Wissenschaftliche Recommendersysteme in Virtuellen Universitäten

Andreas Geyer-Schulz[1], Michael Hahsler[2] und Maximillian Jahn[3]

[1]Informationdienste und elektronische Märkte
Universität Karlsruhe (TH)
D-76128 Karlsruhe
[2,3] Informationswirtschaft
Wirtschaftsuniversität Wien
Augasse 2-6, A-1090 Wien
Andreas.Geyer-Schulz@em.uni-karlsruhe.de
Michael.Hahsler@wu-wien.ac.at
Maximillian.Jahn@wu-wien.ac.at

Abstract: In diesem Beitrag wird die Rolle von Recommendersystemen und ihr Potential in der Lehr-, Lern- und Forschungsumgebung einer Virtuellen Universität untersucht. Die Hauptidee dieses Beitrags besteht darin, die Informationsaggregationsfähigkeiten von Recommendersystemen in einer Virtuellen Universität auszunutzen, um Tutoren- und Beratungsdienste in einer Virtuellen Universität automatisch zu verbessern, um damit Betreuung und Beratung von Studierenden zu personalisieren und für eine größere Anzahl von Teilnehmern bei gleichzeitiger Entlastung der Lehrenden verfügbar zu machen. Im zweiten Teil dieses Beitrags werden die Recommenderdienste von myVU, der Sammlung der personalisierten Dienste der Virtuellen Universität (VU) der Wirtschaftsuniversität Wien und ihre nicht-personalisierten Varianten beschrieben, die im Wesentlichen auf beobachtetem Benutzerverhalten und, in der personalisierten Variante, zusätzlich auf Selbstselektion durch Selbsteinschätzung der Erfahrung in einem Fachgebiet beruhen. Anschließend wird gezeigt, daß diese Varianten eine erfolgversprechende Lösung der Mechanism Design Probleme, die alle Recommendersysteme inhärent haben, darstellen. Abschließend wird noch der innovative Einsatz solcher Systeme diskutiert und an einigen Szenarien beschrieben.

1 Einführung

Weltweit stehen heute Universitäten vor der Herausforderung, eine wachsende Anzahl von Studenten zu unterrichten, lebenslanges Lernen für immer größere Teile der Bevölkerung zu unterstützen und gleichzeitig damit auch mit einer immer stärkeren Heterogenität der Studierenden im Unterricht zu kämpfen. Gleichzeitig müssen sie ein hochwertiges, auch im internationalen Vergleich konkurrenzfähiges Forschungsprofil, trotz Budgetkürzungen und zunehmendem Wettbewerb auf dem Bildungsmarkt, aufrechterhalten. Experten empfehlen Universitäten jedenfalls radikale Umstrukturierungen. Ein solches Szenario besteht zum Beispiel aus einer Reorientierung des Unterrichts in Richtung Fernstudium begleitet

durch massiven Ausbau und Einsatz von Informations- und Telekommunikationstechnologien an Universitäten bei gleichzeitiger Automatisierung der Leistungsbeurteilung und -verfolgung (siehe Dearing Bericht [De97]). Wesentlich dramatischer ist das folgende Szenario, das von D. Tsichritzis vorausgesagt wird [Ts99]: das Universitätssystem wird dabei durch eine radikale Transformation der Geschäftsprozesse durch Arbeitsteilung entlang der Wertschöpfungskette analog zum Mediensektor aufgebrochen und arbeitsteilig in Form kooperierender Spezialorganisationen organisiert.

Überraschenderweise spielen in diesen Überlegungen marktorientierte Ideen, wie das Konzept eines Marktes als dezentraler Koordinationsmechanismus mit dem Preissystem als Informationskanal oder das Konzept die Universität als Marktplatz für Ideen zu organisieren, überhaupt keine Rolle. Im folgenden konzentrieren wir uns auf die Metapher einer Virtuellen Universität als Informationsmarkt mit einem Recommendersystem als Marktinformationskanal. Die Tatsache, daß staatliche Universitätssysteme üblicherweise aus Steuern (indirekt) finanziert werden, sollte kein Hindernis für einen solchen Ansatz darstellen. Sogar wenn ein direktes Preissystem – zum Beispiel für den Besuch von Vorlesungen – fehlt, sind trotzdem ständig Marktkräfte am Werk. Sie wirken durch Vertragsänderungen, Anpassung der Produktqualität und durch Informationskanäle (Mundpropaganda), wie dies bereits von F. Hayek beobachtet wurde [Ha45].

2 Das Potential von Recommendersystemen in Forschung und Lehre

Angenommen, Sie sind gerade als Studienanfänger an der Wirtschaftsuniversität das erstemal in Ihrem Leben angekommen. Welche Vorlesungen und Übungen planen Sie in Ihrem ersten Semester? Normalerweise besuchen Sie die Studienberatung, oder Sie fragen Ihren Mentor in Ihrem Orientierungstutorium, oder Sie erforschen die Gerüchte über Vorlesungen, Übungen und Prüfungen in Gesprächen mit Ihren Freunden in der Mensa. Natürlich könnten Sie auch in der Lehrevaluationsbroschüre der Hochschülerschaft nachlesen, ...

Im täglichen Leben tritt eine solche Situation häufig auf. Sie treffen ständig Entscheidungen ohne genügend eigene Erfahrung über mögliche Alternativen zu haben. Ein Recommendersystem hilft dabei und verbessert diesen sozialen Prozess. In einem typischen Recommendersystem liefern Benutzer Empfehlungen als Inputs, die das System dann aggregiert und an entsprechende Empfänger weiterleitet.

Heute stehen Universitäten weltweit vor folgenden Herausforderungen:

1. Mehr Studenten mit gleich vielen oder weniger Universitätslehrern auszubilden.
2. Lebenslanges Lernen für immer mehr Bürger zu sozial akzeptablen Kosten zu ermöglichen.
3. Als Konsequenz daraus Studentengruppen mit sehr heterogenem Vorwissen betreuen zu müssen.
4. Mit dem exponentiellen Wachstum an im Internet verfügbaren Informationsquellen und der daraus resultierenden Informationsflut fertig zu werden. (90 Prozent aller Forscher leben heute.)

5. Betreuungsintensive Lernformen, wie Team- und Projektarbeit, wie sie heute weltweit in der Industrie vorausgesetzt und erwartet werden, ins Studium zu integrieren und damit die universitäre Lehre radikal zu verändern.

Der Versuch diese Herausforderungen anzunehmen, führt dazu, daß Zeitdruck und Streß den Alltag von Lehrern, Forschern und Studierenden an Universitäten prägt. Recommendersysteme in Lehre und Forschung helfen die Zeitnot zu lindern, indem sie zahlreiche, zeitraubende Routineempfehlungen automatisiert erzeugen oder automatisiert an die richtigen Empfänger senden. In den folgenden Beispielen wird das Potential solcher Systeme in einer Universitätsumgebung gezeigt (siehe [GHJ01b]):

2.1 Informationsüberlastung

Warum ist der Nutzen des Internet in Lehre und Forschung nicht so hoch, wie ursprünglich erwartetet? Der Grund dafür liegt hauptsächlich in den Such- und Evaluationskosten für Informationsprodukte, wie multimediale Lehrmodule oder wissenschaftliche Arbeiten. Ein Recommendersystem im Lehr- und Forschungsbereich reduziert diese Kosten vor allem durch Sammeln und Aggregieren der Erfahrungen großer Benutzergruppen.

2.2 Kommunikation Lehrer/Student

Trotz modernster Kommunikationstechnologie bleibt die Zeit, die der Lehrer in jedes einzelne Beratungsgespräch investieren muß, der Engpaßfaktor, der die Anzahl der Studenten, die ein Lehrer betreuen kann, begrenzt. Ein Recommendersystem für Lehre und Forschung hilft hier, indem Routinefragen, wie “Ich schreibe eine Seminararbeit zum Thema Recommendersysteme. Welche Literatur können Sie mir empfehlen?”, vom System beantwortet werden. Gleichzeitig bietet es auch dem Lehrer die Chance, zu sehen, welche Unterlagen von Studenten besonders häufig verwendet werden.

2.3 Heterogene Studentengruppen

Durch die Ausweitung und Öffnung des Bildungsangebots einer Universität für immer weitere Bevölkerungskreise werden die Studenten der Zukunft zunehmend durch völlig unterschiedliches Vorwissen, Berufsausbildung und -erfahrung und verschiedene Arbeits- und Lernstile geprägt sein. In diesem Forschungsprojekt wurde zur Lösung dieses Problems ein Recommendersystem, das auf der Selbsteinschätzung der Erfahrung in einem bestimmten Bereich basiert, implementiert und einem ersten erfogreichen Test unterzogen.

2.4 Teambildung

Eines der Probleme einer Massenuniversität ist die zunehmende soziale Isolation des einzelnen Studenten/Lehrers/Forschers. Mit ihrer Fähigkeit Personen mit gemeinsamen Interessen zu gruppieren, bieten Recommendersysteme die Chance, sowohl die Bildung von kleinen Lerngruppen als auch von kleinen, spezialisierten Forschungsgruppen zu fördern.

3 Eine Klassifikation von Recommendersystemen

Recommendersysteme zählen heute zu den wichtigsten Softwaresystemen: Alexa, ein Browser-Plugin für ”Related Links” gewann den PC-Magazine-Preis für das beste Produkt 1998. Firefly Network Inc. wurde bereits kurz nach seiner Gründung von Microsoft aufgekauft - vor allem wegen seiner Marktführerschaft im Bereich Personalisierung und Recommenderdienste. “Relationship Tech” und hier vor allem Recommendersysteme stellen für Visionäre wie Kevin Kelly die Schlüsseltechnologie zur Netzwerkökonomie und zur Wissensgesellschaft dar. Zahlreiche Internet-Unternehmen haben bereits Recommendersysteme in der einen oder anderen Form in ihre Informationssysteme integriert. Bei Amazon.com findet der Kunde bereits mehr als 15 verschiedene Recommenderdienste in den Produktkatalog integriert. Internetmarketingagenturen, wie z.B. ActiveAgent Werbenetz verwenden Recommenderdienste, um zielgruppengerecht Bannerwerbung zu betreiben. Web-Mining in allen seinen Formen zählt zu den derzeit aktivsten Forschungsgebieten.

Historisch gesehen stammen Recommenderdienste aus dem Bereich Information Filtering. Ziel dabei war, Information Retrieval Verfahren stärker an individuelle Bedürfnisse anzupassen, um Ergebnisse von Abfragen in personalisierter Form dem Benutzer zu präsentieren. Frühe Beispiele von solchen Recommendersystemen sind beispielsweise Tapestry, Group Lens und Fab. Begriffe wie “collaborative filtering” und “social filtering” wurden in diesen Projekten geprägt - meist mit Groupware-Anwendungen als Ziel.

P. Resnick und H. R. Varian klassifizieren Recommendersysteme in folgenden 5 Dimensionen (siehe [RV97]):

1. Was stellt den Inhalt einer Empfehlung dar?
2. Basiert die Empfehlung auf der Meinung von Benutzern oder auf beobachtetem Benutzerverhalten?
3. Ist der Benutzer anonym?
4. Wie werden Empfehlungen aggregiert?
5. Wie werden Empfehlungen im System verwendet?

Bei der Wahl eines konkreten Recommendersystems aus diesem (riesigen) Designraum sollten außerdem noch bestimmte Eigenschaften der zu bewertenden Produkte und der Benutzer eines solchen Recommendersystems berücksichtigt werden. Auf Produktseite sind dies:

1. Welche Produkte sollen bewertet werden?
2. Wieviele Produkte müssen bewertet werden?
3. Wie groß ist ist die Lebensdauer eines Produkts?
4. Welche Gewinne oder Verluste entstehen durch “falsche” Produktauswahl für Benutzer von Recommendersystemen?

Für die Teilnehmer in einem Recommendersystem sollten folgende Fragen geklärt werden:

1. Wer sind die Produzenten von Empfehlungen?
2. Wer sind die Konsumenten von Empfehlungen?
3. Welche “Dichte” von Empfehlungen wird erreicht?
4. Wie schnell ändern sich die Erfahrungen, Bedürfnisse, und die Geschmäcker von Konsumenten?

Die Antworten auf diese Fragen beeinflussen das technische Design von Recommendersystemen stark. In einer Massenuniversität besitzt ein Recommendersystem, das Studenten mit ähnlichen Studieninteressen in Teams bündelt, wesentlich mehr Potential als in einer kleinen Forschungsuniversität, in der jeder jeden kennt.

4 Mechanism Design Probleme bei Recommendersystemen

Betrachten Sie nun zunächst die Effekte, die ein anonymes Recommendersystem auf den Ruf von Lehrern oder Forschern haben kann. Als Teil der Lehrevaluation, die an österreichischen Universitäten gesetzlich vorgeschrieben ist, wurde eine sogenannte Feedback-Box, die anonyme Vorschläge zur Verbesserung von Lehrveranstaltungen ermöglichen sollte, an der WU-Wien eingerichtet. In der ersten Version dieses Systems wurden alle Vorschläge sofort öffentlich im Web publiziert, ohne dem betroffenen Universitätslehrer eine Chance zu einer Reaktion zu geben. Nach mehreren Fällen, in denen dieses System zum Rufmord durch anonyme Benutzer verwendet wurde, mußte die Art, wie dieses System betrieben wurde, mehrmals grundlegend geändert werden (nur mehr Zugang für WU-Mitglieder, Antworten eines Universitätslehrers nur mehr auf nicht-anonyme Vorschläge, Redaktion und Editor für solche Diskussionen, ...).

Das Design von Recommendersystemen erfordert die Analyse einiger interessanter und schwieriger Anreizprobleme – kombiniert mit Fragen des Schutzes der Privatsphäre Betroffener (siehe [ARZ99]). Natürlich finden sich alle diese Probleme auch im Kontext der universitären Lehre und Forschung:

1. Free-Riding.
2. Falsche Empfehlungen.
3. Schutz der Privatsphäre und Ruf der Empfehlenden.
4. Glaubwürdigkeit.
5. Positive und negative Feedback Effekte.
6. Economies of Scale und Monopole.

Um diese Anreizprobleme und Probleme beim Schutz der Privatsphäre, die jedes Recommendersystem inhärent besitzt, besser lösen zu können, sollten folgende Fragen beim Design von Recommendersystemen im Detail analysiert werden:

1. Welche Beziehungen (und Macht- und Abhängigkeitsverhältnisse) existieren zwischen den Eigentümern der Informationsprodukte, den Betreibern des Recommendersystems, den Produzenten und schließlich auch den Konsumenten von Empfehlungen?

2. Welche Arten von Anreizproblemen und Problemen beim Schutz der Privatsphäre können für ein spezielles, technisches Design eines Recommendersystems identifiziert werden?

3. Welche Risiken, Bedrohungen und Gewinne werden von den einzelnen Parteien erwartet? Welche Verlustfunktion besitzen alle teilnehmenden Parteien?

5 Designprinzipien für Recommenderdienste für Forschung und Lehre

Die Designprinzipien der Recommenderdienste, die in der Virtuellen Universität derzeit eingesetzt werden, basieren auf folgenden Ideen:

Informationskanäle. Wie bereits in der Einleitung diskutiert, wird die Virtuelle Universität als Informationsmarkt mit Recommendersystemen als Informationskanäle betrachtet.

Beobachtetes Benutzerverhalten. Empfehlungen werden auf der Basis von beobachtetem Benutzerverhalten für Informationsprodukte erzeugt. Immer wenn der Benutzer einem externen Link vom Informationsbroker der Virtuellen Universität zu einem Informationsprodukt (Lehrveranstaltung, Lehreinheit, Artikel, Bibliographie, Unternehmensspiel, Multiple-Choice Test, ...) folgt, wird dies als ein "Kauf" des Informationsprodukts interpretiert. Die Benutzung interner Links im Broker zeigt Benutzerpräferenzen für Dienste des Brokers. Informationsprodukte in der Virtuellen Universität besitzen eine reiche Meta-Daten Beschreibung, die auch Klassifikationen nach mehreren Klassifikationssystemen inkludieren kann. Dies entspricht der Sortimentsgliederung im Einzelhandel. Abhängig vom Grad der Anonymität des Benutzers lassen sich verschiedene, weit verbreitete Analysemethoden aus dem Handel auf eine Virtuelle Universität übertragen.

Selbsteinschätzung der Erfahrung. Der Heterogenität von Benutzern wird mit Hilfe einer Variante der inkrementellen Selbstklassifikation der Benutzer nach ihrer Erfahrung in einem bestimmten Fachgebiet Rechnung getragen. Benutzer (z.B. Studenten, Universitätslehrer) beurteilen dabei ihre eigene Erfahrung in Fachgebieten, aus denen sie bereits Informationsprodukte konsumiert haben (Anfänger, Durchschnitt, Fortgeschrittener oder Experte). Im Unterschied zur klassischen Selbstselektion, wie sie z.B. bei der Entwicklung von Produktlinien in gewinnmaximierender Form seit langem angewendet wird, werden hier die Informationsprodukte nicht für die einzelnen Gruppen vorklassifiziert. Gleichzeitig müssen auch Universitätslehrer nicht beurteilen, welche Erfahrung ein Student in einem Fachgebiet besitzt, da der Student seine Erfahrung durch seine Selbsteinschätzung enthüllt. Jeder Studierende hat

in diesem System einen starken Anreiz, seine wahre Erfahrung in einem Fachgebiet zu enthüllen, da dies gleichzeitig den Wert der ihm gegebenen Empfehlungen maximiert. Die im vorigen Abschnitt angesprochenen Analyseverfahren erhalten auf diese Art eine fachgebietsabhängige, persönliche Erfahrungsdimension, die den Nutzen von Empfehlungen speziell in der Lehre wesentlich verbessert, da automatisch jeder Student bei korrekter Selbsteinschätzung auf seine Erfahrung zugeschnittene Empfehlungen erhält. Zusätzlich kann auf Basis eines solchen Systems auch die Überwachung von Lernfortschritten oder auch eine Korrektur der Selbsteinschätzung bei Unter- oder Überschätzungen erfolgen, sobald genügend Benutzer in einem Fachgebiet damit arbeiten.

Evolution. Evolutionäre Algorithmen basieren auf dem Zusammenspiel von Fitness getriebener Selektion und zufälligen Mutationen. Der Selektionsprozeß ist dabei für die Auswertung der Information, die in der Stichprobe (einer Generation) enthalten ist, verantwortlich, der Mutationsprozeß für die Exploration und die Entdeckung weiterer Informationen. Effizientes Suchen erfordert eine Balance zwischen dem Aufwand für die Suche nach neuer Information und der Verwertung bereits bekannter Information. Recommendersysteme fördern die Verwertung bereits bekannter Information. Im universitären Umfeld besteht die Gefahr, daß durch Recommendersysteme das Such- und Rechercheverhalten von Studenten auf wenige (in der Regel ausgezeichnete) Quellen reduziert wird. Dieser Gefahr wird durch die Integration von Mutationsoperatoren in Recommendersysteme in Form von zufällig erzeugten Bannern und zufällig erzeugten Listen begegnet, die die Neugier des Benutzers stärken und ihn zu mehr Recherchen verführen sollen.

Skalierbarkeit. Die Recommenderdienste einer Virtuellen Universität sollen möglichst große Benutzergruppen unterstützen. Benutzerverhalten in Real-Time zu überwachen, Benutzerpräferenzen zu aggregieren und Benutzerschnittstellen on-the-fly dynamisch zu erzeugen, sind rechenintensive Prozesse, die nur schwer auf ein Rechnernetzwerk verteilt werden können. Die Strategie, die hier vorgeschlagen wird, besteht darin, durch Abschwächen von Konsistenzanforderungen in Real-Time nur mehr 2 Prozesse durchzuführen, nämlich erstens das Benutzerverhalten zu protokollieren, und zweitens nur jene Antworten auf Benutzeraktionen sofort zu berechnen, bei denen der Benutzer dies erwarten würde. Alle anderen Prozesse, wie die Aggregation von Benutzerpräferenzen, das Berechnen statistischer Modelle, die Anpassung der Benutzerschnittstellen, erfolgt im Hintergrund in entsprechenden Intervallen. Diese Strategie ist vor allem deshalb ein akzeptabler Kompromiß, da Benutzer in der Regel ständige, automatische Änderungen der Benutzerschnittstelle während einer einzigen Session eher als irritierend empfinden, da sie sich ständig am Bildschirm neu orientieren müssen und dies ihre Arbeitsgeschwindigkeit verlangsamt und gleichzeitig erhöhte Aufmerksamkeit erfordert.

6 Die Architektur von Recommendersystemen.

Die Architektur eines Recommendersystems folgt aus den im vorigen Abschnitt besprochenen Designprinzipien. Abbildung 1 zeigt die Architektur der Recommenderdienste der

Virtuellen Universität der WU-Wien in Form einer Agency von lose-gekoppelten Softwareagenten.

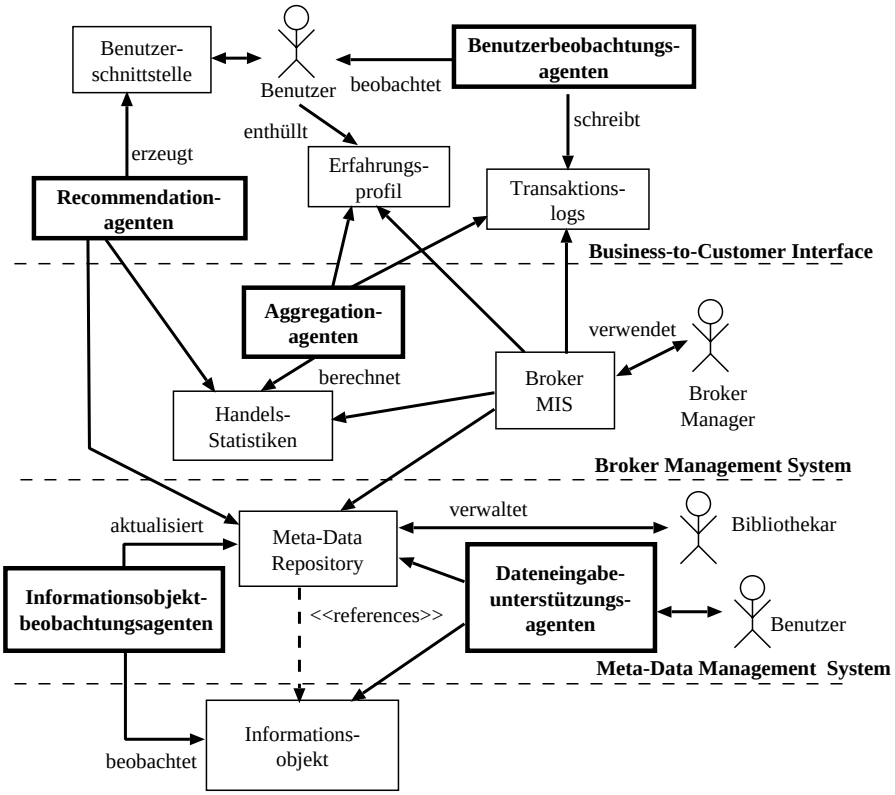

Abbildung 1: Die Architektur eines Informationsbrokers mit Recommendersystemen

Die Recommenderdienste in der Agency bestehen aus 3 Typen von Agenten, nämlich Beobachtungsagenten, Aggregationsagenten, Recommendationagenten, die mit dem Benutzer interagieren (siehe [GH01]):

Beobachtungsagenten. Diese Agenten beobachten das Verhalten der Benutzer und erzeugen die Transaktionslogs. Derzeit werden diese Agenten mit zwei Techniken realisiert:

1. Standard http-Transaktionslog mit Sessioncookies oder
2. Aufzeichnen der Kaufakte samt Cookies durch das Brokersystem.

Pseudoanonymität wird mit Hilfe eines zusätzlichen Cookies, das eine Pseudoidentität enthält, erreicht. Damit können mehrere Sessions demselben Benutzer zugerechnet und damit zu einer Kaufgeschichte zusammengefaßt werden.

Aggregationsagenten. Diese Agenten analysieren die Transaktionslogs, die die Beobachtungsagenten erzeugt haben. Typische Ergebnisse solcher Aggregationsagenten sind Kauffrequenzlisten für Informationsprodukte, eine ABC-Analyse von Informationsprodukten, Warenkörbe, bedingte Wahrscheinlichkeitsverteilungen zur Analyse von Cross-Selling Verhalten, Kaufgeschichten auf unterschiedlichen Aggregationsniveaus (1 Person, Gruppen mit gleichem Erfahrungsniveau, ...)

Recommendationagenten. Diese Agenten verwandeln die entsprechenden Statistiken, die die Aggregationsagenten erzeugt haben, in Empfehlungen. Solche Empfehlungen werden in Form von Labels, sortierten Listen, speziellen Navigationsstrukturen und Statistiken dem Benutzer präsentiert.

Benutzer. Diese erhalten Empfehlungen verschiedenster Form, enthüllen inkrementell ihre Erfahrung in den sie interessierenden Fachbereichen und erzeugen mit ihren Kaufakten Transaktionslogs, wenn sie die Brokersysteme verwenden.

7 Verfügbare Recommenderdienste in VU und myVU

Folgende Recommenderdienste sind derzeit bereits in die Virtuelle Universität der WU-Wien integriert:

1. Recommenderdienste für anonyme Benutzer:
   (a) Nach Benutzungsfrequenz oder Relevanz sortierte Listen von Informationsprodukten.
   (b) Labels, wie “hot”, für Informationsprodukte, die im Rahmen einer ABC-Analyse in Klasse A gereiht wurden.
   (c) Balkendiagramme, die die relative Benutzungshäufigkeit zum meistbenutzten Informationsprodukt zeigen.
   (d) “Others also use” bietet für jedes Informationsprodukt eine Liste jener Produkte, die wie im nächsten Abschnitt beschrieben, wahrscheinlich häufiger als zufällig zu erwarten gewesen wäre, gemeinsam in Sitzungen verwendet werden.
   (e) ”Link of the day” ist ein zufällig ausgewählter Link aus dem Ergebnis der letzten Abfrage.
2. Personalisierte Recommenderdienste:
   (a) “my top 10” ist eine Liste der 10 vom Benutzer am häufigsten besuchten Web-Sites.
   (b) ”top entries” ist eine nach der persönlichen Nutzungshäufigkeit absteigend sortierte Liste von Web-Sites mit allen anderen verfügbaren Empfehlungen.
   (c) ”top categories” ist eine nach der persönlichen Nutzungshäufigkeit absteigend sortierte Liste von Kategorien (Fachgebieten).

(d) Empfehlungen nach Fachgebiet und Erfahrung (z.B. Experts also use, ...)

(e) ”discover VU” ist ein Banner, das immer einen zufällig gewählten Link auf eine Web-Site anzeigt.

(f) ”discover more entries” liefert eine Liste zufällig gewählter Links auf Web-Sites.

(g) ”discover more categories” liefert eine Liste zufällig gewählter Kategorien (Fachgebiete).

In Tabelle 1werden noch einmal die Eigenschaften der derzeit in der Virtuellen Universität der WU-Wien verfügbaren Recommenderdienste zusammengefaßt.

	VU	myVU
Inhalt?	Link auf Web-Site	Link auf Web-Site
Benutzermeinungen oder -Verhalten?	Benutzerverhalten	Benutzerverhalten, Selbsteinschätzung der Erfahrung
Anonym?	Ja, Sitzungsdaten	Pseudonym
Aggregationsverfahren?	ABC-Analyse, Warenkorbanalyse, Repeat-Buying Analyse	Kontextabhängige Warenkorb- und Repeat-Buying Analyse basierend auf Selbsteinschätzung, Analyse der Kaufgeschichte
Verwendung?	Sortierte Listen, Labels, Sortierte Warenkörbe	zusätzlich gruppenspezifische Warenkörbe und Profile verwandter Informationsproduktgruppen, persönliche Favoriten

Tabelle 1: Eigenschaften der VU und myVU-Recommenderdienste

Die hier präsentierten Recommenderdienste wurden unter Berücksichtigung der oben vorgestellten Mechanism Design Probleme ausgewählt und implementiert. Alle derzeit eingesetzten Recommenderdienste basieren auf beobachtetem Benutzerverhalten.

Dies reduziert das Free-Riding Problem, da durch jede Verwendung der Virtuellen Universität automatisch neue Daten gesammelt werden. Derzeit (7.5.2001) existieren für fast 50 % von 9498 Informationsprodukten Daten, für knappe 20 % können sogar statistisch signifikante Modelle geschätzt werden. Das bedeutet, daß ein vom Benutzer keine Arbeit erforderndes, unaufdringliches Recommendersystem durchaus im universitären Umfeld akzeptiert und einsetzbar ist. Die personalisierten Dienste in myVU werden nur auf einer tit-for-tat Basis angeboten, das bedeutet, daß der Benutzer diese Dienste nur erhält, wenn er mit dem personalisierten System arbeitet und über sich Information (die Selbsteinschätzung seiner Erfahrung) preisgibt. Gruppenspezifische Recommenderdienste sind nur nach Selbsteinschätzung der eigenen Erfahrung für ein Fachgebiet verfügbar.

Weiter wird dadurch die Gefahr falscher Empfehlungen weitgehend reduziert, da keine negativen Empfehlungen explizit vom System verarbeitet werden und es damit unmöglich wird den Ruf von Konkurrenten zu schädigen. Das Erzeugen positiver Empfehlungen

bleibt weiterhin möglich, indem z.B. dasselbe Informationsprodukt wiederholt besucht wird. Auf Grund des Einsatzes der Repeat-Buying Theorie, wie im nächsten Abschnitt beschrieben, ist aber diese Form der Verfälschung von Empfehlungen mit einem hohen Preis verbunden, da der "Fälscher" für jeden Kaufakt eine neue Sitzung aufbauen muß und dies jeweils einen Neustart seines Browsers erfordert.

Alle anonymen Recommenderdienste werden nur aus aggregierten Sitzungsdaten gewonnen, sodaß die Privatsphäre der Benutzer gewahrt wird. Im Bereich der personalisierten Dienste wird mit pseudoanonymen Daten gearbeitet. Dies bedeutet, daß die wahre Identität des Benutzers von der Pseudoidentität getrennt verarbeitet wird, damit die Privatsphäre des Benutzers gewahrt bleibt.

Auf Grund der Tatsache, daß sowohl VU als auch myVU aus Forschungsmitteln einer Universität finanziert werden, existieren keine finanziellen Anreize, Empfehlungen für Informationsprodukte zu verfälschen.

Positive und negative Feedback Effekte werden mit den auf Mutation beruhenden Empfehlungsdiensten, sowie durch die Bevorzugung neuer Informationsprodukte, berücksichtigt (siehe [GHJ00]). Eine Analyse dieser Dienste ist derzeit allerdings noch nicht verfügbar.

8 Repeat-Buying Theorie für anonyme Recommendersysteme

Als theoretische Basis anonymer Recommendersysteme wird in diesem Abschnitt Ehrenberg's Repeat-Buying Theorie in [Bö01] und [GHJ01a] vorgeschlagen. Diese Systeme verwenden Warenkorbdaten (Sitzungen) mit dem Benutzungsverhalten von Informationsprodukten, die als Kaufgeschichten mit nicht beobachteter Identität des Konsumenten interpretiert werden. Wir betrachten dazu einen Informationsbroker mit klar definierter Systemgrenze. Anklicken eines externen Links wird mit dem "Kauf eines Informationsprodukts" gleichgesetzt. Im Marketing wird angenommen, daß ein Konsument ein Produkt oder eine Produktkombination nur dann mehrfach kauft, wenn er damit zufrieden ist. Diese Grundüberlegung gilt auch für freie (gratis) Informationsprodukte, da auch der Konsum solcher Informationsprodukte mit Such-, Auswahl- und Bewertungskosten verbunden ist. Empfehlungen werden also aus wiederholt gemeinsam (gleiche Sitzung = gleiche Kaufgelegenheit) benutzten (= gekauften) Produkten abgeleitet. Damit nun nicht zufällig gemeinsam gekaufte Informationsprodukte für Empfehlungen verwendet werden, ist es notwendig folgende Probleme zu lösen:

1. Welche gemeinsam gekauften Informationsprodukte sind nicht zufällig gemeinsam gekauft worden?

2. Wieviele Informationsprodukte sollen empfohlen werden?

Ehrenbergs Repeat-Buying Theorie (siehe [Eh88]) liefert dafür ein Referenzmodell, mit dem wir auf nicht zufällige Ausreisser (= Empfehlungen) testen können, da diese Theorie starke Stationaritäts- und Unabhängigkeitsannahmen trifft. Für diese Theorie spricht, daß sie in mehreren hundert Konsumgütermärkten seit mehr als 40 Jahren immer wieder

erfolgreich getestet wurde. Die Theorie beschreibt, wie sich Konsumenten verhalten, nicht warum.

Im Prinzip trifft jeder Konsument 2 Entscheidungen, nämlich wann er ein Produkt aus einer bestimmten Produktklasse kauft (Wahl des Kaufzeitpunkts) und welches Produkt er kauft (Markenwahl). Ehrenberg behauptet nun, daß fast alle Aspekte des Repeat-Buying Verhaltens durch Formalisierung des ersten Entscheidungsprozesses für jedes Produkt und einer Integration dieser Einzelprozesse adäquat beschrieben werden: Konsumenten kaufen ein Produkt in einem stationären Poissonprozeß, der von allen anderen Kaufprozessen unabhängig ist. Aggregation dieser Kaufprozesse unter der (ziemlich allgemeinen) Annahme, daß der Mittelwert der Poissonverteilung einer abgeschnittenen Gamma-Verteilung folgt, führt zu einer logarithmischen Reihenverteilung, die beschreibt, mit welcher Wahrscheinlichkeit ein Produkt 1, 2, 3, ... mal gekauft wird. Folgende Annahmen über das Verhalten von Konsumenten führen zu einem solchen Modell:

1. Der Anteil der Nichtkäufer in der Grundgesamtheit ist nicht bekannt. Im Falle einer Virtuellen Universität mit anonymen Benutzern ist diese Annahme sicherlich gerechtfertigt.

2. Die Käufe eines Konsumenten in aufeinanderfolgenden Perioden folgen einer Poisson-Verteilung mit einem langfristig stabilen Mittelwert. Dies ist dann der Fall, wenn ein Kauf unabhängig von vorangegangenen Käufen erfolgt, und wenn Käufe so unregelmäßig erfolgen, daß sie als zufällig betrachtet werden können.

3. Die Annahme einer abgeschnittenen Gamma-Verteilung für die Mittelwerte der Poisson-Verteilungen der Konsumenten wird durch folgende Annahmen gerechtfertigt: Für eine Reihe von Produkten ist die durchschnittliche Kauffrequenz unabhängig von der Kauffrequenz für andere Produkte und die Kauffrequenz eines Produktes unabhängig von der gesamten Kauffrequenz für alle Produkte.

4. Der Markt ist im Gleichgewicht (stationär).

Für fast 20 % aller Informationsprodukte (9498) in der Virtuellen Universität der WU-Wien ließ sich ein solches Kaufverhaltensmodell schätzen und entsprechende Ausreisser als Empfehlungen identifizieren. Eine erste Validierung ergab einen Anteil von mehr als 75 % guten Empfehlungen, bei ungefähr zwei Drittel der Listen, wurden mehr als 85 % der Empfehlungen als gut eingestuft.

9 Szenarien für den innovativen Einsatz von Recommenderdiensten in einer Virtuellen Universität

Im letzte Abschnitt werden noch 3 Szenarien für den zukünftigen Einsatz von Recommenderdiensten im universitären Bereich vorgeschlagen, nämlich:

1. Teambildung und Teamorganisation für Lern- und Forschungsgruppen

2. Integrierte Geschäftsprozeßanalyse und adaptive Geschäftsprozesse

3. Virtuelle Universitäten und Digitale Bibiotheken

9.1 Teambildung und Teamorganisation für Lern- und Forschungsgruppen

Studieren in Teams verspricht folgende positive Auswirkungen im Lehrbereich einer Universität: Reduktion der Zahl der Studienabbrecher und eine Verkürzung der Studienzeiten, Schulung von Studierenden in Teamarbeit und eine Förderung der persönlichen Kommunikations- und Teamfähigkeit, sowie praktische Erfahrungen in Kooperationen und in Projektarbeit, wie sie vor allem in der Informationsgesellschaft erforderlich ist. Recommendersysteme unterstützen vor allem die automatisierte Bildung von Teams und die Zuordnung von Studierenden zu bestehenden Teams auf mehrere Arten: Aufbauend auf beobachtetem Benutzerverhalten beim Konsum von Informationsprodukten werden mit geeigneten Clusteralgorithmen potentielle Teamteilnehmer vorgeschlagen. Clustering von Benutzern wird aber auch über benutzerdefinierte Profile oder mittels Präferenzprofilen unterstützt.

9.2 Integrierte Geschäftsprozeßanalyse und adaptive Geschäftsprozesse

Im Bereich der internen Verwaltungs- und Servicebereiche einer Universität lassen sich durch Kombination der im vorigen Abschnitt vorgeschlagenen Recommenderdienste mit Informationen über die Stellenbeschreibungen von Benutzern sowohl Geschäftsprozesse ständig analysieren als auch adaptiv anpassen. Auch Knowledgemanagement wird durch solche Recommenderdienste unterstützt: Welche Empfehlungen würden zum Beispiel aus dem Nutzungsverhalten für Informationsprodukte von Systemadministratoren folgen? (Außer, daß sie alle Dilbert-Fans sind und User Friendly Comic-Strips lesen.)

9.3 Virtuelle Universitäten und Digitale Bibiotheken

Wissenschaftliche Bibliotheken und Virtuelle Universitäten lassen sich mit Hilfe von Recommenderdiensten sehr rasch in kundenorientierte Serviceportale nach dem Vorbild von amazon.com reorganisieren. Für Studierende, Universitätslehrer und Forscher liegt ein wesentlicher Vorteil in einer Reduktion der Such- und Evaluationskosten für Informationsprodukte. Gleichzeitig wird ein permanenter Marktforschungs- und Produktevaluationsprozess in das Serviceportal integriert und dessen Ergebnisse den Betroffenen in verschiedener Weise direkt zugänglich gemacht.

Danksagung

Die Autoren danken dem Jubiläumsfonds der Österreichischen Nationalbank, der das Projekt Nr. 7925, das in dieser Arbeit beschrieben wird, finanziert hat. Ohne diese Unterstützung wäre dieses Projekt ein Traum geblieben.

Literaturverzeichnis

[ARZ99] Avery, C.; Resnick, P.; Zeckhauser, R.: The Market for Evaluations. American Economic Review, 89(3):564-583, 1999.

[Bö01] Böhm, W. et al.: Repeat-Buying Theory and Its Application for Recommender Services. Submitted.

[De97] Dearing, R. et al.: Higher Education in the Learning Society. The Report of the National Committee of Inquiry into Higher Education (NCIHE). http://www.leeds.ac.uk/educol/ncihe/

[Eh88] Ehrenberg, A. S. C.: Repeat-Buying: Facts, Theory and Applications. Charles Griffin & Company Limited, London, 1988.

[GH01] Geyer-Schulz, A.; Hahsler, M.: Recommendations for Virtual Universities from Observed User Behavior. In (Gaul, W. Hrsg.): Tagungsband der 24. Jahrestagung der GfKl, Karlsruhe, Germany, 2001, erscheint.

[GHJ00] Geyer-Schulz, A.; Hahsler, M.; Jahn, M.: myVU: A Next Generation Recommender System Based on Observed Consumer Behavior and Interactive Evolutionary Algorithms. In (Gaul, W.; Opitz, O.; Schader, M. Hrsg.): Data Analysis: Scientific Modeling and Practical Applications, Studies in Classification, Data Analysis, and Knowledge Organization, Vol. 18. Springer, Heidelberg, 2000; S. 447-457.

[GHJ01a] Geyer-Schulz, A.; Hahsler, M.; Jahn, M.: A Customer Purchase Incidence Model Applied to Recommender Services. WEBKDD'2001 Conference. http://robotics.stanford.edu/~ronnyk/WEBKDD2001/WEBKDD2001Accept.html

[GHJ01b] Geyer-Schulz, A.; Hahsler, M.; Jahn M.: Educational and Scientific Recommender Systems: Designing the Information Channels of the Virtual University. International Journal of Engineering Education. 17(2):153-163, 2001.

[Ha45] Hayek, F. A.: The Use of Knowledge in Society. The American Economic Review, 35(4): 519-530, 1945.

[RV97] Resnick, P.; Varian, H. R.: Recommender Systems. Communications of the ACM, Vol. 40(3), 1997, 56-58.

[Ts99] Tsichritzis, D.: Reengineering the University. Communications of the ACM, Vol. 42(6), 1999, 93-100.