

Improving News Recommendations with Implicit Negative Feedback

Ján Suchal

Supervisor: Professor Pavol Návrát

News recommendations

- dynamic domain with fast item aging
- limited recommendation space (TOP 3 or TOP1)
- users very sensitive to bad recommendations
- content & collaborative approaches
- source data: server access logs & article metadata

Detecting implicit negative feedback

- mining articles that users do not like
- mining average usage patterns in server logs
- detecting unclicked items probably seen by user
=> filtering uninteresting articles

Algorithm 1: Finding seen and not clicked articles for given user

```

Input: User visits  $V$ , threshold  $\tau$ 
Data: Top-k articles for each time window period
Result: Collection of articles

seenArticleCounts  $\leftarrow []$ ;
for  $visit \in V$  do
     $period \leftarrow \text{FindPeriod}(visit)$ ;
    for  $article \in \text{FindTopArticlesForPeriod}(period)$  do
         $seenArticleCounts[article] \leftarrow seenArticleCounts[article] + 1$ ;

notClicked  $\leftarrow \emptyset$ ;
for  $(article \rightarrow count) \in seenArticleCounts$  do
    if  $count > \tau$  then
         $notClicked \leftarrow article$ 
return notClicked
    
```



Experiments

- collaboration with largest slovak online news portal sme.sk
- k-nearest neighbor collaborative recommender with negative implicit feedback filtering
- real traffic (1M+ pageviews/day, ~600 recommendations/min)
~2.8M recommendations generated
~25K recommendation clicks
~12K rejected recommendations
- measured click-through rate and recommendation rejection rate
- simultaneous A/B testing of different setups of algorithms

k	click-through rate (CTR)			reject rate (RR)		
	baseline	with NIF	difference (%)	baseline	with NIF	difference (%)
5	0.730	0.851	16.67	0.424	0.412	-2.97
10	0.823	0.920	11.76	0.409	0.415	1.68
15	0.858	0.857	-0.13	0.353	0.284	-19.5
20	0.872	0.911	4.56	0.479	0.537	11.9
25	0.807	0.904	12.05	0.409	0.307	-25
30	0.839	0.949	13.16	0.353	0.505	43.15
35	0.812	0.892	9.78	0.390	0.312	-20.06
40	0.956	0.955	-0.09	0.549	0.397	-27.65
45	0.891	0.964	8.17	0.425	0.483	13.66
50	0.949	1.039	9.52	0.586	0.435	-25.86
average	0.854	0.924	8.55	0.438	0.409	-5.07

Future work

- hybrid recommender approaches
 - metadata (geo, facebook likes, referers)
- long term user preferences
 - long-term user model
- comparison with additional algorithms (bounce visits...)
- more evaluation metrics (diversity, novelty)
- more filtering approaches

Acknowledgment

This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG1/0508/09.

Institute of Informatics and Software Engineering

Faculty of Informatics and Information Technologies, Slovak University of Technology in Bratislava, Slovakia