TagRec: Towards a Standardized Tag Recommender Benchmarking Framework

Dominik Kowald, Emanuel Lacic and Christoph Trattner
dkowald@know-center.at, elacic@know-center.at, ctrattner@know-center.at

Abstract

The purpose of *TagRec* is to provide researchers with a framework that **supports all steps of the development process of a new tag recommendation algorithm in a reproducible way**, including methods for

- data pre-processing
- data modeling
- data analysis
- recommender evaluation against state-of-the-art baseline approaches

Algorithms

Algorithm	Name
MP	Most popular tags
MP_u	Most popular tags by user
MP_r	Most popular tags by resource
$MP_{u,r}$	Mixture of MP_u and MP_r
CF_u	User-based Collaborative Filtering
CF_r	Resource-based Collaborative Filtering
$CF_{u,r}$	Mixture of CF_u and CF_r
APR	Adapted PageRank
FR	FolkRank
FM	Factorization Machines
PITF	Pairwise Interaction Tensor Factorization
LDA	Latent Dirichlet Allocation
LDA&LM	Mixture of LDA and $MP_{u,r}$
GIRP	Temporal Tag Usage Patterns
GIRPTM	Mixture of GIRP and MP_r
BLL	**Base Level Learning Equation**
BLL+C	**Mixture of BLL and MP_r**
3L	**3Layers**
$3LT_{topic}$	**Time-dependent 3L on topic level**
$3LT_{tag}$	**Time-dependent 3L on tag level**

Includes two novel approaches based on a model of **human memory called ACT-R** (BLL and BLL+C) and a model of **human categorization called MINERVA2** (3L, $3LT_{topic}$ and $3LT_{tag}$).

Future Work

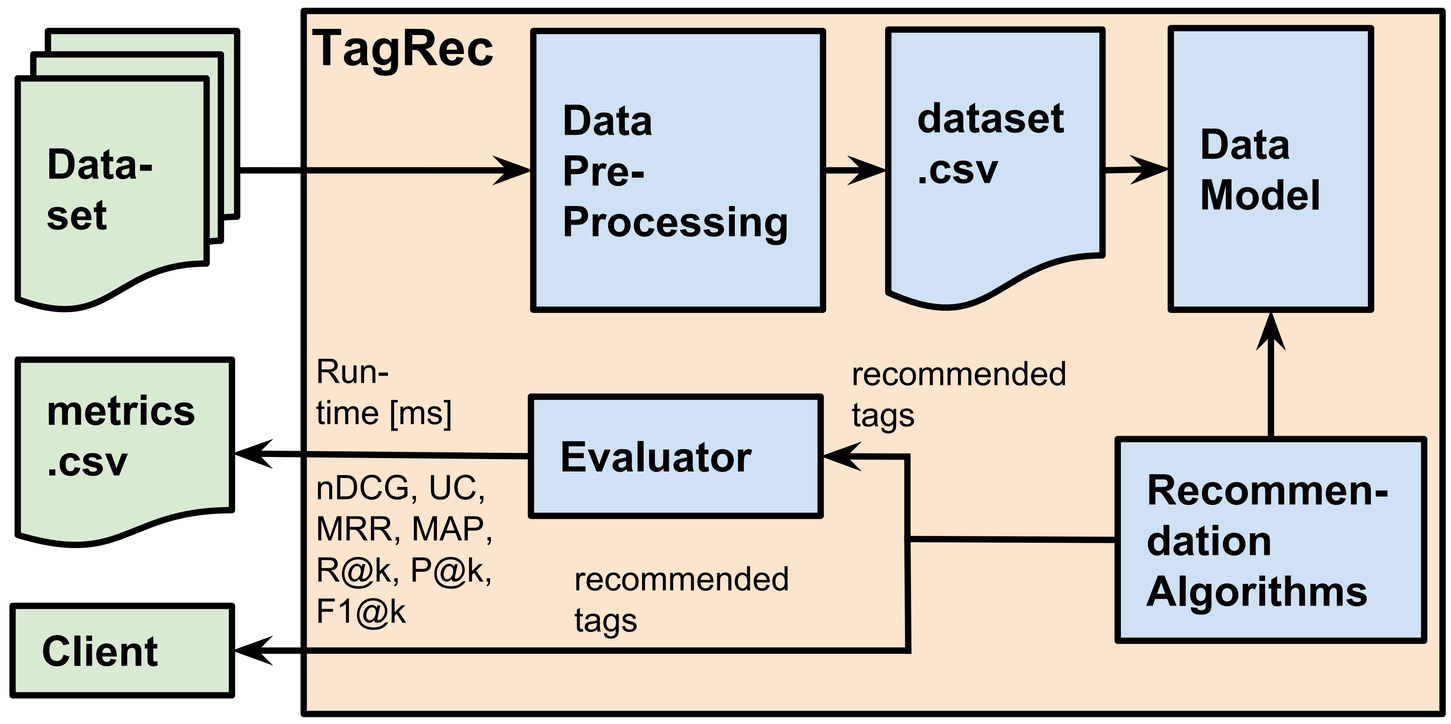
Plans for the future:

- Content-based tag recommendations
- Recommendation of resources and users as well
- Development of more algorithms based on human cognition

Reference

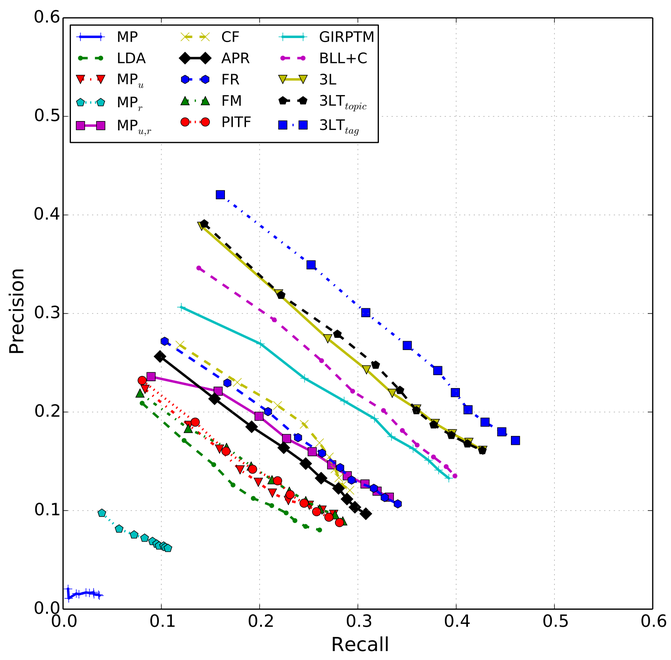
[1] D. Kowald, E. Lacic, and C. Trattner. Tagrec: Towards a standardized tag recommender benchmarking framework. In *Proceedings of the 25th ACM Conference on Hypertext and Social Media*, HT'14, New York, NY, USA, 2014. ACM.

System Overview

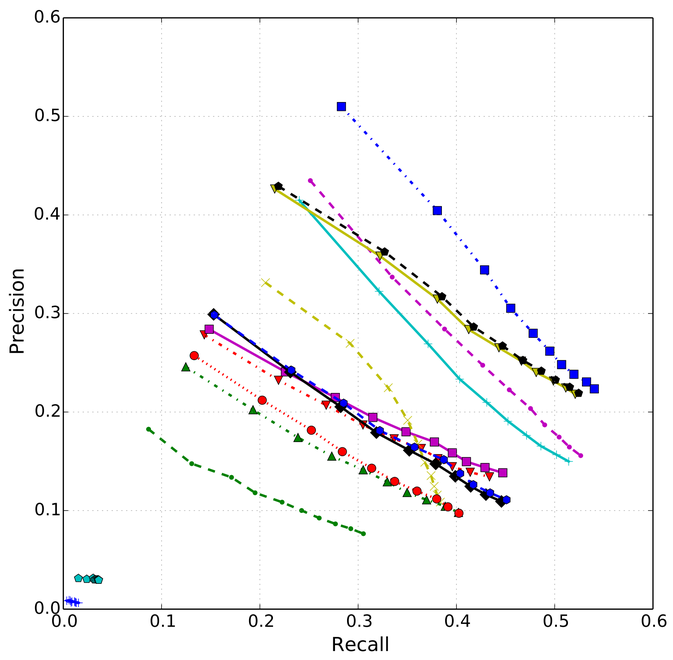


- **Data pre-processing**: (1) processing of social tagging datasets, (2) *p*-core pruning, (3) training/test set splitting and (4) creating LDA topics
- **Data model**: is fully object-oriented (Java) and is created from simple *.csv* files that contain the bookmarks of a folksonomy
- **Recommendation algorithms**: contains state-of-the-art tag recommender approaches to benchmark new methods against or to directly send tag suggestions to a client
- **Evaluator**: Recall, Precision, F1-score, Mean Reciprocal Rank, Mean Average Precision, Normalized Discounted Cumulative Gain, User Coverage and runtime

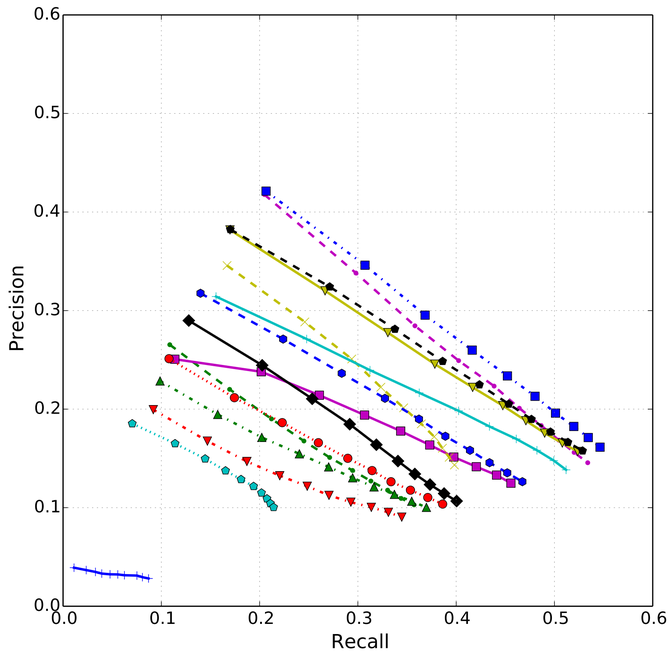
Results



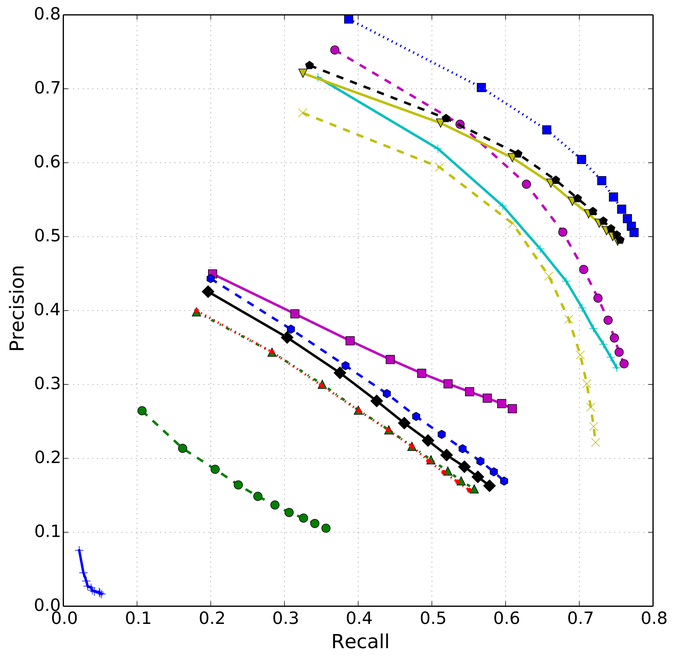
(a) BibSonomy



(b) CiteULike



(c) Delicious



(d) Flickr

The experiments conducted on four real-world folksonomy datasets show that the time-dependent approaches based on **models of human memory and human categorization** perform best in terms of recommender accuracy.