

MyMovieMixer: Ein hybrider Recommender mit visuellem Bedienkonzept

Katja Herrmann, Sandra Schering, Ralf Berger, Benedikt Loepp, Timo Günter, Tim Hussein, Jürgen Ziegler

Universität Duisburg-Essen

Zusammenfassung

In diesem Beitrag stellen wir ein neuartiges, auf direkter Manipulation beruhendes Bedienkonzept für komplexe hybride Empfehlungssysteme anhand des von uns entwickelten Film-Recommenders MyMovieMixer vor. Der Ansatz ermöglicht es den Nutzern, ein hybrides Recommender-System mit einem komplexen Zusammenwirken verschiedener Filtermethoden durch interaktive und visuelle Methoden intuitiv zu steuern. Gleichzeitig wird die Transparenz der Empfehlungsgenerierung deutlich erhöht. Die Ergebnisse einer empirischen Evaluation des Systems zeigen, dass der Ansatz in Bezug auf Usability, User Experience, Intuitivität, Transparenz, wahrgenommene Empfehlungsqualität und somit letztlich im Hinblick auf die Nutzerzufriedenheit vielversprechend ist.

1 Einleitung

Durch das Internet haben Nutzer heutzutage einen nahezu unbegrenzten Zugang zu Informationen, gleichzeitig kommt es aber auch zu einer wahren Informationsüberflutung durch täglich neue Inhalte. Eine damit einhergehende Herausforderung für Entwickler von Empfehlungssystemen (kurz Recommender) ist es, die Informationsmenge, die dem Nutzer präsentiert wird, so einzugrenzen, dass sie seinen individuellen Präferenzen bestmöglich entspricht (Ricci et al., 2010). Heute gängige Recommender verwenden zumeist vorab gewonnene Nutzerprofile und voll automatisierte Empfehlungsmethoden, was für den Nutzer zu einem minimalen Interaktionsaufwand und – bei guter Passung – geringer kognitiver Belastung führt (Pu et al., 2012). Allerdings sind sie meist auch wenig flexibel und nicht oder nur eingeschränkt durch den Nutzer steuerbar. Oft mangelt es zudem an einer nachvollziehbaren, transparenten Empfehlungsgenerierung (Tintarev & Mastoff, 2010), was sich negativ auf Akzeptanz und Vertrauen in die Empfehlungen auswirken kann (Xiao & Benbasat, 2007). Weiterhin ist in vielen Situationen kein Nutzerprofil verfügbar (Cold-Start-Problem) oder die Nutzer wünschen keine Speicherung ihrer Präferenzen. In der Vergangenheit lag der Fokus der Forschung überwiegend auf der Optimierung der Empfehlungsalgorithmen, obwohl der Anteil, zu dem diese über den Erfolg eines Recommenders entscheiden, zunehmend schwin-

det (Konstan & Riedl, 2012). Die typischerweise offline ermittelte Passung der Empfehlungen mit den Nutzerpräferenzen lässt sich durch verbesserte Algorithmen nur noch marginal steigern (Pu et al., 2012). Zuletzt rückte daher die Frage der User Experience zunehmend in den Vordergrund (Knijnenburg et al., 2012), denn für den Anwender sind außer der Präzision der Empfehlungen und der Aufwandsminimierung bei der Suche oft weitere Aspekte von gleicher oder eventuell sogar höherer Bedeutung: So wünschen sich Nutzer eine aktivere Rolle im Empfehlungsprozess (Xiao & Benbasat, 2007) und sind häufig zufriedener, wenn sie den Recommender beeinflussen können. Dafür nehmen sie weniger präzise Vorschläge und sogar zusätzlichen Aufwand in Kauf (Konstan & Riedl, 2012). Die Berücksichtigung der Bedürfnisse und Zufriedenheit des Nutzers ist daher von großer Bedeutung.

Der vorliegende Beitrag stellt einen neuartigen Ansatz vor, der Nutzern die Möglichkeit bietet, die Empfehlungsgenerierung zu steuern, die Wirkung verschiedener Empfehlungsmethoden und deren Kombination zu explorieren und den Prozess besser nachvollziehen zu können. Die exemplarisch für das Anwendungsfeld der Filmempfehlungen konzipierte, implementierte und evaluierte Anwendung *MyMovieMixer* verbindet eine visuelle Darstellung verschiedener Filtermethoden mit direkten Interaktionsmöglichkeiten durch ein Drag-and-Drop-Konzept und nutzt diverse Feedback-Mechanismen zur Transparenzsteigerung: Dem Nutzer stehen verschiedene Kriterien zur Verfügung, mit denen er seine Wünsche hinsichtlich der vorzuschlagenden Filme angeben kann. Zur Empfehlungsberechnung nutzt *MyMovieMixer* einen hybriden Ansatz, bei dem die Ergebnisse der für die verschiedenen Kriterien genutzten Empfehlungsverfahren verrechnet und darauf basierend Vorschläge angezeigt werden. Um die Verständlichkeit bei komplexen, hybriden Empfehlungssystemen zu gewährleisten, ist es wichtig, dass der Nutzer den Ursprung der Empfehlungen kennt und die Kombination der Recommender nach seinen Wünschen konfigurieren kann (Parra et al., 2014). Eine solche Möglichkeit erhält der Anwender bei *MyMovieMixer* über Slider, mit denen er die Gewichtung der einzelnen Kriterien jederzeit modifizieren kann. Die Arbeit zeigt somit einen Lösungsansatz auf, dem Nutzer eine intuitive Steuerung eines hybriden Recommenders und eines komplexen Zusammenspiels von Filtermethoden zu ermöglichen.

2 Visuelle, interaktive Empfehlungsgenerierung

Der Einsatz interaktiver Methoden in Empfehlungssystemen ist ein bislang wenig untersuchtes Gebiet, dessen Potenzial aber zunehmend erkannt wird. So geben manche etablierte Systeme dem Anwender die Möglichkeit der Verfeinerung von Ergebnismengen basierend auf dem Prinzip des Relevance Feedback (Salton & Buckley, 1997). Während sich durch die Anwendung dieses Prinzips auf Recommender die wahrgenommene Kontrolle steigern lässt, werden dabei die bestehenden Interessensprofile der Nutzer zumeist lediglich weiter verfeinert. Solche langfristig gesammelten Informationen über die Nutzerpräferenzen sind als Eingabe für viele Empfehlungsalgorithmen von zentraler Bedeutung, obwohl dies etwa erschwert, auf situative Bedürfnisse zu reagieren. Auch mangelt es häufig an Daten für den aktuellen Nutzer, um akkurate Vorschläge zu generieren. Diesem Cold-Start-Problem durch vermehrten Einfluss des Nutzers zu begegnen, wird bisher jedoch selten in Betracht gezogen.

Kritikbasierte Recommender (Chen & Pu, 2012) steigern die Interaktivität, indem sie den Anwender die Empfehlungen nach ihrer Erzeugung bezüglich bestimmter Features bewerten lassen. Dies beruht auf der Annahme, dass es leichter fällt, die präsentierten Items zu kritisieren als selbst ein Suchziel zu formulieren. Durch die Kritik kann der Nutzer die Empfehlungsmenge jederzeit beeinflussen und kontinuierlich sein Suchziel verfeinern, indem er z. B. explizit Interesse an günstigeren Produkten äußert. Die Verwendung mehrerer, dynamisch ausgewählter Features kann seine Effizienz dabei weiter steigern (McCarthy et al., 2005). Meist beruhen die Features auf einer eingeschränkten Menge vorab modellierter Dimensionen, teilweise aber auch, wie etwa bei *MovieTuner* (Vig et al., 2011), auf einer großen Menge von durch die gesamte Nutzerbasis vergebenen Tags, die automatisch vom System gewählt und präsentiert werden. Dies erlaubt es, explizit Empfehlungen von z. B. Filmen mit mehr Action zu verlangen. Meist kommen Tags jedoch unabhängig von anderen inhaltlichen Kriterien zum Einsatz. Eine individuelle Auswahl und Gewichtung, um sie in kombinierter Form in den Empfehlungsprozess einfließen zu lassen, ist so bis dato nicht möglich.

Zhang et al. (2008) konnten zeigen, dass visuelle Unterstützung und direkte Interaktion mit den Inhalten in kritikbasierten Recommendern positiv in Bezug auf Nachvollziehbarkeit, wahrgenommene Nutzerfreundlichkeit und Interaktionsaufwand bewertet werden. Bisher existieren aber nur wenige Ansätze, die Empfehlungssysteme mit interaktiven visuellen Darstellungen kombinieren. *SmallWorlds* von Gretarsson et al. (2010) stellt eine graphbasierte interaktive Visualisierung für einen in Facebook eingebetteten Recommender dar, welcher dem Nutzer das Äußern seiner Präferenzen erleichtern und die Transparenz des eingesetzten Empfehlungsalgorithmus erhöhen soll. Entsprechend zeigte sich der Empfehlungsprozess als leichter verständlich und die Zufriedenheit der Anwender konnte gesteigert werden.

Insbesondere hybride Empfehlungssysteme (Burke, 2007), bei denen das Zusammenspiel mehrerer Algorithmen oft zu präziseren Vorschlägen führt, sind bisher kaum vom Nutzer steuerbar. *TasteWeights*, ein interaktiver und hybrider Recommender von Bostandjiev et al. (2012) ist ein erster Schritt in diese Richtung. Die dynamische Aktualisierung der Präferenzen für genau diejenigen Items, die zur Empfehlungsgenerierung herangezogen werden, sowie die Exploration mit Hilfe direkter und visuell gestützter Interaktionsmöglichkeiten werden dabei als wesentliche Unterstützung im Empfehlungsprozess herausgestellt. Obwohl primär auf die Gewichtung des Einflussgrades der Informationen aus verschiedenen sozialen Netzen beschränkt, konnte die wahrgenommene Güte der Empfehlungen und das Verständnis für ihr Zustandekommen gesteigert werden. Mit *SetFusion* zeigen Parra et al. (2014) wie sich ein nach einer klassischen Strategie (Burke, 2007) arbeitender hybrider Recommender vom Anwender beeinflussen lässt. Das für Konferenzen konzipierte System erlaubt es, die drei verwendeten Empfehlungsalgorithmen per Slider individuell zu gewichten. Die Änderungen wirken sich allerdings nicht unmittelbar auf die Ergebnisliste aus. Weitergehende Interaktionen wie eine gezielte Verfeinerung der Vorschläge oder die Gewichtung inhaltlicher Aspekte sind nicht vorgesehen. Der eingesetzte inhaltsbasierte Algorithmus erwartet zudem ein persistentes Nutzerprofil. In Nutzerstudien zeigte sich jedoch bereits ein als hoch empfundener Grad an Kontrolle. Dabei erschienen die Empfehlungen durch eine Visualisierung der Ergebnismenge und Erklärungen, welche Methoden für ihr Zustandekommen verantwortlich waren, nachvollziehbar. Zudem ließ sich ein gesteigertes Engagement der Nutzer verzeichnen. Die tatsächliche Empfehlungsqualität wurde allerdings nicht weiter betrachtet.

Insgesamt existieren bislang nur wenige, häufig limitierte Ansätze, welche Explorationsmechanismen mit Recommendern integrieren, die Interaktivität des Empfehlungsprozesses erhöhen und die Kontrolle des Nutzers steigern. Neben der Konzentration auf die Algorithmen ist in der Literatur oft eine Abhängigkeit von langfristigen Nutzerprofilen zu bemerken. Gestaltungshinweise zu Empfehlungssystemen weisen u. a. deshalb auf eine Minimierung des Aufwandes bei der initialen Präferenzenerhebung sowie auf eine möglichst unterhaltsame und attraktive Gestaltung hin, um den Nutzer zu mehr Eingaben zu motivieren (Pu et al., 2012). Auch sind Erklärungen der Vorschläge ein Mittel, um die wahrgenommene Effektivität des Recommenders, die Zufriedenheit, den empfundenen Einflussgrad, die Akzeptanz und schließlich das Vertrauen in das System zu erhöhen (Pu et al., 2012). All dies verdeutlicht, dass interaktive, visuell unterstützte Empfehlungssysteme noch großes Potenzial bieten.

3 Die Anwendung MyMovieMixer

Mit MyMovieMixer haben wir unter Berücksichtigung der zuvor beschriebenen Aspekte ein webbasiertes Empfehlungssystem implementiert, um Anwendern in verschiedensten Nutzungsszenarien (z. B. Filmabende, Suche nach einem Geschenk) Spielfilme aus dem Bestand des MovieLens-Datensatzes¹ vorschlagen zu können. Da sich das System flexibel für verschiedene Situationen eignen sollte, bei denen sich wechselnde Stimmungen, Personenkonstellationen und Rahmenbedingungen auf die Empfehlungen auswirken können, basiert deren Berechnung ausschließlich auf explizit und ad-hoc während der Nutzungssitzung formulierten Präferenzen. Langfristige Profile sind somit nicht nötig. Die explizite Formulierung der Eingaben ist an klassische Produktfilter angelehnt, geht jedoch durch den Einsatz von Empfehlungsalgorithmen darüber hinaus. Das System zielt dabei auf eine Erhöhung der Nachvollziehbarkeit und Kontrolle durch den Nutzer ab, indem die resultierenden Vorschläge stets vollständig auf visuell wahrnehmbare und direkt manipulierbare Elemente der Benutzeroberfläche zurückgeführt werden können.

3.1 Benutzerinteraktion

Die grafische Nutzerschnittstelle besteht aus drei Bereichen (s. Abbildung 1): Die linke Seite enthält eine Reihe voneinander unabhängiger Widgets, über die der Nutzer seine Präferenzen angeben kann. Beispielsweise können über das Genre-Widget ein oder mehrere bevorzugte Genres gewählt werden. Die Kriterien zur Präferenzeingabe werden innerhalb der Anwendung konsistent durch rechteckige Kacheln, bestehend aus Titel und grafischem Hintergrund, repräsentiert. Die Kacheln sind innerhalb der Widgets angeordnet. Möchte ein Nutzer von einem oder mehreren Widgets Gebrauch machen, so zieht er per Drag and Drop darin enthaltene Kacheln in den mittleren Bereich (Arbeitsbereich) und verwendet diese so als Kriterium für die Empfehlungen. Es ist möglich, mehrere Kriterien aus einem Widget zu nutzen (beispielsweise mehrere gewünschte Schauspieler). Im mittleren Bereich verfügen die Kacheln

¹ MovieLens-10M-Datensatz (<http://grouplens.org/datasets/movielens/>)

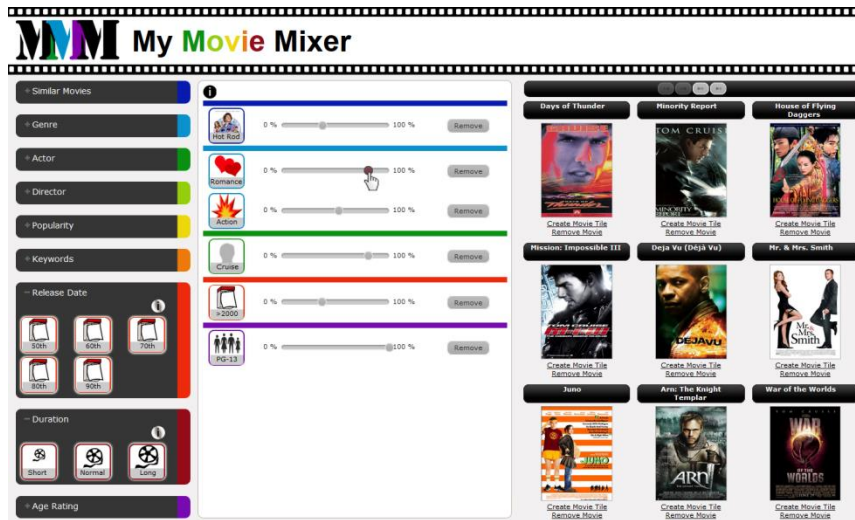


Abbildung 1: Oberfläche von MyMovieMixer: Links die Widgets, deren Kacheln in den mittleren Bereich gezogen werden können. Rechts die nach jeder Interaktion automatisch aktualisierten Empfehlungen.

dann zusätzlich über einen Slider, um dem jeweiligen Kriterium einen Stellenwert für die Empfehlungsgenerierung zuzuweisen. Auf diese Weise kann ausgedrückt werden, dass die Erfüllung eines bestimmten Kriteriums zwar grundsätzlich wünschenswert, dabei im Vergleich zu anderen Kriterien aber mehr bzw. weniger wichtig ist.

Da die zuvor genannten Eingaben, einschließlich der Manipulation der Slider, eine Aktualisierung der Ergebnisse (rechte Seite) auslösen, kann der Nutzer geeignete Wertebereiche für Slider-Einstellungen explorativ erlernen. Darüber hinaus können durch Anklicken einer Empfehlung Filmdetails sowie Erklärungen zu deren Zustandekommen, d. h., welche der Kriterien berücksichtigt werden konnten, als Dialogfenster abgerufen werden. Dies soll die Transparenz erhöhen und zugleich die Anwender befähigen, wie in kritikbasierten Recommendern gezielt mit Feedback zu reagieren und die zugrundeliegenden Kriterien auf- bzw. abzuwerten, um wahlweise mehr oder weniger auf diese zutreffende Vorschläge zu erhalten.

3.2 Algorithmische Umsetzung

Als lose gekoppeltes Hybridsystem (Burke, 2007) behandelt MyMovieMixer alle Kriterien zunächst unabhängig voneinander. Für jeden Film m und jedes Kriterium c_i wird ein Wert zwischen 0 und 1 ermittelt, welcher den Grad angibt, mit dem m das Kriterium erfüllt. Abhängig von der Art des Kriteriums geschieht die Berechnung des Erfüllungsgrades auf unterschiedliche Weise: Für ein Kriterium der Art „Film soll ähnlich zu Film x sein“ wird beispielsweise ein Item-basierter Recommender herangezogen, der die berechnete Ähnlichkeit von m zu x auf den Bereich $[0;1]$ abbildet. Andere Kriterien greifen zur Berechnung des Erfüllungsgrades hingegen auf Fuzzy Logic (Zadeh, 1965) zurück. Als Beispiel sei hier ein präferiertes Jahrzehnt (z. B. 1990er) genannt: Ein Film von 1993 erfüllt dieses Kriterium vollständig und wird dementsprechend auf 1 abgebildet. Ein Film von 1989 sollte aber nicht

komplett unberücksichtigt bleiben, wie es bei binärer Logik der Fall wäre. Daher wird gemäß der Fuzzy Set Theory eine unscharfe Erfüllungsfunktion verwendet, die bei knapper Verfehlung des Kriteriums immer noch einen gewissen, wenn auch niedrigeren Erfüllungsgrad (beispielsweise 0.5) zuweist. Darüber hinaus nutzt MyMovieMixer eine Reihe weiterer, inhaltsbasierter Empfehlungsverfahren: So wird die Suche über Tags unterstützt, wobei die Relevanz eines Tags in Bezug auf einen Film durch die Anwendung der TF-IDF-Heuristik (Baeza-Yates & Ribeiro-Neto, 1999) ermittelt wird. Zudem ist eine binäre Filterung z. B. nach Genres und Regisseuren möglich. Schließlich liegt für einen Film m und jedes Kriterium c_i eine Relevanz $r_i(m, c_i)$ im Bereich $[0;1]$ vor.

Um die aggregierte Gesamtrelevanz r für einen Film gemäß aller Kriterien zu ermitteln, werden die einzelnen Relevanzwerte mit Hilfe der angegebenen Gewichte w_i (die Einstellungen der Slider in Abbildung 1) als gewichtetes Mittel über alle Kriterien zusammengeführt (s. Formel 1).

$$r(m, c_1, \dots, c_n, w_1, \dots, w_n) = \frac{\sum_{i=1}^n w_i \cdot r_i(m, c_i)}{\sum_{i=1}^n w_i} \quad (1)$$

Abschließend wird gemäß der Gesamtrelevanz absteigend sortiert und die Filme mit dem höchsten Wert werden dem Nutzer präsentiert. Tabelle 1 illustriert die Formel anhand eines Beispiels: Gewünscht wird ein Film von Steven Spielberg (Kriterium c_1 mit Gewicht $w_1=100$) aus den 1990er Jahren (Kriterium c_2 mit Gewicht $w_2=50$). Im Sinne einer übersichtlichen Darstellung sei angenommen, dass nur drei Filme in der Datenbasis vorkommen.

Film	$r_1(m, c_1)$ (Regie)	$r_2(m, c_2)$ (Jahr)	$r(m, c_1, c_2, 100, 50)$
Indiana Jones 3 (Spielberg, 1989)	1.0	0.5	0.833
Jurassic Park (Spielberg, 1993)	1.0	1.0	1.000
Pulp Fiction (Tarantino, 1994)	0.0	1.0	0.333

Tabelle 1: Beispiel zur Verrechnung einzelner Relevanzwerte für einige Filme

4 Evaluation des Ansatzes

Bei der Entwicklung von MyMovieMixer verfolgten wir einen nutzerzentrierten Entwicklungsansatz, bei welchem insgesamt drei Nutzerstudien prozessbegleitend durchgeführt wurden. Im Folgenden wird über die letzte dieser Studien berichtet, bei der in einer summativen Evaluation die Interaktions- und Manipulationsmöglichkeiten mit dem System im Hinblick auf die folgenden Fragestellungen untersucht wurden.

1. Wie ist das Nutzungserleben?
2. Welche Probleme zeigen sich während der Nutzung von MyMovieMixer?
3. Wie nehmen die Nutzer die Interaktions- und Manipulationsmöglichkeiten wahr?
4. Ist das System MyMovieMixer intuitiv zu bedienen und schnell erlernbar?
5. Wie benutzerfreundlich ist das System MyMovieMixer insgesamt?
6. Wie wird die Qualität der generierten Empfehlungen wahrgenommen?

4.1 Studienaufbau und Datenerhebung

Die Evaluation wurde als Laborstudie mit mehreren Nutzeraufgaben konzipiert. Insgesamt umfasste die Stichprobe 30 Personen (14 w) im Alter von 21 bis 67 Jahren ($M=28$, $\sigma=10.18$), die Mehrzahl von ihnen Studierende (83.3%). Die erste Aufgabe der Probanden bestand in einer freien Exploration des Systems. Sie sollten dabei MyMovieMixer unter der Annahme nutzen, dass sie einen Film sehen möchten, aber noch nicht wissen, welchen. Nach dieser Aufgabe bewerteten die Teilnehmer in einem Fragebogen die Intuitivität der Bedienung. Nachfolgende Aufgaben umfassten die Erstellung von Kacheln, die Berücksichtigung vorgegebener Präferenzen in unterschiedlicher Anzahl und deren Gewichtung. Auch die Manipulation der Ergebnisse und die Betrachtung des Feedbacks, wie die empfohlenen Filme den eingegebenen Kriterien entsprechen, wurden fokussiert. Durch diese Szenarien wurde sichergestellt, dass die Probanden alle relevanten Funktionen des Systems nutzten. Während der Interaktion erfolgte ein Think Aloud mit einer Verhaltensbeobachtung durch die Versuchsleitung. Außerdem wurde die Bearbeitungszeit erfasst. Im Anschluss an die Aufgabebearbeitung beantworteten die Probanden weitere Fragebogen-Items. Dabei wurde zunächst die Erlernbarkeit, Verständlichkeit und kognitive Anstrengung bezüglich der Nutzung der Widgets, Kacheln und Slider sowie der Interaktion mit den Filmempfehlungen und der Detailanzeige erfragt. Des Weiteren enthielt der Fragebogen Items zur Erfassung der wahrgenommenen Qualität der Empfehlungen. Außerdem setzten wir die *System Usability Scale* (SUS; Brooke, 1996), den *User Experience Questionnaire* (UEQ; Laugwitz et al., 2008) sowie selbst generierte Items zum Gesamteindruck und zur Nutzungsabsicht ein.

4.2 Ergebnisse

Nach der ersten, explorativen Benutzung von MyMovieMixer gaben die Probanden auf einer positiv gepolten Skala von 1 bis 5 an, sich gut zurechtgefunden zu haben ($M=3.63$, $\sigma=1.16$). Die Intuitivität wurde bei der initialen Nutzung im Mittel mit 2.87 ($\sigma=0.97$) bewertet. Die nach Vollendung aller Aufgaben erhobenen Ergebnisse zeigen sehr hohe Werte in Bezug auf Erlernbarkeit und Verständlichkeit der Interaktion mit den verschiedenen Funktionen von MyMovieMixer sowie eine geringe kognitive Belastung (s. Tabelle 2).

	Widgets, Kacheln und Slider		Ergebnis- und Detailanzeige	
	M	σ	M	σ
Erlernbarkeit	4.53	0.819	4.37	0.809
Verständlichkeit	4.07	0.828	4.27	0.980
geringe kogn. Anstrengung	4.13	1.008	4.27	0.944

Tabelle 2: Beurteilung der Funktionen hinsichtlich Erlernbarkeit, Verständlichkeit und kognitiver Anstrengung

Das Zustandekommen der Filmempfehlungen wurde dem Großteil der Probanden klar: Lediglich 4 Probanden (13.3%) gaben an, dies nicht nachvollziehen zu können, was zum einen mit einer fehlenden Darstellung der Gewichtung der Kriterien in der Film-Detailanzeige begründet wurde, zum anderen damit, dass relevante Filme zu weit hinten in der Ergebnisliste erschienen. Die Qualität der Empfehlungen wurde dennoch sehr positiv mit einem Mittelwert von 4.17 ($\sigma=0.42$) bei positiver Polung auf einer Likert-Skala von 1 bis 5 bewertet.

Auch die Usability wurde mit einem Wert von durchschnittlich 77 % auf der SUS positiv bewertet. In Bezug auf die User Experience lagen die durchschnittlichen Bewertungen deutlich im positiven Bereich der siebenstufigen, bipolaren Skala (-3 bis 3) des UEQ ($M=1.94$, $\sigma=0.35$). Außerdem bewerteten die Teilnehmer den Umgang mit dem System auf einer Likert-Skala von 1 bis 5 durchschnittlich als eine spannende Erfahrung ($M=4.13$, $\sigma=0.90$).

Die Auswertung der Ergebnisse des Think Aloud, der Beobachtung sowie der Zeitmessung während der Interaktion mit MyMovieMixer zeigte, dass die meisten Probanden nach einer kurzen Orientierungsphase selbstständig die Interaktionsmöglichkeiten ausprobierten. Der Testung der Systemfunktionalitäten – insbesondere der Widgetleiste sowie dem Drag and Drop der Kacheln – schenkten sie dabei viel Zeit ($M=11.4$ min, $Range=5-26$ min). So erhielten sie bereits in der ersten Aufgabe einen Überblick über die wichtigsten Funktionalitäten. Die folgenden Aufgaben wurden zielgerichteter und deutlich schneller bearbeitet. Es zeigte sich, dass die Funktionen insgesamt schnell und ohne Hilfe der Versuchsleitung gefunden und verstanden wurden. Darüber hinaus wurde beobachtet, dass die Probanden die Drag-and-Drop-Funktionalität zusätzlich zu den bisher implementierten an weiteren Stellen in der Interaktion vermuteten, etwa zum Ziehen einer Filmempfehlung in den Arbeitsbereich.

5 Diskussion und Fazit

Wie eingangs herausgestellt, bilden grafische Visualisierungen, Interaktionsmöglichkeiten und Empfehlungstransparenz wichtige Kriterien für die Nutzerzufriedenheit bei Recommendern. Jedoch mangelt es an konkreten Umsetzungen solch visueller und interaktiver Lösungen. Besonders für komplexe hybride Recommender existieren derartige Bedienkonzepte bis zum jetzigen Zeitpunkt noch nicht. Mit MyMovieMixer haben wir ein visuelles Interaktionskonzept umgesetzt, welches den aufgezeigten Limitationen bisheriger Ansätze entgegenwirken soll. Da Erklärungen der Vorschläge die wahrgenommene Effektivität, die Zufriedenheit, den empfundenen Einflussgrad, die Akzeptanz sowie das Vertrauen in das System erhöhen können (Pu et al., 2012), sollte der Empfehlungsprozess dabei für den Nutzer transparent sein. Durch die interaktive Darstellung geht die in der Evaluation ermittelte hohe Empfehlungsqualität wie nach Zhang et al. (2008) mit gesteigerter Nachvollziehbarkeit einher. Durch direkt auf Aktionen erfolgendes Feedback und die unmittelbare Anpassung der Empfehlungen sollte zudem die empfundene Kontrolle des Nutzers erhöht werden.

Mehr Interaktions- und Kontrollmöglichkeiten können allerdings zu einem Mehraufwand bei der Nutzung führen. Die vorgestellten Evaluationsergebnisse zeigen jedoch, dass MyMovieMixer im Allgemeinen als intuitiv und angenehm zu bedienen bewertet wurde. Dabei konnte eine hohe Empfehlungsqualität erreicht werden und die Nutzer empfanden die Vorschläge als gut nachvollziehbar. In Bezug auf *Fragestellung 1* weisen sowohl die Ergebnisse des UEQ als auch die des Think Aloud und der Verhaltensbeobachtung auf ein positives Nutzungserleben hin. Aufgrund der Beobachtung, dass die langen Bearbeitungszeiten der ersten Aufgabe nicht auf Probleme, sondern auf die intensive Exploration des Systems durch die Teilnehmer zurückzuführen sind, ist anzunehmen, dass der Umgang mit dem System den Nutzern Spaß gemacht und ihr Interesse geweckt hat. Dies deckt sich mit den Angaben der

Probanden, dass der Umgang mit dem System eine spannende Erfahrung gewesen sei. *Fragstellung 2* beleuchtete etwaige Probleme bei der Nutzung von MyMovieMixer: Insgesamt traten wenige Probleme auf. Optimierungen sind im Hinblick auf eine konsequentere Umsetzung des verwendeten Drag-and-Drop-Prinzips jedoch empfehlenswert. Hinsichtlich des Interaktionskonzepts der Anwendung (*Fragstellung 3*) wurden von den Teilnehmern im Rahmen des Think Aloud die direkte Manipulation, die grafische Visualisierung der Empfehlungskriterien als Kacheln und das Drag and Drop als eine der hauptsächlichsten Interaktionsmethoden positiv hervorgehoben. Mit *Fragstellung 4* wurden die intuitive Bedienbarkeit und die Erlernbarkeit von MyMovieMixer untersucht. Die Beobachtungsergebnisse zeigen einen recht selbstverständlichen, anfangs von Exploration geprägten und später zielgerichteten Umgang mit dem System. Dieser Eindruck der intuitiven Handhabung und der einfachen Erlernbarkeit deckt sich mit den Fragebogen-Ergebnissen. Im Hinblick auf *Fragstellung 5* konnte die Benutzerfreundlichkeit des Systems durch die zuvor diskutierten Fragestellungen und die Ergebnisse der SUS gezeigt werden. Die Qualität der Empfehlungen (*Fragstellung 6*) wurde sehr positiv bewertet. Auch wurde die Darstellung, auf welchen Kriterien diese basieren, als transparent wahrgenommen.

Der Ansatz scheint also in Bezug auf Usability, User Experience, Intuitivität, Transparenz und somit letztlich im Hinblick auf die Nutzerzufriedenheit vielversprechend zu sein. Durch das visuelle Konzept konnte eine unterhaltsame, attraktive Gestaltung erreicht werden, was nach Pu et al. (2012) die Nutzer zu mehr Eingaben motiviert. Gleichzeitig wurde der zu leistende kognitive Aufwand als gering bewertet, was sich u. a. durch die Verwendung von Kacheln als Repräsentation für Kriterien und die Interaktion mittels Drag and Drop begründen lässt. Das Drag-and-Drop-Prinzip sollte jedoch noch umfassender und konsequenter eingesetzt werden, etwa beim Hinzufügen empfohlener Filme als Kriterium oder für das Löschen gewählter Kriterien. Weiterhin ist denkbar, die initial in den Widgets angezeigten Kacheln unter Verwendung von Empfehlungsalgorithmen an vorherige Nutzereingaben anzupassen. In zukünftigen Studien planen wir zudem eine detailliertere Untersuchung der Empfehlungsqualität sowie der möglichen Berücksichtigung weiterer relevanter Aspekte, wie z. B. Diversität und Serendipität.

Insgesamt erlaubt das vorgestellte Bedienkonzept für hybride Recommender auf situative Bedürfnisse zu reagieren, ohne langfristige Nutzerprofile zu benötigen. Dadurch wird auch das weitverbreitete Cold-Start-Problem umgangen, da sämtliche Eingaben ad-hoc erfolgen. Zudem bietet die konsistente Visualisierung der Suchkriterien viele Vorteile und hat sich besonders zur Steigerung der Transparenz bei den Empfehlungen bewährt. Die einfach gehaltene Darstellung ermöglicht es dem Anwender dabei in Anlehnung an kritikbasierte Recommender, die Empfehlungen anhand zahlreicher Kriterien kontinuierlich seinem Suchziel anzupassen, ohne dass MyMovieMixer dabei zu komplex oder unübersichtlich wirkt. Durch das angewendete Prinzip der losen Kopplung ist es zudem möglich, beliebige Widgets und somit beliebige Recommender und Filtermethoden zu integrieren.

Danksagung

Wir bedanken uns bei Jennifer Frantz und Katharina Lattenkamp für die Mitarbeit im Entwicklungsprozess von MyMovieMixer.

Literaturverzeichnis

- Baeza-Yates, R. & Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. New York: ACM.
- Bostandjiev, S., O'Donovan, J. & Höllerer, T. (2012). TasteWeights: A visual interactive hybrid recommender system. In *Proc. RecSys 2012*. New York: ACM. 35–42.
- Brooke, J. (1996). SUS – A quick and dirty usability scale. *Usability Evaluation in Industry*, 189–194.
- Burke, R. (2007). Hybrid web recommender systems. In P. Brusilovsky, A. Kobsa & W. Nejdl (Hrsg.), *The Adaptive Web. Methods and Strategies of Web Personalization*. Berlin: Springer. 377–408.
- Chen, L. & Pu, P. (2012). Critiquing-based recommenders: Survey and emerging trends. *User Modeling and User-Adapted Interaction*, 22(1-2), 125–150.
- Gretarsson, B., O'Donovan, J., Bostandjiev, S., Hall, C. & Höllerer, T. (2010). SmallWorlds: Visualizing social recommendations. *Computer Graphics Forum*, 29(3), 833–842.
- Knijnenburg, B. P., Willemsen, M. C., Gantner, Z., Soncu, H. & Newell, C. (2012). Explaining the user experience of recommender systems. *User Mod. and User-Adapted Interaction*, 22(4-5), 441–504.
- Konstan, J. A. & Riedl, J. (2012). Recommender systems: From algorithms to user experience. *User Modeling and User-Adapted Interaction*, 22(1-2), 101–123.
- Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. In *HCI and Usability for Education and Work*. Berlin: Springer. 63–76.
- McCarthy, K., Reilly, J., McGinty, L. & Smyth, B. (2005). Experiments in dynamic critiquing. In *Proc. IUI 2005*. New York: ACM. 175–182.
- Parra, D., Brusilovsky, P. & Trattner, C. (2014). See what you want to see: Visual user-driven approach for hybrid recommendation. In *Proc. IUI 2014*. New York: ACM. 235–240.
- Pu, P., Chen, L. & Hu, R. (2012). Evaluating recommender systems from the user's perspective: Survey of the state of the art. *User Modeling and User-Adapted Interaction*, 22(4-5), 317–355.
- Ricci, F., Rokach, L. & Shapira, B. (Hrsg.). (2010). *Recommender Systems Handbook*. Berlin: Springer. 1–35.
- Salton, G. & Buckley, C. (1997). Improving retrieval performance by relevance feedback. In *Readings in Information Retrieval*. Morgan Kaufmann. 355–364.
- Tintarev, N. & Masthoff, J. (2010). Designing and evaluating explanations for recommender systems. In Ricci, F., Rokach, L., Shapira, B. & Kantor, P. B. (Hrsg.), *Recommender Systems Handbook*. Berlin: Springer. 479–510.
- Vig, J., Sen, S. & Riedl, J. (2011). Navigating the tag genome. In *Proc. IUI 2011*. New York: ACM. 93–102.
- Xiao, B. & Benbasat, I. (2007). E-commerce product recommendation agents: Use, characteristics, and impact. *MIS Quarterly*, 31(1), 137–209.
- Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8(3), 338–353.
- Zhang, J., Jones, N. & Pu, P. (2008). A visual interface for critiquing-based recommender systems. In *Proc. EC 2008*. New York: ACM. 230–239.