
Case study: personalized game recommendations on the mobile Internet

Case studies in recommender systems

- **The MovieLens data set, others**
 - focus on improving the Mean Absolute Error ...
- **What about the business value?**
 - nearly no real-world studies
 - exceptions, e.g., Dias et al., 2008.
 - e-Grocer application
 - CF method
 - short term: below one percent
 - long-term, indirect effects important
- **This study**
 - measuring impact of different RS algorithms in Mobile Internet scenario
 - more than 3% more sales through personalized item ordering

Application platform

- **Game download platform of telco provider**
 - access via mobile phone
 - direct download, charged to monthly statement
 - low cost items (0.99 cent to few Euro)
- **Extension to existing platform**
 - "My recommendations"
 - in-category personalization (where applicable)
 - start-page items, post-sales items
- **Control group**
 - natural or editorial item ranking
 - no "My Recommendations"



Study setup

- **6 recommendation algorithms, 1 control group**
 - CF (item-item, SlopeOne), Content-based filtering, Switching CF/Content-based hybrid, top rating, top selling
- **Test period:**
 - 4 weeks evaluation period
 - about 150,000 users assigned randomly to different groups
 - only experienced users
- **Hypothesis (H1 – H4)**
 - H1: Pers. recommendations **stimulate more users to view items**
 - H2: Person. recommendations **turn more visitors into buyers**
 - H3: Pers. recommendations stimulate **individual users to view more items**
 - H3: Pers. recommendations stimulate **individual users to buy more items**

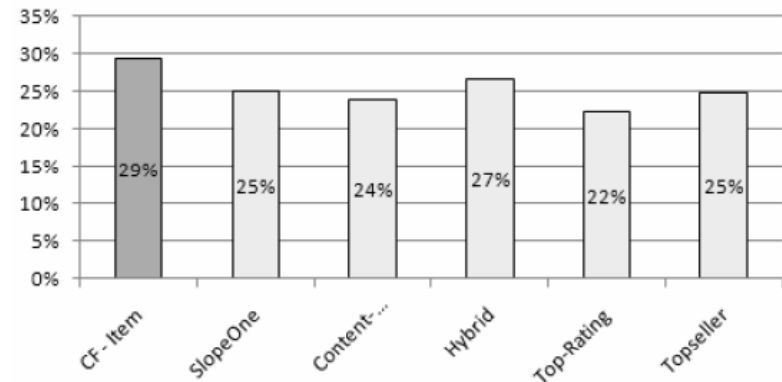
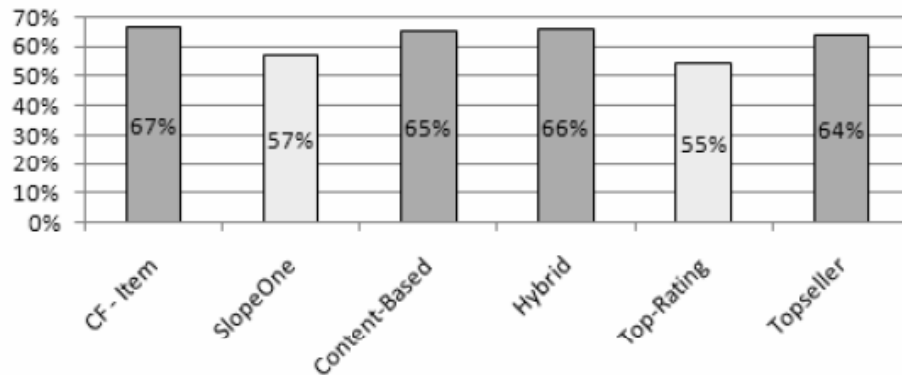
Measurements

- **Click and purchase behavior of customers**
 - customers are always logged in
 - all navigation activities stored in system

 - **Measurements taken in different situations**
 - my Recommendations, start page, post sales, in categories, overall effects
 - metrics
 - item viewers/platform visitors
 - item purchasers/platform visitors
 - item views per visitor
 - purchases per visitor

 - **Implicit and explicit feedback**
 - item view, item purchase, explicit ratings
-

My Recommendations conversion rates



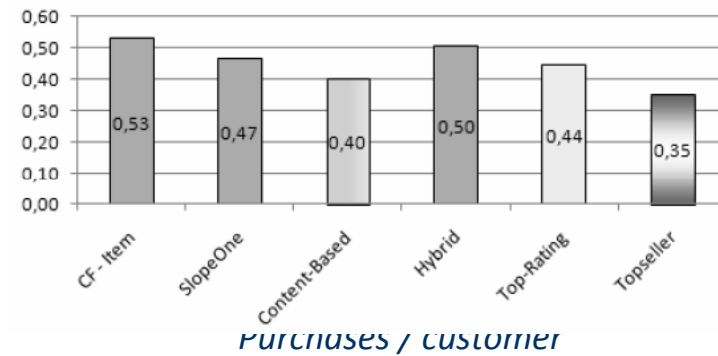
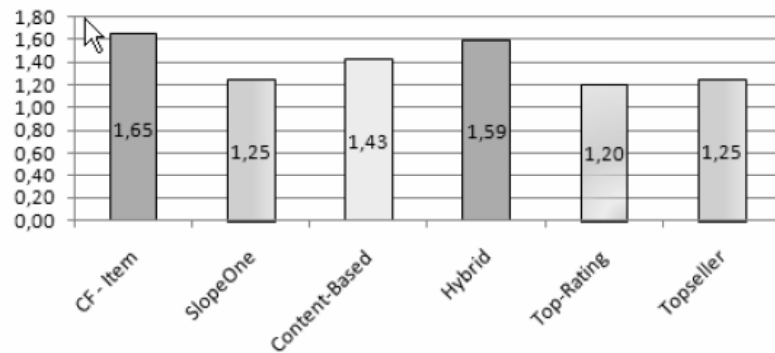
- **Conversion rates**

- top-rated items (SlopeOne, Top-Rating) appear to be non-interesting
- only CF-Item able to turn more visitors into buyers ($p < 0.01$)

- **Overall on the platform**

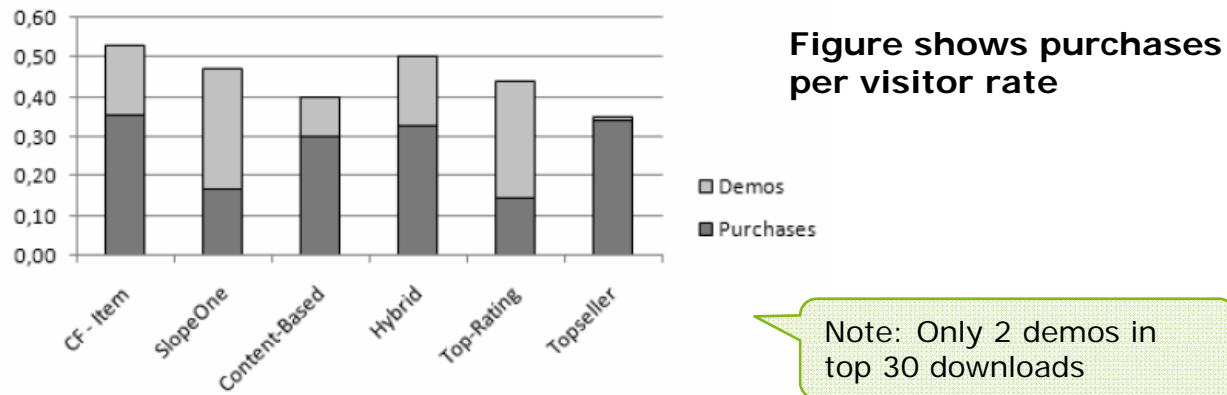
- no significant increase on both conversion rates (for frequent users!)

My Recommendations sales increase (1)



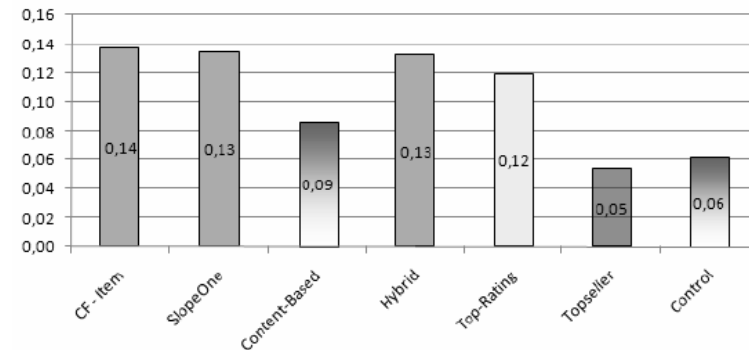
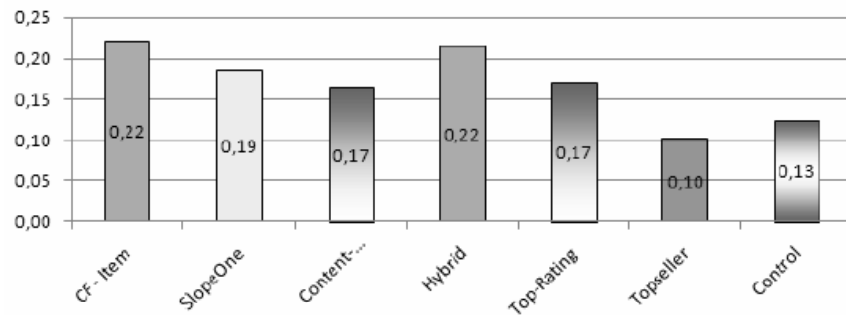
- **Item views:**
 - except SlopeOne, all personalized RS outperform non-personalized techniques
- **Item purchases**
 - RS measurably stimulate users to buy/download more items
 - content-based method does not work well here

My Recommendations sales increase (2)



- **Demos and non-free games:**
 - previous figures counted all downloads
 - figure shows
 - personalized techniques comparable to top seller list
 - however, can stimulate interest in demo games
 - **Note, Rating possible only after download**
-

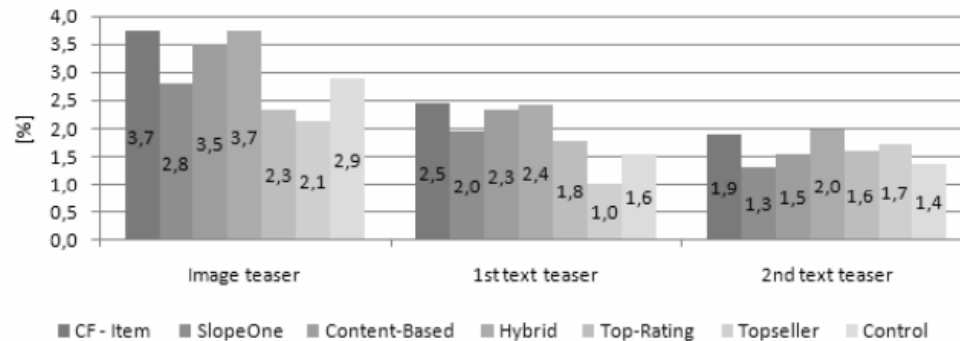
Post-sales recommendations



Findings

- recommending "more-of-the-same", top sellers or simply new items does not work well
- top-Rating and SlopeOne nearly exclusively stimulate demo downloads (Not shown)
- top-Seller und control group sell no demos

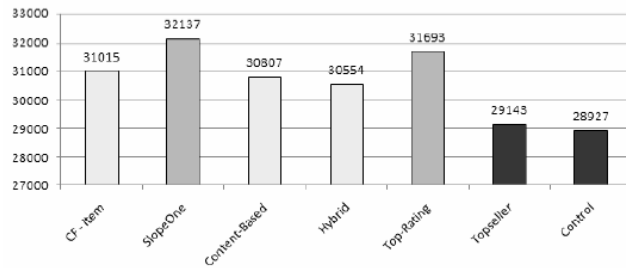
Start page recommendations



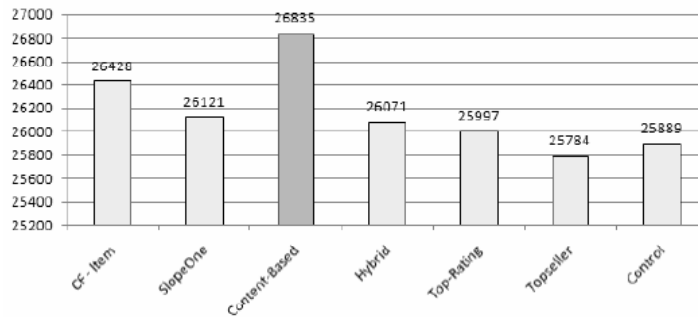
- **Example:**
 - purchasers / visitors conversion rate
- **Findings:**
 - visual presentation is important, click distribution as expected (omitted here)
 - personalization raises attraction also on text links
 - proposing **new items** works also very well on the start pages

Overall effects

- Overall number of downloads (free + non-free games)



- Pay games only



Notes

In-category measurements not shown in paper.

Content-based method outperforms others in different categories

(half price, new games, erotic games)

Effect: 3.2 to 3.6% sales increase!

Further observations: ratings on the Mobile Internet

- **Only 2% of users issued at least one rating**
 - most probably caused by size of displays
 - in addition: Particularity of platform; rating only after download
 - insufficient coverage for standard CF methods
- **Implicit ratings**
 - also count item views and item purchases
 - increase the coverage of CF algorithms
 - MAE however not a suitable measure anymore for comparing algorithms

Summary

- **Large case study on business effects of RS**
 - significant sales increase can be reached! (max. 1% in past with other activities)
 - more studies needed
 - value of MAE measure ...

- **In addition**
 - recommendation in navigational context
 - acceptance of recommendation depends on situation of user

- **Further work**
 - comparison of general sales behavior
 - more information in data to be found