

Content Recommendation in eLearning-enabled Online Social Networks

Andreas Winschu

andreas.winschu@haw-hamburg.de

iNET RG, Department of Computer Science
Hamburg University of Applied Sciences

January 9, 2013



Hochschule für Angewandte
Wissenschaften Hamburg

Hamburg University of Applied Sciences

- 1 Motivation
- 2 Current State
- 3 Forschung
- 4 Ausblick

Motivation

Bisherige Arbeit



- NoSQL
- Semantic Web / morgen-in-meiner-stadt.de
- HTML-5 Apps

Klassische Elearning Umgebungen

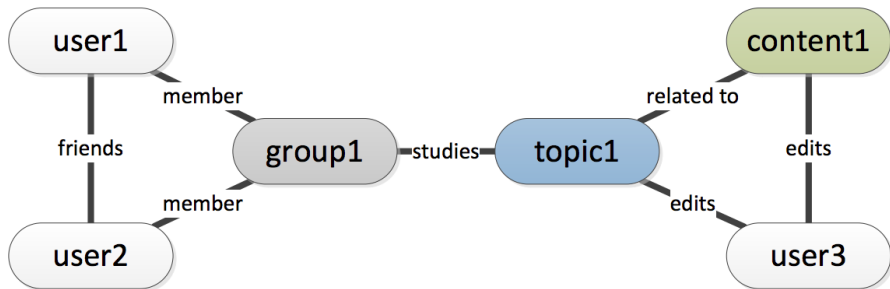
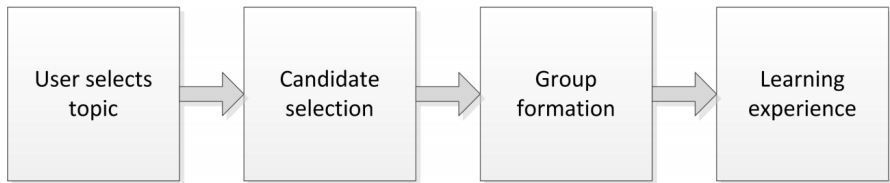
- Intra-group Kommunikation in festgelegten Klassenräumen
- Geleitet durch einen Instruktor
 - Legt Gruppen fest
 - Analysiert Ergebnisse
 - Beobachtet den Lernprozess

Online Social Networks (OSN)

- Soziale Interaktion zwischen Freunden
- Gruppen sind Nutzerinitiiert
- Allgegenwärtige Nutzung im Alltag

- Fokus der Arbeit auf Integration von OSN und eLearning Umgebung
- Verzicht auf einen Instruktor führt zu folgenden Herausforderungen:
 - **1** Wie unterstützt man einen effektiven Teambildungsprozess?
 - **2** **Wie liefert man relevante Inhalte für die Lerngruppe?**
 - **3** Wie erleichtert man einen konsistenten Lernprozess inklusive Feedback und Korrektur?

Vorgehen



Automatische Empfehlung

- Bewertung der Items durch User
- Vorhersage der Bewertung für neue Items
- Empfehlung der höchst vorhergesagten Items

	Avatar	Escape From Alcatraz	K-Pax	Shawshank Redemption	Usual Suspects	
Alice	1	2	?	5	4	
Bob	3	?	2	5	3	→ $R(u)$
Clint	?	?	?	?	2	
Dave	5	?	4	4	5	
Ethan	4	?	1	1	?	

$R(u,i)$ $R(i)$

Bewertungsfunktion: $r : U \times I \rightarrow V$

$I(u)$: Items $i \in I$ bewertet durch User u

$U(i)$: User u , welche Bewertungen an i abgaben

Automatische Empfehlung

- r nicht definiert für den Gesamten Raum $U \times I$
- r meist enorm klein
- U und I haben Millionen Elemente
- r Extrapolierung auf den gesamten Raum

	Avatar	Escape From Alcatraz	K-Pax	Shawshank Redemption	Usual Suspects	
Alice	1	2	?	5	4	
Bob	3	?	2	5	3	→ $R(u)$
Clint	?	?	?	?	2	
Dave	5	?	4	4	5	
Ethan	4	?	1	1	?	

$R(u,i)$ $R(i)$

Bewertungsfunktion: $r : U \times I \rightarrow V$

$I(u)$: Items $i \in I$ bewertet durch User u

$U(i)$: User u , welche Bewertungen an i abgaben

Recommendation

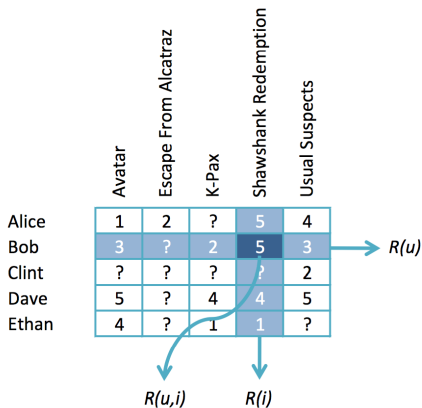
Inputs: a user u

Outputs: an item i^* to be recommended

1. For all unrated items i of user u , compute a prediction $R(u, i)$ using some algorithm.
2. Recommend the item i with the highest prediction.

	Avatar	Escape From Alcatraz	K-Pax	Shawshank Redemption	Usual Suspects	
Alice	1	2	?	5	4	
Bob	3	?	2	5	3	→ $R(u)$
Clint	?	?	?	?	2	
Dave	5	?	4	4	5	
Ethan	4	?	1	1	?	

$R(u, i)$ $R(i)$



Bewertungsfunktion: $r : U \times I \rightarrow V$

$I(u)$: Items $i \in I$ bewertet durch User u

$U(i)$: User u , welche Bewertungen an i abgaben

Matrix R kann auf verschieden Weisen vorhergesagt werden:

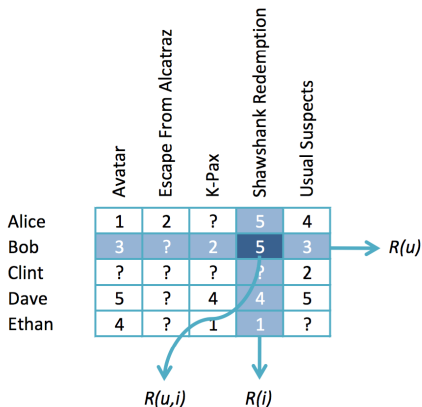
- Machine Learning
- Wahrscheinlichkeitstheorie
- Heuristische Verfahren

Verfahren unterteilen sich anhand Strategie:

- content-based
- collaborative filtering

	Avatar	Escape From Alcatraz	K-Pax	Shawshank Redemption	Usual Suspects	
Alice	1	2	?	5	4	
Bob	3	?	2	5	3	→ $R(u)$
Clint	?	?	?	?	2	
Dave	5	?	4	4	5	
Ethan	4	?	1	1	?	

$R(u,i)$ $R(i)$



Bewertungsfunktion: $r : U \times I \rightarrow V$

$I(u)$: Items $i \in I$ bewertet durch User u

$U(i)$: User u , welche Bewertungen an i abgaben

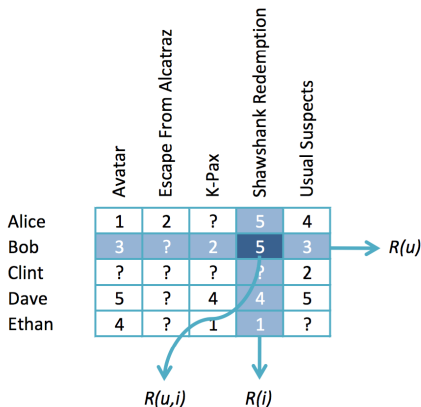
Content-based Recommendation

Empfehlung anhand Item
Ähnlichkeit

- Information Retrieval
- Item zu Item Beziehungen
- Gemeinsame Item Charakteristiken
- Ähnlichkeitsfunktion
- Attribute im Benutzerprofil, durch Interaktion an Items
- Beziehung des neuen Items zu der Profilhistorie

	Avatar	Escape From Alcatraz	K-Pax	Shawshank Redemption	Usual Suspects	
Alice	1	2	?	5	4	
Bob	3	?	2	5	3	→ $R(u)$
Clint	?	?	?	?	2	
Dave	5	?	4	4	5	
Ethan	4	?	1	1	?	

$R(u,i)$ $R(i)$



Bewertungsfunktion: $r : U \times I \rightarrow V$

$I(u)$: Items $i \in I$ bewertet durch User u

$U(i)$: User u , welche Bewertungen an i abgaben

Content-based Recommendation

Formell:

$R(u, i^*)$ ermittelt durch
Bewertungen $R(u, ik)$ von dem
User u an allen mit i^*
ähnlichen Items $ik \in I(u)$

	Avatar	Escape From Alcatraz	K-Pax	Shawshank Redemption	Usual Suspects	
Alice	1	2	?	5	4	
Bob	3	?	2	5	3	→ $R(u)$
Clint	?	?	?	?	2	
Dave	5	?	4	4	5	
Ethan	4	?	1	1	?	

$R(u, i)$ $R(i)$

Bewertungsfunktion: $r : U \times I \rightarrow V$

$I(u)$: Items $i \in I$ bewertet durch User u

$U(i)$: User u , welche Bewertungen an i abgaben

Empfehlung anhand
Bewertungen ähnlicher User

- User zu User Beziehungen
- Ähnliche Menschen haben ähnliche Interessen
- Empfehlung restlicher Interessen
- Profile spielen keine Rolle

	Avatar	Escape From Alcatraz	K-Pax	Shawshank Redemption	Usual Suspects	
Alice	1	2	?	5	4	
Bob	3	?	2	5	3	→ $R(u)$
Clint	?	?	?	?	2	
Dave	5	?	4	4	5	
Ethan	4	?	1	1	?	

$R(u,i)$ $R(i)$

Bewertungsfunktion: $r: U \times I \rightarrow V$

$I(u)$: Items $i \in I$ bewertet durch User u

$U(i)$: User u , welche Bewertungen an i abgaben

Collaborative Filtering

Formell:

$R(u, i^*)$ ermittelt durch
Bewertungen $R(u, ik)$ von allen
mit dem User u ähnlichen
Usern $uk \in U$ welche das Item
 i^* bewertet haben.

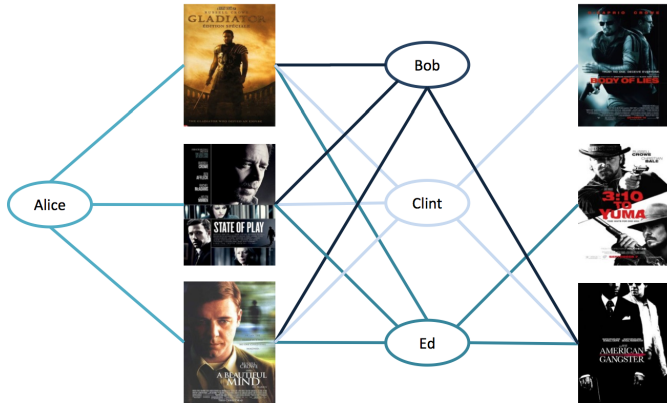
	Avatar	Escape From Alcatraz	K-Pax	Shawshank Redemption	Usual Suspects	
Alice	1	2	?	5	4	
Bob	3	?	2	5	3	→ $R(u)$
Clint	?	?	?	?	2	
Dave	5	?	4	4	5	
Ethan	4	?	1	1	?	

$R(u, i)$ $R(i)$

Bewertungsfunktion: $r : U \times I \rightarrow V$

$I(u)$: Items $i \in I$ bewertet durch User u

$U(i)$: User u , welche Bewertungen an i abgaben

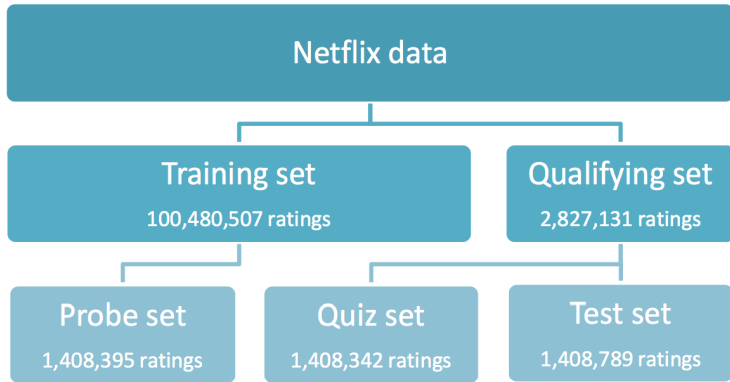


Cinematch recommender System

- Entwickelt 6 Jahre einen robusten Recommendation Algorithmus
- Interessante Verbindungen zwischen Filmen und Usern
- 2006 keine Ideen mehr
- Public Contest an data mining, machine learning scientists
- Training Set mit Ratings von 1-5
- Qualifying Set
- Root Mean Square Error als Maß
- 1 Million Dollars Preis
- 10 % Verbesserung

Research

Netflix Prize



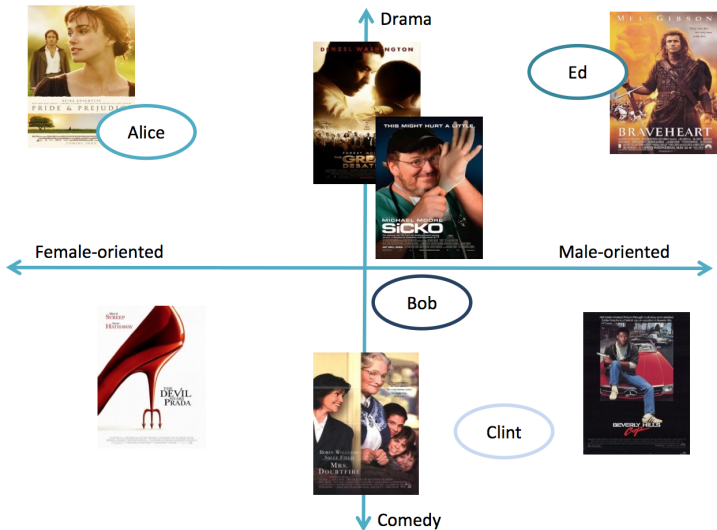
Ergebnis:

- Bellkor gewinnt den Preis
- Keins der Algorithmen ist eindeutig besser
- Bellkor nutzt eine Kombination aus 107 Verfahren
 - neighborhood für Nahe Beziehungen
 - Latent Factor Approach für Allgemeine Trends
- Unzählige Publicationen

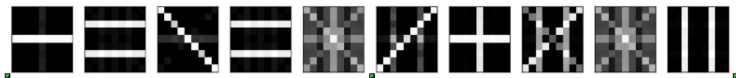
Probleme:

- sparsity
- diversity
- scalability
- implicit feedback

Machines



Machines

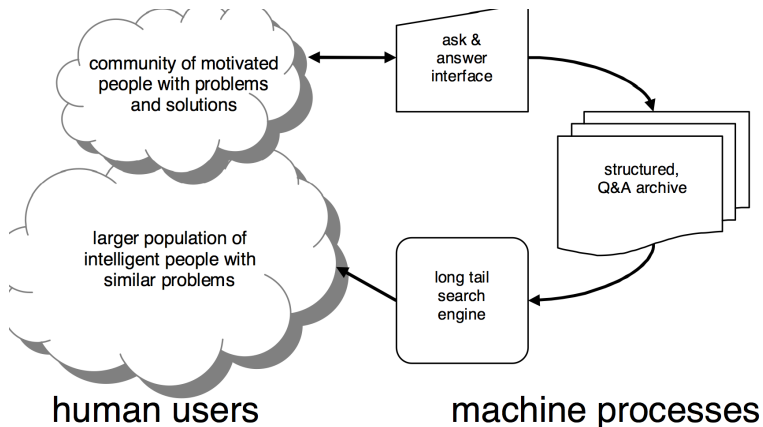


- Ruby on Rails
- Data Privacy
- Contacts
 - Groups, Followers etc.
- Sharing
 - private messaging, share a selection of aspects
- Diaspora Connect
 - sign into other services
- P2P, Pods
- Pfofile, Settings ...

- intuitive, using a graph model for data representation
- reliable, with full ACID transactions
- durable and fast, using a custom disk-based, native storage engine
- massively scalable, up to several billion nodes/relationships/properties
- highly-available, when distributed across multiple machines
- expressive, with a powerful, human readable graph query language
- fast, with a powerful traversal framework for high-speed graph queries
- simple, accessible by a convenient REST interface or an object-oriented Java APIs
- RDF compatible

- Im Web meist collected intelligence
- Mass authoring anstatt mass authority
 - most popular, buzzwords
- Collective Intelligence als lange Vision des Semantic Web
- Erzeugung von Expertensystemen
 - Menschen als Poducer, lernen durch Kommunikation und erzeugen Daten
 - Maschinen sammeln Daten und leisten Inferenz
- Bottleneck in Wissensbeschaffung
 - Aufwand um Daten in verarbeitbare Form zu bekommen

FAQ Sphere



- Reinvent the wheel (Recommendation)
- Collaborative Filtering nicht auf DB Ebene möglich
- Algorithmen lassen sich nicht auf Group Recommendation erweitern
- Große Auswahl an Verfahren
- Diaspora nicht erweiterbar