

# Next Generation Recommender Systems

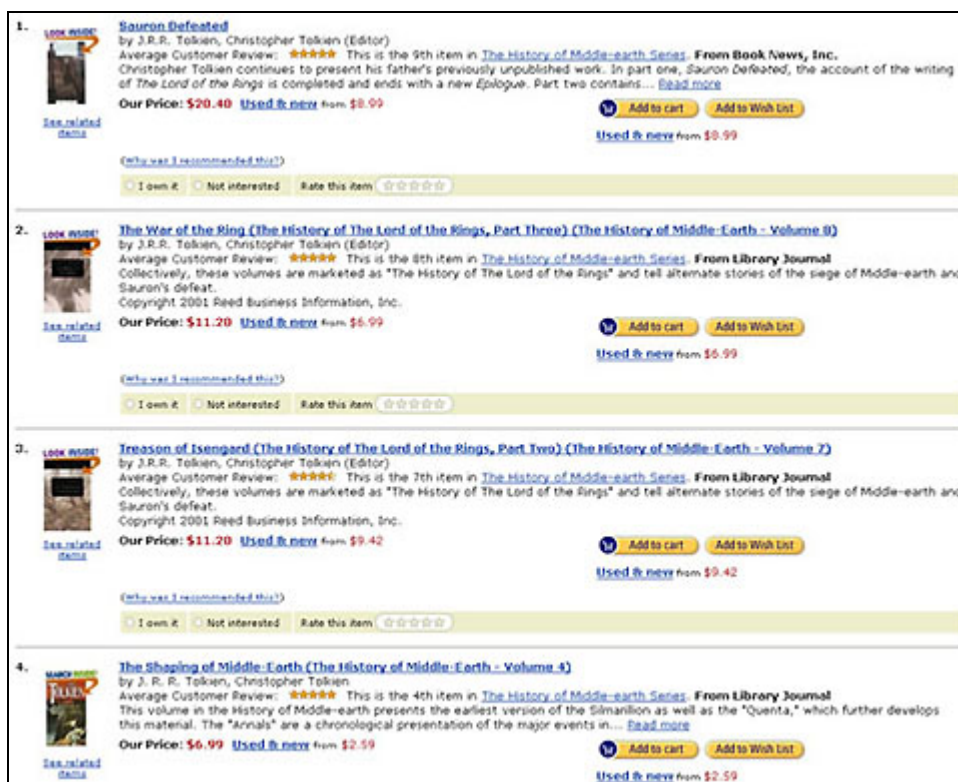
## Overview

Recommender systems are personalization tools that intend to provide people with lists of suggestions that best reflect their individual taste. In order to create profiles of the users' behavioral patterns, explicit ratings (e.g., user  $a$  rates book  $b$  with 4 of 5 stars) and purchase data are collected.

However, in the past, recommender systems research has primarily concentrated on accuracy metrics only, i.e., increasing the performance along precision/recall and mean absolute error (MAE) benchmarks.

However, some issues have remained unnoticed, for instance the lack of *diversity* in recommendation lists. Overfitting of personal user profiles may lead to situations where most recommendations come from one single field of interest only, as can be seen in the image below, which shows [C.-N. Ziegler's](#) top-4 recommendations on [Amazon.com](#).

All 4 recommendations are from the Tolkien universe, though the user's interest also span other fields of interest, e.g., travel, medieval romance, and science. Unfortunately, these are not contained in the recommended products list. Particularly commercial recommender systems suffer from lack of diversity and overfitting.



The screenshot displays a list of four book recommendations on Amazon.com. All four books are part of 'The History of Middle-earth' series by J.R.R. Tolkien and Christopher Tolkien. The recommendations are:

- Sauron Defeated**: Average Customer Review: ★★★★★. Price: \$20.40 (Used & new from \$8.99).
- The War of the Ring (The History of The Lord of the Rings, Part Three) (The History of Middle-earth - Volume 9)**: Average Customer Review: ★★★★★. Price: \$11.20 (Used & new from \$6.99).
- Treason of Isengard (The History of The Lord of the Rings, Part Two) (The History of Middle-earth - Volume 7)**: Average Customer Review: ★★★★★. Price: \$11.20 (Used & new from \$9.42).
- The Shaping of Middle-earth (The History of Middle-earth - Volume 4)**: Average Customer Review: ★★★★★. Price: \$6.99 (Used & new from \$2.59).

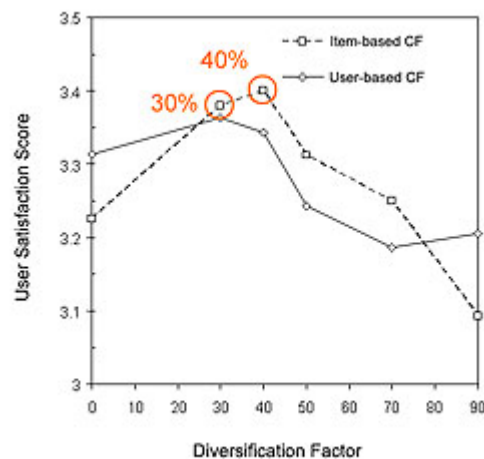
Each item includes a 'Look inside' button, a 'See related items' link, a 'Who was I recommended this?' link, and a 'Rate this item' section with a star rating.

Overfitting effect in recommendation lists

In order to overcome the diversity issue and have recommendation lists reflect all major fields of interest in a proportional weighting, we have developed the novel **Topic Diversification** method, in cooperation with [GroupLens Research](#) labs in Minneapolis, MN, USA.

In order to demonstrate that our approach bears some actual benefit for the user, increasing his level of satisfaction with recommendations made, we collected extensive user data from [BookCrossing](#) (see below) and computed personalized surveys that were completed by more than 2,100 BookCrossing members. The results confirmed our hypothesis. Users like to have diversified recommendation lists spanning their entire spectrum of interests rather than merely one or two topics.

Hereby, tests have indicated that diversification levels around 30-40% are best for the item-based recommender algorithm that commercial systems (e.g., [Amazon.com](#), [TiVo](#)) are typically using (see below).



### Effect of increasing diversification on user satisfaction

## Publications

The following list identifies papers published in prestigious conferences that are relevant for the project. Please feel free to make inquiries in order to obtain most recent, still unpublished work.

- [Improving Recommendation Lists Through Topic Diversification](#); Cai-Nicolas Ziegler, Sean M. McNee, Joseph A. Konstan, Georg Lausen; *Proceedings of the 14th International World Wide Web Conference (WWW '05)*, May 10-14, 2005, Chiba, Japan (acceptance rate: 14%, 77/550).

Download: [ [PDF](#) ] [ [BibTeX](#) ]

- [Taxonomy-driven Computation of Product Recommendations](#); Cai-Nicolas Ziegler, Georg Lausen, Lars Schmidt-Thieme; *Proceedings of the 2004 ACM International Conference on Information and Knowledge Management (CIKM '04)*, November 8-13, 2004, Washington, D.C., USA (acceptance rate: 19%, 59/303).

Download: [ [PDF](#) ] [ [BibTeX](#) ]

# BookCrossing Dataset

Collected by Cai-Nicolas Ziegler in a 4-week crawl (August / September 2004) from the Book-Crossing community with kind permission from Ron Hornbaker, CTO of Humankind Systems. Contains 278,858 users (anonymized but with demographic information) providing 1,149,780 ratings (explicit / implicit) about 271,379 books. Feel free to [download](#) the full dataset.

## Project Members

- [Dr.-Ing. Cai-Nicolas Ziegler](#),  
Post-doctoral research fellow, `chiegler (at) informatik.uni-freiburg.de`.
- [Dipl.-Inf. Kai Simon](#),  
PhD student and research fellow, `ksimon (at) informatik.uni-freiburg.de`.
- [Prof. Dr. Georg Lausen](#),  
Head of databases group, `lausen (at) informatik.uni-freiburg.de`.

Moreover, several students are engaged in the project.

## Cooperations

- [Prof. Joseph A. Konstan](#),  
GroupLens Research faculty, University of Minnesota.
- [Sean M. McNee](#),  
PhD student, University of Minnesota.