehlungssysteme stellen oder auch neue Möglichkeiten bieten. Das zentrale Ziel dieser Arbeit ist es, Anforderungen, Probleme, Lösungsvorschläge sowie neue Möglichkeiten für Empfehlungssysteme in kleinen Online-Communities aufzuzeigen. Dazu wird ein neuer Ansatz vorgestellt und an einem Prototyp evaluiert.

Einen Einstieg bietet das zweite Kapitel, das sich mit dem Begriff *Online-Community* beschäftigt. Es erläutert diesen und vergleicht zwei unterschiedliche Online-Communities. Kapitel 3 behandelt den „State of the Art“ von Empfehlungssystemen und zeigt mögliche Probleme und Lösungsvarianten auf, die besonders im Kontext dieser Arbeit von Interesse sind. Das vierte Kapitel stellt einen Ansatz für ein Empfehlungssystem einer kleinen Online-Community mit regionalem Bezug vor, der im darauf folgenden Kapitel evaluiert wird, um speziell die Frage der Einsetzbarkeit zu klären.

Abschließend wird ein Ausblick über weitere mögliche Erweiterungen bzw. Untersuchungen und eine Zusammenfassung über den Inhalt der Arbeit gegeben.

2 Online-Communities

In diesem Kapitel wird der Begriff *Online-Community* erläutert und historisch beleuchtet. Ferner werden beispielhaft zwei spezielle Online-Communities, eine große mit vielen Mitgliedern und eine kleinere, vorgestellt und miteinander verglichen.

2.1 Definition „Community“

Zur Erläuterung des Begriffs *Online-Community* wird zunächst versucht, den Begriff der *Community* abzugrenzen. Es gibt eine große Anzahl verschiedener Definitionen mit unterschiedlichen Perspektiven, von denen einige vorgestellt werden, um einen Überblick zu ermöglichen.

Einen ersten Eindruck der Bedeutung kann man erhalten, wenn man diesen Begriff in einem Wörterbuch nachschlägt. Das Englisch-Wörterbuch „dictionary.cambridge.org“ der Universität Cambridge liefert z. B. die folgende Definition:

„the people living in one particular area or people who are considered as a unit because of their common interests, social group or nationality“¹

Aus dem Englischen ins Deutsche übersetzt (PONS Wörterbuch), bedeutet „Community“ allgemein „Gemeinschaft“ bzw. „Gemeinde“. Der Duden enthält für „Gemeinschaft“ bzw. „Gemeinde“ ähnliche Definitionen:

„Gruppe von Personen, die durch gemeinsame Anschauungen o. Ä. untereinander verbunden sind“.

„Gesamtheit der Mitglieder, Angehörigen einer Gemeinde“

„Gruppe von Menschen mit gleichen [...] Interessen; Anhängerschaft“

Somit kann ganz allgemein unter einer Community/Gemeinschaft „eine Anzahl von Personen, welche gemeinsame Interessen teilen, gemeinsame Ideen und Vorstellungen [...] oder einen gemeinsamen Kontext haben“ (Gross u. Koch 2007, S. 18) verstanden werden (z. B. Sportverein, Dorfgemeinschaft). Im Folgenden wird das Verständnis dieses noch recht vagen Begriffs weiter verfeinert.

Bereits sehr früh wurden Gemeinschaften in der Soziologie betrachtet. Hillery (1955) klassifizierte eine Vielzahl von Definitionen des Community-Begriffs und hob soziale Interaktion gefolgt von enger Verbundenheit und räumlicher Nähe als zentrale Merkmale hervor. Mynatt u. a. verstehen unter einer Community „a social grouping that includes, in varying degrees: shared spatial relations, social conventions, a sense of

¹abgerufen am 13. September 2011

membership and boundaries, and an ongoing rhythm of social interaction“ (Mynatt u. a. 1997, S. 211), also eine soziale Gruppierung von Personen mit räumlichen Beziehungen, sozialen Konventionen (in der Art einer Community-Kultur), dem Gefühl einer Abgrenzung zu Nicht-Mitgliedern sowie aktive soziale Interaktion. Interaktion darf in diesem Kontext nicht nur als reine Kommunikation verstanden werden, sondern auch als gemeinsame Aktivitäten wie Wissensaustausch und gegenseitige Hilfe. Auch eine gemeinsame Kultur (wie z. B. Rituale, Sprache und „Wir-Gefühl“) zählt dazu (Ishida 1998; Gross u. Koch 2007). Wenger hingegen hebt den informellen Charakter dieser Gruppierung im Kontext von Problemlösung, gemeinsamem Lernen bzw. Aktivitäten und Engagement der Mitglieder hervor: „Members of a community are informally bound by what they do together from engaging in lunchtime discussions to solving difficult problems and by what they have learned through their mutual engagement in these activities“ (Wenger 1998, S. 1). Für Carotenuto u. a. (1999) sind die Freiwilligkeit und Unabhängigkeit der Einzelnen weitere wichtige Merkmale einer Community: „a community is a voluntary association of people who are not directly dependent on each other for success“ (Carotenuto u. a. 1999, S. 2).

Da die Anzahl der Mitglieder in einer Community nicht begrenzt ist, müssen sich nicht alle Mitglieder gegenseitig kennen. Es können auch Sub-Gruppierungen innerhalb einer Community ent- bzw. bestehen. Folglich sind Communities meist nur sehr wenig koordiniert, dafür aber relativ dynamisch in Bezug sowohl auf Mitglieder als auch auf Kurs der Community (Carotenuto u. a. 1999). Whittaker u. a. (1997) identifizieren darüber hinaus die wiederholte aktive Teilnahme eines Einzelnen mit engen emotionalen Bindungen in der Community als Kerneigenschaften. Für Preece (2000) sind zudem ein gemeinsames Ziel, Zweck, verschiedene Rollen und eine gemeinsame Kultur wichtige Eigenschaften von Communities.

Carotenuto u. a. (1999) unterscheiden Community-Typen nach „Fokus“ und „Gemeinsamkeiten“, die sich jedoch nicht gegenseitig ausschließen (zusammengefasst nach Koch 2003):

Community of Interest Die Community hat einen breiten Fokus auf einer Menge gemeinsamer Interessen. Zweck der Community ist es, bei bestimmten Themengebieten auf dem Laufenden zu bleiben, Wissen auszutauschen und Wissensträger zu identifizieren.

Community of Purpose Eine *Community of Purpose* hat einen engen Fokus auf einem gemeinsamen Interesse. Im Gegensatz zur *Community of Interest* geht es hier nicht um den Wissensaustausch auf einem Themengebiet, sondern um die Erfüllung eines gemeinsamen Ziels, wie z. B. die gemeinsame Erstellung von Software.

Community of Practice Die Community hat einen engen Fokus auf einer Menge von Aktivitäten, die entweder gemeinsam oder von Einzelnen unabhängig durchgeführt werden. Diese Art von Communities ist häufig im Kontext von Organisationen anzutreffen und besteht aus Praktikern, die alle ein Mindestmaß an Kompetenz in dem gemeinsamen Fachgebiet besitzen müssen, um sich in die Community einbringen zu können. Geprägt wurde diese Variante von Wenger u. a. (2002), die drei charakteristische Dimensionen identifizieren: Ein gemeinsames Unternehmen, Engagement der Mitglieder sowie eine gemeinsame Kompetenz (vgl. Gross u. Koch 2007, S. 116). *Communities of Practice* haben ein gemeinsames

Fachgebiet, auf dem die einzelnen Mitglieder der Community ihr Wissen mit anderen teilen bzw. andere beim Wissenserwerb, z. B. durch eine gemeinsame Wissensbasis (Tipps, Erfahrungsberichte sowie „best practices“), für bestimmte Aktivitäten unterstützen (vgl. Koch 2003, S. 20). Diese können in der Praxis durch den Zusammenschluss von Personen mit einem gemeinsamen Hintergrund, ähnlichen Problemen oder Personen, die normalerweise nicht in Teams zusammen agieren, auftreten (vgl. Gross u. Koch 2007, S. 116).

Insgesamt lässt sich zusammenfassen, dass sich Personen (i. d. R. freiwillig) in Communities zusammenschließen, „um gegenseitig von der Gemeinschaft mit anderen auf die eine oder andere Art zu profitieren“ (Koch 2003, S. 29). Im Zentrum stehen die Menschen mit gemeinsamen Interessen, Leidenschaften, Problemen oder Aktivitäten, die in einer Community meist mit einer gemeinsamen Community-Kultur, jedoch ohne starke Strukturen, relativ lose miteinander interagieren bzw. sich austauschen. Ein Mitglied ist nicht mehr auf seine eigenen Fähigkeiten beschränkt, sondern kann auf die Interaktion bzw. die Erfahrungen der Community zurückgreifen.

2.2 Definition „Online-Community“

Bei Online-Communities (bzw. virtuellen Gemeinschaften) handelt es sich nicht um eine Erfindung aus dem Bereich der EDV. Vielmehr können Online-Communities als eine („Weiter-“)Entwicklung von Gemeinschaften verstanden werden, bei denen „der Hauptteil der Kommunikation und Interaktion zwischen den Mitgliedern über elektronische Medien abgewickelt“ (Carotenuto u. a. 1999; Gross u. Koch 2007, S. 117), quasi ein virtueller Ort zur Verfügung gestellt wird und es möglich ist, Menschen mit ähnlichen Interessen zu finden und mit ihnen zu kommunizieren (Gupta u. Kim 2004). Der Begriff „virtuelle Community“ selbst geht auf Rheingold zurück, der ihn bei der Beschreibung einer virtuellen Gemeinschaft Namens „The Well“ 1993 für virtuelle Nachrichtensboxen und Gruppendiskussionen zuerst benutzt hat (Gross u. Koch 2007; Koch 2003).

Die Idee hinter Online-Communities ist so alt wie die der Vernetzung von Computern. Licklider u. Taylor, zwei Pioniere des Internets, schrieben 1968 bereits ihre oft zitierte Vorhersage:

„life will be happier for the on-line individual because the people with whom one interacts most strongly will be selected more by commonality of interests and goals than by accidents of proximity.“ (Licklider u. Taylor 1968, S. 40)

Bei der Entstehung des Internets, genauer gesagt des Advanced Research Projects Agency Networks (ARPANet), wurde der zweite Dienst, das File Transfer Protocol (FTP) (Bhushan 1971), kurz nach dessen Einführung zum asynchronen Nachrichtenaustausch „missbraucht“ (Hafner u. Lyon 1998). Das Simple Mail Transfer Protocol (SMTP) (Postel 1981), Mailinglisten, das Usenet (Horton 1983) und das HyperText Transfer Protokoll (HTTP)² (Berners-Lee u. a. 1996) entstanden erst mindestens zehn Jahre später.

²Berners-Lee entwickelte HTTP ab 1989. 1991 wurde Version 0.9 freigegeben. Das für dynamische Webseiten verantwortliche Common Gateway Interface (CGI) wird erstmals 1996 in einem Requests for Comments (RFC)-Entwurf (<http://tools.ietf.org/html/draft-robinson-www-interface-00>) erwähnt.

Die von Licklider u. Taylor getroffene Vorhersage ist durch den Siegeszug des Internets als virtueller Ort für Interaktion eingetreten: „Gemeinschaften waren nicht länger regional begrenzt, so dass geographisch getrennte Individuen mit gleichen Interessen, Persönlichkeiten und Wertvorstellungen sich zu Gemeinschaften zusammenschließen konnten“ (Hartleb 2009, S. 9).

Wie bereits bei den realen Communities gibt es keine eindeutige Definition von Online-Communities. Zusätzlich zur Zusammenfassung des letzten Abschnitts sowie der einleitenden Definition von Carotenuto u. a. (1999) werden nochmals speziellere Definitionen vorgestellt, die die Vernetzung bzw. Computer-Unterstützung deutlicher hervorheben, um weitere Facetten des Begriffs besser fassbar zu machen.

Carroll u. Rosson (1998) definieren eine Online-Community als eine Gruppe von Personen, die über Netzwerke sowohl kommunizieren als auch kollaborieren und damit ihre gemeinsamen Ziele sowie Anliegen stärken. Nach einer Zusammenfassung mehrerer Definitionen sind für Lee u. a. Online-Communities „cyberspace(s) supported by computer-based information technology, centered upon communication and interaction of participants to generate member-driven content, resulting in a relationship being built“ (Lee u. a. 2003, S. 50). Die Psychologin Blanchard definiert Online-Communities als „groups of people who interact primarily through computer-mediated communication and who identify with and have developed feelings of belonging and attachment to each other“ (Blanchard 2004, S. 55).

Sämtliche angeführten Definitionen decken sich im Kern mit der Zusammenfassung, ergänzt um die Verwendung elektronischer Medien. Es wird jedoch oftmals keine aktive direkte Interaktion zwischen den Mitgliedern gefordert, sondern vielmehr die generelle Möglichkeit dafür (Gross u. Koch 2007, S. 19).

Trotz der im Grundsatz gleichen Definition von Communities und Online-Communities weisen letztere auf Grund der elektronischen Kommunikation besondere Charakteristika auf (nach Koch 2003; Gross u. Koch 2007):

Kommunikation Durch die Nutzung elektronischer Kommunikation ist die Interaktion miteinander oft rein text-basiert. D. h. es fehlen oftmals Bild und Ton – damit auch die nonverbale Kommunikation. Zur Kompensation werden oft spezielle Symbole, Emoticons bzw. Smileys, oder andere Awareness-unterstützende Techniken eingesetzt. Daneben hat die elektronisch vermittelte Kommunikation jedoch den Vorteil, dass one-to-one-, one-to-many- und many-to-many-Kommunikation relativ einfach ermöglicht wird, so dass z. B. die in einem Forum geführten Konversationen für andere Mitglieder auch deutlich später nachvollzogen werden können.

Anonymität Online-Communities bieten den Mitgliedern die Möglichkeit, anonym oder unter Verwendung anderer Identitäten bzw. Pseudonymen (für die Wiedererkennbarkeit) aufzutreten. Anonymität kann sich insgesamt sowohl positiv als auch negativ auf eine Gemeinschaft auswirken: Sie kann die Hemmschwelle senken, überhaupt Kontakt herzustellen oder z. B. über (sozial-)problematische Themen offen zu sprechen. Durch ein rein text-basiertes Medium können zudem Geschlecht, Alter, Aussehen usw. in den Hintergrund treten und so eine vorurteilsfreie bzw. von diesen Faktoren unabhängige Mitgliedschaft ermöglichen. Andererseits kann Anonymität den Aufbau von sozialen Bindungen zwischen

Community-Mitgliedern erschweren oder zum Überschreiten allgemeiner sozialer Konventionen führen.

Unabhängigkeit von realen Örtlichkeiten Wie bereits mehrfach erwähnt, kann die Ortsunabhängigkeit einen enormen Vorteil gegenüber einer normalen Community bieten: Durch die räumliche Unbeschränktheit können sich Communities bilden, die in der realen Welt so nicht entstanden wären. Beispielsweise können immobile, kranke Personen so an Communities teilnehmen. Mitglieder einer Online-Community müssen aber nicht zwingend weit verstreut sein. Online-Communities werden in einem solchen Szenario „lediglich“ zur Erleichterung bzw. Bereicherung der Interaktion eingesetzt.

Die Community-Typen von Carotenuto u. a. (1999) aus dem vorherigen Abschnitt lassen sich eins zu eins auf Online-Communities übertragen.

In der praktischen Arbeit werden folgende Klassen von Anwendungen unterschieden (Gross u. Koch 2007, S. 117):

- asynchrone (textbasierte) Mailinglisten bzw. Diskussionsforen
- gemeinsame Informationsräume (zur indirekten Kommunikation)
- Social Bookmarking (als Erweiterung zu gemeinsamen Kommunikationsräumen)
- Expertenfinder, „Gelbe Seiten“-Anwendungen
- (Soziale) Netzwerke

2.3 Vergleich von Online-Communities

Es gibt eine unermessliche Anzahl von Online-Communities. Selbst wenn zwei Communities der gleichen Kategorie (vgl. Abschnitt 2.1) angehören, können diese z. B. auf Grund ihrer Größe Unterschiede aufweisen. In diesem Abschnitt werden zwei relativ ähnliche Communities aus dem Bereich Eltern- bzw. frühkindlicher Bildung miteinander verglichen, die auf den ersten Blick relativ ähnlich erscheinen.

2.3.1 Beispiel einer großen Online-Community

Als Beispiel für eine große Community (of Practice) mit ca. 110.000 registrierten Mitgliedern³ wird an dieser Stelle *eltern.de* vorgestellt. Dabei handelt es sich um eine Community, die eng mit der Webseite der Zeitschrift *Eltern* verwoben ist und sich an alle Eltern bzw. Familien richtet. „eltern.de“ bezeichnet sich selbst als Deutschlands größte Familien-Seite⁴ und kombiniert redaktionelle Inhalte mit einer Online-Community: Artikel können kommentiert werden, es gibt eine große Anzahl verschiedener Foren, Mitglieder können über ein angeschlossenes soziales Netzwerk neue Bekanntschaften herstellen, sich in Gruppen organisieren, ein eigenes Blog/Tagebuch führen oder auch Bilder in Fotoalben hochladen.

³Stand: 14. September 2011

⁴<http://www.eltern.de/service/partnerprogramm.html>, abgerufen am 16. September 2011

Eltern.de Sam Glück Familie

Nutzername / E-Mail: **Login** Suchbegriff: **Suchen**

angemeldet bleiben [Passwort vergessen?](#) [Registrieren](#)

Information **Community** **Shops** [Abo](#) | [Top-Themen](#) | [Bilder](#) | [Videos](#) | [Services](#) | [Gewinnspiele](#) | [Extras](#) | [Newsletter](#)

[Kinderwunsch](#) [Schwangerschaft](#) [Baby](#) [Kleinkind](#) [Kindergarten](#) [Schulkind](#) [Beauty & Style](#) [Familie & Urlaub](#) [Beruf & Geld](#) [Gesundheit & Ernährung](#)

Sie sind hier: [Information](#) [Eltern.de auf Facebook](#) [RSS](#) [Mobil](#) [Archiv](#)



BABYMODE
So wird's im Herbst kuschelig
 Kuschelige Pullunder, süße Teddyjacken und warme Schühchen - mit der neuen Herbstmode für Babys werden die kühleren Tage so richtig gemütlich.
[Herbstliches fürs Baby](#)

- [Schöne Mützen und Pullover zum Stricken](#)
- [Handarbeiten für die kühle Jahreszeit](#)
- [Herbstmode im Eltern.de BabyShop](#)

Schwangerschaft & Arbeit [> Alle Top-Themen](#)

TOP THEMA



Schwangerschaft & Arbeit
In welchen Fällen gilt das Mutterschutzgesetz? Wie Sie sich Chef und Kollegen gegenüber ...

Umstandsmode



[zum BabyShop](#)



GEBURT
Was zieh ich an zur Geburt?
 Einfach irgend ein T-Shirt? Etwas, das sich vorn aufknöpfen lässt? Ein Teil, das Bauch und Po bedeckt, damit man sich nicht so nackt fühlt? Oder gibt es sogar Vorschriften der Geburtskliniken? Gar nicht unwichtige Fragen kurz vor dem Termin. Eltern.de hat alle Antworten für Sie.
[T-Shirt für die Geburt](#)

ELTERN Heft & Abo



Das neue Eltern-Heft
 Jetzt online ins aktuelle ELTERN-Heft reinblättern.
[> Heft-Vorschau](#)
[> Heft abonnieren](#)



POSITIV DENKEN
Test: Wie glücklich ist mein Kind?
 Alle Eltern wollen, dass ihre Kinder glücklich sind. Aber was bedeutet das eigentlich? Und wie können Eltern erkennen, dass ein Kind ein positives Lebensgefühl hat oder nicht? Unser Test gibt Ihnen einen ersten Eindruck über die Gemütslage Ihres Kindes. Und: Wir geben Tipps zum Glückselin!
[Test und Tipps](#)



GEWINNSPIEL
Kindergeschirr-Sets von Emsa zu gewinnen
 So macht essen und trinken lernen Spaß! Denn die Kindergeschirr-Serie "Farm Family" von Emsa ist



Aboprämie
 Abo, Probeabo, Geschenkabo oder Empfehlung: immer mit Prämie!
[> Alle Abo Prämien](#)

Abbildung 2.1: Screenshot der Startseite von eltern.de, Abruf: 14. September 2011

Die enge Verzahnung mit der Zeitschrift *Eltern* bzw. die kommerzielle Ausrichtung der Webseite ist deutlich erkennbar (siehe Abbildung 2.1): An vielen Stellen werden Sonderausgaben, die *Eltern*-Zeitschrift bzw. entsprechende Abonnements beworben. Zudem ist in die Webseite ein eigener Online-Shop integriert, über den nicht nur Erzeugnisse des Verlages erworben werden können, sondern u. a. auch Kinderspielzeug, Kleidung für Kinder und Erwachsene sowie Kindermöbel. Darüber hinaus befindet sich klassische Online-Werbung Dritter auf einigen Unterseiten. Über Rabatte im Shop für Community-Mitglieder werden auch Anreize geschaffen, sich zu registrieren.

Die redaktionellen Inhalte werden vermutlich von den gleichen Autoren erstellt, die auch im Print-Medium Artikel publizieren. Dabei bietet die Web-Plattform nicht nur textlastige Artikel, sondern auch interaktive Selbsttests, Umfragen, redaktionell kommentierte Bildergalerien und kurze Videoclips. Alle Inhalte sind in zehn Kategorien, wie z. B. „Kinderwunsch“, „Schwangerschaft“, „Baby“, „Schulkind“ oder „Beruf und Geld“, eingeteilt (auf der Startseite direkt unter den Tabs „Information“, „Community“ und „Shop“ zu finden, siehe Abbildung 2.1). Jede dieser Kategorien besitzt wiederum weitere Unterkategorien, in denen schließlich die Inhalte, nach Aktualität geordnet, zu finden sind. Die Artikel, Selbsttests und Bildergalerien können von Mitgliedern der Community mit Kommentaren versehen werden, wobei das Lesen aller Inhalte auch für Nicht-Mitglieder möglich ist.

Kern der Community scheint das Foren-Portal zu sein (vgl. Abbildung 2.2, zu finden unter dem Link „Community“): Wie auch die Inhalte werden die Foren über zwei Ebenen verteilt. Die Einteilung der obersten Ebene entspricht in etwa den zehn Kategorien der Artikel, erweitert um Oberkategorien für private Foren, Foren über kindgerechten Umgang mit Medien, Partnerschaft, „Familie allgemein“ (enthält Foren für Alleinerziehende, Neurodermitis, „Gesundheit allgemein“, Patchworkfamilien, Großeltern, „Armut und Haushalt“ sowie ein türkisches und französisches Forum), einen Flohmarkt, regionale Foren (eines für jedes Bundesland) und die Online- sowie Print-Redaktion (bis 2009 gab es darüber hinaus ein Expertenforum). Die Erweiterungen in der obersten Ebene wurden größtenteils für die Unterstützung und den Aufbau von Sub-Communities angelegt. Insgesamt soll es über 200 Foren⁵ mit ca. 647.000 Themen sowie ca. 11,1 Millionen Beiträgen⁶ geben.

Daneben werden spezielle Funktionen für die Community in der Art eines sozialen Netzwerkes angeboten: Mitglieder können ein sog. Familienprofil anlegen. Dazu muss man sich einen Familienbenutzernamen ausdenken und seine komplette Anschrift hinterlegen. Danach kann man seine Rolle (Partner, Vater, Verwandter, Großmutter, ...) in der Familie angeben, weitere *eltern.de*-Mitglieder sowie Einträge für Kinder (mit Geburtsdatum und eigenem Benutzernamen) der Familie hinzufügen und die Inhalte der Familienprofilseite, die standardmäßig öffentlich einsehbar ist, bearbeiten (Motto, Freitext für die Vorstellung der Familie, Hobbies, Interessen). Auf der Profilseite kann man, wie in sozialen Netzen üblich, einen Blog bzw. ein Gästebuch (für Besucher des Profils) erstellen, Fotoalben einrichten und Kontakte sowie Gruppenzugehörigkeiten einsehen. Dadurch soll die Interaktion sowohl mit Benutzern innerhalb der Community als auch mit Verwandten und Freunden der Familien, die nicht auf *eltern.de* registriert sind, ermöglicht werden, um z. B. die neusten Fortschritte des eigenen Nachwuchses

⁵<http://www.eltern.de/service/partnerprogramm.html>, abgerufen am 14. September 2011

⁶<http://www.eltern.de/foren/>, abgerufen am 16. September 2011

The screenshot shows the homepage of Eltern.de FamilienNetz. At the top, there is a navigation bar with 'Information', 'Community', and 'Shops' highlighted. Below this is a search bar and a login section with fields for 'Nutzername / E-Mail' and a password field, along with a 'Login' button. A 'Suchen' button is also present. Below the navigation bar, there are links for 'Foren', 'Gruppen', 'Fotoalben', 'Familiennetz', and 'Kontakt'. The main content area is titled 'Eltern.de Forum' and includes a welcome message for non-logged-in users. Below the message is a table of forum topics with columns for 'Forum', 'Letzter Beitrag', 'Themen', and 'Beiträge'.

Forum	Letzter Beitrag	Themen	Beiträge
Kinderwunsch Neu: Plan B - Leben mit unerfülltem Kinderwunsch Kinderwunsch und Partnerschaft, Kinderwunsch nach Fehlgeburt Wir planen ein Baby, Langzeitkinderwunsch, 1. Hilfe-Forum von Usern für User, Unerfüllter Kinderwunsch, Letzter Ausweg künstliche Befruchtung, Kinderwunsch-Forum für Männer- wie fühlt Ihr Euch dabei?	NMT 01.09. - 15.09. von Fralla Heute 18:04	67.551	1.811.442
Schwangerschaft Neu: Schwanger nach IUI/IVF/ICSI Hurra, ich bin schwanger, Das Zweitel!, Sex in der Schwangerschaft, Schwanger nach Fehlgeburt, Haben Sie Sorgen? Pränatale Diagnostik, Monatsforen.....	Bauchbilder-Thread von Bonnie87 Heute 18:05	187.200	2.694.447
Geburt Neu: Hausgeburt Vorbereitung auf die Geburt, Geburtserlebnisse, Kaiserschnitt, Angst vor der Geburt, Fehlgeburt, Alles gut auf der Wochenstation? (Klinikberichte).....	5 Tage nach As von Grinsekatzemami Heute 14:47	6.556	70.710
Baby Mein Baby ist da und jetzt? (mit Monats-Unterforen und Babygalerie), Plauderforen 1998-2011, Brust oder Flasche? Flasche! Stillforum, Beikost, Endlich ruhige Nächte, Schreibabys, Frühchen.....	Kind lässt alles liegen von Martha74 Heute 18:05	171.922	2.954.638
Kleinkind Auf die Knie, fertig, los, Wie Kinder schlafen lernen, Religion, Kleinkindforum, Zweisprachige Erziehung, Trotz, Sauberwerden.	"Es hat nicht geklappt" -... von rastamma Heute 18:01	12.713	184.292
Krippe und Kindergarten	KiGa-Eingewöhnung klappt nicht		

Abbildung 2.2: Screenshot der eltern.de Foren, Abruf: 14. September 2011

darzustellen. Für die Kommunikation innerhalb der Community gibt es neben den öffentlichen Foren ein internes Nachrichtensystem. Speziell für das Finden anderer Familien bzw. neuer Kontakte innerhalb der Online-Community gibt es im sog. Familien-netz eine Familiensuche, mit der man Familien über allgemeine Suchbegriffe, Angaben zu Kindern, Interessen und regional suchen kann.

Darüber hinaus bietet *eltern.de* verschiedene Newsletter an: Dazu zählen drei allgemeine, wobei der erste alle zwei Wochen über Neuigkeiten berichtet, der zweite täglich ein Rezept vorstellt und der dritte auf Angebote im *eltern.de*-Shop hinweist. Weitere Newsletter sind personalisiert und in der Art eines Ratgebers aufgebaut. Hierzu zählt u. a. ein Schwangerschaftsguide, der auf den Empfänger abgestimmt ist und diesen durch die 40 Wochen einer Schwangerschaft begleitet.

2.3.2 Beispiel einer kleinen Online-Community

Eine kleine Online-Community (of Practice) ist die *Mobile2Learn.de* Online-Bildungs-plattform (vgl. Abbildung 2.3) mit ca. 200 Mitgliedern (Strickroth u. a. 2011). Zielgruppe sind Eltern mit Kleinkindern im Alter von 0 bis 6 Jahren. Hinter der Online-Plattform steht ein Projekt der Volkshochschule Goslar, der Technischen Universität Clausthal und der ländlichen Erwachsenenbildung Niedersachsen, in dem (bildungsferne) Eltern kostenneutral motiviert, befähigt und unterstützt werden sollen, ihre Kleinkinder altersgemäß zu bilden und zu fördern.

Bei der Online-Plattform handelt es sich um keine reine Online-Community, sondern um eine Community mit regionalem Bezug: Da insbesondere bildungsferne Eltern selten an Informationsveranstaltungen teilnehmen (Rupp 2003; Bauer u. Bittlingmayer 2005), wurde eine Kombination aus Präsenzveranstaltungen, themenbasierten Mitmachveranstaltungen für Eltern mit Kindern, die in Kindergärten im Landkreis Goslar stattfinden, und einer Online-Plattform gewählt. Damit sollen die Vorteile der Präsenzveranstaltungen mit denen moderner Medien verbunden werden, so dass sich beide gegenseitig ergänzen und wechselseitig unterstützen. Eltern können online Bilder sowie Bildungsinhalte der Veranstaltungen niederschwellig abrufen, die Zeit zwischen den Veranstaltungen überbrücken, sich austauschen und miteinander in Kontakt bleiben. Hierfür existieren ein Forum und ein internes Nachrichtensystem auf der Online-Plattform. Neben der Vermittlung von Bildungsinhalten, ist es ein Ziel, dass Eltern, die bisher an keiner Präsenzveranstaltung teilgenommen haben, über die Online-Bildungsplattform motiviert werden, sich zu informieren und an einer späteren Veranstaltung teilzunehmen.

Die Bildungsplattform bietet neben der Community ein Angebot von redaktionellen Inhalten (Bildungsinhalte, Erziehungstipps, Spiele und Bilder), die die Inhalte der Präsenzveranstaltungen zielgruppengerecht aufgreifen und vertiefen. Bei den Online-Bildungsinhalten handelt es sich nicht nur um textlastige Artikel, sondern auch um interaktive Quizze und Rätsel. Das Projekt hat sich vorgenommen, sechs Bildungsthe-men in sechs Veranstaltungsreihen zu präsentieren (ausgerichtet am Orientierungsplan für Bildung und Erziehung im Elementarbereich niedersächsischer Tageseinrichtungen für Kinder). Jede Veranstaltungsreihe steht unter einem eigenen Motto und beinhaltet sechs Präsenzveranstaltungen in wechselnden Kindergärten. Diese Struktur der Veran-staltungsreihen findet sich in der Organisation der Artikel wieder: Die Artikel sind als

MOBILE LEARN

Registrieren | Anmelden

Willkommen Gast!

Suchen

Lernen mit allen Sinnen | Alte Kinderlieder - neu entdecken!

Willkommen auf Mobile2Learn.de

Herzlich willkommen auf der **Community-Website Mobile2Learn..**
Der Online-Bildungsplattform für frühkindliche Bildung und Förderung.

Auf Mobile2Learn.de finden Sie:

- ↘ Spiele-Ideen
- ↘ Erziehungstipps für Eltern
- ↘ Erfahrungs- und Ideen-Austausch
- ↘ Bildungstipps/Spiele-Ideen per SMS

Alles kostenlos und ohne Verpflichtungen!
Registrieren? oder einloggen?

Hauptmenü

- Startseite
- Über uns
- Termine
- Netzwerk-Partner
- Bildungsinhalte / Spiele
- Rätsel/Quiz
- Anmelden
- Forum
- Für Kindergärten

Wer ist online

0 registrierte Benutzer und 6 Gäste online.

Noch nicht registriert oder angemeldet. Hier [registrieren](#).

Neues auf Mobile2Learn.de

- ↘ Mit Mobile2Learn in die Natur - die Termine
- ↘ Neue Inhalte und Spiele finden Sie unter:
 - ↘ Tasten -> ab 3. Lebensjahr -> "Knopfkönig"
 - ↘ Tasten -> ab 4. Lebensjahr -> "Das fliegende Geldstück"
 - ↘ Mit Kindern in die Natur -> "Beim Waldkindergarten Lutter"
 - ↘ Die Fotos für sämtliche Veranstaltungen bis 30.06.2011 finden Sie in unserer Galerie.

Gefördert von:

TU Clausthal | vhs | ZEB | nifbe | Niedersächsisches Institut für frühkindliche Bildung und Entwicklung | Gefördert durch Niedersächsisches Ministerium für Wissenschaft und Kultur

© 2010 - 2011 Mobile2Learn
Nutzungsbedingungen
Impressum | Nach oben

Abbildung 2.3: Screenshot der Startseite von Mobile2Learn.de, Abruf: 2. September 2011

Baum strukturiert. An oberster Stelle steht ein Übersichtsartikel und direkt (im Baum) darunter befinden sich die Einleitungsartikel der einzelnen Veranstaltungsreihen/Themen, die schließlich die Bildungsinhalte bzw. Spiele enthalten, evtl. auch über mehrere Ebenen verteilt. Insgesamt wird bei den Inhalten besonderer Wert auf die „Betonung von Erlebnis-, Aktions- und Mitmachelementen gegenüber ‚kopflastigeren‘ theoretischen (...) Bildungselementen“ (Stöhr 2007, S. 130) gelegt. Ein kleiner Teil dieser Inhalte ist auch für Nicht-Mitglieder abrufbar. Dadurch ist es möglich, das Angebot anonym kennenzulernen; gleichzeitig werden Anreize geschaffen, sich zu registrieren und auch aktiver Teil der Community zu werden. Der Großteil hingegen, z. B. die Galerie, erfordert eine Anmeldung. Hierfür ist die Angabe eines Pseudonyms, Name, Vorname, Straße, Wohnort, Alter, Geschlecht, E-Mail-Adresse und/oder Handynummer sowie optional die Geschlechter und Geburtsdaten der eigenen Kinder erforderlich. Mitgliedern der Community ist es möglich, in das Forum zu schreiben, alle Artikel/Bilder abzurufen, diese zu bewerten (auf einer Stern-Skala von 1, gefällt mir nicht, bis 5 Sterne, gefällt mir sehr) und zu kommentieren sowie sich eine eigene Profilseite zu erstellen. Obwohl die Inhalte redaktionell bearbeitet werden, ist es vorgesehen, dass auch „normale“ Mitglieder eigene Inhalte einreichen.

Zur Akquise der Zielgruppe wird ein breites regionales Netzwerk mit Institutionen, Einrichtungen und Diensten gebildet, die typischerweise von der Zielgruppe aufgesucht werden, um familiäre und soziale Leistungen oder Hilfen in Anspruch nehmen zu können, bzw. die im Rahmen ihres Tätigkeitsbereiches Zugang zur Zielgruppe haben. Dazu gehören diverse Fachdienste des Landkreises, Kirchengemeinden, Wohlfahrtsverbände, Familienhebammen, Kinderärzte, Gynäkologen, ein Mütterzentrum, die Gleichstellungsstelle des Landkreises, der Kinderschutzbund, die Kreismusikschule, die Agentur für Arbeit und (in hervorgehobener Rolle) die Kindertagesstätten. Mit diesem Netzwerk sollen speziell Eltern der Zielgruppe über eine direkte und persönliche Ansprache eingeladen werden, an den Veranstaltungen teilzunehmen und sich in der Online-Bildungsplattform anzumelden. Darüber hinaus werden die Netzwerkpartner mit in die Community integriert. Sie können ebenso wie die Eltern eine Profilseite anlegen, aber dort zusätzlich für sich werben und ihre Expertise kostenlos darstellen, so dass professionelle Stellungnahmen von Pädagogen bzw. Fachleuten in den Foren erkannt werden können.

Wesentliche Punkte der Online-Bildungsplattform sind eine kontinuierliche pädagogische Betreuung und regelmäßige personalisierte Nachrichten mit Bildungsbezug, die an die Eltern per SMS bzw. E-Mail verschickt werden. Hierbei werden aktuelle Veranstaltungshinweise oder Bildungsinhalte in „SMS-Spots“ verkürzt, mit Humor und Appellcharakter versehen, so dass prägnante Botschaften entstehen und vermittelt werden (z. B. bei Regenwetter: „Hallo XY, bewegen auch bei Regen! Regenspiele für Kinder auf Mobile2Learn.de“). Zudem soll dieses „auf die Eltern zugehen“ die Bindung zur Community stärken und Eltern immer wieder auf das Projekt und die Online-Plattform hinweisen, so dass der Kontakt nicht abreißt.

Mit einem Empfehlungssystem (Inhalt dieser Arbeit, vgl. Kapitel 4 ab Seite 41) soll es Mitgliedern ermöglicht werden, schnell und einfach Inhalte zu finden, die sie persönlich interessieren.

2.3.3 Unterschiede und Gemeinsamkeiten

Die beiden Online-Communities *eltern.de* und *Mobile2Learn.de* gehören zur Kategorie der Communities of Practice, in denen Praktiker, in diesem Fall Eltern, Informationen finden und sich austauschen können. Dazu unterstützen beide Plattformen die Community durch Foren und ein internes Nachrichtensystem, wobei *eltern.de* Ansätze eines sozialen Netzwerkes aufweist.

Aber bereits bei der genauen Zielgruppe und Ausrichtung unterscheiden sich beide Communities: *eltern.de* richtet sich an alle deutschsprachigen Eltern, die für den Betreiber auch (potentielle) Kunden sind. *Mobile2Learn* hingegen setzt auf ein werbefreies und kostenloses Bildungsportal für alle deutschsprachigen Eltern, speziell für bildungsferne Schichten aus dem Landkreis Goslar. Bei *Mobile2Learn* gibt es also eine feste regionale Verortung mit Präsenzveranstaltungen, bei denen sich Mitglieder in der realen Welt treffen können. Zudem sind bei *Mobile2Learn* unterschiedliche Profiltypen für Eltern und Experten/Institutionen vorgesehen, so dass Experten bei Benutzer-Beiträgen einfach identifiziert werden können und gleichzeitig die Möglichkeit haben, sich als Experte bzw. Institution darzustellen.

Beide Communities bieten neben den Community-Funktionen auch redaktionelle Inhalte und Newsletterdienste an. Der markanteste Unterschied besteht hier in der Anzahl der Inhalte, die bei *eltern.de* deutlich überwiegt. Jedoch scheinen einige Inhalte dort lediglich einen Kontext zu bereiten, um kommerzielle Ziele zu verfolgen (z. B. Artikel über Kindermode mit einem Link zum Online-Shop). Artikel scheinen bei dieser Plattform ausschließlich von Autoren außerhalb der Community zu stammen, während *Mobile2Learn* auch auf Beiträge aus den Reihen der Community-Mitglieder setzt. Darüber hinaus versendet *Mobile2Learn.de* neben einem personalisierten Newsletter auch Informationen per SMS.

Resultierend aus dem Größenunterschied der Mitgliederzahlen bietet *eltern.de* Foren für Sub-Communities, Gruppenunterstützung sowie weitere spezielle Funktionen zum Finden neuer Kontakte an, die in einer kleineren Community in diesem Umfang nicht unbedingt nötig sind.

Bezogen auf die Definitionen von Online-Communities (vgl. Abschnitt 2.2) können die folgenden Unterschiede und Gemeinsamkeiten festgestellt werden: Die grundlegende Definition von Carotenuto u. a. (Hauptteil der Kommunikation wird über elektronische Medien abgewickelt) trifft vornehmlich auf *eltern.de* zu; bei *Mobile2Learn* spielt neben der elektronischen auch die Offline-Kommunikation während der Präsenzveranstaltungen eine besondere Rolle. Eine analoge Aussage kann für die sich bildenden sozialen Beziehungen aus den Definitionen von Blanchard und Lee u. a. (2003) getroffen werden: bei *eltern.de* entstehen Beziehungen hauptsächlich online, bei *Mobile2Learn* werden diese auch bei Präsenzveranstaltungen geknüpft. Die Unabhängigkeit von realen Örtlichkeiten (vgl. Aussage von Hartleb) ist bei *eltern.de* deutlich ausgeprägter, da sich die Leser der Zeitschrift *Eltern* überall befinden können. *Mobile2Learn* konzentriert sich konzeptbedingt auf eine abgegrenzte Region (Landkreis Goslar). Dies schränkt jedoch die Möglichkeit der Teilnahme an der Community für Personen, die z. B. immobil sind oder unerkannt bleiben möchten, nicht ein. Eine Kollaboration (vgl. Definition von Carroll u. Rosson) im Sinne gemeinsamer Erstellung und Bearbeitung von Inhalten bzw. Dokumenten auf den Plattformen ist auf keiner der beiden Communities auszumachen

(ohne Berücksichtigung der redaktionellen Autorentätigkeiten). Jedoch findet auf beiden Plattformen eine Zusammenarbeit der einzelnen Mitglieder im Rahmen gemeinsamer Problemlösung (z. B. durch die Beantwortung von Fragen im Forum) statt. In Bezug auf „member-driven content“ kommt *Mobile2Learn.de* der Definition von Lee u. a. (2003) am nächsten, da dort, neben Benutzer-Eintragungen in Foren, Blogs oder Gästebücher, auch Artikel und Bilder von Mitgliedern beigesteuert werden können.

2.4 Zusammenfassung

In diesem Kapitel wurde der Begriff *Online-Community* beleuchtet. Dabei wurden die Definitionen einer normalen *Community*, die allgemein als Gruppierung von Menschen mit gleichen Interessen und räumlicher Nähe definiert sind, zu Definitionen einer *Online-Community* erweitert, die durch die hauptsächliche Nutzung elektronischer Medien zur Interaktion und weniger durch räumliche Nähe charakterisiert ist. Zudem wurden weitere Charakteristika (mit Vor- und Nachteilen) sowie eine Kategorisierung für (Online-)Communities vorgestellt.

Anhand eines Vergleiches zweier Online-Communities aus dem Bereich Elternbildung wurde herausgearbeitet, wie diese individuellen Communities in der Praxis aussehen können und wie sie sich zu den einzelnen Definitionen verhalten. Weiterhin wurde deutlich, dass Communities, auch wenn sie der gleichen Kategorie angehören, deutliche Unterschiede durch ihre Größe/Mitgliederzahl, Zielgruppe und genauen Kontext (z. B. regionale Bindung oder Einbeziehung der Benutzer) aufweisen.

3 Empfehlungssysteme

Dieses Kapitel gibt einen Überblick über existierende Konzepte und Verfahren für Empfehlungssysteme. Es handelt sich dabei um eine Auswahl von Ansätzen, die im Rahmen dieser Arbeit bzw. dieses Themenbereiches von Interesse sind.

3.1 Definition

Ein Empfehlungssystem (ES) hat die Aufgabe, aus einer Menge von Alternativen, in diesem Kontext *Artefakte* genannt, eine oder mehrere auszuwählen, die für einen bestimmten Benutzer („aktiver Benutzer“ genannt) sinnvoll, interessant bzw. nützlich sind: Ein Filter-Problem, bei dem sich die Kriterien für „sinnvoll“ bzw. „nicht sinnvoll“ bei jedem Benutzer unterscheiden können.

Genau die Kriterien „individualisiert“ und „interessant“ bzw. „nützlich“ unterscheiden Empfehlungssysteme von (einfachen) Suchmaschinen oder Information-Retrieval-Systemen (IRS) (nach Burke 2002). Suchmaschinen arbeiten nach einem Anfrage- bzw. „Matching“-Ansatz: Ein Benutzer stellt eine Anfrage an ein IRS und dieses listet alle auf das Suchwort bzw. die Anfrage passenden Artefakte sortiert (z. B. nach der Häufigkeit) auf. Informationsfiltersysteme können zwar auf gleichen Techniken basieren, sind jedoch speziell darauf ausgelegt, den Präferenzen der Benutzer (oftmals automatisch) zu entsprechen.

Frühe ES haben auf sog. „Association Rules“ aufgesetzt (vgl. Schafer u. a. 2007). Dabei handelt es sich um Verfahren, die versuchen, Regeln für oft gemeinsam auftretende Aktionen zu bestimmen und diese Regeln als Grundlage für Empfehlungen zu nutzen: Wurden z. B. von vielen Kunden häufig Artefakt *A* und *B* zusammen gekauft, lassen sich die Regeln $A \Rightarrow B$ sowie $B \Rightarrow A$ herleiten und für Empfehlungen nutzen, sobald ein Kunde *A* oder *B* in den Warenkorb legt, um das jeweils andere Artefakt zu empfehlen. Jedoch ist die Bestimmung der Regeln und die Berechnung von Empfehlungen, falls sehr viele Regeln berücksichtigt werden müssen, nicht trivial. Andere frühe ES haben versucht, Präferenzen von Benutzern zu lernen. Es hat sich allerdings gezeigt, dass die Lernphasen sehr lang sind, bis sinnvolle Empfehlungen generiert werden können (Balabanović 1997). Im weiteren Verlauf der Forschung wurden ab ca. 1990 speziellere Empfehlungs- bzw. Vorhersagealgorithmen entwickelt, die die Grundlage dieser Arbeit bilden.

Wichtige Kriterien bei der Entwicklung eines Empfehlungssystems sind der Modus, wie Empfehlungen berechnet werden, und die Aufgabe, die das System damit erfüllen soll.

Herlocker u. a. (2004) sowie Schafer u. a. (2007) differenzieren zwischen drei domänen-unabhängigen Aufgaben von Empfehlungssystemen (Auswahl¹):

„Annotation in Context“/„Advise me on a particular item“ Das Empfehlungssystem soll für ein bestimmtes Artefakt vorhersagen, ob es für den aktiven Benutzer interessant/gut ist oder nicht (z. B. SPAM-Filter, Goldberg u. a. 1992)

„Find Good Items“/„Help me find new items I might like“ Einem aktiven Benutzer werden Artefakte vorgeschlagen, die ihn vermutlich interessieren. Dies ist die Standard- bzw. Kern-Aufgabe für ein ES und vermutlich, zusammen mit „Annotation in Context“, auch die in der Literatur am häufigsten untersuchte Aufgabe (Schafer u. a. 2007).

„Find All Good Items“ Einem aktiven Benutzer werden alle „guten“ Artefakte vorgeschlagen. Die meisten ES beschränken sich darauf, einige ausgewählte „gute“ Artefakte vorzuschlagen. Dabei können aber weitere „gute“ Artefakte übersehen werden, um dafür die mit größerer Sicherheit „schlechten“ auszusortieren. Für die meisten Anwendungsfälle ist dies der richtige Ansatz. Es gibt jedoch Ausnahmen, wie z. B. die Rechtswissenschaften. Dort dürfen bei der Suche nach Präzedenzfällen keine relevanten Artefakte übersehen werden (Schafer u. a. 2007).

Jede dieser Aufgaben stellt besondere Anforderungen an die Verfahren, Algorithmen und Evaluationstechniken. Im Folgenden, speziell bei der Auswahl der vorgestellten Verfahren, wird der Fokus auf „Find Good Items“ gelegt, da diese Aufgabe für Empfehlungssysteme der meisten Communities zutrifft.

Ein Empfehlungssystem, das aus einer Menge von Artefakten für einen aktiven Benutzer die wahrscheinlich besten bzw. interessantesten N Artefakte heraussucht, die er noch nicht bewertet bzw. abgerufen hat, wird „TOP- N -Empfehlungssystem“ genannt, wobei die Empfehlung als „TOP- N -Ranking“ bezeichnet wird. Dies kann wie folgt formalisiert werden (angelehnt an Deshpande u. Karypis 2004, Definition 2.1):

Sei \mathcal{I} eine Menge von m Artefakten, \mathcal{U} eine Menge von n Benutzern, \mathcal{R} eine $n \times m$ Benutzer-Artefakt-Bewertungsmatrix und $B \subseteq \mathcal{I}$ die Menge der vom aktiven Benutzer $u \in \mathcal{U}$ bewerteten Artefakte. Gesucht ist eine (geordnete) Menge $X \subseteq \mathcal{I}$ mit $X \cap B = \emptyset$ und $|X| \leq N$.

In der $n \times m$ Bewertungsmatrix \mathcal{R} werden die Benutzer $u \in \mathcal{U}$, $|\mathcal{U}| = n$ (traditionsgemäß) auf den Zeilen, die Artefakte $i \in \mathcal{I}$, $|\mathcal{I}| = m$ auf den Spalten aufgetragen und die Bewertungen $r_{u,i}$ in die jeweiligen Zellen eingetragen. Ein Zeilenvektor r_u enthält damit sämtliche Bewertungen eines Benutzers und ein Spaltenvektor r_i die Bewertungen aller Benutzer für ein Artefakt.

Oftmals bestehen ES aus Vorhersagemodellen bzw. -Algorithmen, die um einen Empfehlungsmodus erweitert wurden und Artefakte empfehlen, deren vorhergesagte Bewertungen einen Schwellenwert übersteigen. Allgemein muss aber zwischen der Berechnung von Vorhersagen und Empfehlungen unterschieden werden (Deshpande u. Karypis 2004; Schafer u. a. 2007): Um Artefakte empfehlen zu können, ist es hinreichend, wenn

¹Darüber hinaus gibt es Abwandlungen dieser Aufgaben, die statt für einen einzelnen Benutzer für eine Gruppe operieren.

das ES auf einer Teilmenge der Artefakte operiert. Für eine Vorhersage von Bewertungen hingegen müssen einige Algorithmen Informationen über jedes einzelne Artefakt vorhalten und verarbeiten. Daher kann die Berechnung einer einzigen Vorhersage viel aufwändiger sein als der Aufwand für eine Empfehlung.

Empfehlungssysteme für die Aufgabe „Find Good Items“ lassen sich wie folgt kategorisieren (siehe Abbildung 3.1): Auf oberster Ebene lassen sich ES in „personalisiert“ und „nicht personalisiert“ aufteilen. Eine weitere Unterteilung auf der nächsten Ebene besteht in der grundlegenden Art der Empfehlungsberechnung, die wiederum unterschiedliche Varianten besitzen kann. Den in Abbildung 3.1 abgebildeten Verfahren wird jeweils ein eigener Abschnitt innerhalb dieses Kapitels gewidmet.

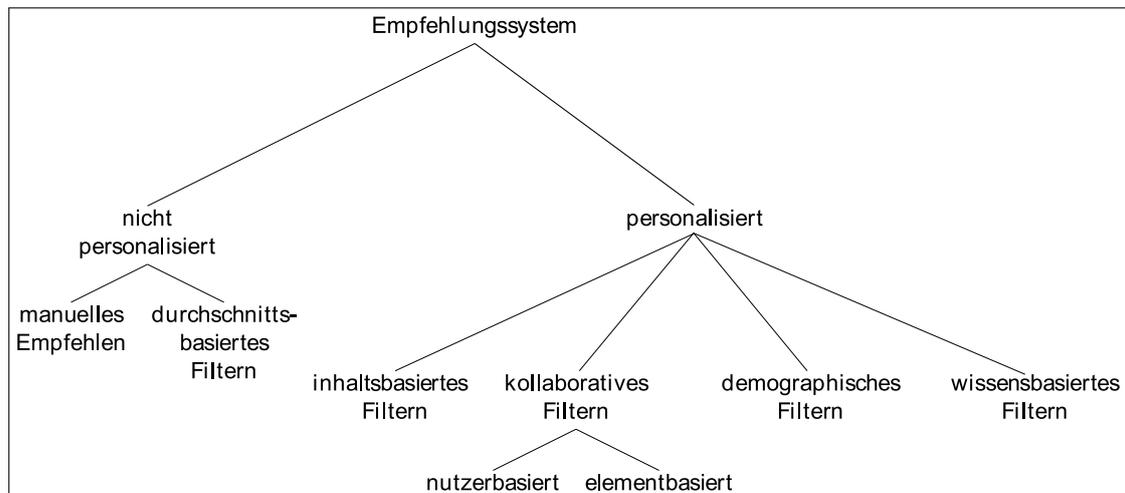


Abbildung 3.1: Einordnung Empfehlungssysteme²

Automatische Empfehlungssysteme stellen, abhängig vom genauen Vorgehen, einige Anforderungen an die Umgebung bzw. Domänen, in denen sie im Allgemeinen besonders einfach eingesetzt werden können (vgl. Schafer u. a. 2007). Zum einen gibt es Voraussetzungen an die Artefakte:

- Es sollte eine gewisse Anzahl von Artefakten vorhanden sein. Gibt es nur sehr wenige, könnte ein Benutzer „leicht“ alle Artefakte ohne Unterstützung sichten.
- Die Artefakte sollten hinreichend homogen und müssen subjektiv bewertbar sein.
- Zudem sollte es für alle Artefakte eine gewisse Anzahl von Bewertungen geben, damit das System sinnvoll Ähnlichkeiten und Empfehlungen bestimmen kann.
- Artefakte sollten möglichst beständig sein: Das Hinzufügen oder Entfernen von Artefakten sollte nicht allzu häufig vorkommen, damit für alle Artefakte (genug) Bewertungen vorliegen können.

²nach Runte 2000, S. 10

Andererseits werden auch Anforderungen an die Benutzer gestellt:

- Die Benutzer sollten mehr als nur ein einziges Artefakt bewerten, damit das System den Geschmack der Benutzer „lernen“ kann.
- Es muss eine gewisse Anzahl von Benutzern geben und darunter auch einige ähnliche.
- Die Vorlieben der einzelnen Benutzer sollten sich nicht kurzfristig ändern, da sich das System sonst evtl. nicht schnell genug anpassen kann.

Zudem sollte es insgesamt mehr Benutzer als Artefakte geben, damit es zu mehr Überschneidungen (Benutzer haben gleiche Artefakte bewertet) bei den Daten kommen kann. Sind Anforderungen verletzt, so bedeutet dies nicht, dass kein Empfehlungssystem in einer Domäne erfolgreich etabliert werden kann, es müssen nur besondere Vorkehrungen getroffen werden (vgl. Abschnitt 3.10).

3.2 Definition „Bewertung“

Im Kontext von Empfehlungssystemen spielen Bewertungen eine Schlüsselrolle, denn sie sind die Grundlage vieler Algorithmen. Voraussetzung für eine automatische Verarbeitung ist eine numerische Repräsentation der Bewertungen der einzelnen Nutzer. Konzeptbedingt gibt es dabei zwei Dimensionen (Nutzer und Artefakte), die durch die Bewertungen miteinander „verbunden“ werden.

Intuitiv bzw. in der Literatur am weitesten verbreitet sind *explizite Bewertungen*. Dabei handelt es sich um aktiv vorgenommene Bewertungen der Nutzer eines Systems. Bewertungen können unterschiedlich ausfallen (Schafer u. a. 2007, Abschnitt 9.4.3): Als unäre Wertung, als binäre Wertung, als Integer-Wertung mit ein bis fünf Sternen/Punkten, Schulnoten oder komplexere Benutzermeinungen (vgl. Tabelle 3.1 sowie Abbildung 3.2). Benutzern ist es damit möglich, ihre Vorlieben sehr genau abzubilden, jedoch erfordert dies zusätzlichen Aufwand vom Benutzer.

unäre Wertung	binäre Wertung	Webseiten-Empfehlung
Gefällt 1	Gefällt/„Daumen hoch“ 1	Zustimmung 1
	Gefällt nicht/„Daumen runter“ 0	Keine Wertung 0
		Ablehnung -1

Tabelle 3.1: Mögliche Abbildungen von Benutzermeinungen auf numerische Werte³



Abbildung 3.2: Bewertungsskalen großer Webportale: Facebook.com, unär; YouTube.com, binär; Amazon.com, Skalar

³angelehnt an Segaran (2008, S. 11, Tabelle 2-1)

Eine weitere Form neben den expliziten sind *implizite Bewertungen* (Oard u. Kim 1998). Dabei handelt es sich um Bewertungen, die (meist) ohne Aufwand für den Benutzer z. B. durch eine Analyse des Benutzerverhaltens bzw. deren Interaktion mit einem System hergeleitet werden können. Studien lassen vermuten, dass sich deutlich mehr implizite als explizite Bewertungen gewinnen lassen (Nichols 1997), da i. d. R. jede Interaktion mit einem System Daten generiert. Beispiele dafür sind die Browser- und die Kauf-Historie, Maus-Bewegungen, Verweildauern auf Webseiten oder auch komplexere Vorgänge (vgl. Tabelle 3.2).

Konzertkarten		Online-Shopping	
Gekauft	1	Gekauft	2
Nicht gekauft	0	Angeschaut	1
		Nicht gekauft	0

Tabelle 3.2: Mögliche Abbildungen von Benutzeraktionen auf numerische Werte⁴

Oft wird aus Mangel an *expliziten Bewertungen* auf *implizite* zurückgegriffen oder eine Kombination von beiden benutzt. Dies kann folgende Gründe haben:

- Das System nimmt keine Eingabe entgegen (z. B. ein Serienempfehlungssystem für den Fernseher, Hu u. a. 2008).
- Das Bewerten von Artefakten scheint keinen Mehrwert für die Benutzer zu haben, ist zu aufwändig oder kompliziert.
- Benutzer tendieren vornehmlich dazu Artefakte zu bewerten, die sie mögen und Artefakte, die sie nicht mögen, eher nicht zu bewerten (Nguyen u. a. 2007).

Allerdings können hergeleitete bzw. *implizite Bewertungen* ungenau oder stärker voraussetzt sein als *explizite Bewertungen* (Hu u. a. 2008): Nimmt man die Kauf-Historie, ist daraus nicht zu erkennen, ob ein Kauf für einen selbst oder einen Dritten getätigt wurde, wobei Bewertungen sicherlich nach den eigenen Ansichten vergeben werden. Ähnliches gilt für die Besuchsdauer einer Webseite: Es ist ausschließlich anhand der Verweildauer nicht zu erkennen, ob der Nutzer die ganze Zeit die Webseite gelesen oder sich zwischenzeitlich „einen Kaffee geholt hat“. Dennoch konnte von Konstan u. a. (1997) sowie Morita u. Shinoda (1994) eine positive Korrelation zwischen Verweildauern und expliziten Bewertungen (bei Usenet-Beiträgen) gefunden werden. Auf der anderen Seite wird jedoch auch argumentiert, dass implizite Bewertungen in bestimmten Situationen (deutlich) besser sein können als explizite: In einem Online-Musik-Shop kann die Anzahl der Abspielungen eines Liedes (bei hinreichender Zahl von Daten) sicherlich bessere Differenzierungen bieten als eine Bewertungsskala von eins bis fünf (Schafer u. a. 2007). Zudem können Benutzer bei expliziten Bewertungen absichtlich oder unabsichtlich falsche Angaben machen (Lee u. Park 2006) – bei impliziten Bewertungen ist dies i. d. R. deutlich aufwändiger. In einer Studie von Morita u. Shinoda (1994) führten implizite Bewertungen aus dem genannten Grund zu besseren Ergebnissen als explizite.

⁴nach Segaran 2008, S. 11, Tabelle 2-1

3.3 Ähnlichkeitsmaße

In diesem Abschnitt werden einige Ähnlichkeitsmaße vorgestellt, die im Bereich der Empfehlungssysteme Anwendung finden und zum „State of the Art“ zählen. Die Vorstellung erfolgt an dieser Stelle, da Ähnlichkeitsmaße Teil verschiedener Algorithmen sind, die in den folgenden Abschnitten erläutert werden.

3.3.1 Definition

Ein Ähnlichkeitsmaß ist eine Funktion $s : I \times I \rightarrow [0, 1] \subset \mathbb{R}$ (I endliche Menge), die insbesondere die folgenden Eigenschaften besitzt:

- $s(i, j) \leq s(i, i) \forall i, j \in I$ (Diagonaldominanz)
- $s(i, j) \geq 0 \forall i, j \in I$ (Positivität)
- $s(i, i) = 1 \forall i \in I$ (Normiertheit)

Im Allgemeinen wird noch die Symmetrie ($s(i, j) = s(j, i) \forall i, j \in I$) gefordert - für Empfehlungssysteme ist diese nicht zwingend erforderlich bzw. kann in bestimmten Fällen kontraproduktiv sein (vgl. Deshpande u. Karypis 2004, Abschnitt 4.1.1). Die folgenden Ähnlichkeitsmaße erfüllen jedoch alle die Symmetrie-Eigenschaft.

Einige der in diesem Abschnitt vorgestellten Ähnlichkeitsmaße (Pearson-, Cosinus- und die angepasste Cosinus-Korrelation) können Werte im Intervall $[-1; 1]$ annehmen, jedoch wurde von Schafer u. a. (2007) angemerkt, dass negative Ähnlichkeiten wenig zielführend bei der Verbesserung der Vorhersagen seien und möglichst nicht verwendet werden sollten. Dies konnte in eigenen Tests bestätigt werden (die Einbeziehung negativer Ähnlichkeiten verschlechterte die Güte der Vorhersagen).

3.3.2 Euklidischer Abstand

Der euklidische Abstand gehört nicht direkt zum „State of the Art“ bei Ähnlichkeitsmaßen, ist jedoch ein sehr einfacher und naheliegender Weg, um den (kürzesten) Abstand zweier Punkte in einem mehrdimensionalen Raum zu bestimmen und eignet sich daher gut als Einstieg. Es handelt sich dabei um eine Metrik (sie erfüllt die Definitheit, Symmetrie, Dreiecksungleichung und Nicht-Negativität); d. h. insbesondere, dass zwei identische Punkte den Abstand null haben und der Wert größer wird, je weiter zwei Punkte voneinander entfernt sind.

Sind $x = (x_1, x_2, \dots, x_d)$ und $y = (y_1, y_2, \dots, y_d)$ zwei Punkte (bzw. Bewertungsvektoren) in einem d -dimensionalen Raum, so berechnet sich der Abstand $d(x, y)$ wie folgt:

$$d(x, y) = \|x - y\|_2 = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_d - y_d)^2} = \sqrt{\sum_{i=1}^d (x_i - y_i)^2}$$

Für die Verwendung als Ähnlichkeitsmaß (vgl. Definition in Abschnitt 3.3.1) ist folgende „Umrechnung“ erforderlich:

$$s(x, y) = \frac{1}{1 + d(x, y)}$$

Die Addition der 1 ist notwendig, damit bei identischen Objekten ($d(x, x) = 0$) eine Division durch null vermieden wird.

3.3.3 Jaccard-Ähnlichkeitsmaß

Der Jaccard-Koeffizient ist ein weiteres einfaches Ähnlichkeitsmaß. Er wird durch das Verhältnis der Anzahl der Elemente des Schnitts und der Vereinigung zweier Mengen A, B bestimmt (vgl. Marmanis u. Babenko 2009):

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

Auf Vektoren identischer Dimension angewandt ist die Ähnlichkeit durch das Verhältnis der Anzahl der Komponenten mit identischem Wert zur Dimension der Vektoren bestimmt:

$$s(x, y) = \frac{|\{(x_i, y_i) : x_i = y_i \forall i \in \{1, 2, \dots, d\}\}|}{d} \text{ mit } x, y \in \mathbb{R}^d$$

3.3.4 Pearson-Korrelation

Der Pearson-Korrelationskoeffizient (oder auch Produkt-Moment-Korrelation, vgl. Jones u. Furnas 1986) ist ein Maß für den Grad des linearen Zusammenhangs zwischen zwei Datenmengen. Er kann Werte zwischen -1 und $+1$ annehmen. Bei den Werten $+1$ und -1 besteht ein vollständiger linearer Zusammenhang, bei 0 gibt es keinen linearen Zusammenhang. Er kann auf den Vektoren x und y wie folgt berechnet werden:

$$s(x, y) = \frac{\text{Cov}(x, y)}{\sqrt{\text{Var}(x)} \cdot \sqrt{\text{Var}(y)}} = \frac{\sum_{i=1}^d (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^d (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^d (y_i - \bar{y})^2}}$$

mit den Mittelwerten $\bar{x} = \frac{1}{d} \sum_{i=1}^d x_i$ sowie \bar{y} , der Standardabweichung $\sqrt{\text{Var}(x)}$ von x und der Kovarianz $\text{Cov}(x, y)$ von x und y .

Nach Segaran (2008) ist diese Metrik besser für nicht-normalisierte Daten geeignet als das auf dem euklidischen Abstand basierende Ähnlichkeitsmaß: Es gibt tolerantere Menschen, die im Schnitt bessere Bewertungen als andere vergeben. Die Pearson-Korrelation kann dieses Phänomen ausgleichen (Ding u. a. 2006).

3.3.5 Cosinus-Korrelation

Die Cosinus-Korrelation vergleicht, ähnlich zum euklidischen Abstand, zwei Vektoren (siehe Jones u. Furnas 1986); genauer gesagt, bestimmt sie den Cosinus des Winkels zwischen beiden Vektoren. Damit gehört sie zu den Ähnlichkeitsmaßen und berechnet sich für die Vektoren $x, y \in \mathbb{R}^d$ wie folgt:

$$s(x, y) = \cos(x, y) = \frac{x \cdot y}{\|x\|_2 \cdot \|y\|_2} = \frac{\sum_i^d x_i y_i}{\sqrt{\sum_{i=1}^d x_i^2} \sqrt{\sum_{i=1}^d y_i^2}}$$

3.3.6 Angepasste Cosinus-Korrelation

Die angepasste Cosinus-Korrelation (Adjusted Cosine Similarity) wurde von Sarwar u. a. (2001) speziell für elementbasierte kollaborative Empfehlungssysteme (vgl. Abschnitt 3.9.2) vorgeschlagen. Da bei der elementbasierten Methode keine zwei Benutzer, sondern Bewertungen zweier Artefakte eines Benutzers verglichen werden, ist es nötig bzw. ratsam, die verschiedenen Bewertungsskalen der Benutzer herauszurechnen. Dies geschieht durch die Subtraktion der Durchschnittsbewertung von den Bewertungen eines Benutzers.

Zur Beschreibung sind spezielle Notationen erforderlich: Die Menge \mathcal{U} bestehe aus allen Benutzern, \mathcal{I} sei die Menge der Artefakte ($i, j \in \mathcal{I}$) und \mathcal{R} sei die Matrix der Bewertungen, wobei $r_{u,i}$ die Bewertung des Benutzers u für das Artefakt i und \bar{r}_u der Mittelwert aller Bewertungen von u sei. Dann ist die angepasste Cosinus-Korrelation definiert als:

$$s(i, j) = \frac{\sum_{u \in \mathcal{U}} (r_{u,i} - \bar{r}_u)(r_{u,j} - \bar{r}_u)}{\sqrt{\sum_{u \in \mathcal{U}} (r_{u,i} - \bar{r}_u)^2} \sqrt{\sum_{u \in \mathcal{U}} (r_{u,j} - \bar{r}_u)^2}}$$

3.4 Manuelles Empfehlen

Die einfachste Methode (sinnvolle) Empfehlungen zu geben, besteht darin, diese von einem Redakteur bzw. Experten erstellen zu lassen. Auf Grund des (personellen) Aufwandes werden diese Empfehlungen i. d. R. nicht oder nur in begrenztem Umfang personalisiert. In einer Film-Community könnten so allgemeine Empfehlungen (wie z. B. der Film des Tages) oder auch Empfehlungen für einzelne Genre abgegeben werden. Vorteile dieses Vorgehens sind die Unabhängigkeit von Benutzerbewertungen und die Möglichkeit, ausführliche bzw. textbasierte Empfehlungen geben zu können.

3.5 Durchschnittsbasiertes Filtern

„Was der Mehrheit gefällt, wird wahrscheinlich auch dem Einzelnen gefallen“ - unter diesem Motto stehen die sehr einfachen und nicht-personalisierten Empfehlungssysteme dieses Typs. Artefakte mit den besten durchschnittlichen Bewertungen über alle Nutzer, die einen Schwellenwert $t_{\min. \text{Bewertung}}$ übersteigen, werden vorgeschlagen. Damit basieren Empfehlungen auf den Meinungen bzw. Aktionen anderer Benutzer und setzen folglich eine gewisse Datenbasis voraus. Um Verzerrungen zu verhindern, wird oft zusätzlich gefordert, dass eine Mindestanzahl von Bewertungen (*agreement* genannt) für ein Artefakt vorhanden sein muss. Ohne diese Einschränkung würde ein Artefakt, das von 99 der 100 Benutzer auf einer Skala von 1 bis 5 im Schnitt mit 4,9 bewertet wurde, hinter einem Artefakt stehen, das nur von einer einzigen Person (evtl. einer Einzelmeinung) mit 5 bewertet wurde. Das durchschnittsbasierte Filtern ist vor allem in Fällen hilfreich, in denen keine Bewertungen von bzw. Information über den aktiven Benutzer vorliegen.

3.6 Inhaltsbasiertes Filtern

Das inhaltsbasierte bzw. eigenschaftsbasierte Filtern (CB, engl. content-based resp. feature-based) basiert auf Inhalt oder ausgezeichneten Eigenschaften der Artefakte (Pazzani u. Billsus 2007). Annahme ist hier, dass Personen wahrscheinlich Artefakte mögen, die den für sie relevanten Artefakten (inhaltlich) ähnlich sind. Bei Texten oder Artikeln können die einzelnen Wörter (nach Entfernung von Stop-Words) als Grundlage benutzt werden; andere Artefakttypen, wie z. B. Bilder, müssen mit Meta-Daten (z. B. Schlüsselwörter) versehen werden. Auf diesen Daten operieren dann die Algorithmen dieses Filtertyps, die sich in der genauen Funktionsweise teilweise stark unterscheiden:

Eine pragmatische Methode, um Empfehlungen auf diesen Daten zu ermitteln, ist es, den Benutzer selbst ein Profil von für ihn interessanten Schlüsselwörtern anlegen zu lassen oder die für einen Benutzer interessanten Eigenschaften zu lernen (z. B. mit Hilfe eines Bayes-Classifiers, siehe Pazzani u. Billsus 2007), um schließlich Artefakte zu empfehlen, die diese Eigenschaften enthalten. Eine weitere Möglichkeit besteht darin, die Ähnlichkeiten zwischen den einzelnen Artefakten zu ermitteln und Empfehlungen auf Grundlage der Bewertungen von Artefakten zu berechnen, die der aktive Benutzer bereits bewertet hat (Marmanis u. Babenko 2009, vgl. Abschnitt 3.9.2 über elementbasiertes kollaboratives Filtern): Das Schlüsselement ist die Abbildung der textuellen Informationen (der Text der Artikel bzw. die Schlüsselwörter) auf numerische Werte. Ein einfacher Weg besteht darin, die N meist genutzten Wörter der Texte zu extrahieren und (deren Auftreten) in einem Koordinatensystem abzubilden (auch „bag-of-words“-Modell genannt). Dazu kann ein Text in einzelne Wörter aufgespalten, die Wörter normalisiert und das absolute Auftreten der verbliebenen Wörter gezählt bzw. deren Auftrittshäufigkeiten bestimmt werden. Mit Normalisierung sind in diesem Zusammenhang Verfahren aus dem Natural-Language-Processing (NLP) gemeint, wie das Entfernen von Stopp-Wörtern, evtl. Stemming sowie Rauschreduktion (Wörter, die nur in einem einzigen Text vorkommen, sind für Ähnlichkeiten nicht nützlich).

Beispiel mit drei Texten und jeweils vier Wörtern:

- $T_1 = \{\text{Informatik, Recommender, Filter, Webseite}\}$
- $T_2 = \{\text{Informatik, Internet, Browser, Webseite}\}$
- $T_3 = \{\text{E-Mail, Internet, Konversation, SPAM}\}$

In einem Koordinatensystem mit neun Dimensionen (für [Informatik, Recommender, Filter, Webseite, Internet, Browser, E-Mail, Konversation, SPAM]) können die drei Texte als Binärvektoren wie folgt dargestellt werden:

- $T_1 = (1, 1, 1, 1, 0, 0, 0, 0, 0)$
- $T_2 = (1, 0, 0, 1, 1, 1, 0, 0, 0)$
- $T_3 = (0, 0, 0, 0, 1, 0, 1, 1, 1)$

Die Ähnlichkeiten der einzelnen Texte können mit einem Ähnlichkeitsmaß (vgl. Abschnitt 3.3) berechnet werden. Mit der Cosinus-Korrelation gilt: $s(T_1, T_2) = \frac{2}{2 \cdot 2} = 0,5$, $s(T_1, T_3) = \frac{0}{2 \cdot 2} = 0$ und $s(T_2, T_3) = \frac{1}{2 \cdot 2} = 0,25$. Weitere Möglichkeiten, die Ähnlichkeiten von Artefakten zu bestimmen, bietet die normalisierte Informationsdistanz (ein generisches Abstandsmaß (Li u. a. 2003; Cilibrasi u. Vitanyi 2003), das auch für Nicht-Text-Artefakte geeignet ist und bereits erfolgreich zum Clustering von MIDI- und GIF-Dateien eingesetzt wurde (Cebrián u. a. 2007)) oder die Nutzung von String-Kernen (nur für textbasierte Artefakte, vgl. Lodhi u. a. 2002).

Zur Berechnung der Empfehlungen kann auf die Formel(n) vom elementbasierten kollaborativen Filtern aus Abschnitt 3.9.2 mit den Artefakt-Ähnlichkeiten als Ähnlichkeitsmatrix S und dem Schwellenwert t_{\min} . Bewertung zurückgegriffen werden.

Der Vorteil dieses Vorgehens liegt darin, dass die Berechnungen nur auf den Artefakten sowie den Benutzerbewertungen des aktiven Benutzers basieren. Ähnlichkeiten können zwischen allen, also auch sehr neuen (unbewerteten), Artefakten berechnet und auch (länger) zwischengespeichert werden. Zu den Nachteilen zählt, dass Empfehlungen ohne Veränderungen an den Artefakten bzw. den Bewertungen sehr statisch sind und dem Benutzer größtenteils identische Artefakte angezeigt werden – es ist unwahrscheinlich, dass nicht so offensichtliche bzw. inhaltlich „neue“ Artefakte empfohlen werden, die den Benutzer interessieren und die er nicht selbst gefunden hätte (Sarwar u. a. 1998; Sollenborn u. Funk 2002). Zudem sind diese Verfahren sehr von den Daten abhängig: Werden z. B. Schlüsselwörter uneinheitlich vergeben, leidet darunter die Empfehlungsqualität.

3.7 Wissensbasiertes Filtern

Das wissensbasierte Filtern („knowledge-based filtering“, Burke 2000) ist eng verwandt mit dem inhaltsbasierten Filtern, jedoch ist der Nutzer hier stark in den Filterprozess mit eingebunden. Ziel dieses Ansatzes ist es, mit Wissen über Produkte bzw. Benutzer die möglichst am besten passende Alternative (iterativ) zu bestimmen. Dazu wird dem Benutzer zu einer Empfehlung eine Schnittstelle bereitgestellt, mit der er innerhalb der Datenbank gemäß seinen Anforderungen navigieren kann. Da dieses Verfahren

unabhängig von Benutzer-Bewertungen ist, hat es keine Kaltstart-Probleme und kann auch auf sich schnell ändernde Benutzervorlieben reagieren. Dennoch müssen sämtliche Eigenschaften und Features der Artefakte durch einen Experten (objektiv) bestimmt und in einer Datenbank abgelegt werden, bevor das Verfahren eingesetzt werden kann.

Um dieses Vorgehen besser verstehen zu können, wird es am Beispiel eines Restaurant-Empfehlungssystems (vgl. Burke 2000) beschrieben: Auf einer Eingangsseite kann sich ein Benutzer zu einem bekannten Restaurant ein ähnliches oder eines z. B. anhand der Kriterien Küche (Italienisch, Spanisch, Französisch, ...), Preis und Atmosphäre herausuchen. Daraufhin wird er zu einem oder mehreren Vorschlägen geleitet. Direkt unter jedem Vorschlag befinden sich die Navigationsbuttons „günstiger“, „teurer“, „schöner“, „leiser“ und „kreativer“ mit deren Hilfe der Benutzer durch den Datenbestand zu dem oder den ähnlichsten Restaurants mit dieser Eigenschaft navigieren kann. Dieser Vorgang kann wiederholt werden, bis ein passendes Restaurant gefunden wurde.

3.8 Demographisches Filtern

Empfehlungssysteme dieses Typs versuchen, Benutzer abhängig von ihren persönlichen Eigenschaften zu klassifizieren und anhand der Klassen/Stereotypen Empfehlungen zu berechnen (vgl. Burke 2002). Dieses Vorgehen basiert auf der Annahme, dass Personen mit gleichen Eigenschaften sehr wahrscheinlich gleiche Interessen haben. Die Klassen/Eigenschaften werden vom Entwickler vorgegeben oder durch Clustering bzw. Learning-Verfahren gewonnen. Zu den typischen Klassen gehören Alter („jung“ / „mittleres Alter“ / „alt“), Geschlecht, Wohnort (verschiedene Detailgrade: Staaten, Länder, Städte, ...), Schulbildung, Einkommen und Religionszugehörigkeit. Eine besondere Schwierigkeit dieser Systeme liegt im Datenschutz bzw. der Notwendigkeit einer Erhebung und Speicherung persönlicher Daten, zu der nicht alle Menschen bereit sind.

α -Community-Räume (AC) (Nguyen u. a. 2007) sind eine dynamische Variante für demographische Filter, obwohl diese nicht explizit für diese Kategorie vorgestellt wurden, sondern um einem Problem der kollaborativen Filter (vgl. Abschnitt 3.9 sowie „New User“-Problem in Abschnitt 3.10) mit „cold-user data“ bzw. Benutzerprofil-Daten entgegenzuwirken. Das α -Community-Modell-Framework gruppiert explizit Benutzer mit gleichen Eigenschaften in sog. α -Communities bezüglich einer Ähnlichkeitseigenschaft α . Nguyen u. a. führten die α -Community-Räume als Partition der Benutzermenge ein, jedoch sind paarweise disjunkte Gruppen nicht unbedingt notwendig. Speziell bei dynamischen Gruppen, wie z. B. bei der Gruppe aller Nutzer mit Alter im Intervall *Alter des aktiven Benutzers* $\pm x$ Jahre, ist diese Eigenschaft zu strikt.

Die Generierung von Empfehlungen erfolgt in drei Schritten mit der sog. „Level of Agreement“-Methode: Zuerst wird die α -Community, d. h. die Gruppe der Benutzer mit der Ähnlichkeitseigenschaft α , bestimmt. Im zweiten Schritt wird für alle Artefakte, die von mindestens $t_{\text{agreement}}$ (z. B. $t_{\text{agreement}} = 0,25 = 25\%$) der Benutzer in der Gruppe bewertet wurden, die Durchschnittsbewertung innerhalb der gesamten Gruppe bestimmt. Davon werden schließlich bis zu N Artefakte empfohlen, die im Schnitt mit mindestens $t_{\text{min. Bewertung}}$ bewertet wurden.

Es ist möglich, verschiedene α -Communities zu kombinieren: Nguyen u. a. schlugen hier eine (gewichtete) Linearkombination der einzelnen vorhergesagten Bewertungen verschiedener α -Communities für ein Artefakt vor. Damit können nutzerbasierte kollaborative Filter ganz allgemein als eine einzige große α -Community mit α als Übereinstimmung in den Bewertungen angesehen werden.

3.9 Kollaboratives Filtern

Das kollaborative Filtern (CF) setzt auf die kollektive Intelligenz der Benutzer eines Systems statt nur auf das Wissen über einen bzw. eines einzelnen. Vereinfacht gesagt werden die Daten anderer Nutzer benutzt, um eine personalisierte Vorhersage bzw. Empfehlung zu erstellen (unter Berücksichtigung der Bewertungen des aktiven Benutzers).

Erstmals wurde dieser Begriff von Goldberg u. a. (1992) in Zusammenhang mit dem „Tapestry“-System verwendet. Das System ermöglichte es Benutzern, Dokumente (Usenet-Einträge) als interessant bzw. nicht interessant zu bewerten und anderen Benutzern anhand dieser Informationen Dokumente zu filtern. Jedoch erforderte das Filtern bzw. das Erstellen der Filter (z. B. „sortiere alle Beiträge aus, die Person X auch ausgefiltert hat“) viele manuelle Eingriffe.

Heutzutage wird der Begriff „kollaboratives Filtern“ in der Fachliteratur (Deshpande u. Karypis 2004; Candillier u. a. 2007; Segaran 2008; Marmanis u. Babenko 2009) meist synonym für „nachbarschaftsbasierte kollaborative Filter“ mit einem k -Nächste-Nachbarn-Ansatz verwendet. Dieses Verständnis wird im Folgenden adaptiert.

Nach Herlocker u. a. (1999) haben CF-basierte Systeme drei entscheidende Vorteile gegenüber inhaltsbasierten:

- Es kann mit Artefakten gearbeitet werden, deren Inhalt bzw. Eigenschaften sich nicht (leicht) automatisch verarbeiten lassen (CF ist domänenunabhängig, es werden nur Benutzer, Artefakte und Bewertungen benötigt).
- Die Fähigkeit, Artefakte nach Qualität, Geschmack und Vorlieben zu filtern sowie
- unerwartete, nicht offensichtliche Empfehlungen (z. B. für Artefakte, die sich ein Benutzer sonst nicht angesehen hätte, aber interessant für ihn sind) zu berechnen.

3.9.1 Nutzerbasiertes Filtern

Das erste System unter der Flagge „kollaborativer Filter“, das automatisch konkrete Empfehlungen berechnet hat, wurde 1994 von Resnick u. a. sowie Konstan u. a. (1997) vorgestellt: „GroupLens“, ein System, das Usenet-Beiträge nachbarschaftsbasiert empfehlen kann.

Beim nutzerbasierten kollaborativen Filtern werden ähnliche Benutzer gesucht, um einem Benutzer Artefakte vorzuschlagen, die er selbst noch nicht bewertet und ein ähnlicher Benutzer als gut bewertet hat. Dahinter verbirgt sich die Annahme, dass Menschen mit ähnlichen Vorlieben Artefakte ähnlich bewerten bzw. Benutzer, die in

der Vergangenheit Artefakte ähnlich bewertet haben, wahrscheinlich auch weiterhin Artefakte ähnlich bewerten werden.

Im Allgemeinen besteht der Algorithmus aus vier Schritten (nach Herlocker u. a. 1999):

1. Gewichtung aller Benutzer in Abhängigkeit von deren Ähnlichkeit zum aktuellen Benutzer
2. Auswahl einer Untermenge von Benutzern (Nachbarn, $\mathcal{U}' \subseteq \mathcal{U}$), die für die Vorhersage herangezogen werden sollen (i. d. R. alle oder die k ähnlichsten Benutzer)
3. Normalisierung der Bewertungen und Berechnung der Vorhersagen aus einer gewichteten Kombination der Bewertungen sowie Ähnlichkeiten der ausgewählten Nachbarn
4. Auswahl der N Artefakte mit den besten vorhergesagten Bewertungen mit mindestens t_{\min} Bewertung, die der aktive Benutzer noch nicht bewertet hat

Wenn die Menge der ähnlichen Nachbarn \mathcal{U}' bestimmt ist, berechnet sich eine Vorhersage $p_{u,i}$ für einen Benutzer u und ein Artefakt i wie folgt:

$$p_{u,i} = \bar{r}_u + \frac{\sum_{u' \in \mathcal{U}'} (r_{u',i} - \bar{r}_{u'}) s(u, u')}{\sum_{u' \in \mathcal{U}'} s(u, u')} \quad \text{mit } \bar{r}_u = \frac{1}{|\mathcal{I}_u|} \sum_{i \in \mathcal{I}_u} r_{u,i}$$

Es werden für alle ähnlichen Benutzer $u' \in \mathcal{U}'$ die mit der Benutzerähnlichkeit $s(u, u')$ gewichteten Abweichungen von deren Durchschnittsbewertungen aufsummiert. Diese Summe wird, nach einer Normalisierung mit einer Division durch die Summe aller Benutzer-Ähnlichkeiten, zum Durchschnitt der Bewertungen des Benutzers u addiert, was schließlich die Vorhersage ergibt. Die Ähnlichkeiten können mit einem beliebigen Ähnlichkeitsmaß (vgl. Abschnitt 3.3) berechnet werden. Bei der Ähnlichkeitsberechnung werden jedoch ausschließlich Artefakte betrachtet, die beide zu vergleichenden Benutzer bereits bewertet haben. Durch die Verwendung der Durchschnittswerte bzw. der Abweichungen davon sollen die Unterschiede verschiedener Benutzer ausgeglichen bzw. an den aktuellen Nutzer angepasst werden.

Die Vorhersage entspricht einer Skalarmultiplikation $r', s \in \mathbb{R}^k$ sowie der Addition des Bewertungsdurchschnitts des aktiven Benutzers: Der auf Länge eins normalisierte Vektor $s = (s(u, u'_1), \dots, s(u, u'_k))$ besteht aus den Ähnlichkeiten der k Benutzer zum aktiven Nutzer und der Vektor $r' = (r_{u'_1,i} - \bar{r}_{u'_1}, \dots, r_{u'_k,i} - \bar{r}_{u'_k})$ aus der Differenz der Bewertungen zum Durchschnitt der Bewertungen der k Nachbarn für das Artefakt i .

$$p_{u,i} = \bar{r}_u + r' \cdot s^T$$

Zur Berechnung von Empfehlungen müssen schließlich alle m' Artefakte bestimmt werden, die der aktive Benutzer nicht, aber mindestens ein Nachbar bewertet hat. Für diese werden dann Vorhersagen berechnet, um die Artefakte mit den höchsten vorhergesagten Bewertungen zu empfehlen. Dies kann durch eine einfache Erweiterung auf Matrizen erzielt werden, die simultan alle m' Vorhersagen bestimmt:

$$\begin{pmatrix} p_{i_1,u} \\ \vdots \\ p_{i_{m'},u} \end{pmatrix} = \begin{pmatrix} \bar{r}_u \\ \vdots \\ \bar{r}_u \end{pmatrix} + \mathcal{R}'^T \cdot s^T$$

mit der $k \times m'$ Matrix \mathcal{R}' , die die Abweichungen von den Bewertungen der k Nachbarn für die m' Artefakte enthält.

Wie man an den Berechnungsvorschriften leicht erkennen kann, müssen immer alle Benutzer mit deren Bewertungen betrachtet werden, um die k ähnlichsten Benutzer zu bestimmen (Komplexität $\mathcal{O}(nm + n \log n)$). Zudem müssen die k ähnlichsten Benutzer für (fast) jede Empfehlung (mit einer Komplexität von $\mathcal{O}(km)$ bzw. $\mathcal{O}(km')$ mit $m' < m$ für die Berechnung einer Empfehlung) bestimmt werden, da die Ähnlichkeiten nicht sinnvoll über längere Zeit zwischengespeichert werden können: Bewertet ein beliebiger Benutzer ein Artefakt, kann sich die Ähnlichkeit verändern und eine zwischengespeicherte Version wäre nicht mehr aktuell. Daher wird dieses Verfahren auch oft „memory-based“ genannt.

Darüber hinaus liegen von einem Benutzer i. d. R. nur für einen (sehr) kleinen Teil der Artefakte Bewertungen vor: Der Benutzervektor r_u besteht aus überwiegend leeren Komponenten bzw. null-Werten. Daraus ergibt sich das sog. „Sparsity-Problem“. Für die Berechnung der Ähnlichkeit zwischen zwei Benutzern ist es notwendig, dass sich deren Daten überschneiden bzw. diese Benutzer gleiche Artefakte bewertet haben. Liegen Überschneidungen nur in wenigen oder gar keinen Komponenten vor, so kann sich dies sehr negativ auf die Vorhersagen auswirken bzw. eine Vorhersage verhindern. Zudem können mit dem reinen nutzerbasierten kollaborativen Filtern keine Abhängigkeiten zwischen den Artefakten abgebildet werden.

Um diese Probleme zu entschärfen wurde ein Verfahren vorgeschlagen, dass auf der Singulärwertzerlegung (Singular Value Decomposition, SVD) basiert (Sarwar u. a. 2000b): Dabei wird eine sehr dünn besetzte Matrix auf eine Matrix kleinerer Dimension, die nicht so dünn besetzt ist, durch mehrere Matrixmultiplikationen und Entfernung kleiner Singulärwerte herunter gebrochen. Zudem „findet“ dieses Verfahren versteckte Abhängigkeiten zwischen einzelnen Artefakten (Latent Semantic Indexing Technik). Es hat sich jedoch gezeigt, dass dieses Verfahren schwer zu debuggen, aufwändig bzw. langsam ist und nur für eine kleine Verbesserung des Ergebnisses sorgt (Tang u. a. 2005; Schafer u. a. 2007).

Beispiel:

Gegeben ist die folgende Bewertungsmatrix \mathcal{R} (Tabelle 3.3) für fünf Benutzer (u_1, \dots, u_5) und vier Artefakte (i_1, \dots, i_4). Dabei wurden von u_2 und u_4 nicht alle Artefakte bewertet (i_3 resp. i_2 fehlen). Es soll die Vorhersage für Benutzer u_1 und Artefakt i_4 berechnet werden.

Zuerst müssen die Ähnlichkeiten der Benutzer u_2, \dots, u_5 mit u_1 bestimmt werden. Da die Berechnung mit der Pearson-Korrelation (vgl. Abschnitt 3.3.4) erfolgen soll, werden die Durchschnittsbewertungen der einzelnen Benutzer benötigt (vgl. Tabelle 3.4).

\mathcal{R}	i_1	i_2	i_3	i_4
u_1	5	1	2	?
u_2	1	5	n/a	3
u_3	4	1	3	4
u_4	3	n/a	2	2
u_5	4	4	3	3

Tabelle 3.3: Bewertungsmatrix

$$\begin{aligned}
 \bar{r}_{u_1} &\approx 2,67 \\
 \bar{r}_{u_2} &= 3 \\
 \bar{r}_{u_3} &= 3 \\
 \bar{r}_{u_4} &\approx 2,34 \\
 \bar{r}_{u_5} &= 3,5
 \end{aligned}$$

Tabelle 3.4: Bewertungsmittelwerte der Benutzer

Die Ergebnisse der Ähnlichkeitsberechnung sind in der folgenden Tabelle dargestellt (Tabelle 3.5):

	$s(u_1, \cdot)$
u_2	0
u_3	0,8609
u_4	0,8102
u_5	0,2615

Tabelle 3.5: Ähnlichkeitsvektor

Der Wert von $s(u_1, u_2)$ beträgt rein rechnerisch $-0,99$. Da jedoch negative Ähnlichkeiten, wie in Abschnitt 3.3.1 angedeutet, nicht hilfreich sind, wird die Ähnlichkeit auf 0 gesetzt.

Für dieses Beispiel sollen lediglich die zwei nächsten Nachbarn ($k = 2$) betrachtet werden. Die Benutzer mit den zwei höchsten Ähnlichkeiten sind u_3 und u_4 (in dieser Reihenfolge). Die Berechnung der Vorhersage lautet dann:

$$\begin{aligned}
 p_{u_1, i_4} &= \bar{r}_{u_1} + \frac{s(u_1, u_3)(r_{u_3, i_4} - \bar{r}_{u_3}) + s(u_1, u_4)(r_{u_4, i_4} - \bar{r}_{u_4})}{s(u_1, u_3) + s(u_1, u_4)} \\
 &= 2,67 + \frac{0,8609 \cdot (4 - 2,75) + 0,8102 \cdot (2 - 2,34)}{0,8609 + 0,8102} \approx 2,8
 \end{aligned}$$

Damit ergibt sich in diesem sehr vereinfachten Beispiel eine Vorhersage von 2,8 für Benutzer u_1 und Artefakt i_4 . In realen Umgebungen ist die Bewertungsmatrix i. d. R. deutlich spärlicher besetzt.

3.9.2 Elementbasiertes Filtern

Abhilfe für das Problem, Ähnlichkeiten nicht für längere Zeit zwischenspeichern zu können, schafft eine leicht veränderte Herangehensweise: das element- bzw. modellbasierte kollaborative Filtern (IB) (Sarwar u. a. 2001).

Statt der Berechnung der sich relativ oft ändernden Ähnlichkeiten der Benutzer (Zeilenvektoren der Bewertungsmatrix) werden die Ähnlichkeiten zwischen den Artefakten (Spaltenvektoren) berechnet. Diese können bereits im Voraus oder auch verteilt berechnet (Zeitkomplexität $\mathcal{O}(nm^2)$, Speicherkomplexität $\mathcal{O}(mk)$) und für eine längere Zeit zwischengespeichert werden. Zudem wird das bereits in Abschnitt 3.9.1 erwähnte „Sparsity-Problem“ entschärft, da das Verfahren nicht auf Überlappungen zwischen Benutzern, sondern auf häufiger vorkommenden Überlappungen bei Artefakten basiert.

Möchte man nun für einen aktiven Benutzer u eine Vorhersage treffen, sucht man die ähnlichen Artefakte zu den Artefakten, die der aktive Benutzer u bereits bewertet hat und bildet die gewichtete Summe der Bewertungen des Benutzers mit der Ähnlichkeit der Artefakte (Komplexität $\mathcal{O}(km)$ bzw. $\mathcal{O}(kb)$ mit $b \ll m$ Anzahl der Bewertungen des aktiven Benutzers). Für eine Empfehlung müssen hier nicht mehr alle Benutzer bzw. Artefakte betrachtet werden – der Algorithmus skaliert besser.

Im Allgemeinen besteht der Algorithmus aus vier Schritten, wobei der erste Schritt nicht für jede Empfehlung, sondern nur in regelmäßigen Abständen durchgeführt werden muss:

1. Berechnung der Ähnlichkeiten der Artefakte
2. Auswahl einer Untermenge von Artefakten (Nachbarn, $\mathcal{I}' \subseteq \mathcal{I}$), die zu den durch den aktiven Benutzer bewerteten ähnlich sind (i. d. R. alle oder die k ähnlichsten)
3. Normalisierung der Bewertungen und Berechnung der Vorhersagen aus einer gewichteten Kombination der Bewertungen sowie Ähnlichkeiten der Artefakte aus \mathcal{I}'
4. Auswahl der N Artefakte mit den besten vorhergesagten Bewertungen mit mindestens t_{\min} Bewertung

Eine Vorhersage $p_{u,i}$ für einen Benutzer u und ein Artefakt i berechnet sich wie folgt:

$$p_{u,i} = \frac{\sum_{i' \in \mathcal{I}'} s(i, i') r_{u,i'}}{\sum_{i' \in \mathcal{I}'} s(i, i')}$$

Die Linearkombination der Bewertungen wird mit einer Division durch die Summe der Ähnlichkeiten normalisiert, damit sich die Vorhersage im gleichen Bereich wie alle Bewertungen befindet. Dies entspricht einer Matrixmultiplikation mit der zeilenweise normalisierten $m \times m$ Ähnlichkeitsmatrix S und dem $1 \times m$ Bewertungsvektor r_u (nach Deshpande u. Karypis 2004, S. 153, Algorithm 4.2):

$$p = S \cdot r_u^T$$

wobei im $m \times 1$ Vorhersagevektor p nur Artefakte betrachtet werden müssen, die der aktive Benutzer noch nicht bewertet hat.

Verglichen mit den nutzerbasierten Filtern hat sich das elementbasierte Filtern nicht nur als effizienter erwiesen, sondern die Ergebnisse sind in verschiedenen Studien (vgl. Sarwar u. a. 2001; Candillier u. a. 2007; Deshpande u. Karypis 2004) als vergleichbar oder besser evaluiert worden.

Beispiel:

Gegeben ist die Bewertungsmatrix \mathcal{R} aus Tabelle 3.3 (auf Seite 31) des vorherigen Abschnitts. Es soll eine Empfehlung für den Benutzer u_1 bestimmt werden.

Zuerst müssen die Ähnlichkeiten der einzelnen Artefakte untereinander bestimmt werden. Die Berechnung erfolgt mit der angepassten Cosinus-Korrelation (Abschnitt 3.3.6), deren Ergebnisse in der folgenden Ähnlichkeitsmatrix \mathcal{S} aufgeführt sind (Tabelle 3.6). Negative Korrelationen werden, wie bereits erläutert, auf 0 gesetzt. Auf Grund der Symmetrie genügt es, lediglich die obere (bzw. untere) Dreiecksmatrix zu bestimmen. Da neue Artefakte empfohlen werden sollen, ist die Berechnung bzw. Speicherung der Selbstähnlichkeiten ($s(i, i) = 1$ auf der Hauptdiagonalen) nicht notwendig.

\mathcal{S}	i_1	i_2	i_3	i_4
i_1	-	0	0	0,1896
i_2	-	-	0,3898	0
i_3	-	-	-	0,5151
i_4	-	-	-	-

Tabelle 3.6: Ähnlichkeitsmatrix

Da der Benutzer u_1 die Artefakte i_1 , i_2 und i_3 bewertet hat, aber nur i_1 und i_3 zu i_4 ähnlich sind, lautet die Formel zur Berechnung der Vorhersage wie folgt:

$$p_{u_1, i_4} = \frac{s(i_1, i_4)r_{u_1, i_1} + s(i_3, i_4)r_{u_1, i_3}}{s(i_1, i_4) + s(i_3, i_4)} = \frac{(0,1896 \cdot 5) + (0,5151 \cdot 2)}{0,1896 + 0,5151} \approx 2,81$$

Damit ergibt sich in diesem sehr vereinfachten Beispiel eine Empfehlung von Artefakt i_4 mit einer Vorhersage von 2,81 für Benutzer u_1 . Zufällig stimmt dieser Wert in diesem Beispiel mit der des UB-Ansatzes überein. Im Allgemeinen trifft dies nicht zu.

3.10 Allgemeine Probleme und Lösungsansätze

Im Folgenden werden Probleme von Empfehlungssystemen mit einigen Lösungsvorschlägen beschrieben. Die meisten der angesprochenen Probleme beziehen sich vornehmlich auf kollaborative Filter, treffen teilweise jedoch auch auf andere Verfahren zu (sofern dies in den jeweiligen Abschnitten nicht ausgeschlossen wurde).

3.10.1 Sparsity

Ein allgemeines Problem der kollaborativen Filter ist die Spärlichkeit der Bewertungsmatrix, d. h. die Mehrzahl der Werte sind 0 oder nicht besetzt. Dies wirkt sich besonders stark auf die nutzerbasierte Variante (vgl. Abschnitt 3.9.1) aus – die elementbasierte kann mit diesem Problem etwas besser umgehen, jedoch nicht mit beliebig wenigen Bewertungen. Eine wichtige Kennzahl ist das Sparsity-Level der Daten, da die Qualität der Empfehlungen stark davon abzuhängen scheint (Sarwar u. a. 2001; Huang u. a. 2007): Für eine Bewertungsmatrix ist dieses definiert als $1 - \frac{\text{nichtleere Einträge}}{\text{Anzahl Einträge}} = 1 - \frac{\text{nichtleere Einträge}}{nm}$.

Sind insgesamt nur wenige Bewertungen vorhanden und ist das Sparsity-Level nahe an eins, kann versucht werden, eine Pseudo-Bewertungsmatrix aufzustellen, die ein geringeres Sparsity-Level besitzt (aber dafür etwas ungenauer wird). Hierfür werden i. d. R. domänenspezifische Eigenschaften ausgenutzt. Im Folgenden werden einige Ansätze vorgestellt:

Zusammenfassen von Artefakten Gibt es strukturelle oder inhaltliche Zusammenhänge zwischen den Artefakten, können diese dazu genutzt werden, die Dimension der Bewertungsmatrix zu reduzieren: Beziehen sich mehrere Artefakte auf das gleiche Thema oder sind sie auf eine andere Art und Weise eng verwandt, ist es möglich, diese Artefakte zu einer Gruppe bzw. einem Cluster zusammenzufassen. Li u. Kim (2003) beschreiben ein Verfahren, in dem erst Ähnlichkeiten zwischen Artefakten berechnet (vgl. Abschnitt 3.6) und schließlich sehr ähnliche Artefakte gleichen Clustern zugeordnet werden. Die Bewertungen der Cluster können dabei über alle in der Gruppe enthaltenen Artefakte (z. B. nach der Ähnlichkeit) gewichtet oder ungewichtet gemittelt werden. Durch dieses Vorgehen kann die Anzahl der Artefakte im Empfehlungssystem auf die Anzahl der eindeutigen Gruppen, die wiederum oftmals besser gefüllte Bewertungsvektoren besitzen, reduziert werden. Eine ähnliche Methode ergibt sich, falls Artefakte in einem Baum verwaltet werden, bei denen die hauptsächlichen Informationen in den Blättern zu finden und die übergeordneten Knoten lediglich Übersichtsseiten sind. Dann bietet es sich an, die inneren Knoten als „natürliche“ Cluster und die Übersichtsseiten bei Empfehlungen zu nutzen. Dadurch wird es den Benutzern ermöglicht, die für sie interessanten Artikel innerhalb des Clusters (schnell) zu finden (der Google PageRank (Page u. a. 1999) priorisiert aus diesem Grund ebenfalls Übersichtsseiten). Der Nachteil ist natürlich, dass (ohne Durchführung weiterer Schritte) nur noch Empfehlungen für die einzelnen Gruppen statt auf der Ebene der einzelnen Artefakte berechnet werden können.

Filterbots Sarwar u. a. (1998) sowie Park u. a. (2006) entwickelten sog. naive Filterbots: Agenten/Programme, die automatisiert (alle) Artefakte nach gewissen Kriterien bewerten, also inhaltsbasierte Informationen in den kollaborativen Filter integrieren. Filterbots für nutzerbasiertes kollaboratives Filtern wurden im Bereich Usenet von Sarwar u. a. vorgeschlagen. Dabei handelte es sich um sehr einfache Agenten, die Bewertungen abhängig von der Länge bzw. der Anzahl der Schreibfehler von Usenet-Beiträgen vergeben. Dadurch werden Überlappungen in den Bewertungsvektoren der Benutzer geschaffen. Park u. a. haben das Konzept auch auf elementbasierte Systeme übertragen, bei denen die Filterbots lediglich für die

Vergrößerung der Nachbarschaft sorgen. Dadurch ist es möglich, explizit Ähnlichkeiten auch zwischen Artefakten (z. B. Elemente aus Buch- oder DVD-Reihen) in kollaborative Filter zu integrieren. Allgemein sind Filterbots einfach zu implementieren und auszuführen. Bereits einfache Varianten können einen signifikanten Einfluss auf die Empfehlungsqualität haben (Konstan u. a. 1998). Es muss jedoch darauf geachtet werden, dass nicht zu viele Filterbots eingesetzt oder die Filterbotbewertungen in der Ähnlichkeitsberechnung anders gewichtet werden, da sonst Benutzerbewertungen im Verhältnis immer weniger in die Ähnlichkeiten eingehen.

Boosting Anstatt durch automatisierte Bewertungen oder das *Zusammenfassen von Artefakten* das Sparsity-Level zu verringern, wird beim Boosting versucht, über einen anderen Algorithmus aus den bestehenden Bewertungen eines Nutzers neue zu generieren. Melville u. a. (2002) schlugen das „content-boosting“ vor: Dabei werden mit Hilfe eines inhaltsbasierten Vorgehens (vgl. Abschnitt 3.6) Bewertungen vorhergesagt und in eine Pseudo-Bewertungsmatrix eingefügt. Dieser Ansatz macht sich den Nachteil der inhaltsbasierten Filter zunutze, dass diese vornehmlich identische Artefakte empfehlen. Dadurch wird das Sparsity-Level gesenkt und ermöglicht, die Qualität der Empfehlungen zu steigern. Eine weitere Boosting-Methode, die bisher nicht in der Literatur gefunden wurde, ist das Boosting mittels nutzerbasiertem kollaborativen Filtern für einen elementbasierten Filter (oder vice versa). Auf diese Weise, speziell in Kombination mit demographischen Daten (z. B. Alter oder Geschlecht der Benutzer als Bewertung von Pseudo-Artefakten, die für Überlappungen sorgen; vgl. Vozalis u. Margaritis (2004)), könnten die Vorteile von nutzerbasierten und elementbasierten Filtern theoretisch kombiniert werden: Es ließen sich das nutzerbasierte Filtern in die Modellbildung verlagern, demographische Informationen sowie Zusammenhänge zwischen Artefakten bei Empfehlungen nutzen und das Sparsity-Level verringern.

3.10.2 Cold-Start-Probleme

Darüber hinaus existieren die „cold-start“- bzw. Kaltstart-Probleme (Schafer u. a. 2007; Middleton u. a. 2002). Dies sind spezielle Situationen, in denen keine oder nur wenige Daten verfügbar sind (das im vorherigen Abschnitt angesprochene Sparsity-Problem ist allgemeiner zu verstehen). Kaltstartprobleme können in drei unterschiedlichen Ausprägungen auftreten:

New System/New Community Hierbei handelt es sich um das gravierendste Kaltstartproblem: Ein neues Empfehlungssystem wird aufgebaut; aber ohne bzw. mit sehr wenigen Bewertungen können keine oder nur unterdurchschnittlich gute Vorausagen berechnet und Empfehlungen abgegeben werden. Dieser Umstand kann zu einem „Teufelskreis“ führen, da die Benutzer nicht von den Bewertungen profitieren und deshalb keine (expliziten) Bewertungen abgeben oder die Benutzung des Systems einstellen. Damit Empfehlungen ein gutes Niveau erreichen, ist eine kritische Masse an Daten erforderlich (Maltz u. Ehrlich 1995). Diesem Problem kann entgegengetreten werden, indem externe Bewertungen zugeführt werden oder zunächst Bewertungen (z. B. implizit oder nur von einer kleinen Teilmenge

der Benutzer) gesammelt werden und das Empfehlungssystem später (für alle Nutzer) dazu geschaltet wird.

New User Das „New User“-Problem ist ein Spezialfall des „New Community“-Problems, bei dem im Allgemeinen genug Daten vorhanden sind, aber für einen (neuen) Benutzer keine oder nur sehr wenige Daten vorliegen – es also nicht möglich ist, für diesen speziellen Benutzer (gute) Empfehlungen zu berechnen. Beim nutzerbasierten Vorgehen tritt dieses Problem z. B. auf, falls es keine Überlappungen mit anderen Nutzern gibt. Rashid u. a. (2002) schlugen vor, einen neuen Benutzer ausgewählte Artefakte bewerten zu lassen, so dass ein größtmöglicher Informationsgewinn für das Profil des Nutzers entsteht, um so eine Grundlage für gute Empfehlungen zu schaffen. Auch Filterbots können automatisiert viele Artefakte bewerten und so für Benutzer, die bisher sehr wenige Artefakte bewertet haben, Überlappungen erzeugen. Eine weitere Möglichkeit besteht darin, unpersonalisiertes durchschnittsbasiertes Filtern (vgl. Abschnitt 3.5) einzusetzen oder weitere Daten (z. B. demographische, vgl. Abschnitt 3.8) mit in den Empfehlungsalgorithmus einfließen zu lassen.

New Item Ein weiterer Spezialfall des „New Community“-Problems ist das „New Item“-Problem. Es tritt jedes Mal auf, sobald ein neues Artefakt hinzugefügt wird. In diesem Moment besitzt es noch keine Bewertungen und kann folglich (speziell bei CF-Verfahren) nicht empfohlen werden. In Systemen, in denen häufig neue Artefakte eingestellt werden und die Aktualität nicht lange gewährleistet ist, z. B. auf einem Nachrichtenportal, kann dies kritisch sein. In anderen Systemen, in denen die Artefakte auch über andere Wege gefunden werden können, besteht die Möglichkeit, dass diese nach und nach bewertet werden. Ebenso können Filterbots in diesem Kontext eingesetzt werden, da durch sie auch für neue Artefakte Bewertungen vergeben werden können. Weitere Lösungsmöglichkeiten bestehen in der Nutzung von Daten inhaltlicher Natur (vgl. Abschnitt 3.6) oder Benutzer zufällig unbewertete Artefakte bewerten zu lassen.

3.10.3 Transparenz-, Blackbox- bzw. Vertrauens-Problem

Das Blackbox-Problem beschreibt die fehlende Transparenz des Empfehlungsprozesses – unabhängig von der Qualität der Empfehlungen. Im Gegensatz zu Empfehlungen durch Bekannte ist es Benutzern nicht möglich, das Ergebnis zu hinterfragen und die Gründe bzw. Intentionen zu erfahren. Das kann Benutzer daran hindern, das System richtig zu verstehen und optimal einzusetzen oder „Fehler“ zu melden. Da dies zudem zu einem Vertrauens- oder Zufriedenheitsverlust in das System führen kann, ist dieses Problem Gegenstand aktueller Forschung (Sinha u. Swearingen 2002; Tintarev u. Masthoff 2007). Ergebnisse zeigen, dass die Präsentation der Erklärungen eine entscheidende Rolle spielt, Benutzer aber auch durch zu viele Informationen leicht überfordert werden können (Herlocker u. a. 2000).

Zusätzlich können durch mangelnde Transparenz im Empfehlungsprozess bösartige Benutzer Empfehlungen für andere Nutzer manipulieren, indem sie z. B. bei einem Produkt-Bewertungsportal versuchen, die Produkte ihrer eigenen Firma besser und

die anderen Firmen schlechter erscheinen zu lassen. UB-Verfahren bieten größere Angriffsflächen für Manipulationen als IB-Verfahren, aber insgesamt sind neue bzw. wenig bewertete Artefakte besonders anfällig (Schafer u. a. 2007).

3.10.4 Datenschutz und Privatsphäre

Empfehlungssysteme basieren i. d. R. auf sehr detaillierten Profilen, um ihrer Aufgabe, der Berechnung von personalisierten Empfehlungen, nachzukommen. Eine Faustregel lautet: Je mehr ein Empfehlungssystem über eine Person weiß, desto bessere Empfehlungen kann es generieren. Diese Profile sind zudem meist an einer zentralen Stelle gespeichert, damit das ES einfach darauf zugreifen kann. Speziell bei demographischen Filtern sind neben den Interessen, die durch Bewertungen repräsentiert werden, zusätzlich noch persönliche Daten vorhanden. Überall, wo persönliche Daten bzw. Profile vorliegen, können diese prinzipiell ausgenutzt bzw. nicht im Sinne der Benutzer verarbeitet werden. Die Benutzer müssen also darauf vertrauen, dass die Daten ausschließlich für die angegebenen Zwecke benutzt werden. Daher ist es wichtig, Datenschutzrichtlinien zu etablieren und diese Daten besonders zu schützen bzw. so weit wie möglich zu anonymisieren.

3.10.5 Fehlende Diversifikation

Mangelnde Diversifikation bedeutet „fehlende Vielfalt“ und kann als „Überspezialisierung“ verstanden werden. Das Problem betrifft speziell CB-Verfahren, die ziemlich oft nur identische Artefakte empfehlen, aber auch andere Verfahren (wie CF) sind anfällig, sehr homogene bzw. ähnliche Artefakte zu empfehlen, die nicht mehr „überraschend“ sind (McNee u. a. 2006). ES fokussieren sich i. d. R. darauf, bei Empfehlungen die Artefakte mit den besten bzw. sichersten vorhergesagten Bewertungen vorzuschlagen. Es konnte in mehreren Studien gezeigt werden, dass die Benutzerzufriedenheit nicht immer mit einer hohen Vorhersagegenauigkeit korreliert (McNee u. a. 2002; Ziegler u. a. 2005). Im CF-Bereich kann sich das Problem auf mehreren Ebenen manifestieren:

Hat ein aktiver Benutzer viele ähnliche Benutzer, von denen die meisten den Themenbereich *A* gut finden, aber zahlenmäßig weniger ähnliche Benutzer, die zu Themenbereich *B* tendieren, so können CF (speziell UB-Verfahren) dazu neigen, vornehmlich Artefakte aus dem Themenbereich *A* zu empfehlen, obwohl für den aktiven Benutzer beide Themenbereiche gleich interessant sind.

Auch das IB-Verfahren ist nicht gegen Überspezialisierung immun. Meist wird ein TOP-*N*-Ranking als Empfehlungsliste generiert. Dabei kann es vorkommen, dass von diesen *N* empfohlenen Artefakten für den aktiven Benutzer alle *N* Artefakte aus der gleichen Reihe (z. B. Star Wars) stammen (vgl. Boim u. a. 2011). Selbst wenn der Benutzer diese Reihe mag, könnte eine größere Vielfalt bei der Empfehlung wünschenswert sein: Nur ein Artefakt stammt aus der Star Wars-Reihe und daneben werden auch andere ähnliche Artefakte, wie z. B. Star Trek oder E. T., vorgeschlagen, um einen breiteren Überblick zu ermöglichen.

Erste Ansätze versuchen, Attribute (z. B. Genre, Schauspieler, ...) der Artefakte bzw. existierende Strukturen für eine Themendiversifikation (Ziegler u. a. 2005) zu nutzen

oder gehen einen etwas anderen Weg und optimieren die Benutzerschnittstelle (Boim u. a. 2011, Empfehlungen werden zweistufig angezeigt: auf der obersten Ebene die Reihen. Erst wenn man diese anklickt, werden die einzelnen zugehörigen Artefakte angezeigt).

3.11 Zusammenfassung

Bei den in diesem Kapitel angesprochenen Verfahren handelt es sich jeweils um deren Reinformen. Abbildung 3.3 zeigt zur Übersicht noch einmal die Datenquellen der vorgestellten Verfahren (links Datenquellen, die vom aktiven Benutzer unabhängig sind; rechts die abhängigen). Es lassen sich leicht die nicht personalisierten Verfahren erkennen, da diese auf nur einer einzigen benutzerunabhängigen Datenquelle basieren.

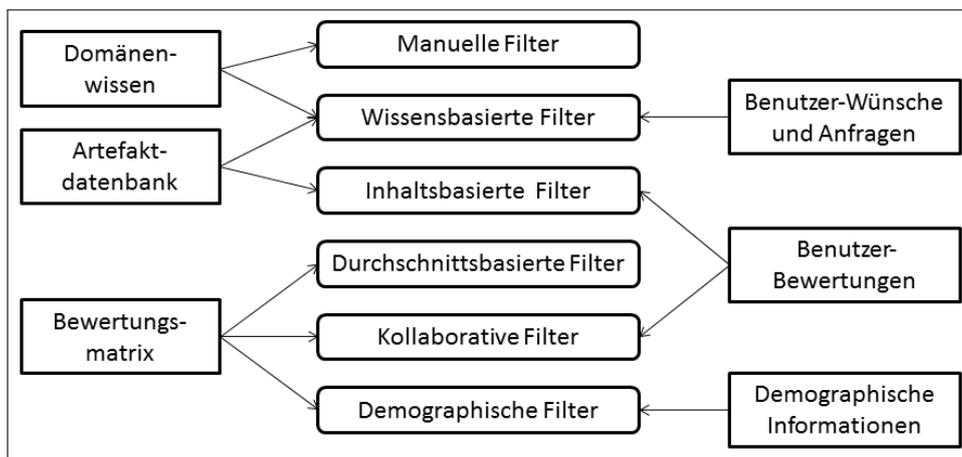


Abbildung 3.3: Übersicht der Datenquellen aller vorgestellten Filterverfahren⁵

Es wurden allgemeine und charakteristische Stärken, mögliche Probleme sowie Lösungsansätze der einzelnen Verfahren angesprochen, die bei einem Einsatz in einer (kleinen) Community von Interesse sein können. Tabelle 3.7 zeigt diese Probleme und Stärken nochmals zusammengefasst. Probleme bzgl. Transparenz betreffen alle Verfahren (auch beim manuellen Filtern sind die Gründe für eine Empfehlung nicht unbedingt klar erkennbar).

Kollaborative Filter bzw. darauf aufbauende ES werden aktuell als Stand der Technik angesehen. Es existieren zwar sehr viele Ähnlichkeitsmaße, jedoch gibt es keine allgemeine Empfehlung, welches Maß man in einem bestimmten Kontext einsetzen sollte (Segaran 2008). Für jeden Anwendungsfall muss eine Auswahl mit Fingerspitzengefühl getroffen werden.

Die Verfahren werden jedoch selten in ihrer Reinform eingesetzt: Die „besten“ Empfehlungssysteme sind Kombinationen aus unterschiedlichen Verfahren, um jeweils deren charakteristische Vorteile zu nutzen und Schwächen auszugleichen (Burke 2002; Segaran 2008). Diese Systeme werden als *hybrid* bezeichnet. Einen guten Überblick über

⁵nach Burke 2007, Fig. 12.1 auf S. 379

mögliche Kombinationen sowie eine Taxonomie für hybride Empfehlungssysteme ist bei Burke (2002) zu finden.

Zu Beginn der Untersuchungen von Empfehlungssystemen standen Verfahren und Genauigkeitsverbesserungen im Mittelpunkt. Aktuell werden in der Forschung vornehmlich die sog. kontext-beachtenden („context-aware“) Filter und Erweiterungen für Diversifikation untersucht. Bei kontext-beachtenden Filtern (Woerndl u. a. 2007) handelt es sich um dynamische Verfahren, die z. B. abhängig vom aktuellen Aufenthaltsort unterschiedliche Empfehlungen generieren und vereinfacht als orthogonal zu den in diesem Kapitel vorgestellten Verfahren angesehen werden können.

Verfahren	Probleme	Stärken
Manuelle Filter	hoher personeller Aufwand; keine oder nur begrenzte Personalisierung	Unabhängigkeit von Bewertungen, textbasierte Empfehlungen
Wissensbasierte Filter	keine automatischen Empfehlungen, Bestimmung der Eigenschaften notwendig	Unabhängigkeit von Bewertungen
Inhaltsbasierte Filter	evtl. aufwändige, einheitliche Annotationen notwendig, relativ statische Empfehlungen nahezu identischer Artefakte, Diversifikation	nicht anfällig für „New Item“-Kaltstartproblem
Durchschnittsbasierte Filter	nicht personalisiert, „New Community“- und „New Item“-Kaltstartproblem, Diversifikation	Empfehlungen für Benutzer mit keinen oder wenigen Bewertungen möglich
Kollaborative Filter	Sparsity- sowie alle Kaltstartprobleme, Diversifikation	Cross-Category-Empfehlungen
Demographische Filter	Datenschutz, Diversifikation	nicht anfällig für „New User“-Kaltstartproblem

Tabelle 3.7: Übersicht über Probleme und Stärken aller vorgestellten Filterverfahren

4 Hybrider Empfehlungssystem-Prototyp für Mobile2Learn

In diesem Kapitel wird exemplarisch die Anforderungsanalyse und Architektur eines Empfehlungssystems für eine kleine Online-Community mit regionalem Bezug am Beispiel von *Mobile2Learn* (vgl. Abschnitt 2.3.2) erläutert, so dass das Vorgehen auch auf Empfehlungssysteme anderer Communities übertragen werden kann. Die Anforderungen werden zusammen mit den Entwurfsentscheidungen vorgestellt, da dies den Prozess, die Auswahl der Lösungsoptionen sowie die Entstehung des allgemeinen Ansatzes am besten widerspiegelt.

4.1 Anforderungen und Entwurfsentscheidungen

Genauso unterschiedlich wie einzelne Online-Communities (vgl. Abschnitt 2.3) sind auch die Anforderungen an Empfehlungssysteme, die in diesem Kontext eingesetzt werden sollen. Für den Prototyp legt die *Mobile2Learn.de* Online-Community einen entscheidenden Teil der Anforderungen mit fest. Das ES soll insgesamt für eine verbesserte Personalisierung für die Mitglieder der Plattform sorgen und möglichst gut auf die Interessen der einzelnen Benutzer eingehen.

Die Kernanforderung an das ES ist das sinnvolle, interessenabhängige Empfehlen von Artefakten (Bildungsinhalte, Spielideen und Galerien) an die Mitglieder der Community (Aufgabe „Find Good Items“ aus Abschnitt 3.1 mit möglichst personalisierten Verfahren).

Speziell für die stark strukturierten Spielideen (durch Anforderungen an z. B. Alter der Kinder, Aufbau der Beschreibung und Art der Spiele) eignen sich wissensbasierte Filter (vgl. Abschnitt 3.7), in denen Eltern schnell passende Spiele finden oder empfohlen bekommen können. Jedoch wurde dieses Verfahren für den Prototyp nicht ausgewählt, da es für die spezielle Zielgruppe besonders wichtig ist, dass Empfehlungen ohne Aufwand für den Benutzer automatisch generiert werden. Darüber hinaus wäre dies nur eine partielle Lösung, die sich nur auf eine bestimmte Domäne (wie die Spiele) anwenden ließe.

Alle anderen Verfahren (außer dem manuellen Filter) basieren auf Bewertungen der Artefakte durch die Benutzer. Im Kontext von *Mobile2Learn* waren jedoch nur sehr wenige explizite Bewertungen (60 bei 151 Artefakten und 175 Benutzern zum Zeitpunkt der Anforderungsanalyse) vorhanden, so dass keines der personalisierten Verfahren ohne weitere Ergänzungen eingesetzt werden kann (nicht einmal durch Zusammenlegen der Artefakte zu Artefaktgruppen oder Boosting). Webbasierte Online-Communities

können auf eine Vielzahl von möglichen Eingabedaten für Empfehlungsalgorithmen zurückgreifen, um daraus z. B. implizite Bewertungen zu generieren:

- Benutzerprofil (in unterschiedlich detaillierten Ausprägungen)
- Referrer (Verweisende Seiten oder Zugriff über Lesezeichen)
- Weiterempfehlungen von Artefakten durch Benutzer
- Meta-Daten oder Strukturen der Artefakte
- Anzahl der Abrufe einzelner Artefakte
- Lese- bzw. Verweildauer
- Kategorien oder Tags (Schlüsselwörter)
- Informationen über (Teilnahmen an) Präsenzveranstaltungen (sofern vorhanden)
- Angeklickte Empfehlungen

Diese Liste erhebt keinen Anspruch auf Vollständigkeit, soll aber einen Überblick über mögliche und in Betracht gezogene Datenquellen geben.

Zum Zeitpunkt der Anforderungsanalyse betrug das Sparsity-Level für die expliziten Bewertungen $1 - \frac{60}{175 \cdot 151} = 0,9977 = 99,77\%$ (175 Benutzer und 151 Artefakte). Durch die Hinzunahme impliziter Bewertungen, basierend auf Lesedauern der Artikel bzw. Verweildauer auf den zugehörigen Seiten, konnte dieser Wert auf $1 - \frac{1438}{175 \cdot 151} = 0,9456 = 94,56\%$ verringert werden, damit es überhaupt möglich wird, (sinnvolle) Empfehlungen zu generieren. Referrer und Weiterempfehlungen haben sich im *Mobile2Learn*-Kontext nicht angeboten, da sich aus beiden auf Grund weniger Daten keine verlässlichen Informationen ableiten ließen. Schafer u. a. (1999) sowie Sollenborn u. Funk (2002) schlugen vor, aus dem Klickverhalten der Benutzer auf Empfehlung weitere implizite Bewertungen herzuleiten und die nicht angeklickten Empfehlungen, die dem Benutzer wahrscheinlich nicht genug gefallen haben, leicht abzuwerten. Dieser Vorschlag wird nicht umgesetzt, da ein Benutzer bereits durch die TOP-N-Darstellung ein Ranking (wie bei einer Suchmaschine) annimmt, somit die hinteren Empfehlungen seltener angeklickt werden und bei jeder TOP-N-Empfehlung bestraft würden (vgl. Herlocker u. a. 2004).

Ein rein inhaltsbasierter Filter (Abschnitt 3.6) wäre zwar in der Lage, automatisch Empfehlungen zu berechnen, weist jedoch Einschränkungen in der Anwendbarkeit auf: Da Mitgliedern auch Galerien empfohlen werden sollen, müssten neben den Galerien auch die Artikel (aufwändig) mit Meta-Daten bzw. Attributen annotiert werden, damit Ähnlichkeiten übergreifend berechnet werden können. Zudem haben inhaltsbasierte Filter das Problem, dass sie oft nur Artefakte empfehlen, die identisch oder sehr ähnlich zu den Artefakten sind, die der Benutzer bereits gut bewertet hat. Auf Grund der stärkeren Struktur der Spiele verschärft sich das gerade genannte Problem ohne die Verwendung besonderer Stopp-Wort-Listen.

Da sich kollaborative Filter im Gegensatz zu demographischen Filtern sehr gut an die Vorlieben eines Benutzers anpassen können, von keinen domänenspezifischen Daten abhängen und besonderer Wert auf unerwartete bzw. Cross-Category-Empfehlungen gelegt wird, stehen diese im Fokus. Das System soll jedoch auch Empfehlungen für hinzukommende Artefakte und für Benutzer generieren können, die sich neu registriert

bzw. bisher keine oder nur sehr wenige Artefakte bewertet haben. Zum Zeitpunkt der Anforderungsanalyse gehörten fast 50 % der Mobile2Learn-Community dieser Gruppe an (siehe Abbildung 4.1). Das bedeutet, dass der Algorithmus die „cold-start“-Probleme für „New Item“ sowie „New User“ (vgl. Abschnitt 3.10) bewältigen muss.

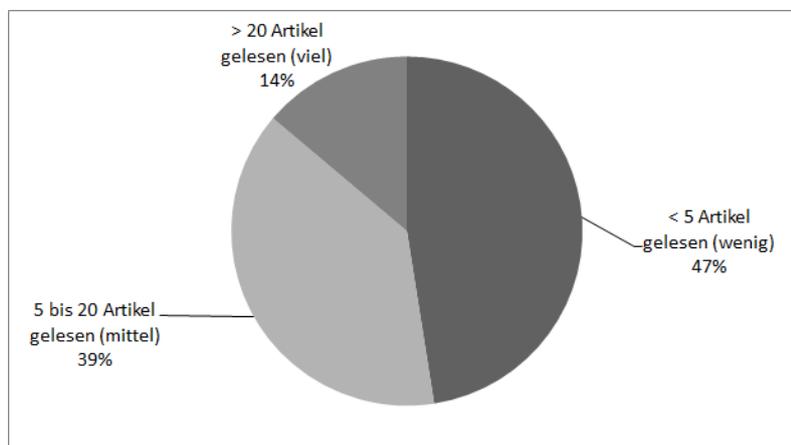


Abbildung 4.1: Artikel-Leseverhalten der Mobile2Learn-Benutzer

Ein Lösungsvorschlag für das „New User“-Kaltstartproblem ist, dem neuen Benutzer eine Auswahl von Artefakten zur Bewertung zu geben, um damit schnell ein persönliches Profil aufzubauen. Jedoch ist diese Methode für die Zielgruppe nicht adäquat, da von den Benutzern erst eine zusätzliche Arbeit verrichtet werden muss, bevor Empfehlungen generiert werden können. Da Benutzer bei der Registrierung Wohnort, Alter, Geschlecht und Kinder angeben müssen bzw. können, liegt es nahe, diese Daten zu nutzen. Somit fiel die Wahl auf demographische Filter (α -Communities, vgl. Abschnitt 3.8) als Ergänzung zum kollaborativen Filter. Die Annahme dafür ist, dass Personen mit gleichen Eigenschaften ähnliche Interessen haben: Eltern, die in ländlichen Gegenden wohnen, haben z. B. gleiche Interessen für Bildungsthemen bzw. Spiele, die sich mit der Natur beschäftigen, im Unterschied zu Eltern, die in eher städtischen Gebieten wohnen und damit z. B. Spielplatz-Spiele bevorzugen. Gleiches gilt für unterschiedliche Altersstrukturen der Eltern und Kinder: Junge Eltern haben sicherlich andere Anforderungen an (z. B. Bewegungs-)Spiele als Großeltern. Familien mit Kindern im Alter von zwei Jahren haben andere Anforderungen an Spiele oder Artikel als Familien mit sechsjährigen Kindern. Neben diesen klassischen Daten für demographische Filter existieren bei *Mobile2Learn*, als regionaler Community mit Präsenzveranstaltungen, weitere charakterisierende Informationen über die Mitglieder, die mit in den Empfehlungsprozess einbezogen werden können: Eltern, die gleiche Veranstaltungen besucht haben, bilden eine eigene Klasse und haben vermutlich ähnliche Interessen an den behandelten Themen.

Das „New Item“-Problem ist im Kontext von *Mobile2Learn* nicht kritisch, da neue Artikel über die Startseite bzw. über Kampagnen speziell angekündigt werden und auch keinem Aktualitätsverlust unterliegen. Jedoch soll dieses Problem über die Nutzung struktureller Informationen weiter abgemildert werden (vgl. Ganesan u. a. 2003). Die Wahl fiel auf Filterbots (vgl. Abschnitt 3.10.1), da diese eine Lösung für verschiedene Teilprobleme darstellen: Sie können Abhängigkeiten zwischen Artefakten (Nähe im

redaktionellen Artikelbaum, Artefakte in den gleichen Kategorien und Zusammenhänge zwischen verschiedenen Artefakttypen, z. B. Galerien und Artikeln, die zu einer gleichen Veranstaltungsreihe gehören) in den kollaborativen Filter integrieren und gleichzeitig das „New Item“-Problem abschwächen. Die Kombination des UB- und IB-Verfahrens kann nicht eingesetzt werden, da die Pseudobewertungen für Kinder und Präsenzveranstaltungsbesuche nicht atomar sind (Eltern können mehrere Kinder haben und mehrere Präsenzveranstaltungen besuchen).

Automatische Empfehlungen sollen auch für Benutzer möglich sein, über die der Online-Plattform fast keine Informationen vorliegen. Durchschnittsbasierte Filter bieten diese Möglichkeit.

Neben den Filtern, die auf Bewertungen basieren, können Empfehlungen auch auf anderen Annahmen und Daten beruhen:

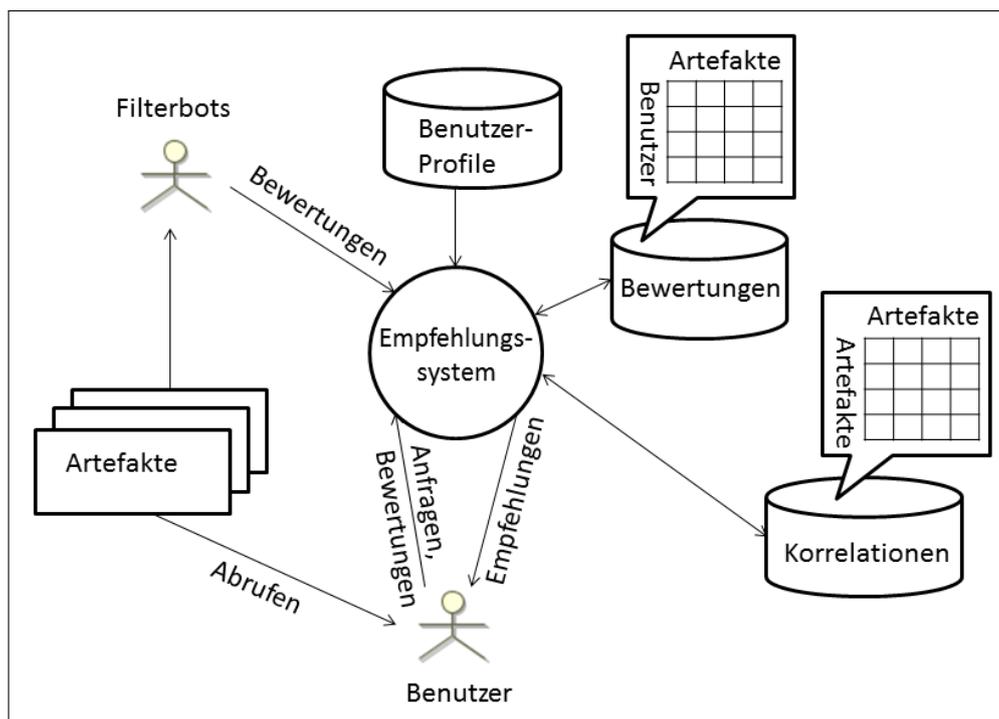
- Mitglieder ohne Abruf eines einzigen Artikels haben vielleicht die Übersicht noch nicht gefunden.
- Benutzer, die bisher keine einzige Präsenzveranstaltung besucht haben, fehlt evtl. die Information, dass regelmäßig Veranstaltungen angeboten werden.
- Benutzer, die eine Präsenzveranstaltung besucht haben, sind vermutlich (stark) daran interessiert, sich die Bilder im Nachhinein anzusehen bzw. sich auf Bildern wiederzuerkennen oder die Themen online nachzulesen.
- In Kürze stattfindende Termine können gezielt Mitgliedern empfohlen werden, die in der Nähe wohnen oder bereits an einem Veranstaltungsort waren, da sie vermutlich eher an diesen Präsenzveranstaltungen teilnehmen werden.

Da es sich hier um die Kombination unterschiedlicher Ansätze und Methoden handelt, wird der Prototyp als hybrid bezeichnet. Die genaue Funktionsweise des ES wird im nächsten Abschnitt erläutert. Allgemein handelt es sich um einen Ansatz, der über mehrere Stufen versucht, möglichst personalisierte Empfehlungen unter dem Einsatz vieler domänen-spezifischer Daten und Annahmen zu generieren.

Darüber hinaus ist es entscheidend, wie den Benutzern die Ergebnisse präsentiert werden. Die Verbreitung von Empfehlungen ist über verschiedene Kanäle möglich (Webseite, SMS, E-Mail), hierbei ist vorgesehen, hauptsächlich TOP-*N*-Rankings einzusetzen. Da auf den Artefaktseiten im unteren Bereich bereits viel Platz für die Weiterempfehlungs-, Bewertungs- und Kommentarfunktionen belegt wird, sollen dort im Prototyp keine „Ähnlichen Artikel“ angezeigt werden. Die genaue Umsetzung und weitere Details folgen im Abschnitt 4.4.

4.2 Architektur- und Algorithmusbeschreibung

Die Gesamtarchitektur ist in Abbildung 4.2 dargestellt. An zentraler Position befindet sich der Kern des Empfehlungssystems. Dieser greift dabei auf die Bewertungen, Benutzerprofile und Korrelationen bzw. Ähnlichkeiten der Artefakte zurück und verwaltet sie. Bewertungen stammen von den Benutzern und Filterbots des Systems. Abhängig davon werden für Benutzer Empfehlungen möglichst personalisiert generiert.

Abbildung 4.2: Gesamtarchitektur des hybriden Empfehlungssystems¹

Aus architektonischer Sicht besteht das Empfehlungssystem aus zwei aufeinander aufbauenden Komponenten. Die erste Komponente ist für vorbereitende Schritte und die Modellbildung des elementbasierten kollaborativen Filters zuständig. Die zweite Komponente berechnet schließlich die konkreten Empfehlungen für einen aktiven Benutzer.

4.2.1 Vorbereitende Schritte und Modellbildung

Die Modellbildung verläuft in vier Schritten (vgl. Abbildung 4.3): Im ersten Schritt wird die Verweildauer-Analyse mitsamt Normalisierung durchgeführt. Der zweite Schritt besteht aus der Transformation der normalisierten Verweildauern in Bewertungen. Im dritten Schritt werden die Filterbots ausgeführt. Der letzte Schritt berechnet schließlich das Modell (Ähnlichkeitsmatrix der Artefakte) für den elementbasierten kollaborativen Filter.

Für die Vorbereitung der Verweildauer-Analyse (vgl. Abbildung 4.4) aller angebotenen Artikel, die auf den Log-Files des HTTP-Servers basiert, wird eine Tabelle für eine Abbildung der Uniform Resource Locator (URL) auf die internen Artikel-IDs aufgebaut. Diese wird benötigt, um die Artikel unter den abgerufenen Ressourcen zu erkennen und zuordnen zu können.

Die Verweildauer-Analyse wird ausschließlich für Artikel durchgeführt – nicht für die Foto-Galerien. Bei den Fotos hat sich bei Testläufen gezeigt, dass sich implizite

¹nach Sarwar u. a. 1998, Fig. 1

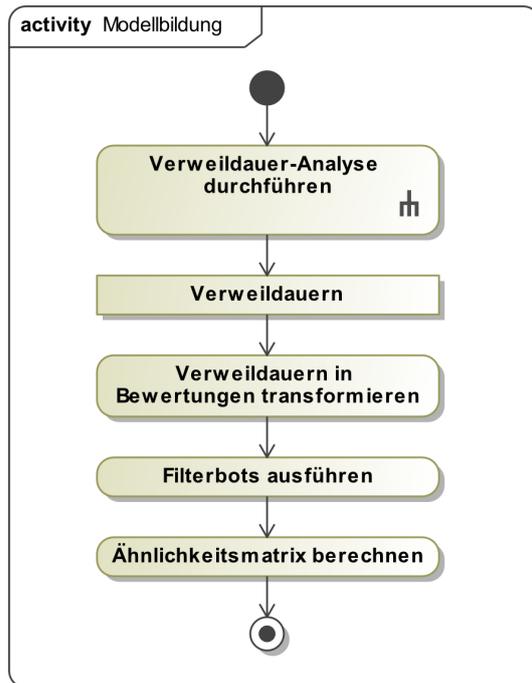


Abbildung 4.3: Vorgehen bei der Modellbildung

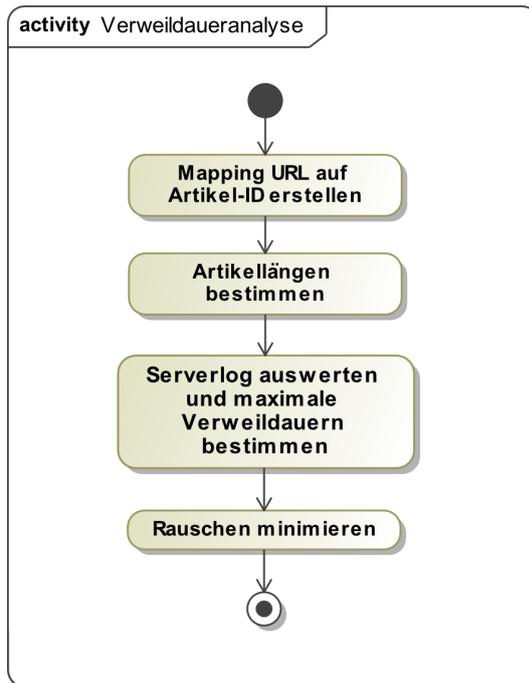


Abbildung 4.4: Aktivität: Verweildauer-Analyse durchführen

Bewertungen nicht sinnvoll aus der Betrachtungsdauer von Bildern generieren lassen. Ein Grund dafür ist, dass gleiche Bilder über mehrere URLs (unterschiedliche Views) abrufbar sind und Bilder in der Dia-Show automatisch nach einem festen Zeitintervall (standardmäßig zehn Sekunden) gewechselt werden.

Verweildauern wurden von Morita u. Shinoda (1994) sowie Konstan u. a. (1997) unabhängig von den Artikellängen betrachtet. Morita u. Shinoda (1994) haben sogar nachgewiesen, dass die Verweildauern in ihrem Kontext unabhängig von der Artikellänge sind. Dies scheint in diesem Umfeld nicht zu stimmen, da die Artikel nicht homogen genug sind: Ohne eine Normalisierung der Verweildauern mit der Artikellänge werden längere Artikel (z. B. Erziehungstipps) deutlich bevorzugt vorgeschlagen. Daher wird eine Normalisierung der Verweildauern mit der Anzahl der Wörter eines Artikels durchgeführt (die Anzahl der Wörter korreliert mit der Länge eines Artikels). Zur Bestimmung der Anzahl der Wörter werden alle Artikel normalisiert, d. h. HyperText Markup Language (HTML)-Elemente, -Kommentare, Steuerzeichen sowie Satzzeichen entfernt, und dann die Anzahl der verbliebenen Wörter gezählt. Da diese Informationen lediglich für die Normalisierung der Verweildauern benötigt werden, kann auf aufwändige NLP-Verfahren, wie Stemming und Entfernung von Stop-Wörtern, verzichtet werden. Bevor die Ermittlung der Verweildauern beginnt, wird die Anzahl der Wörter für alle Artikel berechnet und zwischengespeichert.

Die Bestimmung der Verweildauern selbst basiert auf einer sequentiellen Analyse der Zugriffslog-Datei des Web-Servers. Jede Anfrage an den Webserver erzeugt eine eigene Zeile mit der folgenden Struktur (Combined Log Format² des Apache HTTP-Servers):

²<http://httpd.apache.org/docs/2.2/logs.html#accesslog>, abgerufen am 20. August 2011

```
{IP-Adresse} {RFC 1413 Identität} {Benutzername} {Zeitstempel des Zugriffs}
"{HTTP-Methode} {Ressource} HTTP/{HTTP-Version}" {HTTP-Antwort-Code}
{Dateigröße} {Referrer} {User-Agent}
```

Beispiel für einen Log-Eintrag:

```
139.174.242.81 - sstrickroth [20/Feb/2011:17:38:28 +0100]
"GET /bildungsinhalte.html HTTP/1.1" 200 6037 "http://www.mobile2learn.de/"
"Mozilla/5.0 (Windows; U; Windows NT 5.1; de; rv:1.9.2.13) Gecko/20101203 Firefox/3.6.13"
```

Interessant für die Ermittlung einer Verweildauer sind der *Benutzername*, der *Zeitstempel des Zugriffs* und die angeforderte *Ressource*. Da HTTP zustandslos ist, müssen zusammenhängende Anfragen, Sitzung genannt, meist aufwändig ermittelt werden. Da die Verweildauern jeweils für einen bekannten Benutzer berechnet werden sollen, ist die Sitzungserkennung einfacher und der *Benutzername* zusammen mit dem *Zeitstempel des Zugriffs* ausreichend. Für die Ermittlung der Verweildauer eines Artikels wird die Zeit von der Abfrage bis zum Aufruf einer weiteren Seite (z. B. weiterer Artikel, Galerie oder Log-out) innerhalb einer Sitzung betrachtet; alle anderen Zugriffe, ausgelöst durch Asynchronous JavaScript and XML (AJAX) oder Anfragen auf z. B. Bilder, Cascading Style Sheets (CSS)- und JavaScript-Dateien, werden ignoriert. Handelt es sich bei einem Artikel um die letzte Anfrage einer Sitzung, so kann die Verweildauer für diesen Artikel (auf diese Weise) nicht bestimmt werden. Ermittelt wird jeweils die maximale Verweildauer eines Benutzers pro Artikel, da Artikel i. d. R. nur einmal komplett gelesen werden und somit z. B. ein Mittelwert verschiedener Verweildauern eines Artikels nicht sinnvoll ist. Für eine Sitzung wurde ein Timeout von 30 Minuten festgelegt, wobei für die Betrachtung eines Artikels eine Untergrenze von drei Sekunden sowie eine Obergrenze von zehn Minuten festgelegt wurde, um das Rauschen in den Daten zu verringern (vgl. Abschnitt 3.2). Neben den festen Grenzen für die Berücksichtigung von Verweildauern wurde ein dynamisches nutzerabhängiges Limit implementiert: Nach dem Bestimmen der Verweildauern und einer Normalisierung (Division durch die Anzahl der Wörter in einem Artikel) wird für jeden Benutzer die Populationsstandardabweichung bestimmt. Es werden im Folgenden nur Daten berücksichtigt, die nicht mehr als die doppelte Standardabweichung vom Durchschnitt abweichen, um Ausreißer zu entfernen. Da die impliziten Verweildauern auf die gleiche fünfstufige Bewertungsskala wie die expliziten Bewertungen transformiert werden müssen, reicht es nicht aus, eine binäre Wertung „lange gelesen“ / „nur kurz gelesen“ bzw. „gefällt“ / „gefällt nicht“ zu benutzen. Stattdessen wird erneut die Populationsstandardabweichung der Bewertungen eines Nutzers zugrunde gelegt, um die normalisierten Verweildauern auf die Bewertungsskala zu transformieren (vgl. Tabelle 4.1 für die Transformationsbeschreibung):

Bewertung	Intervall der normalisierten Verweildauern um den Mittelwert
1	$(-\infty; -1,64\sigma]$
2	$(-1,64\sigma; -\sigma]$
3	$(-\sigma; \sigma)$
4	$[\sigma; 1,64\sigma)$
5	$[1,64\sigma; \infty)$

Tabelle 4.1: Transformationsvorschrift der normalisierten Verweildauern auf die Bewertungsskala

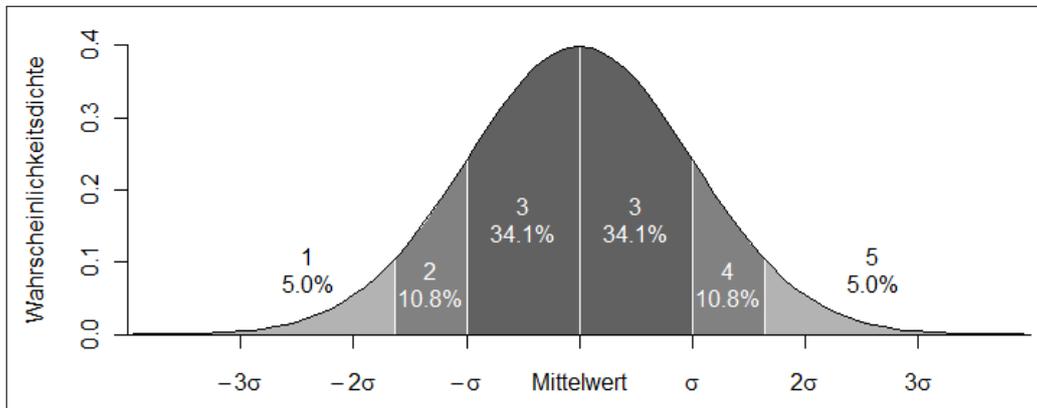


Abbildung 4.5: Wahrscheinlichkeitsdichte mit Bewertungssegmenten und relativer Segmentgröße

Durch diese Transformation (vgl. Tabelle 4.1) wird bei einer Normalverteilung der Daten erreicht, dass ca. 68 % der normalisierten Verweildauern mit 3, jeweils ca. 5 % mit 1 bzw. 5 und die restlichen ca. 22 % der Verweildauern mit 2 bzw. 4 bewertet werden (vgl. Abbildung 4.5). Dies entspricht der Annahme, dass Artikel, die besonders lange gelesen werden, eine hohe Bewertung, Artikel, die nur sehr kurz gelesen bzw. abgebrochen wurden, eine niedrige Bewertung und der Großteil der Artikel, die durchschnittlich lange gelesen werden, eine Bewertung von 3 erhalten.

Hu u. a. (2008) nutzten lediglich nicht-negative implizite Bewertungen (negativ im Sinne von schlechten Bewertungen). Dies wurde damit begründet, dass man aus den Daten, falls jemand eine TV-Serie nicht gesehen hat, kein negatives Feedback erhält, sondern nur hergeleitet werden kann, was ein Benutzer wahrscheinlich mag. Im vorliegenden Prototyp werden aus den Daten auch negative implizite Bewertungen generiert und für die Empfehlungen verwendet, da in diesem Kontext sehr wohl Aussagen über Ablehnung erkennbar sind, wenn die *maximale* Verweildauer einer Seite sehr gering ist: Interessiert einen Benutzer ein Artikel nicht, so wird er sehr wahrscheinlich nie bis zum Ende gelesen, sondern frühzeitig „weggeklickt“. Zudem wirkten sich ausschließlich nicht negativ implizite Bewertungen übermäßig bekräftigend auf einzelne Artefakte (speziell die Erziehungstipps) bei Empfehlungen aus, so dass diese überproportional oft empfohlen wurden (bzw. die Empfehlungen in den Testläufen unnatürlich wirkten).

Damit die potentiell unsicheren impliziten Bewertungen nicht gleichberechtigt mit den expliziten Bewertungen der Benutzer in die Berechnungen eingehen, werden die impliziten mit dem Faktor 0,85 abgewertet.

Danach werden die Filterbots (vgl. Abschnitt 3.10.1) ausgeführt, um die Bewertungen basierend auf dem Inhaltsbaum bzw. den Kategorien zu berechnen. Damit werden strukturelle Informationen in den kollaborativen Filter integriert. Bei elementbasierten Filtern sorgen die Filterbots lediglich für eine Vergrößerung der Nachbarschaft (der ähnlichen Artikel), daher sind die konkreten Bewertungen, die die Filterbots vergeben, unerheblich und können willkürlich gewählt werden (Park u. a. 2006).

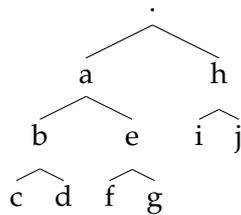


Abbildung 4.6: Generalisierter (Artikel-)Baum

Für den (Artikel-)Baum in Abbildung 4.6 werden vier Filterbots benötigt, die jeweils einen Teilbaum mit einer Tiefensuche traversieren und als eigenständige Pseudo-Benutzer Bewertungen für die Knoten/Artefakte abgeben. In jedem Teilbaum mit mindestens einem Kind wird eine neue Tiefensuche gestartet, die allen Kindknoten die Bewertung 5 gibt. Der erste Filterbot traversiert den im Knoten a beginnenden Teilbaum und bewertet alle Kindknoten. Bei den Knoten b und e werden zwei neue Filterbots erzeugt, die analog vorgehen – ebenso für den Knoten h . Insgesamt ergeben sich folgende Bewertungsvektoren für die Artefakte/Knoten (für die Lesbarkeit transponiert):

Filterbots:	1	2	3	4
$r_a =$	5			
$r_b =$	5	5		
$r_c =$	5	5		
$r_d =$	5	5		
$r_e =$	5		5	
$r_f =$	5		5	
$r_g =$	5		5	
$r_h =$				5
$r_i =$				5
$r_j =$				5

Tabelle 4.2: Filterbotbewertungen: Bewertungsvektoren der Artefakte „a“ bis „j“

Durch dieses Schema ist sichergestellt, dass thematisch verwandte Artikel (also des gleichen Teilbaums) eine höhere Ähnlichkeit haben (z. B. $s(r_b, r_c) = 1$) als un- bzw. weniger verwandte (z. B. $s(r_e, r_i) = s(r_a, r_i) = 0$). Die anderen Ähnlichkeiten sind abhängig von der Berechnungsart und vom Ähnlichkeitsmaß; inspiriert durch Ganesan u. a. 2003, Abschnitt 3.1). Auf analoge Weise vergeben weitere Filterbots, je ein Filterbot pro Kategorie bzw. Veranstaltungsreihe, Bewertungen für Artefakte, die zur gleichen Kategorie („Spiele“, „Spiele 0-3“, „Spiele 4-6“, „Erziehungstipps“, ...) oder Reihe gehören.

Im Fall des IB-Vorgehens kann es hier bei genauerer Betrachtung zu einem Problem kommen: Angelehnt an den Vorschlag von Ganesan u. a. sollten z. B. folgende Ähnlichkeiten bei Abwesenheit weiterer Bewertungen gelten (mit der angepassten Cosinus-Korrelation): $s(r_b, r_e) = \frac{1}{3}$ und $s(r_a, r_b) = s(r_a, r_e) = \frac{1}{2}$. Jedoch werden nur Komponenten der jeweils zwei zu vergleichenden Bewertungsvektoren betrachtet, die in beiden Vektoren besetzt sind. Folglich gilt: $s(r_b, r_e) = s(r_a, r_b) = s(r_a, r_e) = 1$. Daher wurden zwei weitere Implementierungen getestet und evaluiert, die die Ähnlichkeiten der Filterbot-Bewertungen separat behandeln. Genaueres zu den anderen beiden Varianten folgt im

Abschnitt 4.5.2, da sie nicht im Prototyp eingesetzt werden. Der problematische Fall tritt auf Grund der Verwendung der Verweildauern nur sehr kurz bzw. eher selten auf. Zudem gehen die Strukturinformationen über die Anzahl der Komponenten aller Filterbotbewertungen mit gleichem Wert im Verhältnis zu anderen Bewertungen mit in die Ähnlichkeit ein.

Im letzten Schritt wird schließlich die Ähnlichkeitsmatrix der Artefakte ermittelt. Da nur sehr wenige Bewertungen für einzelne Bilder vorliegen, werden diese durch Berechnung des arithmetischen Mittels für die entsprechende Galerie, die allein nicht bewertet und als Übersichtsseite angesehen werden kann, zusammengefasst. Für die Berechnung der Ähnlichkeitsmatrix wird auf die angepasste Cosinus-Korrelation zurückgegriffen (vgl. Abschnitt 3.3.6). Es werden ausschließlich Ähnlichkeiten von mindestens 1 % ($s(\cdot, \cdot) \geq 0,01$) gespeichert.

In der Fachliteratur gibt es keine Anmerkungen zum Umgang mit Bewertungsvektoren der Länge eins. Vermutlich ist dies darin begründet, dass derartige Vektoren bei großen Systemen fast nicht vorkommen. Bei kleinen Online-Communities treten diese jedoch häufiger auf. Vektoren der Länge eins haben bei verschiedenen Ähnlichkeitsmaßen (vgl. Abschnitt 3.3) unterschiedliche Auswirkungen: Bei der Pearson-Korrelation haben sie eine Ähnlichkeit von 0 und bei den cosinus-basierten Ähnlichkeitsmaßen immer eine Ähnlichkeit von 1. Im Abschnitt 4.5.2 wurden die beiden Möglichkeiten, Bewertungsvektoren der Länge eins einzubeziehen bzw. zu ignorieren, evaluiert: Der Prototyp beinhaltet Ähnlichkeiten von Artefakten, die auf Bewertungsvektoren der Länge eins basieren, da diese im vorliegenden Kontext besser geeignet sind.

Damit das Modell und die impliziten Bewertungen stets in einem aktuellen Zustand vorliegen, muss die Modellbildung regelmäßig durchgeführt werden. Im Prototyp wird das Modell täglich aktualisiert.

4.2.2 Berechnung von Empfehlungen

Der allgemeine Ansatz für die Generierung von Empfehlungen besteht darin, verschiedene Stufen der Reihe nach zu durchlaufen (vgl. Abbildung 4.7), deren Personalisierungsgrade abnehmen: Jede Stufe generiert Empfehlungen inkl. berechneter Bewertung und fügt diese zu einer Kandidatenliste hinzu. Sind nach einer Stufe weniger als fünf Kandidaten in der Liste vorhanden, so wird die nächste Stufe durchlaufen. Jede Stufe fügt ausschließlich Artefakte als Kandidaten hinzu, die vom aktiven Benutzer bis zur Empfehlungsgenerierung noch nicht abgerufen wurden.

Enthält die Kandidatenliste mehr als fünf Einträge bzw. wurden alle Stufen durchlaufen, erfolgt die finale Berechnung der Empfehlungsliste, die schließlich als „TOP-5-Ranking“, also als geordnete Liste mit maximal fünf Einträgen als Empfehlungen, angesehen werden kann. Wie bei vielen ES üblich (Marmanis u. Babenko 2009), wird der Schwellenwert $t_{\min. \text{Bewertung}}$ für den Einzug in die Kandidatenliste sowie das finale Ranking in etwa auf die Mitte der Bewertungsskala ($\frac{1}{2} \cdot 5 \approx 3 = t_{\min. \text{Bewertung}}$) festgelegt, so dass einem Benutzer ausschließlich Artefakte empfohlen werden, die ihm mindestens durchschnittlich gut gefallen.

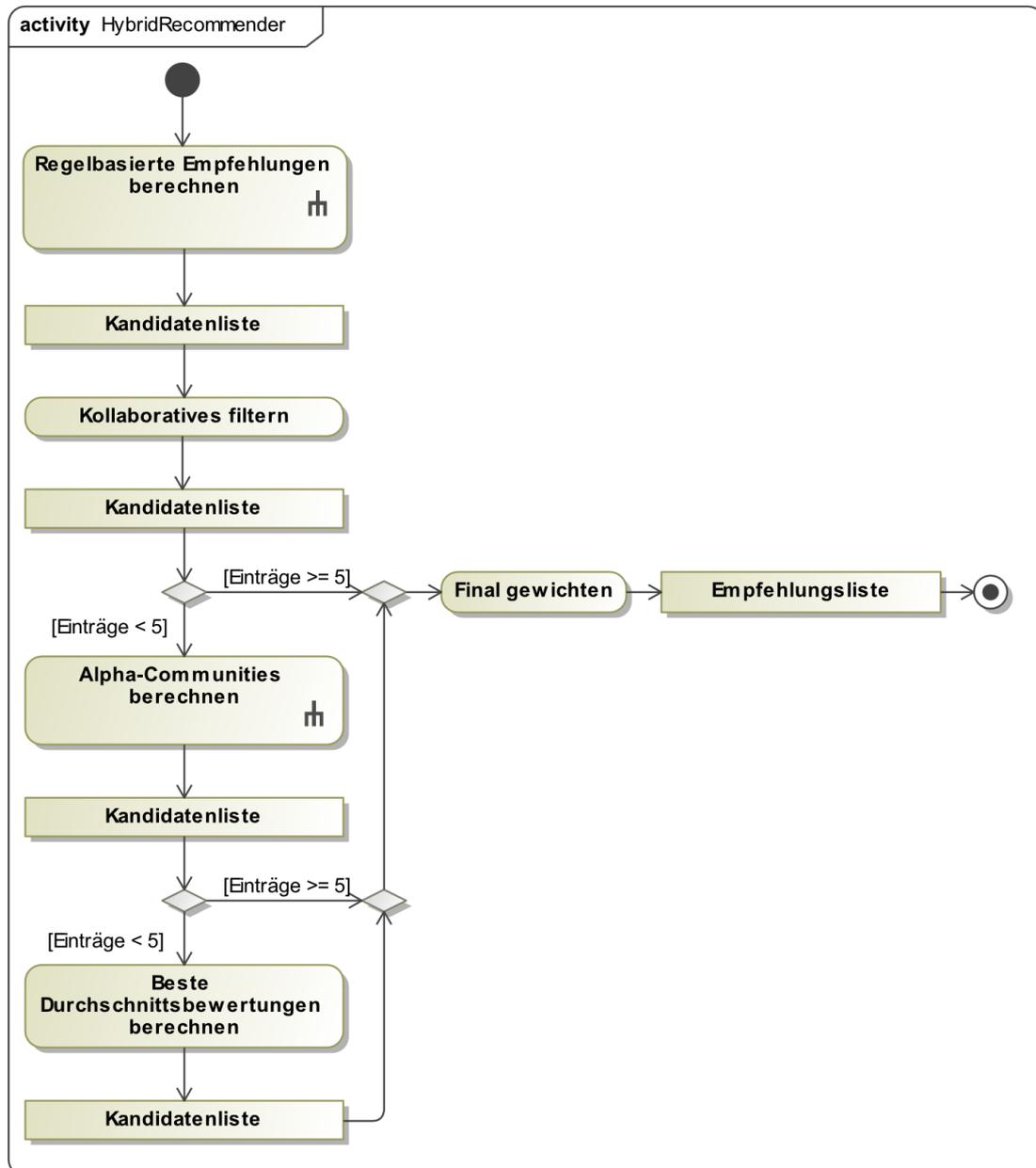


Abbildung 4.7: Architektur des hybriden Empfehlungssystems

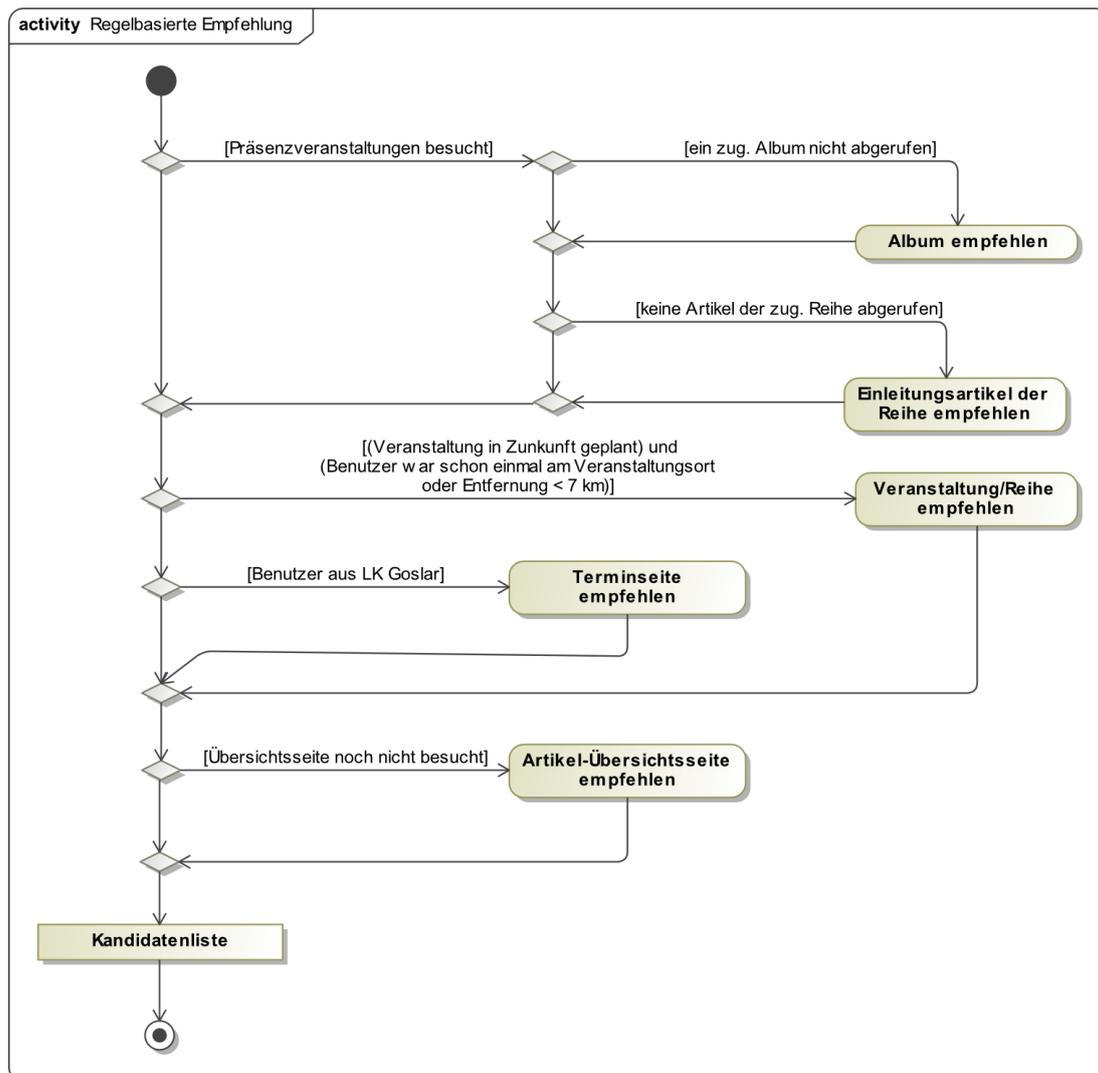


Abbildung 4.8: Aktivität: Regelbasierte Empfehlungen berechnen

Die erste Stufe, *regelbasierte Empfehlungen*, bestimmt Empfehlungen anhand fest vorgelegter Regeln (vgl. Abbildung 4.8). Diese Stufe ist in drei Abschnitte unterteilt, die einzeln ab- und hinzuschaltbar sind: Empfehlungen abhängig von besuchten Präsenzveranstaltungen, für kommende Veranstaltungen und zur Steigerung der Übersicht über alle Inhalte.

Hat der aktive Benutzer eine Präsenzveranstaltung besucht und sich das zugehörige Galerie-Album nicht angesehen oder keinen Artikel aus dieser Reihe gelesen, so wird das Album bzw. der Einleitungsartikel der Reihe in die Kandidatenliste eingefügt. Diese Regel wurde implementiert, da Besucher, die eine Veranstaltung besucht haben, wahrscheinlich ein höheres Interesse haben werden, diese Artikel abzurufen bzw. sich die Bilder, auf denen sie selbst abgebildet sein könnten, anzusehen. Die Bewertung p ergibt sich aus der Zeit wie lange die Veranstaltung zurück liegt (exponentielle Abnahme, $p = 5 \cdot \exp\left(\frac{-t}{h}\right)$ mit der Halbwertszeit h und dem Alter t in Tagen), so dass diese Empfehlungen lediglich eine begrenzte Zeit angezeigt werden. Die Halbwertszeit h wurde auf 810 Tage festgelegt. Dadurch ist für den Prototyp sichergestellt, dass Artefakte von besuchten Veranstaltungen innerhalb der letzten 413 Tage ($5 \cdot \exp\left(\frac{-t}{810}\right) = t_{\min. \text{Bewertung}} = 3 \Rightarrow t \approx 413$) noch in der TOP-5-Empfehlung auftauchen können.

Abhängig von Wohn- und Veranstaltungsort bzw. bereits besuchten Veranstaltungsorten werden dem aktiven Benutzer Empfehlungen für kommende Veranstaltungen gegeben. Wohnt der aktive Benutzer weniger als 7 km (durchschnittliche Entfernung der an einer Präsenzveranstaltung teilnehmenden Mitglieder, plus Varianz gemittelt über alle Präsenzveranstaltungen) vom Ort einer geplanten Veranstaltung entfernt oder war er bereits an einem Veranstaltungsort, so wird eine Empfehlung für diese Reihe mit dem Text „Mobile2Learn-Veranstaltung zum Thema X ab dem x.y.z (wieder) in Ihrer Nähe“ mit einer Bewertung von 5 in die Kandidatenliste eingefügt. Abhängig davon, ob der Benutzer an einem Veranstaltungsort schon einmal war, wird das Wort „wieder“ eingefügt.

Überlegungen, Veranstaltungen auch über einen kollaborativen Ansatz zu empfehlen, scheiterten an einer mangelnden Struktur: Die Veranstaltungsreihen beschäftigen sich mit wechselnden, nicht aufeinander aufbauenden Themen und werden nur ein einziges Mal angeboten. Bei den Veranstaltungen verhält es sich ähnlich. Diese finden innerhalb einer Reihe in kurzen Abständen hintereinander und zudem noch an wechselnden Orten statt. Des Weiteren ist es nicht möglich, über inhaltliche Vorlieben einzelne Reihen bzw. Veranstaltungen zu empfehlen, da die Inhalte der einzelnen Reihen erst nach deren Abschluss über die Online-Plattform verfügbar gemacht werden. Das Belegen der einzelnen Veranstaltungsreihen mit Keywords sowie die Verwendung eines inhaltsbasierten Filters wurde nach Rückfrage mit dem Projektpartner ausgeschlossen. Es wurde als zu aufwändig angesehen, einheitliche und prägnante Schlüsselwörter zu vergeben. Werden die Schlüsselwörter nicht mit größter Sorgfalt zugeordnet, kann es bei den ähnlichen Themen zu großen (im Falle zu ähnlicher Keywords) bzw. zu wenigen (im Falle uneinheitlicher Keywords) Überlappungen der Schlüsselwörter führen.

Hat der aktive Benutzer noch keinen einzigen Artikel gelesen, so wird ihm die redaktionell zusammengestellte Übersicht über alle Bildungsinhalte mit einer Bewertung von 5 vorgeschlagen, so dass er sich einfach eine Übersicht über alle Bildungsinhalte verschaffen kann. Gleiches gilt für die Terminseite, wobei diese nur angezeigt wird, sofern der aktive Benutzer keine direkten Terminempfehlungen erhalten hat.

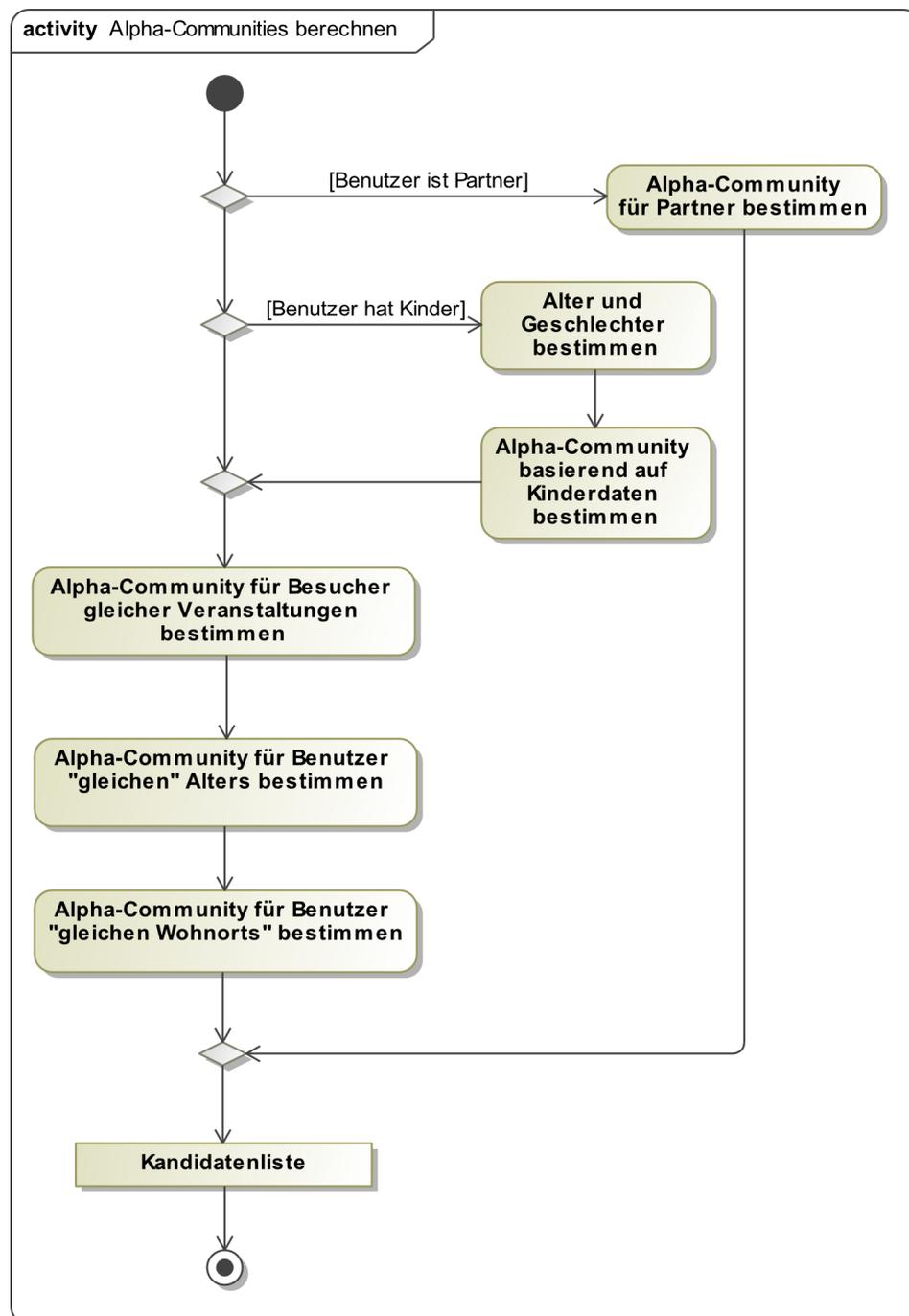


Abbildung 4.9: Aktivität: α -Communities berechnen

Die zweite Stufe versucht, für den aktiven Benutzer eine Empfehlung mittels kollaborativem Filter (elementbasiert) zu berechnen. Dabei wird keine Beschränkung der Nachbarschaftsgröße (d. h. es gilt $k = m$) vorgenommen, die oftmals bei größeren Systemen Anwendung findet, um die Laufzeit zu verbessern, ohne die Qualität merklich zu beeinflussen. Bei einem System mit ca. 160 Artefakten ist dies nicht nötig. Zudem wird keine direkte Mindestanzahl von Bewertungen vom aktiven Benutzer vorausgesetzt, sondern auf eine Signifikanzgewichtung gesetzt (Herlocker u. a. 1999; viele Systeme, wie MovieLens³, setzen voraus, dass der aktive Benutzer mindestens eine gewisse Anzahl Bewertungen (fünfzehn bei MovieLens) vorgenommen hat, bevor überhaupt Empfehlungen generiert werden. So wird vermieden, dass Empfehlungen auf zu lückenhaften Benutzerprofilen basieren). Signifikanzgewichtung bedeutet, dass in der Vorhersage-Berechnungsvorschrift (vgl. Abschnitt 3.9.2) eine Gewichtung abhängig von der Anzahl der ähnlichen Artefakte vorgenommen wird: Der Gewichtungsfaktor $devalue$ berechnet sich wie folgt: Gibt es in der Nachbarschaft \mathcal{I}' des Artefakts i weniger als $devalue$ Artefakte, so wird die Vorhersage mit dem Faktor $\frac{|\mathcal{I}'|}{devalue}$ abgewertet. Durch eine entsprechende Wahl von $devalue$ (\geq maximal mögliche Bewertung) kann indirekt eine Mindestanzahl von Bewertungen des aktiven Benutzers etabliert werden.

Sofern die Kandidatenliste weniger als fünf Einträge enthält, wird versucht über α -Communities (vgl. Abschnitt 3.8) personalisierte Empfehlungen zu berechnen. Dabei wird zwischen Netzwerkpartnern und Eltern unterschieden (vgl. Abbildung 4.9). Netzwerkpartner haben ihre eigene α -Community, für Eltern werden mehrere α -Communities für eine Empfehlung herangezogen: Hat ein aktiver Benutzer Kinder im Profil angegeben, so wird das minimale und maximale Alter der Kinder bestimmt und eine α -Community für Eltern mit Kindern im gleichen Intervall berechnet. Sofern alle Kinder das gleiche Geschlecht haben, wird auch hierauf eine α -Community berechnet. Für alle „normalen“ Benutzer werden noch drei weitere α -Communities herangezogen: Benutzer, die die gleichen Präsenzveranstaltungen besucht haben, etwa gleich alt sind (Intervall von 6 Jahren um das Alter des aktiven Benutzers) und nicht weiter als 7 km voneinander entfernt wohnen.

Sollte die Kandidatenliste weiterhin weniger als fünf Einträge enthalten, wird schließlich auf das unpersonalisierte durchschnittsbasierte Empfehlen (vgl. Abschnitt 3.5) auf Artefakten mit mindestens neun Bewertungen (*agreement*) zurückgegriffen.

Das *finale Ranking* fasst schließlich die Einträge der Kandidatenliste zusammen. Es ist möglich, dass ein Artefakt (Artikel, Galerie Album, ...) von mehreren Stufen als Kandidat vorgeschlagen wird. Das *finale Ranking* bestimmt mittels einer Linearkombination aus Gewichtungsfaktoren und den berechneten Bewertungen jedes Artefakts die finale Bewertung, die schließlich die Priorität bzw. Position im TOP-5-Ranking festlegt. Damit fällt dieses ES, gemäß der Taxonomie von Burke (2002), in die Kategorie „gewichtet“. Damit den einzelnen Stufen unterschiedliches Gewicht entsprechend des Personalisierungsgrades gegeben werden kann, ist folgender Schritt notwendig: Die Empfehlungen des kollaborativen Filterns haben höhere Wertigkeit (1,1) als die der α -Communities (1,05) und diese wiederum als die Empfehlungen der durchschnittlich am besten bewerteten Artefakte (1):

³Online-Filmempfehlungssystem <http://www.movielens.umn.edu>, Konstan u. a. (1998)

$$p_i = 1,1 \cdot p_{i,\text{kollaborativ}} + 1,05 \cdot p_{i,\alpha\text{-communities}} + p_{i,\text{durchschnittsbasiert}}$$

Da ein Großteil der Artikel für eine bestimmte Altersgruppe der Kinder vorgesehen ist, wird darüber hinaus versucht, die Kandidatenliste weiter zu personalisieren, selbst wenn der aktive Benutzer keine Angaben zu Kindern im Profil hinterlegt hat: Dazu werden die Artikel betrachtet, die der aktive Benutzer bisher gelesen hat; sind dies mehr als drei und hat er mehr als 75 % der altersabhängigen Artikel für die Kategorie „bis 3 Jahre“ gelesen, werden Artikel der Kategorie „4 bis 6 Jahre“ abgewertet. Analog wird dies für den komplementären Fall durchgeführt. Beinhaltet das Profil des aktiven Benutzers Angaben zu Kindern, werden diese mit dem genannten Verfahren kombiniert, so dass auch auf Veränderungen reagiert werden kann, falls ein Benutzer ein weiteres Kind bekommt und dies nicht im Profil hinterlegt. Eine Auf- oder Abwertung von Artefakten, die einer Veranstaltungsreihe zugeordnet sind, die der aktive Benutzer besucht bzw. nicht besucht hat, erfolgt explizit nicht, da Eltern auch auf neue oder weitere für sie interessante Inhalte hingewiesen werden sollen. Die Mobile2learn.de-Nutzungsstatistik zeigt, dass Eltern nicht ausschließlich Artikel von besuchten bzw. nicht besuchten Präsenzveranstaltungen lesen (jeweils ca. 40 % der Präsenzveranstaltungsbesucher mit Online-Zugang haben mindestens einen Artikel aus einer besuchten bzw. nicht besuchten Reihe gelesen; jeweils ca. 16 % haben ausschließlich Artikel aus einer besuchten bzw. nicht besuchten Reihe abgerufen).

Veranstaltungsort	Lernen mit allen Sinnen		Alte Kinderlieder neu entdecken		Kinder fördern mit Musik		Naturerleben - Bauernhof		Kinderrohren hören - Kindermünder erzählen	
	Ø	#	Ø	#	Ø	#	Ø	#	Ø	#
Kindergarten Pustebume, Dörnten, Liebenburg	3.42	3	3.27	2	5.00	1	5.00	1	n/a	n/a
Kindergarten Ohlhof, Goslar	3.25	3	3.00	1	5.00	1	n/a	n/a	n/a	n/a
Kindergarten Hahndorf, Goslar	3.37	2	3.00	1	n/a	n/a	n/a	n/a	n/a	n/a
Kita Lilliputt, Goslar	3.41	5	3.13	5	n/a	n/a	3.00	1	n/a	n/a
Kita 'Am Spottberg', Seesen	3.67	1	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Mensa TU Clausthal, Clausthal-Zellerfeld	3.08	4	3.29	3	n/a	n/a	3.00	1	n/a	n/a
Ev. Kita 'Zum Frankenberge', Goslar	3.25	2	4.00	1	3.00	1	n/a	n/a	n/a	n/a
Kath. Kiga St. Jacobi, Goslar	3.13	3	3.50	3	5.00	1	n/a	n/a	n/a	n/a
Bauernhof Broihan, Vienenburg	3.11	8	3.32	7	3.80	4	4.33	3	n/a	n/a
Reiterhof Wetzel, Liebenburg-Dörnten	3.33	5	3.00	4	n/a	n/a	5.00	1	n/a	n/a
Clausthal, Clausthal-Zellerfeld	3.08	4	3.20	2	n/a	n/a	3.00	1	n/a	n/a
Waldkindergarten, Lutter am Barenberge	3.05	1	2.67	1	3.00	1	4.00	2	n/a	n/a
Steinbergspielplatz, Goslar	3.04	2	2.75	2	5.00	1	5.00	1	n/a	n/a

Abbildung 4.10: Übersicht der Besucher-Interessen je Veranstaltungsort

Es liegt nahe, nicht nur den Nutzern Artefakte zu empfehlen, sondern auch „Empfehlungen“ in Gestalt eines Informationssystems über Orte und Veranstaltungsreihen bereitzustellen. Kindergärten besitzen häufig unterschiedliche Schwerpunkte, wie z. B. ein Waldkindergarten oder ein musikalischen Kindergarten im Landkreis Goslar. Es ist leicht vorstellbar, dass verschiedene Veranstaltungsreihen/Themen an diesen Orten unterschiedlich beliebt sind. Abbildung 4.10 zeigt die Ausgabe zu Beginn der Evaluation. Dabei sind für alle bisherigen Veranstaltungsorte sowie Veranstaltungsreihen die Anzahl sowie Durchschnittsbewertungen aller mit den Veranstaltungsreihen assoziierten Artikel aufgeführt, wobei lediglich die Bewertungen der Mitglieder, die eine Veranstaltung an diesem Ort besucht haben, berücksichtigt werden. Für jeden Ort wird die am besten bewertete Themenreihe mit mindestens drei Bewertungen farblich hervorgehoben. Artikel für die Veranstaltungsreihe „Kinderohren hören – Kindermünder erzählen“ waren zu diesem Zeitpunkt noch nicht auf der Online-Plattform verfügbar, daher werden alle Werte mit „not available“ (n/a) gekennzeichnet.

4.3 Implementation

Die Online-Plattform *Mobile2Learn.de* setzt auf dem Zikula⁴-Framework (hervorgegangen aus PostNuke) auf. Dabei handelt es sich um ein in PHP⁵ geschriebenes Open-Source (GNU GPL⁶ Version 2) Web-Application-Framework, das auf der Web-Variante Model View Controller (MVC) Modell 2 des MVC-Musters (vgl. Lahres u. Rayman 2006, Kapitel 8.2) basiert, einen Kern für allgemeine Webanwendungen mitbringt und in das neue Funktionalitäten in Form von einzelnen Modulen eingebunden werden können. Der Kern enthält eine Benutzer-, Gruppen- sowie Rechteverwaltung, eine allgemeine Suchfunktion, die Möglichkeit verschiedene Themes/Layouts zu definieren und diverse Schnittstellen, die für die Entwicklung und Nutzung von Modulen erforderlich sind (Datenbankabstraktion mittels Object Relational Mapping (ORM) und Data Access Objects (DAOs), HTML-Formular-Validierungs-Application Programming Interface (API), Sitzungsmanagement, Mehrsprachigkeit, usw.). Das Framework kann durch das Einbinden von Modulen mit Content-Management- (Content Management System (CMS) für Artikel, Galerien) und Community-Funktionen (Profil-Management, Forum, internes Nachrichtensystem, Newsletter, Bewertungs- und Kommentarfunktionen) ausgestattet werden. Das Framework ermöglicht, dass die einzelnen Module über definierte Schnittstellen miteinander interagieren können, so dass z. B. die Suchfunktion auch die Artefakte der einzelnen Module berücksichtigen kann oder Artikel des CMS-Moduls kommentiert werden können.

Das Empfehlungssystem wurde als eigenständiges Modul entwickelt und greift auf das bestehende Bewertungs- und Profilmanagementmodul zurück. Genauso kann es auch von anderen Modulen (z. B. CMS, Newsletter) genutzt werden, um Empfehlungen dort einzubinden.

Bei der Implementation wurde stets viel Wert auf Erweiterbarkeit gelegt, so dass sich weitere Artefakte anderer Module leicht einbinden und somit auch empfehlen lassen.

⁴Version 1.2, <http://www.zikula.org>

⁵hauptsächlich PHP Version 4; PHP: PHP Hypertext Preprocessor, <http://www.php.net>

⁶General Public License: <http://www.gnu.org/licenses/gpl-2.0.html>, abgerufen am 15. September 2011

4.4 Präsentation der Empfehlungen

Neben einer möglichst hohen Qualität der Empfehlungen spielt die Präsentation eine wichtige Rolle (Cosley u. a. 2003), so dass die Benutzer von den Empfehlungen bestmöglich profitieren können. Dabei geht es nicht nur um die optische Gestaltung, sondern auch um die Kanäle, über die die Empfehlungen zur Verfügung gestellt werden.

Mobile2Learn setzt bereits im Ansatz auf unterschiedliche Kanäle, um die Benutzer zu erreichen: Online-Plattform/Homepage, E-Mail und SMS (nach steigender Invasivität geordnet). Der Empfehlungssystem-Prototyp benutzt alle drei Kanäle, jedoch in unterschiedlicher Intensität. Unabhängig davon werden bei den Präsenzveranstaltungen auch direkt Empfehlungen vom *Mobile2Learn*-Team für die Eltern ausgesprochen. Diese sind jedoch nicht Gegenstand der vorliegenden Arbeit.

Das Darstellen von Empfehlungen auf der Homepage ist die pragmatischste Form, jedoch werden diese so lediglich von Benutzern wahrgenommen, die auch die Webseite (regelmäßig) besuchen. Abbildung 4.11 zeigt ein TOP5-Ranking auf der *Mobile2Learn.de*-Startseite. Durch diese Platzierung sind Empfehlungen für jeden Benutzer leicht zu entdecken bzw. zugänglich, so dass keine spezielle Empfehlungsseite aufgerufen werden muss. Die Darstellung erfolgt als eine geordnete Liste von Hyperlinks auf die empfohlenen Artefakte. Da speziell die Artikelnamen allein, wie z. B. „Erziehungstipps“, nicht eindeutig sind, ist es notwendig, den Pfad im Artikelbaum mit anzugeben, so dass der Kontext des Artikels verständlich wird (z. B. „Lernen mit allen Sinnen » Sehen » Erziehungstipps“). Damit das Hauptaugenmerk auf dem Artikelnamen bzw. Artefakttitel liegt, wird dieser in fett und der Pfad mit kleinerer Schriftgröße über dem Titel dargestellt. Tooltips mit dem ersten Foto des Artikels bzw. der Galerie und, sofern vorhanden, dem Teaser-Text (kurze Zusammenfassung eines Artikels) sollen dafür sorgen, dass Benutzer auch bei weniger aussagekräftigen Titeln einen Eindruck vom Inhalt der Empfehlung erhalten können, ohne die Übersichtlichkeit einzuschränken (Forderung von Sinha u. a. 2001).

Durch das Versenden von E-Mails ist es möglich, auch passive Benutzer zu erreichen und mit Empfehlungen den (erneuten) Besuch der Online-Plattform anzuregen. Dabei gibt es zwei Optionen, die sich nicht gegenseitig ausschließen: Es werden ausgezeichnete Empfehlungs-E-Mails, die lediglich eine oder mehrere Empfehlungen enthalten, verschickt oder Empfehlungen werden in (andere) E-Mails, wie z. B. Kampagnen⁷, eingebettet. Der Prototyp setzt ausschließlich auf das Einbetten von Empfehlungen, da das zusätzliche Versenden von Empfehlungs-E-Mails neben den zweiwöchentlichen E-Mail-Kampagnen schnell als SPAM aufgefasst werden kann bzw. die redaktionell gestalteten E-Mails in den Postfächern der Benutzer untergehen könnten.

Eine aktive Ansprache der Community-Mitglieder ist neben E-Mails mit SMS-Nachrichten möglich, jedoch müssen die Nachrichten hierbei nicht explizit abgerufen werden. Da die Benutzer zudem ihr Mobiltelefon häufig bei sich tragen, ist dieser Kanal mit besonderer Vorsicht zu benutzen, um die Empfänger nicht zu verärgern: Deshalb wird im vorliegenden System darauf gesetzt, für jeden Benutzer, der die automatischen SMS-Empfehlungen nicht deaktiviert hat, maximal einmal pro Woche eine

⁷Aus diesem Grund lassen sich die regelbasierten Empfehlungen deaktivieren, damit z. B. in einer terminbewerbenden Kampagne keine „doppelte“ Empfehlung für diesen Termin enthalten ist.

The screenshot shows the homepage of Mobile2Learn.de. At the top, there is a green banner with the text "Lernen mit allen Sinnen | Alte Kinderlieder - neu entdecken!". Below this, a welcome message reads: "Willkommen Sven Strickroth bei Mobile2Learn! Herzlich willkommen auf der Community-Website Mobile2Learn. Der Online-Bildungsplattform für frühkindliche Bildung und Förderung." To the left is a photo of a child with a green wig playing a xylophone. The main content area is divided into several sections:

- Neues auf Mobile2Learn.de:**
 - ↘ Persönliche Empfehlungen von Mobile2Learn
 - ↘ Mit Mobile2Learn in die Natur - die Termine
 - ↘ Neue Inhalte und Spiele finden Sie unter:
 - ↘ Sehen -> bis 3. Lebensjahr -> "Knopfkönig"
 - ↘ Tasten -> ab 4. Lebensjahr -> "Das fliegende Geldstück"
 - ↘ Mit Kindern in die Natur -> "Beim Waldkindergarten Lutter"
 - ↘ Die Fotos für sämtliche Veranstaltungen bis 30.06.2011 finden Sie in unserer Galerie.
- Unsere Empfehlungen für Sie:**
 - ↘ Musikalische Frühförderung von der Kreismusikschule Goslar "Musik macht Spaß" von Rose Meyer-Zurwelle
 - ↘ Lernen mit allen Sinnen
 - Hören
 - Lernen mit allen Sinnen » Hören
 - Erziehungstipps
 - ↘ Lernen mit allen Sinnen » Sehen - visuelle Wahrnehmung » Spiele bis 3. Lebensjahr
 - Fallschirmspiele
 - ↘ Lernen mit allen Sinnen » Tasten/Bewegen » Spiele ab 4. Lebensjahr
 - [Karate mit Zeitungspapier](#)
- Hauptmenü:**
 - Startseite
 - Über uns
 - Termine
 - Netzwerk-Partner
 - Bildungsinhalte / Spiele
 - Inhalte bearbeiten
 - Rätsel/Quiz
 - Anmelden
 - Administration
 - Galerie
 - Forum
 - Nachrichten-Zentrale
 - Artikel einsenden
 - Freunde einladen
 - Benutzerkonto
 - Für Kindergärten
 - Aktuelle URL hinzufügen
 - Block bearbeiten
- Wer ist online:**
 - 2 registrierte Benutzer und 9 Gäste online.
 - Es gibt keine neuen Nachrichten.
 - Angemeldet als **strickroth (Sven Strickroth)**.

At the bottom, there are logos for "Gefördert von: TUH Clausthal" and "Niedersächsisches Institut für Weiterbildung", along with copyright information: "© 2010 - 2011 Mobile2Learn Nutzungsbedingungen Impressum | Nach oben".

Abbildung 4.11: Screenshot der Mobile2Learn.de-Startseite mit Empfehlungen

Empfehlung für genau ein Artefakt (mit überdurchschnittlich guter Vorhersage, > 4,5) per SMS zu versenden. Allgemeine Empfehlungen für die Übersicht über alle Inhalte werden nicht versandt, da diese zu unspezifisch sind. Eine besondere Schwierigkeit besteht in der begrenzten Textlänge einer SMS (160 Zeichen). Die URL der Artefakte enthalten zur Suchmaschinenoptimierung den gesamten Artikelpfad (Beispiel: <http://www.mobile2learn.de/bildungsinhalte/lernen-mit-allen-sinnen/sehen/spiele-fuer-die-drei-ersten-lebensjahre/fallschirmspiele.html>). Somit kann die URL in dieser Form (im Beispiel 134 Zeichen) nicht per SMS versandt werden, da es zum einen sehr aufwändig ist, diese lange Adresse fehlerfrei im Browser abzutippen und zum anderen so kein Platz für weitere Förmlichkeiten oder Hinweise bleibt. Kurz-URLs (z. B. <http://mobile2learn.de/jmp/J3f47fJH>) sind zwar deutlich kürzer, dafür aber auch schlechter zu lesen und zu verstehen. Darüber hinaus tragen viele Menschen URLs statt in die Adresszeile eines Browsers in eine Suchmaschine ein. Ist die eingegebene URL (Kurz-URLs sind i. d. R. eindeutig, um z. B. Benutzer-Tracking zu ermöglichen) nicht in deren Index vorhanden, wird kein Hyperlink darauf generiert, sondern „Keine Ergebnisse gefunden“⁸ angezeigt. Daher wurde für jeden Artefakttyp eine SMS-Vorlage

⁸Getestet mit dem Marktführer Google, sowie den Suchmaschinen Bing und Yahoo am 14.09.2011.

erstellt, in welcher der Titel des Artefaktes (evtl. abgekürzt) und ein Hinweis zum Abruf der Empfehlung über die *Mobile2Learn.de*-Startseite enthalten sind. Für einen Artikel sieht eine SMS beispielweise wie folgt aus: „Liebe Eltern, der Artikel ‚Fallschirmspiele‘ könnte Sie interessieren. Abruf über *Mobile2Learn.de*“. Trotz dieser relativ starken Einschränkungen (160 Zeichen, keine Bilder und kein Benutzer-Tracking) werden ausschließlich SMS-Textnachrichten eingesetzt, da z. B. Multimedia Messaging Service (MMS)-Nachrichten, die auch Bilder enthalten können, oder spezielle Smartphone-Apps teurer und vor allem nicht auf allen Mobiltelefonen verfügbar sind.

Neben Artikelempfehlungen kann der Personalisierungsgrad auch durch weitere Maßnahmen erhöht werden: So wurde ein dynamischer Block für die Newsletter entwickelt, der personalisiert Empfehlungen bzw. Ankündigungen für neue Präsenzveranstaltungen generiert. Bisher wurden in einer Rund-E-Mail zur Ankündigung von Präsenzreihen immer alle Veranstaltungen tabellarisch aufgelistet. Der neue Block basiert auf der gleichen Logik wie die regelbasierten Empfehlungen für eine Präsenzveranstaltung und passt den Inhalt der Tabelle an den aktiven Benutzer an: Statt für jeden Benutzer alle Veranstaltungen anzukündigen, wird die Liste automatisch gefiltert, so dass nur noch Veranstaltungsorte, die sich in der Nähe befinden oder die der Benutzer bereits besucht hat, angezeigt werden (vgl. Abbildung 4.12). Alle anderen Veranstaltungen können trotzdem über einen Link auf „*Mobile2Learn.de* -> Termine“ abgerufen werden. Ist über einen Benutzer nichts bekannt, so wird die komplette Liste angezeigt (vgl. Abbildung 4.13).

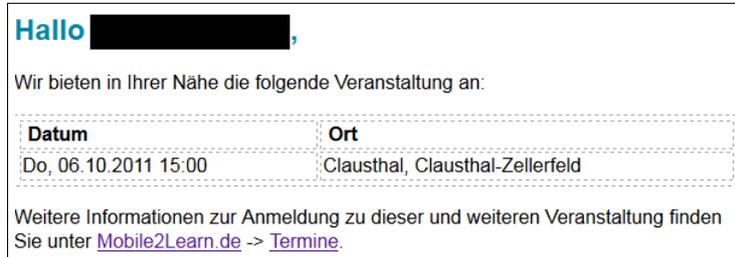


Abbildung 4.12: Ausschnitt einer dynamischen Veranstaltungsankündigungs-E-Mail



Abbildung 4.13: Ausschnitt einer allgemeinen Veranstaltungsankündigungs-E-Mail

4.5 Parameterwahl

Es gibt in diesem hybriden Empfehlungssystem-Prototyp eine Reihe von Parametern, für die möglichst optimale Belegungen zu bestimmen sind, damit das ES gute Voraussagen bzw. Empfehlungen generieren kann. In diesem Abschnitt werden das Vorgehen und die Wahl der Belegungen erläutert.

4.5.1 Evaluationsmetrik

Um die Güte eines ES anhand der Datenbasis zu messen, benötigt man eine Metrik bzw. ein Gütemaß. Herlocker u. a. (2004) beschreiben eine Reihe verschiedener Metriken, die bereits für die Evaluation von Empfehlungssystemen benutzt wurden. Dabei wird eine Metrik auf das Ergebnis einer Kreuzvalidierung angewendet: Die Benutzerdatensätze werden zufällig in zwei disjunkte Mengen, eine Trainings- sowie eine Testmenge, eingeteilt. I. d. R. werden 80 % bis 95 % der Datensätze der Trainingsmenge zugeordnet. Danach wird der Algorithmus auf der Trainingsmenge ausgeführt. Anschließend werden die Empfehlungen mit Hilfe einer Metrik mit der Testmenge verglichen.

Die Auswahl einer geeigneten Metrik ist sehr schwierig. Der sehr häufig benutzte durchschnittliche absolute Fehler (Mean Absolute Error (MAE)) ist für den Anwendungsfall „Find Good Items“ nicht optimal, da er zwar misst, wie genau die einzelnen Voraussagen sind, jedoch keine Aussage darüber trifft, ob sich die größten Fehler am Anfang oder am Ende der Empfehlungsliste, von der die obersten N Artefakte empfohlen werden, befinden (Herlocker u. a. 2004). Darüber hinaus ist diese Metrik nur sinnvoll, wenn der Empfehlungsalgorithmus auf einem Vorhersagealgorithmus basiert, der darauf ausgelegt ist, möglichst genaue Voraussagen zu treffen (vgl. Abschnitt 3.1).

Da ein Ranking als Empfehlung generiert wird, wurden bereits Rangkorrelationskoeffizienten (wie z. B. der Spearman Rangkorrelationskoeffizient oder Kendalls Tau, vgl. Hartung u. a. 2009) zur Gütemessung eingesetzt. Beide Korrelationen vergleichen zwei rangbasierte Listen und geben an, inwieweit diese in den Reihenfolgen übereinstimmen (unabhängig von konkreten Variablenbelegungen bzw. Vorhersagen). Für die Auswertung wird das TOP- N -Ranking des Algorithmus mit der nach der Benutzer-Bewertung sortierten Testmenge verglichen. Diese Metriken haben jedoch auch den gerade genannten Nachteil, dass der „Ort“ für zwei vertauschte Artefakte im TOP- N -Ranking ebenfalls nicht gemessen wird. Zudem gibt es Probleme mit einer „schwachen“ bzw. partiellen Ordnung (Herlocker u. a. 2004): Das TOP- N -Ranking weist i. d. R. eine totale Ordnung auf. Die sortierte Artefaktliste der Testmenge besitzt bei einer Bewertungsskala mit fünf Möglichkeiten hingegen meist nur eine partielle Ordnung. Die beiden Rangkorrelationskoeffizienten „bestrafen“ nun Listeneinträge, die der Benutzer gleich bewertet hat, aber im TOP- N -Ranking in anderer Reihenfolge auftreten.

Angelehnt an Deshpande u. Karypis (2004) wurde die Hit-Ratio (Trefferquote, Recall) als Maß für die Güte ausgewählt. Dieses Maß beruht auf den für Empfehlungssysteme modifizierten Precision-/Recall-Häufigkeiten aus dem Bereich der Information-Retrieval (Cleverdon u. a. 1966; Basu u. a. 1998; Herlocker u. a. 2004). Die Precision-

/Recall-Häufigkeiten basieren auf den binären disjunkten Klassen „relevant“ sowie „nicht relevant“ und bieten ein Maß, wie gut Klassifizierungen einzelner Artefakte durch einen Algorithmus vorgenommen werden konnten (vgl. Tabelle 4.3). Da die Bewertungen nicht binär, sondern in fünf Klassen (1 bis 5 Sterne) vorliegen, ist eine Transformation erforderlich: Es ist üblich, Bewertungen ≥ 4 als „relevant“ und alle anderen als „nicht relevant“ bzw. „irrelevant“ einzustufen (vgl. Herlocker u. a. 2004). Im vorliegenden Fall gilt zudem, dass die Anzahl der empfohlenen Artefakte N_e nicht vom Algorithmus variiert wird, sondern hinsichtlich des TOP- N -Rankings auf N festgelegt ist.

	Empfohlen	Nicht empfohlen	Gesamt
Relevant	N_{re}	N_{rn}	N_r
Irrelevant	N_{ie}	N_{in}	N_i
Gesamt	N_e	N_n	N_g

Tabelle 4.3: Klassifizierungsmatrix zur Berechnung von Precision und Recall

Der Recall ist die relative Häufigkeit, dass die relevanten Artefakte der Testmenge vorgeschlagen wurden:

$$R = \frac{N_{re}}{N_r}$$

Die Precision hingegen beschreibt den Anteil der relevanten Artefakte der Testmenge von allen empfohlenen Artefakten:

$$P = \frac{N_{re}}{N_e}$$

Diese beiden relativen Häufigkeiten werden oftmals als repräsentativ angesehen und somit als Schätzung der zugehörigen bedingten Wahrscheinlichkeiten interpretiert.

In der Literatur ist der Sinn und Zweck der Precision bei der Evaluation von Empfehlungssystemen (mit fester Anzahl empfohlener Artefakte) umstritten: Auf der einen Seite argumentieren Befürworter, dass sie notwendig ist, da durch einfaches Verlängern der Empfehlungsliste der Recall bis zum Maximum von 1 (durch Ausgabe aller Artefakte) gesteigert werden kann. Der Nutzen (Precision) der Empfehlung nimmt jedoch ab, wenn zwischen den „relevanten“ Artefakten (sehr) viele „nicht relevante“ zu finden sind. Dieses Argument ist aber nur für Systeme stichhaltig, in denen der Algorithmus selbstständig die Anzahl von empfohlenen Artefakten variieren kann. Auf der anderen Seite bestehen für die Ausnutzung des gesamten Wertebereiches $[0; 1]$ sowie die Erreichung des Maximums von 1 Anforderungen: Beispielsweise muss die Anzahl der relevanten Artefakte in der Testmenge N_r mindestens so groß wie die Anzahl von empfohlenen Artefakten N_e sein, da sonst insbesondere der Fall „alle empfohlenen Artefakte sind relevant“ nicht korrekt abgebildet werden kann ($P = 1$ müsste gelten, aber $\frac{N_{re}}{N_e} \neq 1$ mit $N_{re} = N_r < N_e$). Bei den vorliegenden Daten ist dies nicht der Fall.

Die Anzahl der relevanten Artefakte in der Testmenge ist maximal so groß wie die Anzahl der empfohlenen Artefakte ($N_r \leq N_e$). Enthält die Testmenge genau ein relevantes Artefakt, kann die Precision (im Falle von fünf Empfehlungen, $N = N_e = 5$) maximal 0,2 betragen. Eine künstliche Beschränkung der Empfehlungsliste bzw. das Einfügen eines Korrekturfaktors $\frac{N_e}{N_r}$ in die Precision ist nicht sinnvoll bzw. verändert

das Maß erheblich ($P = \frac{N_{re}}{N_e} \cdot \frac{N_e}{N_r} = \frac{N_{re}}{N_r} = R$). Die Anzahl von empfohlenen Artefakten $N_e = N$ ist bei allen Läufen gleich, da er insbesondere nicht vom Algorithmus adaptiert wird. Die Anzahl der relevanten Artefakte N_r ist abhängig von der konkreten Wahl der Testmenge. Für eine einzelne Bestimmung von Precision und Recall sind diese beiden Werte konstant: Precision und Recall hängen linear mit einem festen Skalierungsfaktor voneinander ab. Für eine einzelne Messung ist unerheblich welcher der beiden Werte (oder eine Kombination davon) für ein Gütemaß verwendet wird. Selbst bei verschiedenen Testmengen und der Mittelung der Werte existiert weiterhin ein Zusammenhang über einen Skalierungsfaktor. Da die Anzahl der relevanten Artefakte N_r auf den vorliegenden Daten nicht stark variiert, ist auch der Skalierungsfaktor nach der Mittelung ähnlich zu denen vor der Berechnung des Mittelwertes. Unter dieser Voraussetzung sind Precision und Recall keine voneinander unabhängigen Größen. Aus diesem Grund ist die konkrete Wahl nicht entscheidend und es wurde aus Vergleichbarkeitsgründen der Recall ausgewählt.

Zur Verdeutlichung, dass von der Precision/Recall-Kombination lediglich der Recall Anwendung findet, wird im Bereich der ES stattdessen der Begriff Hit-Ratio verwendet. Dieser berechnet sich als Prozentsatz der relevanten Artefakte in der Testmenge, die auch in der Empfehlung enthalten sind. Folglich bedeutet ein Hit-Ratio von 0 bzw. 1, dass kein einziges relevantes Artefakt bzw. alle relevanten Artefakte in der Empfehlung enthalten waren. Die Hit-Ratio ist jedoch kein absolutes Gütemaß und kann daher nur vergleichend eingesetzt werden.

Insgesamt muss dieses Maß mit besonderer Vorsicht eingesetzt werden: Die Trefferrate gibt ausschließlich den Prozentsatz aller relevanten Artefakte der Testmenge in der Empfehlung an. Jedoch kann über die anderen in der Empfehlung enthaltenen Artefakte keine Aussage getroffen werden, da über diese Artefakte keine (Vergleichs-)Daten vorhanden sind (wie auch bei allen anderen Metriken). Das Maß bewertet somit nur die Fähigkeit des Algorithmus', bereits gut bewertete Artefakte aus der Testmenge zu empfehlen. Es liegt nahe, auch die negative Trefferrate zu bestimmen. Diese gibt an, wie viele „nicht relevante“ Artefakte der Testmenge in der Empfehlung vorhanden sind. Auf den vorliegenden Daten wurde dies auf Grund der vielen abgewerteten impliziten Bewertungen jedoch nicht durchgeführt, da durch die Abwertung Artefakte, die gerade noch „relevant“ sind, in die Kategorie „nicht relevant“ einsortiert werden.

4.5.2 Parameteroptimierung

Zur Bestimmung der einzelnen Parameter wurden systematisch Testläufe im Rahmen von Kreuzvalidierungen mit der Hit-Ratio-Metrik durchgeführt, die sich i. d. R. nur in einem einzigen Parameter unterschieden haben. Diese Einschränkung ist nötig, da sich herausgestellt hat, dass nicht alle Parameter voneinander unabhängig sind (z. B. Ähnlichkeitsmaß und Signifikanzgewichtung). In diesen Fällen wurden alle (sinnvollen) Kombinationen untersucht.

Die Kreuzvalidierungen wurden wie folgt durchgeführt: Für eine Auswertung wurden alle Benutzer betrachtet und für jeden 25 Kreuzvalidierungsläufe angesetzt. Das Hit-Ratio einer Auswertung ergibt sich aus der Mittelung der einzelnen Ergebnisse. Die Daten wurden für jeweils einen Lauf in zwei Mengen (eine Trainings- und eine

Testmenge) aufgeteilt, um zu gewährleisten, dass die Ergebnisse keine Überanpassungserscheinungen (Overfitting) darstellen. Die Trainingsmenge bestand aus 95 % zufällig ausgewählter Bewertungen eines Nutzers – die Testmenge bestand aus den restlichen 5 % der Daten. Zur Sicherstellung der Vergleichbarkeit wurden diese Aufteilungen (je eine Aufteilung pro Lauf und Benutzer) festgehalten und die Vorbereitungsschritte je Lauf auf der Trainingsmenge ausgeführt (vgl. Abschnitt 4.2.1). Befand sich in der Testmenge mindestens ein Artefakt mit einer Bewertung von 4 oder besser, wurde der Empfehlungsalgorithmus ausgeführt und das Hit-Ratio bestimmt.

Ausgehend von einer Implementation eines elementbasierten kollaborativen Filters wurden schrittweise die zwei weiteren Stufen aus Abschnitt 4.2 ergänzt und jeweils Evaluationen durchgeführt. Abbildung 4.14 zeigt eine grobe Übersicht der Steigerungen der Güte über die einzelnen Stufen. Zusätzlich wurden die drei Stufen einzeln getestet, um deren Beitragsmöglichkeit zur Gesamtempfehlung zu bestimmen.

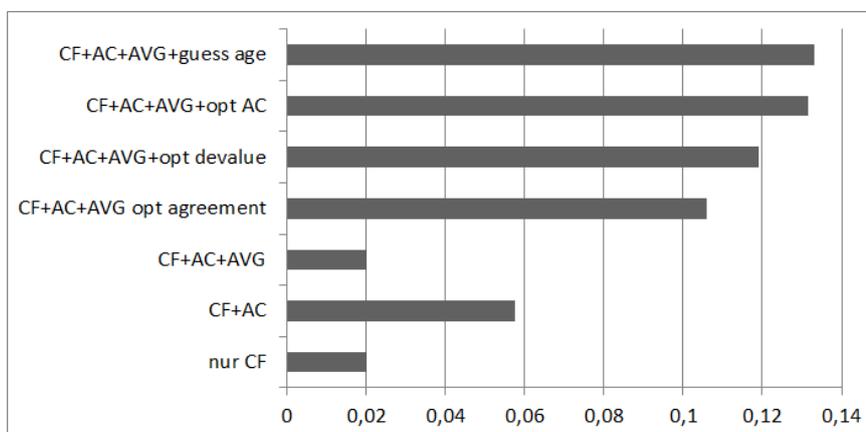


Abbildung 4.14: Übersicht der Parameter-Evaluation

Im ersten Schritt wurde der reine elementbasierte kollaborative Filter mit verschiedenen Ähnlichkeitsmaßen (vgl. Abschnitt 3.3) evaluiert. Nur die angepasste Cosinus-Korrelation (Hit-Ratio: 0,02) und die Pearson-Korrelation (Hit-Ratio: 0,005) waren in der Lage, gemäß der Metrik überhaupt gute Empfehlungen zu geben. Daraus ist ersichtlich, dass es Qualitätsunterschiede bei verschiedenen Ähnlichkeitsmaßen gibt. Am besten scheint die angepasste Cosinus-Korrelation abzuschneiden. Daher wird diese für die weitere Evaluation festgehalten.

Im zweiten Schritt wurden die α -Community Räume einzeln und in Kombination mit dem kollaborativen Filter evaluiert ($t_{\text{agreement}} = 0,3$): Die Gütemessung der einzelnen α -Communities wurde durchgeführt, um herauszufinden, welche α -Communities in der Lage sind, Vorhersagen bzw. Empfehlungen zu generieren. Abbildung 4.15 zeigt die Hit-Ratios der einzelnen α -Communities sowie deren kombiniertes Ergebnis: Keine einzige α -Community ist in der Lage, Empfehlungen mit der Güte des CF-Verfahrens zu generieren; in der Kombination hingegen sind die Empfehlungen in der Güte vergleichbar (CF: 0,02, AC: 0,026). Es hat sich gezeigt, dass die α -Community für die Geschlechter der Eltern nicht in der Lage war, Empfehlungen zu generieren. Daher wurde diese aus der Implementation entfernt. Weiterhin leisten die α -Communities für das Geschlecht der Kinder sowie den Wohnort der Eltern isoliert nur einen sehr kleinen Beitrag, werden jedoch trotzdem beibehalten. In der Kombination der α -Communities mit dem CF zeigt

sich, dass die Kombination insgesamt besser für diese Community geeignet ist, als ein rein kollaborativer Filter (vgl. Abbildung 4.14). Diese Aussage bezieht sich jedoch ausschließlich auf die gesamte Community, also explizit auch auf die Benutzer, die bisher keine Artikel gelesen haben. An dieser Kombination wurden keine weitere Parameteroptimierungen vorgenommen, da Ergebnisse lediglich für diese Zweier-Kombination gültig wären, aber der Ansatz noch eine weitere Stufe enthält, die speziell Empfehlungen für Benutzer generieren soll, über die „so gut wie nichts“ bekannt ist.

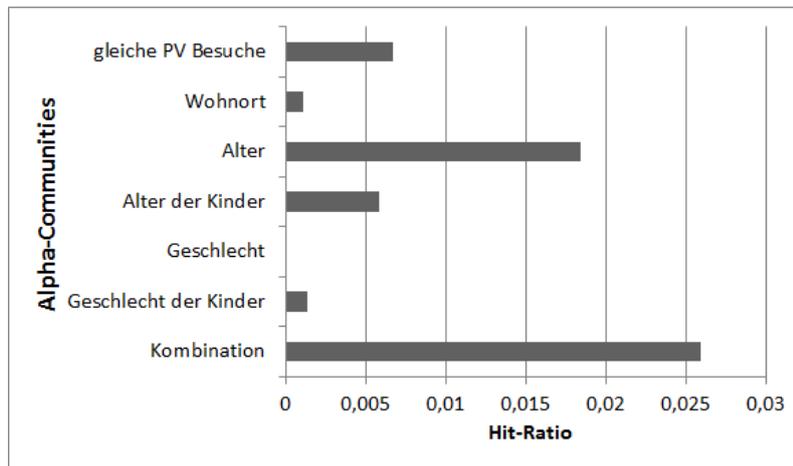


Abbildung 4.15: Hit-Ratio der α -Communities

Im dritten Schritt wurde der Beitrag des durchschnittsbasierten Filters allein sowie die Güte in Kombination mit dem kollaborativen Filter und den α -Communities bestimmt (vgl. Abbildung 4.14, dritter Balken von unten). Die einfache Integration des durchschnittsbasierten Filters hat die Güte erneut verschlechtert. Erst durch die Optimierung der Durchschnittsbildung der Artefakte mit einer Mindestanzahl von Bewertungen („agreement“, vgl. Abschnitt 3.5 und Abbildung 4.16) konnte die Empfehlungsgüte deut-

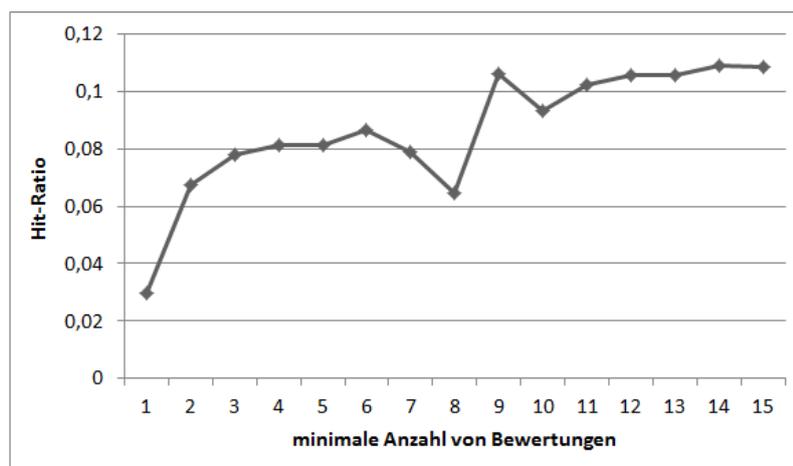


Abbildung 4.16: Hit-Ratio in Abhängigkeit der minimalen Anzahl von Bewertungen beim durchschnittsbasierten Filter

lich gesteigert werden (Hit-Ratio: 0,11 mit $t_{\text{agreement}} = 9$, lokales Optimum im Intervall [1; 13]). Ein Agreement von 9 erscheint zunächst relativ hoch, jedoch merkten Schafer u. a. (2007) an, dass bei impliziten Daten gute Ergebnisse eine größere Datengrundlage erfordern. Darüber hinaus ist dies auch keine allzu große Einschränkung (vgl. Abbildung 4.17), da zum Evaluationszeitpunkt bereits ca. 60 % der Artefakte mindestens 9 Bewertungen aufweisen. Werte größer 12 für das Agreement sind nicht ratsam, da zu diesem Zeitpunkt von den Artefakten weniger als 30 % so häufig bewertet wurden.

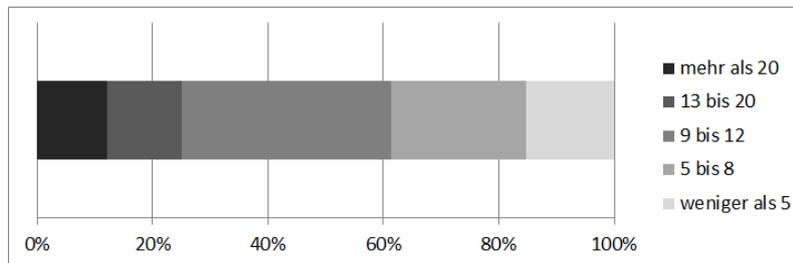


Abbildung 4.17: Prozentualer Anteil der Artefakte sortiert nach Anzahl von Bewertungen

Nach Abschluss der Integration aller drei Stufen des Ansatzes können die weiteren Parameter der einzelnen Bestandteile im Ganzen schrittweise optimiert werden. Die Optimierung der Signifikanzgewichtung („devalue“) ist in Abbildung 4.18 ersichtlich. Bei einem Wert von 5 existiert ein sinnvolles lokales Optimum (Hit-Ratio: 0,12). Bei Werten ab 15 wird es überschritten. Diese Empfehlungen werden in der Güte gemäß der Metrik besser als das lokale Optimum, jedoch können nur noch Empfehlungen für Benutzer generiert werden, die bereits viele Bewertungen (implizit und explizit > 20) vorgenommen haben (kleiner Teil der gesamten Community, vgl. Abbildung 4.1 auf Seite 43).

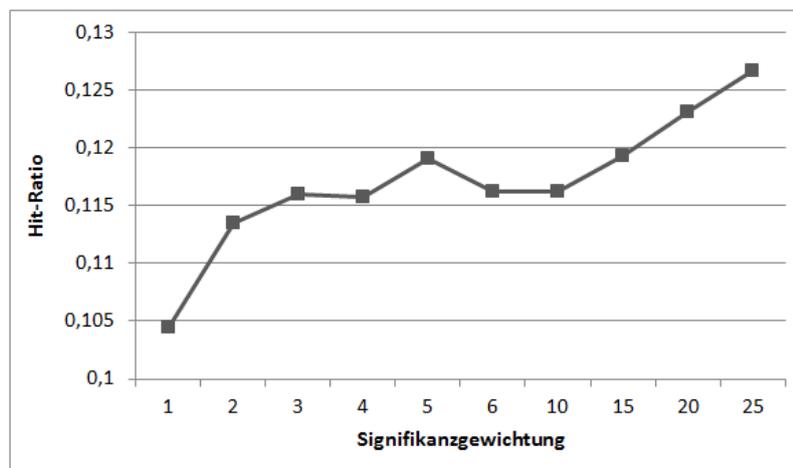


Abbildung 4.18: Hit-Ratio in Abhängigkeit der Signifikanzgewichtung im kollaborativen Filter

Im nächsten Schritt wurde der $t_{\text{agreement}}$ -Parameter der α -Communities optimiert (vgl. Abbildung 4.19). Dieser hat sein globales Optimum mit einer Hit-Ratio von 0,13 bei $t_{\text{agreement}} = 0,2$ (Nguyen u. a., die Entwickler der ACs, schlugen $t_{\text{agreement}} = 0,25$ vor). Mit der Hinzunahme der Alterserkennung ("guess age") kann das Hit-Ratio auf ca. 0,14 gesteigert werden. Die Alterserkennung hat zwar nur einen sehr geringen Einfluss auf die Hit-Ratio, jedoch ist deren Einfluss auf eine Empfehlung sehr deutlich sichtbar.

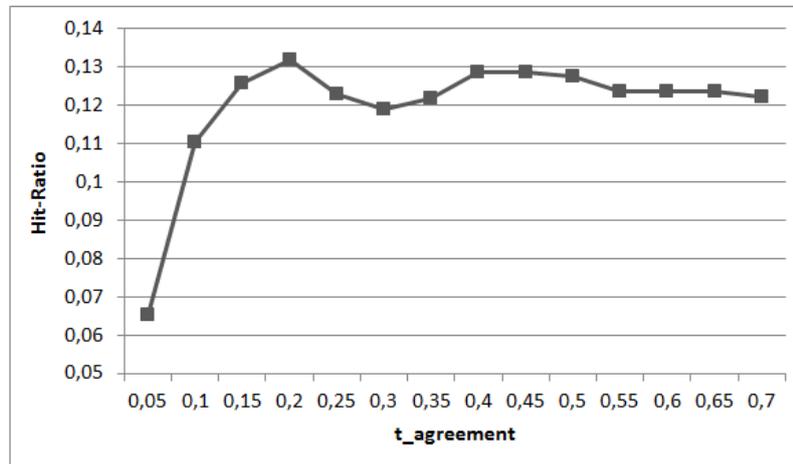


Abbildung 4.19: Hit-Ratio in Abhängigkeit der Variable $t_{\text{agreement}}$ der α -Communities

Schließlich wurden noch einmal die verschiedenen Ähnlichkeitsmaße zusammen mit der Hinzunahme bzw. Auslassung von Ähnlichkeitsvektoren der Länge eins sowie drei Signifikanzgewichtungen (1, 3 und 5) evaluiert (vgl. Abbildung 4.20, diese drei Variablen haben sich als nicht voneinander unabhängig gezeigt). Das Präfix „o1“ bedeutet, dass Bewertungsvektoren der Länge eins ausgeschlossen wurden, „w1“ das Gegenteil.

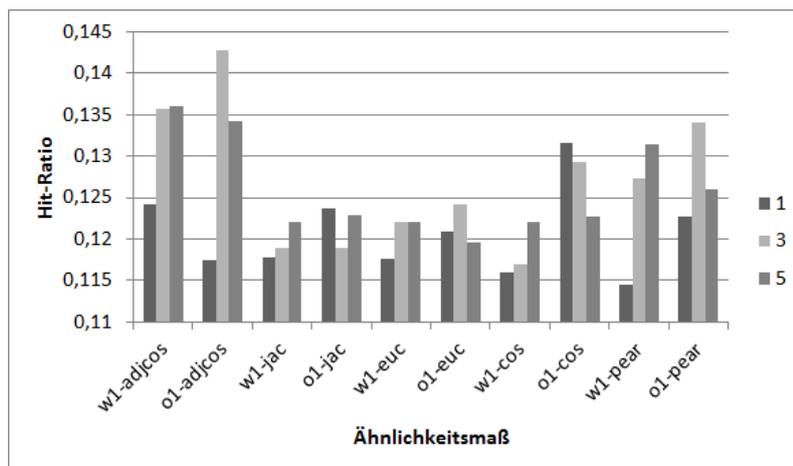


Abbildung 4.20: Hit-Ratio in Abhängigkeit unterschiedlicher Ähnlichkeitsmaße und Signifikanzgewichtungen

Der erst signifikant erscheinende Unterschied der Ähnlichkeitsmaße ist auf dem parameteroptimierten Algorithmus nicht mehr so ausgeprägt erkennbar. Alle Metriken (bei entsprechender Wahl der Signifikanzgewichtung) erreichen ein Hit-Ratio von mehr als 0,12. Bei der auffallenden Spitze vom *o1-adjcos*, handelt es sich um einen Ausreißer (dies hat sich in verschiedenen Testläufen gezeigt). Die Wahl der angepassten Cosinus-Korrelation mit Betrachtung der Bewertungsvektoren der Länge eins (*w1-adjcos*) wurde trotz der geringen Unterschiede auf Grund von Empfehlungen (vgl. Abschnitt 3.3.6, Sarwar u. a. 2001) beibehalten. Zusätzlich wurden für 10 % der Benutzer Empfehlungen der „w1“- und „o1“-Varianten manuell verglichen und anhand der bisherigen Erfahrungen und besser erscheinenden Diversifikation die „w1“-Variante ausgewählt.

Darüber hinaus wurde der Einfluss der Filterbots mit drei unterschiedlichen Implementationen evaluiert (vgl. Abbildung 4.21). Die drei Varianten unterscheiden sich darin wie die Ähnlichkeiten der Artefakte berechnet werden: Bei der ersten Variante handelt es sich exakt um das in Abschnitt 4.2.1 vorgestellte Vorgehen (mit dem angesprochenen Problem), in dem die Ähnlichkeit zweier Artefakte über alle Bewertungen (Benutzer- und Filterbot-Bewertungen) bestimmt wird. Die beiden weiteren Varianten berechnen die Artefakt-Ähnlichkeiten der Filterbot- sowie der Benutzerbewertungen getrennt und kombinieren beide Werte: Die Ähnlichkeiten der Benutzerbewertungen werden gemäß der Beschreibung aus Abschnitt 3.9.2 berechnet, die Ähnlichkeiten der Filterbotbewertungen hingegen auf den vollständigen Bewertungsvektoren mit dem Jaccard-Ähnlichkeitsmaß (vgl. Abschnitt 3.3.3). Die zweite Variante bildet das ungewichtete arithmetische Mittel beider Werte. Die dritte Variante gewichtet die Ähnlichkeit der Benutzerbewertungen doppelt so stark wie die der Filterbotbewertungen. Abbildung 4.21 zeigt, dass die erste Variante trotz des vermeintlichen Fehlers am besten geeignet und die Hit-Ratio dieser Variante mit der von Empfehlungen ohne Filterbots vergleichbar ist, also trotz der Vorteile der Filterbots bei neuen Artefakten und der Integration der Artikelstruktur in den CF keine Verschlechterung eintritt. Eine mögliche Erklärung, warum die beiden anderen Varianten schlechtere Ergebnisse als die erste Variante liefern, könnte die zu hohe Gewichtung der Filterbotbewertungen im Vergleich mit denen der Benutzer in der Ähnlichkeitsberechnung sein (1:1 bzw. 2:1, bei der ersten Variante sind diese dynamisch bzw. abhängig von der Anzahl der Benutzerbewertungen).

Insgesamt sind in Abbildung 4.22 die einzelnen Stufen allein und in verschiedenen Kombinationen dargestellt. Isoliert betrachtet sind der CF- und der AC-Ansatz (Hit-Ratio: 0,02 resp. 0,03) deutlich schlechter als der AVG-Ansatz (Hit-Ratio: ca. 0,12) oder andere Kombinationen (Hit-Ratio von CF+AVG: ca. 0,09, Hit-Ratio von CF+AVG: ca. 0,12). Lediglich der AVG-Ansatz sticht mit einer Hit-Ratio von ca. 0,12 heraus. Der Grund dafür könnte die vorliegende Community mit sehr vielen Benutzern sein, die nur sehr wenige Artikel bewertet bzw. gelesen haben. Erst durch die Kombination aller drei Ansätze kann eine noch höhere Hit-Ratio von ca. 0,13 erreicht werden.

Da die Hit-Ratio-Metrik sich darauf beschränkt zu messen, ob bereits gut bewertete Artefakte aus der Testmenge wieder empfohlen werden, gibt sie keine Auskunft über die Qualität der anderen empfohlenen Artefakte. Darüber hinaus ist nicht unbedingt klar, ob ein Benutzer bereits sehr kleine Verbesserungen der Güte bemerkt bzw. ob dies die Zufriedenheit erhöht. Sicherlich ist dies situationsabhängig, denn bei der Erweiterung des Algorithmus' um die Alterserkennung für die Kinder eines Benutzers verbesserte sich die Hit-Ratio nur sehr marginal, jedoch war der Einfluss auf die Empfehlungen

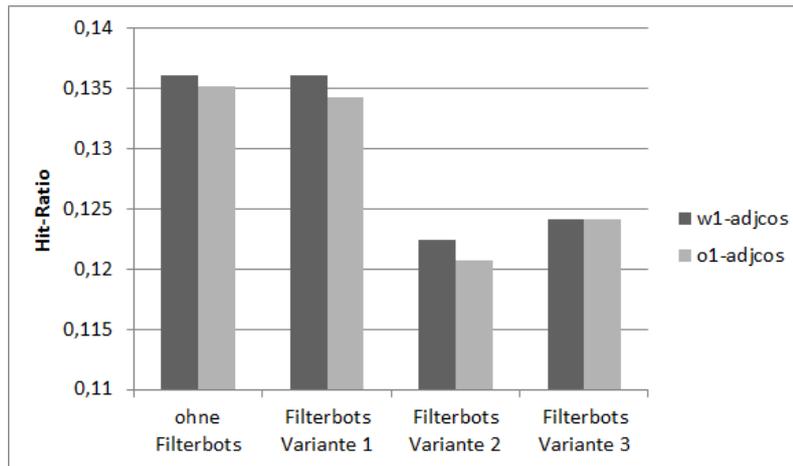


Abbildung 4.21: Hit-Ratio in Abhängigkeit unterschiedlicher Filterbot-Implementationen

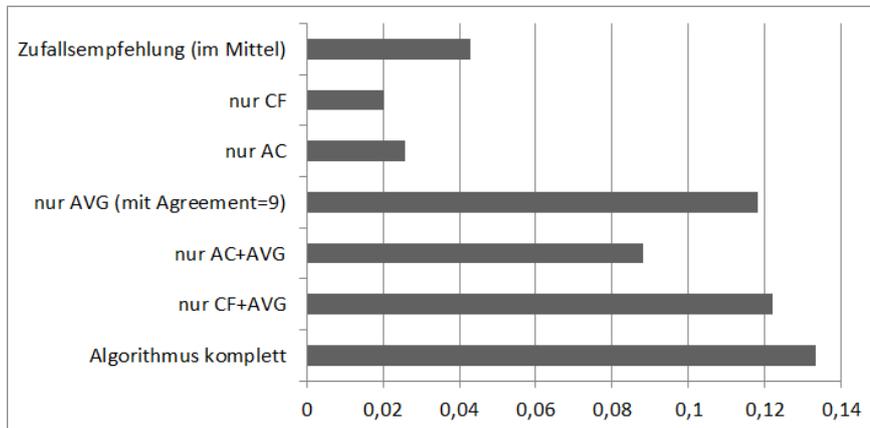


Abbildung 4.22: Hit-Ratios der einzelnen Verfahren und Kombinationen

deutlich sichtbar. Insgesamt konnte die Güte der Empfehlungen durch die Wahl der Parameter und Verfahren von ca. 0,02 auf ca. 0,13 um das 6,5-fache gesteigert werden (vgl. Abbildung 4.14, Seite 64). Hierdurch wurden während der Testläufe deutlich sichtbare Veränderungen bei den Empfehlungen hervorgerufen. Mit einer genaueren Beurteilung der Empfehlungen beschäftigt sich das nächste Kapitel.

4.6 Übertragbarkeit des Ansatzes

Dieser Ansatz ist durch die Mehrstufigkeit sehr flexibel einsetzbar: Er ist nicht auf kleine Online-Communities beschränkt, sondern kann auch für wachsende bzw. größere Communities eingesetzt werden, bei denen die Aufgabe „Find Good Items“ im Vordergrund steht. Bei ausreichender Datenmenge wird die Zurückschaltung auf den

durchschnittsbasierten bzw. den α -Community-Ansatz seltener vollzogen, da das IB-Verfahren wahrscheinlich deutlich häufiger eine vollständige Empfehlung generiert. Jedoch können die tieferen Stufen auch bei neuen Benutzern („New User“-Problem) aktiv werden und so Empfehlungen für diese ermöglichen. Die Filterbots, α -Communities, Regeln für regelbasierte Empfehlungen und Parameter müssen ggf. an die speziellen Gegebenheiten (kein Artikelbaum bzw. andere Struktur, unterschiedliche Datenfelder des Benutzerprofils, keine Präsenzveranstaltungen) angepasst oder erweitert werden. Bei größeren Communities kommen weitere Faktoren, wie Performanzoptimierungen, hinzu und gewinnen je nach Größe immer mehr an Gewicht.

Die wichtigste Voraussetzung für die Anwendbarkeit dieses Ansatzes wurde bereits im Abschnitt 4.1 genannt: Es muss eine Mindestanzahl von Bewertungen vorliegen, so dass die zugrunde liegenden Verfahren eine ausreichende Datenbasis vorfinden, auf der sie operieren können (Sparsity-Level, vgl. Abschnitt 3.10). Daneben gelten natürlich die allgemeinen Voraussetzungen aus Abschnitt 3.1: Die Artefakte sind hinreichend homogen und subjektiv bewertbar, werden von mehr als nur einem Benutzer bewertet, weisen keine zu große Fluktuation auf, die Benutzer bewerten mehr als nur ein Artefakt und die Vorlieben verändern sich nicht zu schnell.

Konkret lässt sich der Ansatz auf viele kleinere Online-Communities übertragen (u. a. Vereine oder Fan-Clubs), die eine Online-Plattform betreiben: Einerseits könnte beispielsweise ein Kino-Verein, der auf seiner Webseite Beschreibungen und evtl. Trailer von Filmen anbietet, ein ES etablieren, in dem Vorführungen bzw. Filme empfohlen werden. Dazu könnten explizite Bewertungen der Filme, Anzahl der Trailer-Wiedergaben oder Weiterempfehlungen und auch Verweildauern auf Film-Beschreibungsseiten genutzt werden. In diesem Kontext ist es sogar möglich, Empfehlungen für Filmvorführungen anhand der online vorgenommenen Bewertungen zu generieren und die Empfehlungen durch die Hinzunahme von Wohnorten zu verbessern (z. B. mit Hilfe der Annahme: je besser die vorhergesagte Bewertung, desto eher ist ein Mitglied geneigt, eine größere Entfernung bis zum Vorführungsort zurückzulegen). Durch Filterbots ist es auch hier möglich, inhaltliche Aspekte, wie z. B. Genre, Regisseur oder Zugehörigkeit zu Reihen, in den kollaborativen Filter zu integrieren. Als α -Communities würden sich z. B. Geschlecht, Alter, Genre oder Angaben zu favorisierten Hauptdarstellern bzw. Regisseuren anbieten. Die Empfehlungen könnten über die Startseite oder per E-Mail-Newsletter versandt werden. Gleiches gilt z. B. auch für Buch-Clubs mit Lesungen.

Andererseits ist der Ansatz nicht auf die Kombination von Online- und Offline-Artefakten angewiesen: Eine Online-Lyrik-Community, in der Benutzer Gedichte veröffentlichen und bewerten, könnte den mehrstufigen Ansatz verwenden, um Gedichte anderen Nutzern möglichst personalisiert zu empfehlen. Für die Generierung der Empfehlungen bieten sich die Anzahl von Weiterempfehlungen oder Aufrufen (allgemein oder über Lesezeichen) und Verweildauern als weitere implizite Bewertungsquellen an. α -Communities könnten erneut auf demographischen Daten, wie Alter, Geschlecht oder bevorzugten Gattungen bzw. Formen, basieren. Sollen Filterbots eingesetzt werden, liegt es nahe, Gedichte z. B. hinsichtlich der Länge oder Gattung automatisch bewerten zu lassen (Sarwar u. a. 1998). E-Mail-Newsletter, als weiterer Kanal zur Präsentation der Empfehlungen neben der Webseite, sind ebenfalls möglich.

5 Evaluation

In diesem Abschnitt wird die Evaluation des in Abschnitt 4.2 beschriebenen Empfehlungssystems vorgestellt.

Im *Mobile2Learn*-Kontext wurde eine mehrstufige Evaluation des Empfehlungssystem-Prototyps durchgeführt, um ein möglichst genaues Bild von der Güte der Empfehlungen zu erhalten:

1. Offline-Evaluation
2. Expertenbefragung zur Bewertung der fachlichen Güte der Empfehlungen
3. Feldstudie mit der *Mobile2Learn*-Community

Im Rahmen der Evaluation steht die Nutzer-Zufriedenheit im Mittelpunkt. Gleichzeitig ist diese aber auch am aufwändigsten zu ermitteln. Daneben soll die pädagogische Güte der Empfehlungen als Online-Bildungsplattform nicht vernachlässigt werden. Da es keine Möglichkeit gibt, beide Aspekte simultan zu betrachten, wurde dieses mehrstufige Vorgehen gewählt. Jede einzelne Stufe bewertet unterschiedliche Aspekte, weist aber auch charakteristische Schwächen auf:

Die Offline-Evaluation bewertet automatisiert die Empfehlungsgüte gemäß einer gewählten Metrik und dient dazu sicherzustellen, dass die Empfehlungen des Algorithmus', basierend auf bekannten Daten, im Schnitt eine höhere Güte aufweisen als eine zufällige Auswahl von Artefakten. Sie ist aber stark abhängig von der verwendeten Metrik und kann keine Aussagen zur Nutzer-Zufriedenheit treffen.

Die Expertenbefragung hingegen wird mit Menschen durchgeführt, die die fachliche bzw. pädagogische Güte der Empfehlungen einschätzen. Sie basiert jedoch auf Meinungen von einer kleinen Gruppe von Experten zu einem Ausschnitt der Community.

Die Feldstudie dient schließlich zur Bewertung der Nutzer-Zufriedenheit. Sie stützt sich auf Rückmeldungen von Benutzern zu einzelnen Empfehlungen und ist folglich auf eine Vielzahl von Rückmeldungen unterschiedlicher Benutzer angewiesen. Auf Grund der *Mobile2Learn*-Zielgruppe war ungewiss, wie viele Feedbacks abgegeben werden.

Es gibt kein einzelnes Verfahren, das alle diese Aspekte kombiniert und bei dieser Zielgruppe einsetzbar ist. Insgesamt sollen sich die einzelnen Stufen gegenseitig ergänzen, um ein möglichst aussagekräftiges Gesamtbild zu erhalten. Im Folgenden ist jeder Stufe jeweils ein eigener Abschnitt gewidmet.

5.1 Offline-Evaluation

Die Offline-Evaluation wurde analog der Methode zur Parameterfindung (Kreuzvalidierung, vgl. Abschnitte 4.5) durchgeführt. Da diese Methode, wie im Abschnitt 4.5.1 bereits erwähnt, kein absolutes Gütemaß darstellt, kann sie nur vergleichend genutzt werden. Daher wurde der Algorithmus mit einer zufälligen Auswahl von ungelesenen Artefakten gemäß der Metrik verglichen.

Es wurden ca. 80 Auswertungen, bestehend aus jeweils 25 Kreuzvalidierungen pro Benutzer und Algorithmus bzw. Zufallsempfehlung (vgl. Abschnitt 4.5.2), auf zufälligen Test- und Trainingsmengen durchgeführt. Die Anzahl der Auswertungen resultiert daher, dass der Evaluation keine feste Anzahl von Auswertungszyklen vorgegeben und die Ausführung nach etwa sieben Tagen beendet wurde. Die über alle Auswertungen gemittelten Hit-Ratios (Abbildung 5.1) zeigen, dass die berechneten Empfehlungen (Hit-Ratio: ca. 0,13) im Schnitt besser sind als eine rein zufällige Auswahl von Artefakten (Hit-Ratio: ca. 0,04). Zudem wird ersichtlich, dass die Ergebnisse aus dem Parameteroptimierungsabschnitt (4.5.2) keine (deutlichen) Effekte einer Überspezialisierung aufweisen.

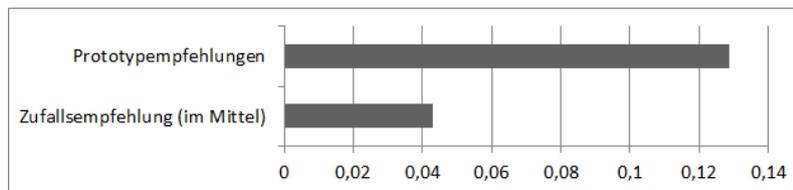


Abbildung 5.1: Hit-Ratio-Vergleich zwischen Prototyp und Zufallsempfehlungen

Darüber hinaus wurden die Hit-Ratios für drei Benutzerklassen, basierend auf deren Leseverhalten („wenige Artikel gelesen“ < 5 , „viele Artikel gelesen“ > 20 sowie „mittel“ dazwischen), untersucht (vgl. Abbildung 5.2). Für alle drei Benutzerklassen bewegen sich die Hit-Ratios im Mittel über denen der Zufallsempfehlungen. Abbildung 5.2 lässt erkennen, dass die Hit-Ratios für Wenigleser besonders hoch sind ($\approx 0,2$, doppelt so hoch wie die der Vielleser) und das Hit-Ratio für Vielleser unter dem Gesamtschnitt liegt. Eine mögliche Erklärung für das hohe Hit-Ratio für Wenigleser besteht darin, dass für viele Benutzer dieser Klasse die Testmenge leer ist, diese somit nicht in die Hit-Ratio eingehen und andere Benutzer dieser Klasse mit wenigen bewerteten Artefakten der Mehrheitsmeinung entsprechen. Speziell für die Vielleser gilt, dass das Hit-Ratio lediglich die Empfehlungsrate von relevanten Artefakten der Testmenge bestimmt und keine Aussage über die Güte der anderen empfohlenen Artefakte geben kann, die auf Grund des genaueren Profils ausgewählt wurden und zu den unerwarteten bzw. Cross-Category-Empfehlungen zählen könnten.

Ein weiteres wichtiges Maß ist die *Coverage* (Abdeckung, vgl. Sarwar u. a. 2001). Sie gibt an, wie viele der Artefakte empfohlen werden können. Theoretisch liegt die Coverage für den Prototyp nahe 100 %, da für alle Artefakte Bewertungen vorliegen und diese somit empfohlen werden können. Praktisch weist die Coverage einen Wert zwischen 60 % und 100 % auf. Für neue Benutzer, über die dem System keine Informationen vorliegen,

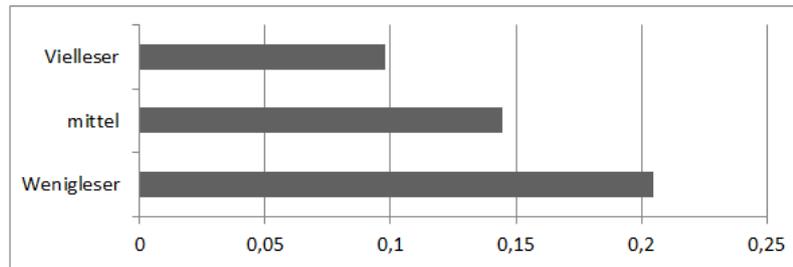


Abbildung 5.2: Hit-Ratios für verschiedene Benutzerklassen

beträgt die Coverage ca. 60 % (vgl. Abbildung 4.17 auf Seite 66), für alle anderen liegt die Coverage darüber. Am höchsten (bis zu 100 %) ist die Abdeckung für Benutzer, die bereits einige Artefakte bewertet haben, da für diese Benutzer alle Artefakte über Ähnlichkeiten beim IB- bzw. AC-Verfahren vorgeschlagen werden können.

5.2 Expertenbefragung

Nach der positiven Offline-Evaluation wurde vor der Feldstudie eine Expertenbefragung zur Bewertung der fachlichen bzw. pädagogischen Güte der Empfehlungen durchgeführt. Dazu wurden vier Mitglieder des pädagogischen Teams von *Mobile2Learn* mit einem Fragebogen nach Ihrer persönlichen Einschätzung befragt.

Bei den Fragebögen handelte es sich konkret um zwölf verschiedene Benutzerprofile zusammen mit jeweils drei unterschiedlichen Empfehlungsvarianten. Die einzelnen Benutzerprofile enthielten jeweils die folgenden Daten eines *Mobile2Learn*-Benutzers in anonymisierter, graphisch aufbereiteter Form (Beispiel im Anhang):

- Wohnort
- Mitgliedsstatus (Netzwerkpartner oder Eltern)
- Alter
- bekannte Informationen über Kinder (Geschlecht und Alter)
- Auflistung besuchter Präsenzveranstaltungen
- Aufgerufene Seiten (zusammen mit einer Bewertung, der Anzahl der Aufrufe und Verweildauer (kurz: < 15 Sekunden, mittel und lang: > 60 Sekunden))

Die Auswahl der Profile erfolgte zufällig, wobei diese möglichst repräsentativ für die *Mobile2Learn*-Community sein sollten. Die zugrunde gelegten Eigenschaften sind Leseverhalten („wenige Artikel gelesen“ < 5, „viele Artikel gelesen“ > 20 sowie „mittel“ dazwischen), Wohnort (innerhalb des Landkreises Goslar oder nicht), Besuch von Präsenzveranstaltungen und Angaben zu Kindern im Profil (Kinder im Alter 0 bis 6 Jahre angegeben vs. keine Kinder angegeben). Diese Klassifikation kann als Baum dargestellt werden (siehe Abbildung 5.3, inkl. der Anzahl der Benutzer, auf die diese Eigenschaften zutreffen), wobei Teilbäume abgeschnitten wurden, in denen nicht beide Kindknoten

mindestens sieben Benutzer enthalten. Aus jedem Blatt wurde ein Benutzer (sowie zusätzlich ein Netzwerkpartner) für die Evaluation zufällig bestimmt:

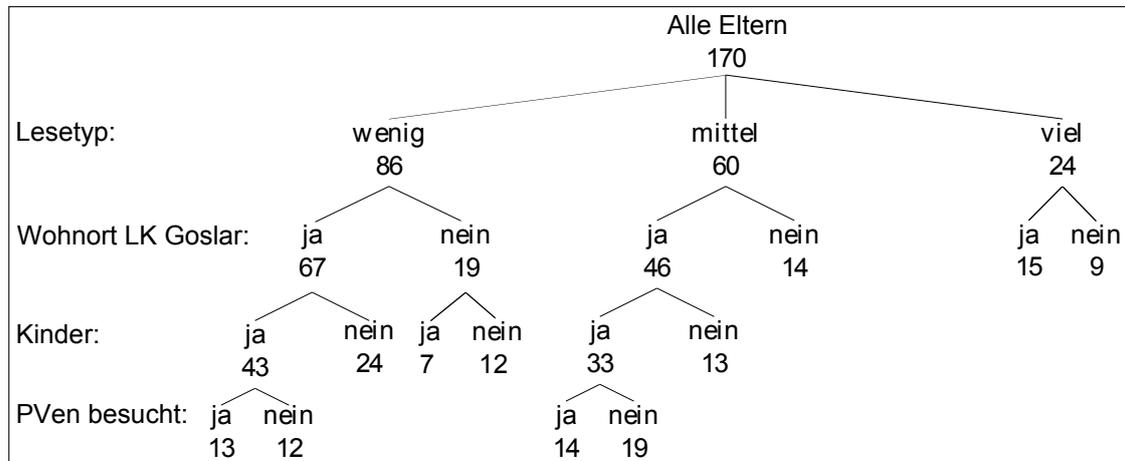


Abbildung 5.3: Expertenbefragung: Klassifikation der Benutzer

Bei den drei Empfehlungsvarianten handelte es sich jeweils um eine Liste mit fünf Empfehlungen (TOP-5-Ranking). Zwei Listen wurden mit fünf zufällig ausgewählten Artefakten erstellt, die der Benutzer noch nicht abgerufen hatte. Die dritte Liste entsprach der Ausgabe des in Kapitel 4 vorgestellten Empfehlungssystems. Die Reihenfolge der drei Listen wurde zufällig mit der Nebenbedingung bestimmt, dass die Ausgabe des Algorithmus' gleich oft auf allen Positionen vertreten ist, um der Bevorteilung eines Musters bei der Bewertung entgegenzuwirken.

Nach einer kurzen Erklärung der Profile an einem Beispiel wurden die Fragebögen in unterschiedlicher Reihenfolge ausgeteilt, damit gerade am Anfang kein Experte durch die Wahl des Sitznachbarn beeinflusst werden konnte. Daneben erhielten die Experten eine Auflistung der zehn im Mittel am besten bewerteten Artikel. Anhand dieser Informationen wurden die Experten für jedes der zwölf Profile gebeten, die ihrer Meinung nach beste TOP-5-Empfehlung auszuwählen.

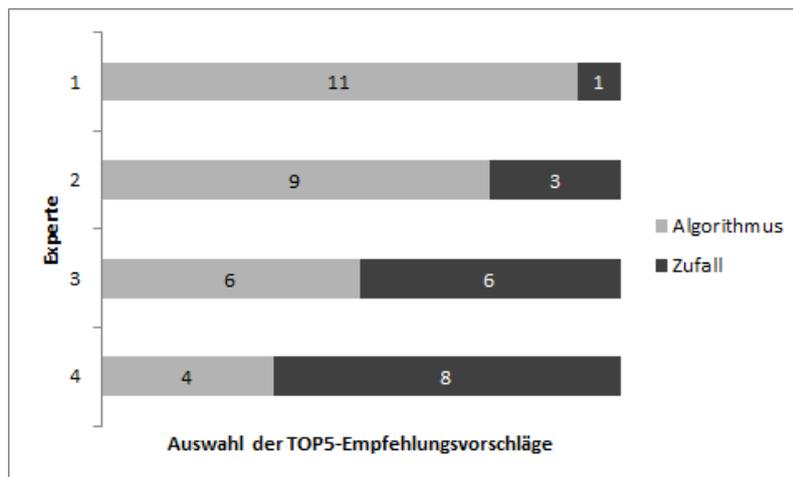


Abbildung 5.4: Auswertung der Expertenbefragung

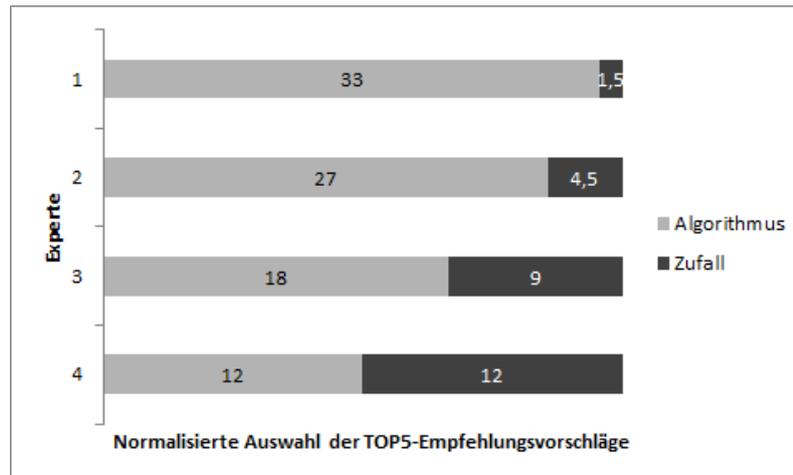


Abbildung 5.5: Normalisierte Auswertung der Expertenbefragung

Abbildung 5.4 und 5.5 zeigen die Ergebnisse der Expertenbefragung absolut resp. normalisiert bzgl. der Auswahlwahrscheinlichkeiten. Experte 1 und 2 bevorzugten klar die Empfehlungen des Systems. Experte 3 tendiert ebenfalls zu den Empfehlungen des Systems, da nach dem Studiendesign die Wahrscheinlichkeit für die Wahl einer Zufallsempfehlung überwiegt. Lediglich Experte 4 scheint keine klare Präferenz zu haben.

In Tabelle 5.1 sind die aggregierten Meinungen der Experten für jedes einzelne Profil dargestellt. Darüber hinaus ist ersichtlich zu welchem Typ die Profile gehören (wenig, mittel oder viele Artikel gelesen bzw. Netzwerkpartner), wie viele Präsenzveranstaltungen der Benutzer bisher besucht hat, wie viele Kinder der Benutzer angegeben hat, ob er aus dem Landkreis Goslar stammt und wie die genauen Wahlen der Experten aussehen. „Oben“, „mitte“ und „unten“ beziehen sich auf die Positionen der drei Empfehlungslisten. Die durch das Empfehlungssystem generierten TOP-5-Empfehlungen sind fett auf grauem Grund dargestellt.

Profil	Typ	aus LK GS	Kinder	PV	oben	mitte	unten
1	wenig	ja	2	1		4	
2	wenig	ja	n/a	keine	4		
3	wenig	nein	1	keine			4
4	wenig	ja	1	1	1	2	1
5	wenig	nein	n/a	keine		4	
6	wenig	ja	1	keine	1		3
7	mittel	ja	n/a	keine	3	1	
8	mittel	ja	1	keine		1	3
9	mittel	nein	1	keine	1	1	2
10	viel	ja	2	2	2	1	1
11	viel	nein	1	keine	1	1	2
12	Partner	ja	nein		1	1	2

Tabelle 5.1: Ergebnisse der Expertenbefragung

Bei jedem Profil entschied sich mindestens ein Experte für die generierten Empfehlungen. Es ist erkennbar, dass sich die Experten genau viermal einig waren. In allen Fällen handelte es sich um die durch den Prototyp generierten Empfehlungen. Lediglich einmal, bei Profil 62, stimmten drei Experten für die gleiche durch den Zufall generierte Liste.

Bei den Profilen, die wenige Inhalte abgerufen haben, gibt es die meisten Übereinstimmungen der Experten mit den durch den Prototyp generierten Empfehlungen. Dieser Umstand könnte auf die kleine vorliegende Stichprobe der Profile mit mittel bzw. vielen abgerufenen Artikeln oder die Empfehlungen selbst zurückgehen. Denn gerade bei den Viellesern weisen diese durch die Personalisierung eine besondere Komplexität auf, die sehr schwer zu bewerten ist (z. B. auf Grund von Cross-Category-Empfehlungen).

Direkt im Anschluss an die papiergebundene Befragung wurde mit den Experten ein offenes Interview geführt. Die Mehrheit der Experten gab an, dass es nicht einfach gewesen sei, sich in die einzelnen Profile hinein zu versetzen und viele Entscheidungen für eine der drei Listen subjektiv „aus dem Bauch heraus“ getroffen wurden. Zudem wurden sowohl aus dem Interview als auch aus der Auswertung teilweise sehr unterschiedliche Ansichten bei den Experten deutlich, wann Foto-Galerien empfohlen werden sollen: Die Hälfte der Experten war der Meinung, dass Foto-Galerien auch bzw. speziell für Benutzer empfohlen werden sollen, die bisher keine einzige Präsenzveranstaltung besucht haben. Dies könnte die häufigere Wahl des Zufalls der Experten 3 und 4 erklären, da bei den Zufallsempfehlungen die Galerien deutlich häufiger auftauchten als bei den algorithmisch generierten. Die andere Hälfte der Experten sah dies anders und präferierte, dass Galerien möglichst nur für Veranstaltungen, die der Benutzer besucht hatte, empfohlen werden sollen. Interessant war auch, dass nur ein einziger Experte (Experte 4) bei Profil 4 die generierte Empfehlungsvariante gewählt hat, in welcher der Benutzer speziell auf eine neue Veranstaltung in seiner Nähe hingewiesen wurde, obwohl alle Experten im Gespräch angaben, dass ihnen gerade diese Art von Empfehlungen besonders wichtig sei.

Insgesamt scheint es eine Tendenz der Experten zu den Empfehlungen des Prototyps zu geben, es ist jedoch zu beachten, dass es sich um eine kleine Stichprobe handelt.

5.3 Feldstudie

Um die Güte eines Empfehlungssystems beurteilen zu können, ist die Nutzer-Zufriedenheit das wichtigste Kriterium (McNee u. a. 2006; Schafer u. a. 2007). Diese ist am aufwändigsten zu bestimmen, da sie nicht über Expertenbefragungen oder Offline-Evaluationen ermittelt werden kann, sondern die einzelnen Benutzer befragt werden müssen.

Am 25. August 2011 wurde das Empfehlungssystem auf *Mobile2Learn.de* freigeschaltet und über eine Kampagne der Community vorgestellt. Damit wurde die Feldstudie gestartet. Erste Empfehlungen wurden direkt mit der ankündigenden Kampagne verschickt. Zusätzlich werden stets aktuelle Empfehlungen für angemeldete Benutzer auf der Startseite angezeigt. Zur Evaluation der Empfehlungen wurde auf der Seite, die

empfohlen wurde und die ein Benutzer angeklickt hat, eine Feedback-Abfrage eingebaut (vgl. Abbildung 5.6). Diese Box hat eine feste Position in der rechten unteren Ecke des Browsers und ragt als Störer zur Aufmerksamkeitsgenerierung leicht in den Text hinein. In der Kampagne wurden die Mitglieder gebeten, Feedback über diese Box abzugeben, um „das Empfehlungssystem weiter zu verbessern“. Es wird angenommen, dass so mehr Benutzer Feedback abgeben, da Rückmeldungen auch für sie einen Vorteil (bessere Empfehlungen) bedeuten. Von einer Belohnung für eine Teilnahme wurde abgesehen, da bereits ähnliche Versuche in der Community nicht sehr fruchtbar waren.



Abbildung 5.6: Screenshot einer abgerufenen Empfehlung mit Feedbackmöglichkeit

Am späten Abend des 25. September 2011 wurde ein Schnappschuss der Evaluation durchgeführt, der im Rahmen dieser Arbeit vorgestellt wird:

Insgesamt wurden 45 Empfehlungen von 29 Benutzern angenommen. 21 empfohlene Artefakte (ca. 49 %) wurden von 18 eindeutigen Benutzern über zwei in diesem Zeitraum versandte Newsletter (an 189 bzw. 199 Mitglieder) abgerufen, die restlichen über die Startseite von *Mobile2Learn.de*. Vier Benutzer riefen weitere Empfehlungen von der Startseite ab, nachdem sie über eine Newsletter-Empfehlung auf die *Mobile2Learn.de*-Plattform gelangt sind. Zusätzlich wurden 14 SMS-Empfehlungen versandt. Wie viele Benutzer durch eine Empfehlung per SMS die *Mobile2Learn.de*-Startseite aufgerufen haben, lässt sich nicht (eindeutig) ermitteln, jedoch haben zwei Benutzer zeitnah *Mobile2Learn.de* besucht, von denen einer die (gleiche) Empfehlung über die Startseite abgerufen hat. Insgesamt haben 17 verschiedene Benutzer 20 Rückmeldungen abgegeben. Die Rückmeldungen waren ausschließlich positiver Natur. Über die Meinungen

der wahrgenommenen Empfehlungen ohne Rückmeldung lässt sich keine Aussage treffen (auch wenn Benutzer i. d. R. negatives Feedback seltener abgeben). Es ist nicht nachvollziehbar, ob diese Benutzer explizit keine Rückmeldungen geben wollten, die Feedbackbox nicht gesehen haben oder ihnen die Empfehlung bzw. die Seite gefallen bzw. nicht gefallen hat. Jedenfalls hat sich bei einer Stichprobe gezeigt, dass auch Mitglieder, die keine Rückmeldung gegeben haben, sich z. B. eine empfohlene Galerie ausgiebig angeschaut oder weitere Empfehlungen abgerufen haben.

Die relativ kleine Anzahl von 17 Benutzern (ca. 9 % der Community) entspricht in etwa der Größe der Kern-Community, die regelmäßig auf die Kampagnen reagiert. Jedoch sind unter diesen 17 sechs neue Benutzer, die sich erst nach dem Beginn der Feldstudie registriert haben.

Die folgende Tabelle 5.2 zeigt die Anzahl von eindeutigen Benutzern, die Anzahl der abgerufenen Empfehlungen und die Anzahl der Rückmeldungen in Abhängigkeit der Lesetypen (wenig, mittel, viel). Zusätzlich enthält die Tabelle für die Wenigleser eine Aufschlüsselung nach Benutzern, die zum Start der Feldstudie bereits registriert waren (alt) und neu registrierte Benutzer (neu). Es ist ersichtlich, dass bei den Lesetypen *wenig* und *mittel* etwa die Hälfte der abgerufenen Empfehlungen bewertet wurden; bei dem Lesetyp *viel* ein Drittel (ein Benutzer hat sechs Empfehlungen angenommen ohne zu einer einzigen Feedback zu geben, die anderen haben mindestens zu einer Empfehlung Feedback abgegeben). Bei den Lesetypen *wenig* und *mittel* haben die Benutzer im Schnitt ca. 1,3 Empfehlungen angenommen; bei dem Lesetyp *viel* 3,25. Anhand der positiven Rückmeldungen lässt sich eine Tendenz ableiten, dass für alle Lesetypen gute Empfehlungen generiert werden konnten.

Lesetyp	# eind. Benutzer	# Artefakte abgerufen	# Rückmeldungen
wenig (gesamt)	17	22	12
wenig (alt)	11	14	7
wenig (neu)	6	8	5
mittel	8	10	4
viel	4	13	4

Tabelle 5.2: Feldstudie: Anzahl der Abrufe empfohlener Artefakte und Rückmeldungen in Abhängigkeit vom Lesetyp

Unter den Empfehlungen befanden sich alle Artefakttypen (Präsenzveranstaltungen, Galerien und Artikel) sowie alle Empfehlungstypen (regel-, durchschnitts-, α -Community- sowie CF-basiert). Die Rückmeldungen zu regelbasierten Empfehlungen betragen ca. 35 % des Feedbacks (darunter eine Terminempfehlung und für besuchte Präsenzveranstaltungen fünf Galerieempfehlungen sowie ein Einleitungsartikel), so dass die getroffenen Annahmen für Regeln (vgl. Abschnitt 4.1) sicherlich nicht falsch waren, sondern sehr wohl einen Nutzen für die rückmeldenden Benutzer hatten.

Die Abfrage von empfohlenen Artefakten des TOP-5-Rankings beschränkte sich nicht auf die ersten Artefakte der geordneten Liste. Jedoch ist eine klare Präferenz für die ersten Empfehlungen des Rankings auszumachen (vgl. Abschnitt 4.1): Allein 42 % der abgerufenen Artefakte wurden als erstes Element der Empfehlung angezeigt, 62 % als die ersten zwei und 73 % als die ersten drei Elemente.

Insgesamt gab es in der Feldstudie nur sehr wenige Rückmeldungen. Dieser Umstand ist sicherlich auch auf die spezielle Zielgruppe zurückzuführen. Dafür handelte es sich aber ausschließlich um positive Meinungen von Benutzern aller drei Lesetypen.

5.4 Zusammenfassung

Zur Klärung der Fragestellung, ob Empfehlungssysteme in kleinen Online-Communities eingesetzt werden können bzw. zur Bestimmung der Güte von Empfehlungen, wurde eine mehrstufige Evaluation durchgeführt. Jede einzelne Stufe der Evaluation zeigt eine klare Tendenz für die Güte der Empfehlungen des vorgestellten Ansatzes.

Ein Vergleich der durchschnittlich von einem Benutzer gelesenen Artikel vor und nach der Feldstudie zeigt eine leichte Steigerung (von 8,9 auf 9,2, trotz der Registrierung zehn neuer Benutzer). Es hat sich in Stichproben gezeigt, dass die (abgerufenen) Empfehlungen bei einigen Benutzern dafür gesorgt haben, dass erst neue Inhalte entdeckt und schließlich in diesen oder verwandten Themengebieten weitere Artikel abgerufen wurden. Gestützt wird dieses dadurch, dass im Rahmen der Feldstudie neun registrierte Besucher einer Präsenzveranstaltung Inhalte (direkt oder indirekt) über eine Empfehlung abgerufen haben, die nicht mit den besuchten Veranstaltungen assoziiert sind.

Anhand der Ergebnisse lässt sich grundsätzlich feststellen, dass Empfehlungssysteme auch in kleinen Online-Communities eingesetzt werden können, sofern man das Empfehlungssystem hinreichend an den Kontext anpasst. Ohne den mehrstufigen Evaluationsansatz hätte diese Aussage nicht getroffen werden können.

Daneben zeigt die Evaluation, wie schwer es ist, ein ES zu beurteilen: Für die Offline-Evaluation muss erst eine passende Metrik als Maß der Empfehlungsgüte gefunden werden (vgl. Abschnitt 4.5.1). Diese Aufgabe ist nicht trivial, speziell bei einer kleinen Datenbasis. Bei der Experten-Befragung hat sich gezeigt, dass diese oftmals sehr abhängig von den subjektiven Meinungen der einzelnen Experten ist. Schließlich ist man bei einer Feldstudie sehr von der Feedback-Bereitschaft der Probanden abhängig.

6 Ausblick

Bei kleinen Online-Communities ist die Anzahl der Artefakte deutlich geringer als bei großen. Wenn ein Benutzer bereits sehr viele Artefakte bewertet hat, könnte sich das negativ auf die Qualität der Empfehlungen auswirken, da Empfehlungen dann lediglich aus einer noch kleineren Menge von Artefakten gewählt werden können. Durch die Vorhersage der Bewertungen für einen aktiven Benutzer werden zuerst die Artefakte mit sehr hohen Bewertungen vorgeschlagen, später auch Artefakte, die eine schlechtere vorhergesagte Bewertung (aber trotzdem mindestens von $t_{\min. \text{Bewertung}}$) aufweisen. Im Kontext von *Mobile2Learn* ist nicht klar, wie sich dieses mögliche Problem auswirkt, da Eltern wahrscheinlich nur eine begrenzte Zeit (solange sie Kinder im Alter von 0 bis etwa 6 Jahren haben) in der Community aktiv sind. Unabhängig davon bietet dieses mögliche Problem die Grundlage für weitere Untersuchungen in einer kleinen Community mit einem etablierten Empfehlungssystem.

Bewertungen für Artefakte, die in der Vergangenheit getätigt wurden, werden von den vorgestellten Verfahren als konstant angesehen. Jedoch verändern sich die Vorlieben der Benutzer im Allgemeinen mit der Zeit (Ding u. Li 2005; Tang u. a. 2005), so dass Benutzer in der Vergangenheit bereits bewertete Artefakte später anders bewerten (Hill u. a. 1995). Im Kontext dieser Arbeit ist vor auszusehen, dass sich die Vorlieben der Eltern z. B. mit dem Alter der Kinder verändern: Haben Eltern in der Vergangenheit Artikel mit *sehr gut* bewertet, die für ein Kind im Alter von drei Jahren interessant sind, könnten die Eltern diese Artikel jetzt schlechter bewerten, da diese für ein fünfjähriges Kind nicht mehr attraktiv sind. Ding u. Li (2005) schlugen vor, ältere Bewertungen abzuwerten. Andere Systeme nutzen ausschließlich eine feste Anzahl der neuesten Bewertungen (vgl. Tang u. a. 2005), um Empfehlungen zu generieren (und um gleichzeitig die Datenmengen zu reduzieren). Ob eine solche Erweiterung im Laufe der Zeit zusätzlich zur bisherigen Altersbestimmung hilfreich ist bzw. wie sie implementiert und evaluiert werden kann, öffnet Raum für weitere Forschung.

In den Abschnitten 4.2.1 und 4.5.2 wurden verschiedene Filterbot-Typen vorgestellt und getestet. Statt eine feste Gewichtung (von $1/2$ bzw. $1/3$) für die Filterbot-Ähnlichkeiten im Vergleich zu den Artefakt-Ähnlichkeiten (von $1/2$ bzw. $2/3$) vorzugeben, könnte eine weitere Möglichkeit mit dynamischen Gewichtungen (z. B. $\frac{\text{AnzahlFilterbot-Bewertungen}}{\text{GesamtanzahlBewertungen}}$ für die Filterbot-Ähnlichkeiten) bessere Ergebnisse liefern, da damit ein höheres Gewicht auf den Benutzer- als auf den Filterbotbewertungen liegt. Dieser Ansatz konnte aus Zeitgründen nicht mehr evaluiert werden.

Die Parameter des Prototyps wurden auf der Gesamtmenge der Benutzer gewählt bzw. optimiert. Zur weiteren Verbesserung der Empfehlungsgüte könnten für die unterschiedlichen Lesetypen verschiedene Parameterbelegungen gewählt werden, um so jeden Typ individuell möglichst gut unterstützen zu können.

Im weiteren Einsatz könnte bei der TOP-5-Empfehlungsliste ein Fall auftreten, in dem alle fünf empfohlenen Artefakte aus dem gleichen Ast des Artikelbaumes stammen bzw. sogar Nachbarknoten darstellen. Zur Steigerung der Diversität würde sich eine Begrenzung der benachbarten bzw. aus demselben Ast stammenden Artefakte anbieten.

Anknüpfend an den gerade genannten Punkt besteht die Möglichkeit, den Benutzern zur Steigerung der Zufriedenheit bei der Präsentation der Empfehlungen eine Option anzubieten, um einen uninteressanten Themenbereich komplett aus der Empfehlung auszuschließen, so dass Eltern, deren Kinder nicht musikalisch sind, z. B. alle Artikel der „musikalischen Früherziehung“ ausblenden können.

7 Zusammenfassung

Im Rahmen dieser Masterarbeit wurden Empfehlungssysteme in kleinen Online-Communities mit regionalem Bezug untersucht.

Das erste Kapitel beleuchtete die Begriffe Community und Online-Community. Anhand eines Vergleiches zweier verschiedener Online-Communities wurde gezeigt, wie solche Communities in der Praxis aussehen und welche deutlichen Unterschiede sie durch ihren spezifischen Kontext aufweisen können. Dabei wurde die *Mobile2Learn*-Community vorgestellt, die im Rahmen dieser Arbeit eine zentrale Rolle als Prototyp- und Evaluationsumgebung spielt.

Im zweiten Kapitel wurden Empfehlungssysteme definiert und grundlegende Verfahren beschrieben. Neben der reinen Vorstellung der Konzepte wurden allgemeine Probleme, auch speziell für den Einsatz in einer kleinen Online-Community, und Lösungsmöglichkeiten erläutert.

Basierend darauf wurde anschließend ein Ansatz für ein hybrides Empfehlungssystem im Kontext von *Mobile2Learn* präsentiert, das aus mehreren Stufen besteht, die möglichst personalisierte Empfehlungen generieren. Exemplarisch wurden Anforderungen, mögliche Verfahren, spezifische Probleme sowie Entwurfsentscheidungen dargelegt und eine übertragbare Architektur des Prototyps vorgestellt. Neben üblichen expliziten Bewertungen werden auch domänen-spezifische Daten der *Mobile2Learn*-Community (Verweildauern auf einzelnen Seiten der Online-Plattform, Artikelstrukturen, Präsenzveranstaltungsbesuche, Wohnorte und Angaben über eigene Kinder) mit in den Empfehlungsalgorithmus integriert, um die Generierung von Empfehlungen für Artikel und Galerien zu ermöglichen bzw. zu verbessern.

Das Empfehlungssystem beschränkt sich nicht nur auf das Empfehlen von Artefakten, die lediglich innerhalb des Systems existieren, sondern integriert auch domänen-spezifische Empfehlungen für Präsenzveranstaltungen oder damit verbundene Artefakte. Neben der reinen Berechnung von Empfehlungen ist die Präsentationsform ein weiterer wichtiger Aspekt, denn Empfehlungen müssen einen Benutzer nicht nur erreichen, sondern auch sein Interesse wecken. Dazu wird neben der etablierten Präsentation eines TOP-5-Rankings auf der Startseite auch auf eine Einbettung in E-Mail-Kampagnen und SMS-Nachrichten gesetzt.

Dieser Ansatz wurde als Prototyp für die Web-Plattform von *Mobile2Learn* implementiert und evaluiert. Die Evaluation erfolgte mehrstufig, um neben einer Offline-Evaluation verschiedene Aspekte wie fachliche Güte und Nutzerzufriedenheit der Empfehlungen überprüfen bzw. messen zu können. Es hat sich eine klare Tendenz sowohl hinsichtlich der fachlichen Güte als auch der Nutzerzufriedenheit der Empfehlungen gezeigt, die durch den vorgestellten Ansatz generiert wurden. Ferner haben einige Eltern durch die Empfehlungen neue Inhalte entdeckt und in diesen bzw. verwandten

Themenbereichen weitere Artikel abgerufen. Bei der Beschreibung des Evaluationsvorgehens wurde speziell für die Offline-Evaluation die Schwierigkeit der Wahl einer angemessenen Metrik zur Messung der Empfehlungsgüte bei einer kleinen Online-Community herausgestellt.

Zentrales Ergebnis dieser Arbeit ist die Erkenntnis, dass Empfehlungssysteme auch in kleinen Online-Communities eingesetzt werden können. Der Prototyp innerhalb der *Mobile2Learn*-Community bleibt deshalb weiterhin im Einsatz. Wichtig ist, dass Empfehlungssysteme speziell an den Kontext des Einsatzortes angepasst werden müssen und die Generierung von Empfehlungen nur durch die Kombination verschiedener Verfahren und Datenquellen möglich ist. Mit dem Prototyp wurde eine Grundlage geschaffen, diesen Bereich weiter untersuchen zu können.

Literaturverzeichnis

Balabanović 1997

BALABANOVIĆ, Marko: An adaptive Web page recommendation service. In: *Proceedings of the first international conference on Autonomous agents*. New York, NY, USA: ACM, 1997 (AGENTS '97). – ISBN 0–89791–877–0, S. 378–385

Basu u. a. 1998

BASU, Chumki; HIRSH, Haym; COHEN, William: Recommendation as Classification: Using Social and Content-Based Information in Recommendation. In: *In Proceedings of the Fifteenth National Conference on Artificial Intelligence*, AAAI Press, 1998, S. 714–720

Bauer u. Bittlingmayer 2005

BAUER, Ullrich; BITTLINGMAYER, Uwe H.: Wer profitiert von Elternbildung? In: *ZSE: Zeitschrift für Soziologie der Erziehung und Sozialisation* (2005), S. 263–280

Berners-Lee u. a. 1996

BERNERS-LEE, T.; FIELDING, R.; FRYSTYK, H.: *Hypertext Transfer Protocol – HTTP/1.0*. RFC 1945 (Informational). <http://www.ietf.org/rfc/rfc1945.txt>. Version: Mai 1996 (Request for Comments)

Bhushan 1971

BHUSHAN, A.K.: *File Transfer Protocol*. RFC 114. <http://www.ietf.org/rfc/rfc114.txt>. Version: April 1971 (Request for Comments). – Updated by RFCs 133, 141, 171, 172

Blanchard 2004

BLANCHARD, Anita: The Effects of Dispersed Virtual Communities on Face-to-Face Social Capital. In: HUYSMAN, Marleen (Hrsg.); WULF, Volker (Hrsg.): *Social Capital and Information Technology*. Cambridge, MA, USA: MIT Press, 2004. – ISBN 0–262–08331–0, S. 53–73

Boim u. a. 2011

BOIM, Rubi; MILO, Tova; NOVGORODOV, Slava: DiRec: Diversified recommendations for semantic-less Collaborative Filtering. In: *Proceedings of the 2011 IEEE 27th International Conference on Data Engineering*. Washington, DC, USA: IEEE Computer Society, 2011 (ICDE '11). – ISBN 978–1–4244–8959–6, S. 1312–1315

Burke 2000

BURKE, Robin: Knowledge-based recommender systems. In: DEKKER, Marcel (Hrsg.): *Encyclopedia of Library and Information Systems* Bd. 69. New York, NY, USA, 2000, 180–200

Burke 2002

BURKE, Robin: Hybrid Recommender Systems: Survey and Experiments. In: *User*

Modeling and User-Adapted Interaction 12 (2002), S. 331–370. <http://dx.doi.org/10.1023/A:1021240730564>. – DOI 10.1023/A:1021240730564. – ISSN 0924–1868

Burke 2007

BURKE, Robin: Hybrid Web Recommender Systems. Version: 2007. http://dx.doi.org/10.1007/978-3-540-72079-9_12. In: BRUSILOVSKY, Peter (Hrsg.); KOBISA, Alfred (Hrsg.); NEJDL, Wolfgang (Hrsg.): *The Adaptive Web* Bd. 4321. Berlin, Heidelberg: Springer Berlin/Heidelberg, 2007. – DOI 10.1007/978-3-540-72079-9_12. – ISBN 978-3-540-72078-2, Kapitel 12, S. 377–408

Candillier u. a. 2007

CANDILLIER, Laurent; MEYER, Frank; BOULLÉ, Marc: Comparing State-of-the-Art Collaborative Filtering Systems. In: *Proceedings of the 5th international conference on Machine Learning and Data Mining in Pattern Recognition*. Berlin, Heidelberg: Springer-Verlag, 2007 (MLDM '07). – ISBN 978-3-540-73498-7, S. 548–562

Carotenuto u. a. 1999

CAROTENUTO, Linda; ETIENNE, William; FONTAINE, Michael; FRIEDMAN, Jessica; MULLER, Michael; NEWBERG, Helene; SIMPSON, Matthew; SLUSHER, Jason; STEVENSON, Kenneth: Community Space: Toward Flexible Support for Voluntary Knowledge Communities. (1999), April, Nr. 99-04

Carroll u. Rosson 1998

CARROLL, John M.; ROSSON, Mary B.: Network communities, community networks. In: *CHI 98 conference summary on Human factors in computing systems*. New York, NY, USA: ACM, 1998 (CHI '98). – ISBN 1-58113-028-7, S. 121–122

Cebrián u. a. 2007

CEBRIÁN, Manuel; ALFONSECA, Manuel; ORTEGA, Alfonso: The Normalized Compression Distance Is Resistant to Noise. In: *IEEE Transactions on Information Theory* 53 (2007), Nr. 5, 1895–1900. <http://dx.doi.org/10.1109/TIT.2007.894669>. – DOI 10.1109/TIT.2007.894669

Cilibrasi u. Vitanyi 2003

CILIBRASI, Rudi; VITANYI, Paul: Clustering by Compression. In: *IEEE Transactions on Information Theory* 51 (2003), S. 1523–1545

Cleverdon u. a. 1966

CLEVERDON, C. W.; MILLS, J.; KEEN, M.: Factors determining the performance of indexing systems. (1966). <https://dspace.lib.cranfield.ac.uk/handle/1826/863>

Cosley u. a. 2003

COSLEY, Dan; LAM, Shyong K.; ALBERT, Istvan; KONSTAN, Joseph A.; RIEDL, John: Is seeing believing?: how recommender system interfaces affect users' opinions. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA: ACM, 2003 (CHI '03). – ISBN 1-58113-630-7, S. 585–592

Deshpande u. Karypis 2004

DESHPANDE, Mukund; KARYPIS, George: Item-based top-N recommendation algorithms. In: *ACM Trans. Inf. Syst.* 22 (2004), January, S. 143–177. <http://dx.doi.org/10.1145/963770.963776>. – DOI 10.1145/963770.963776. – ISSN 1046–8188

Ding u. Li 2005

DING, Yi; LI, Xue: Time weight collaborative filtering. In: *Proceedings of the 14th ACM international conference on Information and knowledge management*. New York, NY, USA: ACM, 2005 (CIKM '05). – ISBN 1–59593–140–6, S. 485–492

Ding u. a. 2006

DING, Yi; LI, Xue; ORLOWSKA, Maria E.: Recency-based collaborative filtering. In: *Proceedings of the 17th Australasian Database Conference - Volume 49*. Darlinghurst, Australia: Australian Computer Society, Inc., 2006 (ADC '06). – ISBN 1–920682–31–7, 99–107

Ganesan u. a. 2003

GANESAN, Prasanna; GARCIA-MOLINA, Hector; WIDOM, Jennifer: Exploiting hierarchical domain structure to compute similarity. In: *ACM Trans. Inf. Syst.* 21 (2003), January, S. 64–93. <http://dx.doi.org/10.1145/635484.635487>. – DOI 10.1145/635484.635487. – ISSN 1046–8188

Goldberg u. a. 1992

GOLDBERG, David; NICHOLS, David A.; OKI, Brian M.; TERRY, Douglas B.: Using Collaborative Filtering to Weave an Information Tapestry. In: *Commun. ACM* 35 (1992), Nr. 12, S. 61–70. <http://dx.doi.org/10.1145/138859.138867>. – DOI 10.1145/138859.138867

Gross u. Koch 2007

GROSS, Tom; KOCH, Michael: *Computer-Supported Cooperative Work (Interaktive Medien)*. Oldenbourg, 2007

Gupta u. Kim 2004

GUPTA, Sumeet; KIM, Hee-Wong: Virtual Community: Concepts, Implications, and Future Research Directions. In: *Proceedings of the 10th American Conference on Information Systems* (2004), August, S. 2679–2687

Hafner u. Lyon 1998

HAFNER, Katie; LYON, Matthew: *Where wizards stay up late. The origins of the Internet*. New York: Touchstone, 1998

Hartleb 2009

HARTLEB, Vivian: *Brand Community Management: eine empirische Analyse am Beispiel der Automobilbranche*, Universität Münster, Diss., 2009

Hartung u. a. 2009

HARTUNG, Joachim; ELPELT, Bärbel; KLÖSENER, Karl-Heinz: *Statistik – Lehr- und Handbuch der angewandten Statistik (15. Aufl.)*. München: Oldenbourg Wissenschaftsverlag, 2009. – ISBN 978–3–486–59028–9

Herlocker u. a. 1999

HERLOCKER, Jonathan L.; KONSTAN, Joseph A.; BORCHERS, Al; RIEDL, John: An algorithmic framework for performing collaborative filtering. (1999), S. 230–237. <http://dx.doi.org/10.1145/312624.312682>. – DOI 10.1145/312624.312682. ISBN 1581130961

Herlocker u. a. 2000

HERLOCKER, Jonathan L.; KONSTAN, Joseph A.; RIEDL, John: Explaining collaborative filtering recommendations. In: *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. New York, NY, USA: ACM, 2000 (CSCW '00). – ISBN 1-58113-222-0, S. 241–250

Herlocker u. a. 2004

HERLOCKER, Jonathan L.; KONSTAN, Joseph A.; TERVEEN, Loren G.; RIEDL, John T.: Evaluating collaborative filtering recommender systems. In: *ACM Trans. Inf. Syst.* 22 (2004), Januar, Nr. 1, S. 5–53. <http://dx.doi.org/10.1145/963770.963772>. – DOI 10.1145/963770.963772. – ISSN 1046–8188

Hill u. a. 1995

HILL, Will; STEAD, Larry; ROSENSTEIN, Mark; FURNAS, George: Recommending and evaluating choices in a virtual community of use. In: *Proceedings of the SIG-CHI conference on Human factors in computing systems*. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1995 (CHI '95). – ISBN 0-201-84705-1, S. 194–201

Hillery 1955

HILLERY, George A.: Definitions of community: Areas of Agreement. In: *Rural Sociology* 20 (1955), Nr. 2, S. 111–123

Horton 1983

HORTON, M.R.: *Standard for interchange of USENET messages*. RFC 850. <http://www.ietf.org/rfc/rfc850.txt>. Version: Juni 1983 (Request for Comments). – Obsoleted by RFC 1036

Hu u. a. 2008

HU, Yifan; KOREN, Yehuda; VOLINSKY, Chris: Collaborative filtering for implicit feedback datasets. In: *In IEEE International Conference on Data Mining (ICDM 2008, 2008, S. 263–272*

Huang u. a. 2007

HUANG, Zan; ZENG, Daniel; CHEN, Hsinchun: A Comparison of Collaborative-Filtering Recommendation Algorithms for E-commerce. In: *IEEE Intelligent Systems* 22 (2007), September, S. 68–78. <http://dx.doi.org/10.1109/MIS.2007.80>. – DOI 10.1109/MIS.2007.80. – ISSN 1541–1672

Ishida 1998

ISHIDA, Toru: *Community Computing: Collaboration over Global Information Networks*. John Wiley & Sons, 1998. – ISBN 0471979651

Jones u. Furnas 1986

JONES, W. P.; FURNAS, G. W.: Pictures of Relevance: A Geometric Analysis of Similarity Measures. In: *Journal of the American Society of Information Science* 38 (1986), Mai, Nr. 6, S. 420–442

Koch 2003

KOCH, Michael: *Community-Unterstützungssysteme – Architektur und Interoperabilität*, Technische Universität München, Habilitation, 2003

Konstan u. a. 1998

KONSTAN, J. A.; RIEDL, J.; BORCHERS, A.; HERLOCKER, J. L.: Recommender Systems: A GroupLens Perspective. In: *Recommender Systems. Papers from 1998 Workshop. Technical Report WS-98* Bd. 8, AAAI Press, 1998, S. 60–64

Konstan u. a. 1997

KONSTAN, Joseph A.; MILLER, Bradley N.; MALTZ, David; HERLOCKER, Jonathan L.; GORDON, Lee R.; RIEDL, John: GroupLens: applying collaborative filtering to Usenet news. In: *Commun. ACM* 40 (1997), March, S. 77–87. <http://dx.doi.org/10.1145/245108.245126>. – DOI 10.1145/245108.245126. – ISSN 0001–0782

Lahres u. Rayman 2006

LAHRES, Bernhard; RAYMAN, Gregor: *Praxisbuch Objektorientierung. Professionelle Entwurfsverfahren*. Galileo Computing, 2006 <http://openbook.galileodesign.de/oo/>. – ISBN 978–3–89842–624–4

Lee u. a. 2003

LEE, Fion S. L.; VOGEL, Douglas; LIMAYEM, Moez: Virtual community informatics: A review and research agenda. In: *Journal of Information Technology Theory and Application JITTA* 5 (2003), Nr. 1, S. 47–61

Lee u. Park 2006

LEE, Tong-Queue; PARK, Young: A Time-Based Recommender System Using Implicit Feedback. In: *CSREA EEE*, 2006, S. 309–315

Li u. a. 2003

LI, Ming; CHEN, Xin; LI, Xin; MA, Bin; VITÁNYI, Paul: The similarity metric. In: *SODA '03: Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2003. – ISBN 0–89871–538–5, S. 863–872

Li u. Kim 2003

LI, Qing; KIM, Byeong M.: An approach for combining content-based and collaborative filters. In: *Proceedings of the sixth international workshop on Information retrieval with Asian languages - Volume 11*. Stroudsburg, PA, USA: Association for Computational Linguistics, 2003 (AsianIR '03), S. 17–24

Licklider u. Taylor 1968

LICKLIDER, J. C. R.; TAYLOR, Robert W.: The computer as a communication device. In: *Science and Technology* 76 (1968), S. 21–31

Lodhi u. a. 2002

LODHI, Huma; SAUNDERS, Craig; SHAW-TAYLOR, John; CRISTIANINI, Nello; WATKINS, Chris: Text classification using string kernels. In: *J. Mach. Learn. Res.* 2 (2002), March, S. 419–444. <http://dx.doi.org/10.1162/153244302760200687>. – DOI 10.1162/153244302760200687. – ISSN 1532–4435

Maltz u. Ehrlich 1995

MALTZ, David; EHRLICH, Kate: Pointing the way: active collaborative filtering. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1995 (CHI '95). – ISBN 0–201–84705–1, S. 202–209

Marmanis u. Babenko 2009

MARMANIS, Haralambos; BABENKO, Dmitry: *Algorithms of the Intelligent Web*. 1st. Greenwich, CT, USA: Manning Publications Co., 2009. – ISBN 9781933988665

McNee u. a. 2002

MCNEE, Sean M.; ALBERT, Istvan; COSLEY, Dan; GOPALKRISHNAN, Prateep; LAM, Shyong K.; RASHID, Al M.; KONSTAN, Joseph A.; RIEDL, John: On the recommending of citations for research papers. In: *Proceedings of the 2002 ACM conference on Computer supported cooperative work*. New York, NY, USA: ACM, 2002 (CSCW '02). – ISBN 1-58113-560-2, S. 116-125

McNee u. a. 2006

MCNEE, Sean M.; RIEDL, John; KONSTAN, Joseph A.: Being accurate is not enough: how accuracy metrics have hurt recommender systems. In: *CHI '06 extended abstracts on Human factors in computing systems*. New York, NY, USA: ACM, 2006 (CHI EA '06). – ISBN 1-59593-298-4, S. 1097-1101

Melville u. a. 2002

MELVILLE, Prem; MOONEY, Raymod J.; NAGARAJAN, Ramadass: Content-boosted collaborative filtering for improved recommendations. In: *Eighteenth national conference on Artificial intelligence*. Menlo Park, CA, USA: American Association for Artificial Intelligence, 2002. – ISBN 0-262-51129-0, 187-192

Middleton u. a. 2002

MIDDLETON, Stuart E.; ALANI, H.; SHADBOLT, Nigel R.; DE ROURE, David C.: Exploiting Synergy Between Ontologies and Recommender System. In: *Proceedings of the 11th International World Wide Web Conference WWW-2002*, 2002

Morita u. Shinoda 1994

MORITA, Masahiro; SHINODA, Yoichi: Information filtering based on user behavior analysis and best match text retrieval. In: *Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval*. New York, NY, USA: Springer-Verlag New York, Inc., 1994 (SIGIR '94). – ISBN 0-387-19889-X, 272-281

Mynatt u. a. 1997

MYNATT, Elizabeth D.; ADLER, Annette; ITO, Mizuko; O'DAY, Vicki L.: Design for network communities. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA: ACM, 1997 (CHI '97). – ISBN 0-89791-802-9, S. 210-217

Nguyen u. a. 2007

NGUYEN, An-Te; DENOS, Nathalie; BERRUT, Catherine: Improving new user recommendations with rule-based induction on cold user data. In: *Proceedings of the 2007 ACM conference on Recommender systems*. New York, NY, USA: ACM, 2007 (RecSys '07). – ISBN 978-1-59593-730-8, S. 121-128

Nichols 1997

NICHOLS, David M.: Implicit Rating and Filtering. In: *In Proceedings of the Fifth DELOS Workshop on Filtering and Collaborative Filtering*, 1997, S. 31-36

Oard u. Kim 1998

OARD, Douglas; KIM, Jinmook: Implicit Feedback for Recommender Systems. In: *in Proceedings of the AAAI Workshop on Recommender Systems*, 1998, S. 81–83

Page u. a. 1999

PAGE, Lawrence; BRIN, Sergey; MOTWANI, Rajeev; WINOGRAD, Terry: The PageRank Citation Ranking: Bringing Order to the Web. / Stanford InfoLab. Version: November 1999. <http://ilpubs.stanford.edu:8090/422/>. Stanford InfoLab, November 1999 (1999-66). – Technical Report. – Previous number = SIDL-WP-1999-0120

Park u. a. 2006

PARK, Seung-Taek; PENNOCK, David; MADANI, Omid; GOOD, Nathan; DECOSTE, Dennis: Naïve filterbots for robust cold-start recommendations. In: *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*. New York, NY, USA: ACM, 2006 (KDD '06). – ISBN 1–59593–339–5, S. 699–705

Pazzani u. Billsus 2007

PAZZANI, Michael J.; BILLSUS, Daniel: The adaptive web. Version: 2007. <http://portal.acm.org/citation.cfm?id=1768197.1768209>. Berlin, Heidelberg: Springer-Verlag, 2007. – ISBN 978–3–540–72078–2, Kapitel Content-based recommendation systems, 325–341

Postel 1981

POSTEL, J.: *Simple Mail Transfer Protocol*. RFC 788. <http://www.ietf.org/rfc/rfc788.txt>. Version: November 1981 (Request for Comments). – Obsoleted by RFC 821

Preece 2000

PREECE, Jenny: *Online Communities: Designing Usability and Supporting Socialbilty*. 1st. New York, NY, USA: John Wiley & Sons, Inc., 2000 <http://portal.acm.org/citation.cfm?id=517688>. – ISBN 0471805998

Rashid u. a. 2002

RASHID, Al M.; ALBERT, Istvan; COSLEY, Dan; LAM, Shyong K.; MCNEE, Sean M.; KONSTAN, Joseph A.; RIEDL, John: Getting to know you: learning new user preferences in recommender systems. In: *Proceedings of the 7th international conference on Intelligent user interfaces*. New York, NY, USA: ACM, 2002 (IUI '02). – ISBN 1–58113–459–2, S. 127–134

Resnick u. a. 1994

RESNICK, Paul; IACOVOU, Neophytos; SUCHAK, Mitesh; BERGSTROM, Peter; RIEDL, John: GroupLens: an open architecture for collaborative filtering of netnews. In: *Proceedings of the 1994 ACM conference on Computer supported cooperative work*. New York, NY, USA: ACM, 1994 (CSCW '94). – ISBN 0–89791–689–1, S. 175–186

Runte 2000

RUNTE, M.: *Personalisierung im Internet - Individualisierte Angebote mit Collaborative Filtering*, Christian-Albrechts-Universität zu Kiel, Diss., 2000. http://www.runte.de/matthias/publications/personalisierung_im_internet.pdf

Rupp 2003

RUPP, Marina: *Niederschwellige Familienbildung*. Bamberg: ifb Materialien. Staatsinstitut für Familienforschung an der Universität Bamberg (ifb), 2003

Sarwar u. a. 2001

SARWAR, Badrul; KARYPIS, George; KONSTAN, Joseph; REIDL, John: Item-based collaborative filtering recommendation algorithms. In: *Proceedings of the 10th international conference on World Wide Web*. New York, NY, USA: ACM, 2001 (WWW '01). – ISBN 1-58113-348-0, S. 285–295

Sarwar u. a. 2000a

SARWAR, Badrul; KARYPIS, George; KONSTAN, Joseph; RIEDL, John: *Analysis of Recommendation Algorithms for E-Commerce*, ACM Press, 2000, S. 158–167

Sarwar u. a. 2000b

SARWAR, Badrul M.; KARYPIS, George; KONSTAN, Joseph A.; RIEDL, John T.: Application of Dimensionality Reduction in Recommender System – A Case Study. In: *IN ACM WEBKDD WORKSHOP*, 2000

Sarwar u. a. 1998

SARWAR, Badrul M.; KONSTAN, Joseph A.; BORCHERS, Al; HERLOCKER, Jon; MILLER, Brad; RIEDL, John: Using filtering agents to improve prediction quality in the GroupLens research collaborative filtering system. In: *Proceedings of the 1998 ACM conference on Computer supported cooperative work*. New York, NY, USA: ACM, 1998 (CSCW '98). – ISBN 1-58113-009-0, S. 345–354

Schafer u. a. 2007

SCHAFFER, J. B.; FRANKOWSKI, Dan; HERLOCKER, Jon; SEN, Shilad: *The adaptive web*. Version: 2007. <http://portal.acm.org/citation.cfm?id=1768197.1768208>. Berlin, Heidelberg: Springer-Verlag, 2007. – ISBN 978-3-540-72078-2, Kapitel Collaborative filtering recommender systems, 291–324

Schafer u. a. 1999

SCHAFFER, J. B.; KONSTAN, Joseph; RIEDL, John: Recommender Systems in E-Commerce. In: *Proceedings of the 1st ACM conference on Electronic commerce*. New York, NY, USA: ACM, 1999 (EC '99). – ISBN 1-58113-176-3, S. 158–166

Segaran 2008

SEGARAN, Toby: *Kollektive Intelligenz: Analysieren, programmieren und nutzen*. Beijing: O'Reilly, 2008. – ISBN 978-3-89721-780-5

Shardanand u. Maes 1995

SHARDANAND, Upendra; MAES, Pattie: Social information filtering: algorithms for automating „word of mouth“. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1995 (CHI '95). – ISBN 0-201-84705-1, S. 210–217

Sinha u. a. 2001

SINHA, Rashmi; ; SINHA, Rashmi; SWEARINGEN, Kirsten: Comparing Recommendations Made by Online Systems and Friends. In: *DELOS Workshop: Personalisation and Recommender Systems in Digital Libraries*, 2001

Sinha u. Swearingen 2002

SINHA, Rashmi; SWEARINGEN, Kirsten: The role of transparency in recommender systems. In: *CHI '02 extended abstracts on Human factors in computing systems*. New York, NY, USA: ACM, 2002 (CHI EA '02). – ISBN 1-58113-454-1, S. 830-831

Sollenborn u. Funk 2002

SOLLENBORN, Mikael; FUNK, Peter: Category-Based Filtering and User Stereotype Cases to Reduce the Latency Problem in Recommender Systems. In: *Proceedings of the 6th European Conference on Advances in Case-Based Reasoning*. London, UK, UK: Springer-Verlag, 2002 (ECCBR '02). – ISBN 3-540-44109-3, 395-420

Stöhr 2007

STÖHR, Ottmar: Elternbildung und Beratung in Kindertagesstätten in Armutswohngebieten. Kritischer Praxisüberblick, Sensibilisierung und Vorschläge zur Veränderung. In: *Verhaltenstherapie mit Kindern & Jugendlichen. Zeitschrift für die psychosoziale Praxis* 3. Jg. (2007), Nr. 2, S. 113-134

Strickroth u. a. 2011

STRICKROTH, Sven; PINKWART, Niels; MÜLLER, Jörg P.: Neue Medien und Präsenzveranstaltungen: Ein didaktisches Modell für die Elternbildung? In: FRIEDRICH, Steffen (Hrsg.); KIENLE, Andrea (Hrsg.); ROHLAND, Holger (Hrsg.): *DeLFI 2011: Die 9. e-Learning Fachtagung Informatik*. Dresden, Germany: TUDpress, 2011. – ISBN 978-3-942710-36-7

Tang u. a. 2005

TANG, Tiffany; WINOTO, Pinata; CHAN, Keith: Scaling Down Candidate Sets Based on the Temporal Feature of Items for Improved Hybrid Recommendations. In: MOBASHER, Bamshad (Hrsg.); ANAND, Sarabjot (Hrsg.): *Intelligent Techniques for Web Personalization* Bd. 3169. Springer Berlin/Heidelberg, 2005. – ISBN 978-3-540-29846-5, S. 169-186

Tintarev u. Masthoff 2007

TINTAREV, Nava; MASTHOFF, Judith: A Survey of Explanations in Recommender Systems. In: *Proceedings of the 2007 IEEE 23rd International Conference on Data Engineering Workshop*. Washington, DC, USA: IEEE Computer Society, 2007. – ISBN 978-1-4244-0831-3, 801-810

Vozalis u. Margaritis 2004

VOZALIS, Manolis; MARGARITIS, Konstantinos G.: Collaborative filtering enhanced by demographic correlation. In: *Proceedings of the AIAI Symposium on Professional Practice in AI, part of the 18th World Computer Congress, 2004*, S. 293-402

Wenger 1998

WENGER, Etienne: Communities of practice: Learning as a social system. In: *Systems Thinker* (1998). http://ewenger.com/pub/pub_systems_thinker_wrd.doc

Wenger u. a. 2002

WENGER, Etienne; MCDERMOTT, Richard; SNYDER, William M.: *Cultivating Communities of Practice*. Boston, USA: Harvard Business School Press, 2002

Whittaker u. a. 1997

WHITTAKER, Steve; ISAACS, Ellen; O'DAY, Vicki: Widening the net: workshop

report on the theory and practice of physical and network communities. In: *SIGCHI Bull.* 29 (1997), July, S. 27–30. <http://dx.doi.org/10.1145/264853.264867>. – DOI 10.1145/264853.264867. – ISSN 0736–6906

Woerndl u. a. 2007

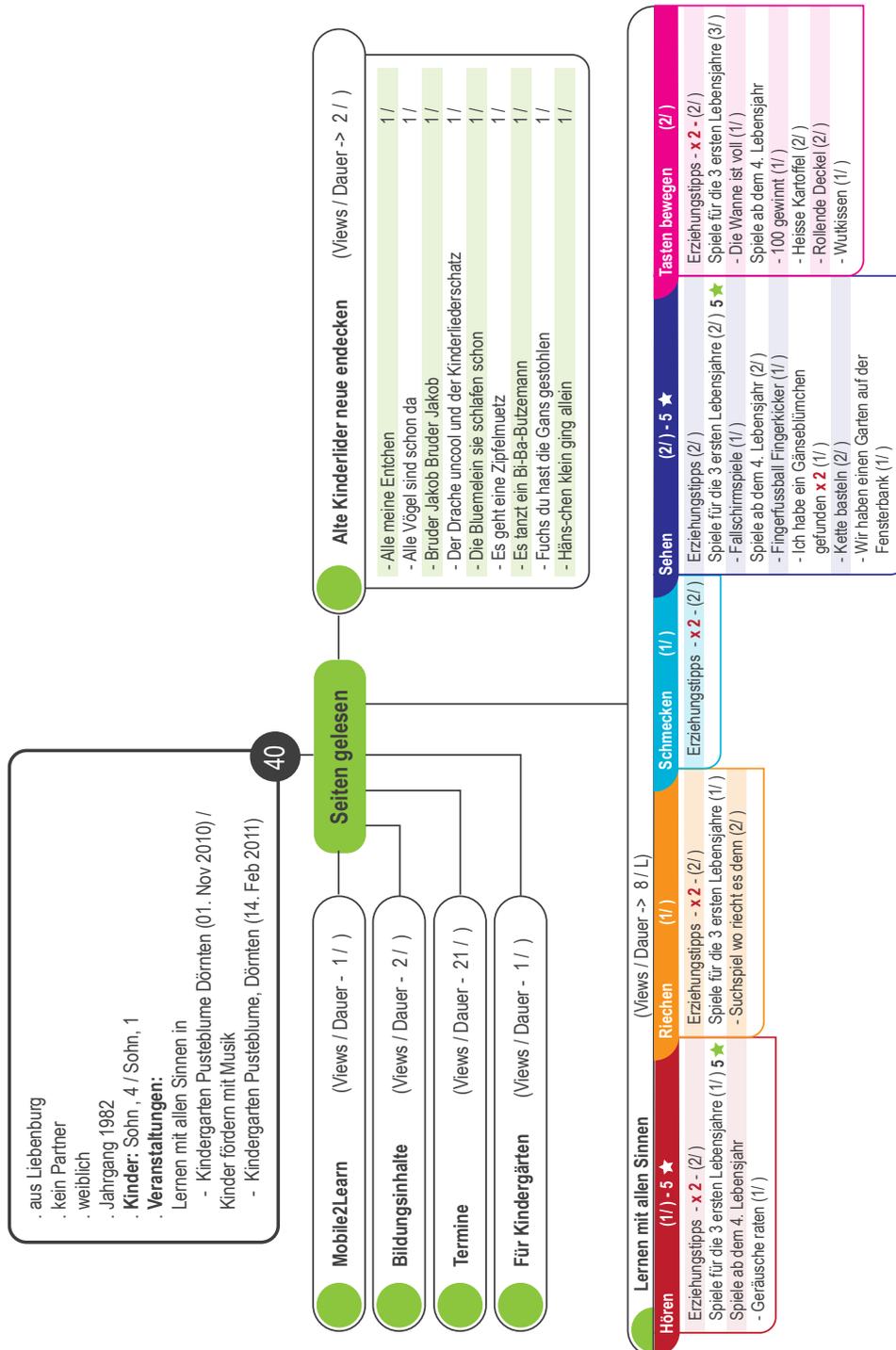
WOERNDL, Wolfgang; SCHUELLER, Christian; WOJTECH, Rolf: A Hybrid Recommender System for Context-aware Recommendations of Mobile Applications. In: *Proceedings of the 2007 IEEE 23rd International Conference on Data Engineering Workshop*. Washington, DC, USA: IEEE Computer Society, 2007. – ISBN 978–1–4244–0831–3, S. 871–878

Ziegler u. a. 2005

ZIEGLER, Cai-Nicolas; MCNEE, Sean M.; KONSTAN, Joseph A.; LAUSEN, Georg: Improving recommendation lists through topic diversification. In: *Proceedings of the 14th international conference on World Wide Web*. New York, NY, USA: ACM, 2005 (WWW '05). – ISBN 1–59593–046–9, S. 22–32

Anhang

Anhang 1: Beispielfragebogen der Expertenbefragung



Anhang 2: Ehrenwörtliche Erklärung

E R K L Ä R U N G

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, dass alle Stellen der Arbeit, die wörtlich oder sinngemäß aus anderen Quellen übernommen wurden, als solche kenntlich gemacht sind, und dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegt wurde.

Desweiteren erkläre ich mich damit einverstanden, dass die vorliegende Arbeit in der Universitätsbibliothek bzw. in der Bibliothek des Instituts für Informatik der TU Clausthal ausgelegt und zur Einsichtnahme aufbewahrt werden darf.

Clausthal-Zellerfeld, den 23. Oktober 2011