Learning Layers

Scaling up Technologies for Informal Learning in SME Clusters

Long Time No See

The Probability of Reusing Tags as a Function of Frequency and Recency

Dominik Kowald, Paul Seitlinger, Christoph Trattner, Tobias Ley

Many Thanks To

Paul Seitlinger
paul.seitlinger@tugraz.at
Graz University of Technology
Austria

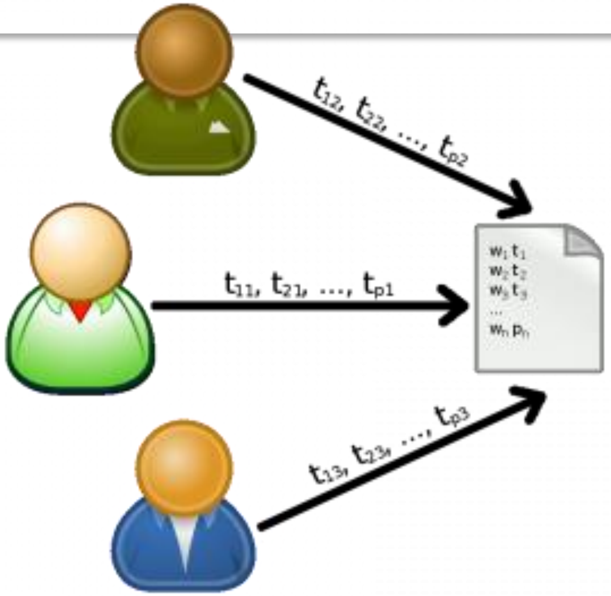
Christoph Trattner
ctrattner@know-center.at
Know-Center Graz
Austria

Tobias Ley
tley@tlu.ee
Tallinn University
Estonia

What will this talk be about?

- Social tags
- Prediction/recommendation of social tags
- Using an equation derived from human memory theory to implement a novel tag recommender

Why are we doing this?



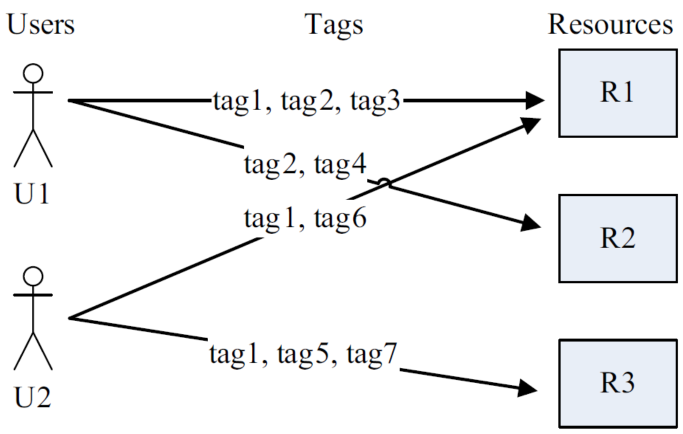
- Social tagging is the process of collaboratively annotating content
- Essential instrument of Web 2.0
- Helps users to
  - classify and structure Web content [Zubiaga et al., 2012]
  - navigate large knowledge repositories [Helic et al., 2012]
  - search and find information [Trattner et al., 2012]

Problem:

People are typically lazy in applying social tags (!)

Solution: Tag Recommenders

- Tag recommendation algorithms support the users in applying appropriate tags for resources and can be based on:
  - Tag Frequencies (MP)
  - Collaborative Filtering (CF)
  - Graph Structures (APR, FR)
  - Factorization Models (FM, PITF)
  - Hybrid approaches



[Huang et al., 2014]

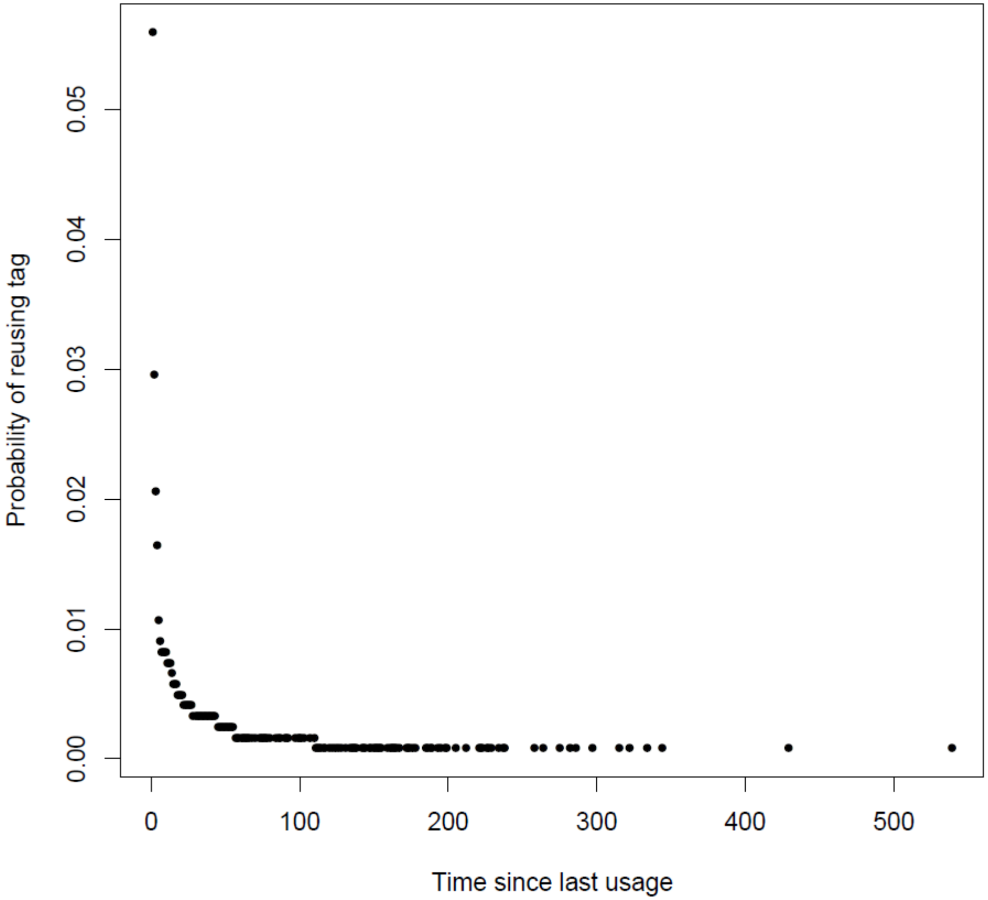
- *Issues*
  - Usually users change their tagging behavior **over time**
  - **BUT** all of these approaches **ignore the time component**

What‘s about the time component?

- Only a few time-based approaches available
- The Temporal Tag Usage Pattern approach (GIRPTM) of Zhang et al. (2012) shows that the **time component is important for tag recommenders**
  - Models the time component using an exponential function
- Empirical research on human memory (Anderson & Schooler, 1991) showed that the reuse-probability of a word depends on its **usage-frequency and recency** in the past
  - Models the time component using a power function

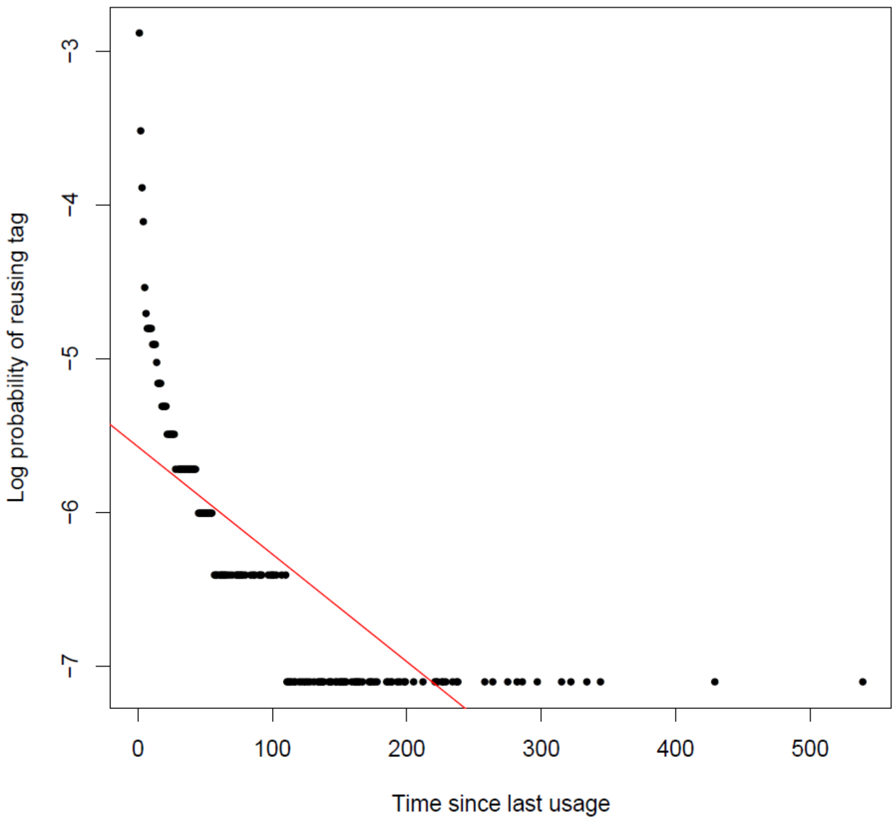
Which function fits better to model the drift of interests in social tagging systems?

Empirical Analysis: BibSonomy (1)



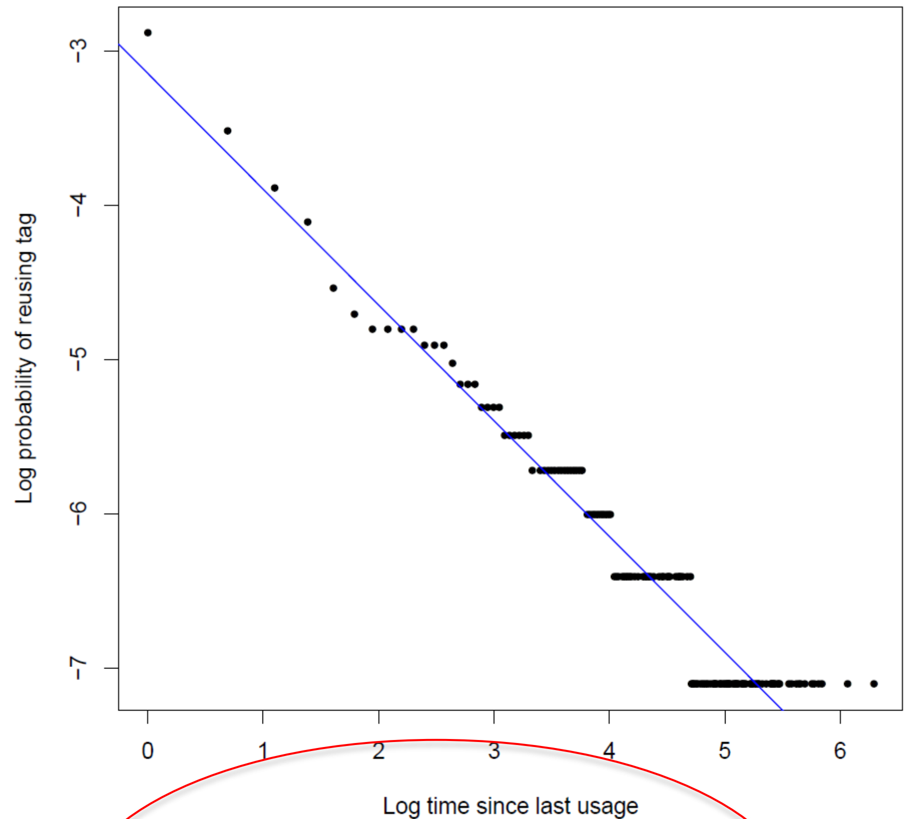
- Linear distribution with log-scale on Y-axis → **exponential function**
- Linear distribution with log-scale on X- and Y-axes → **power function**

Empirical Analysis: BibSonomy (2)



Exponential distribution

$R^2 = 35\%$



Power distribution

$R^2 = 65\%$

Our Approach

- **Base-Level learning (BLL) equation** - part of ACT-R model [Anderson et al., 2004]:

$$BLA(t,u) = ln(\sum_{i=1}^{n}(timestamp_{ref} - timestamp_i)^{-d})$$

- Also the context (resource) is important
  - Modeled with the most frequent tags of the resource (MP_r)

$$\widetilde{T}(u,r) = \underset{t \in T}{\arg\max}\ \underbrace{(\beta\ \underbrace{\|BLA(t,u)\|}_{BLL} + (1-\beta)\|\|Y_{t,r}\|\|)}_{BLL+C}$$

- **Linear runtime:** $O(|Y_{t,u}| + |Y_{t,r}|)$
- **Code:** https://github.com/learning-layers/TagRec/

How does it perform?

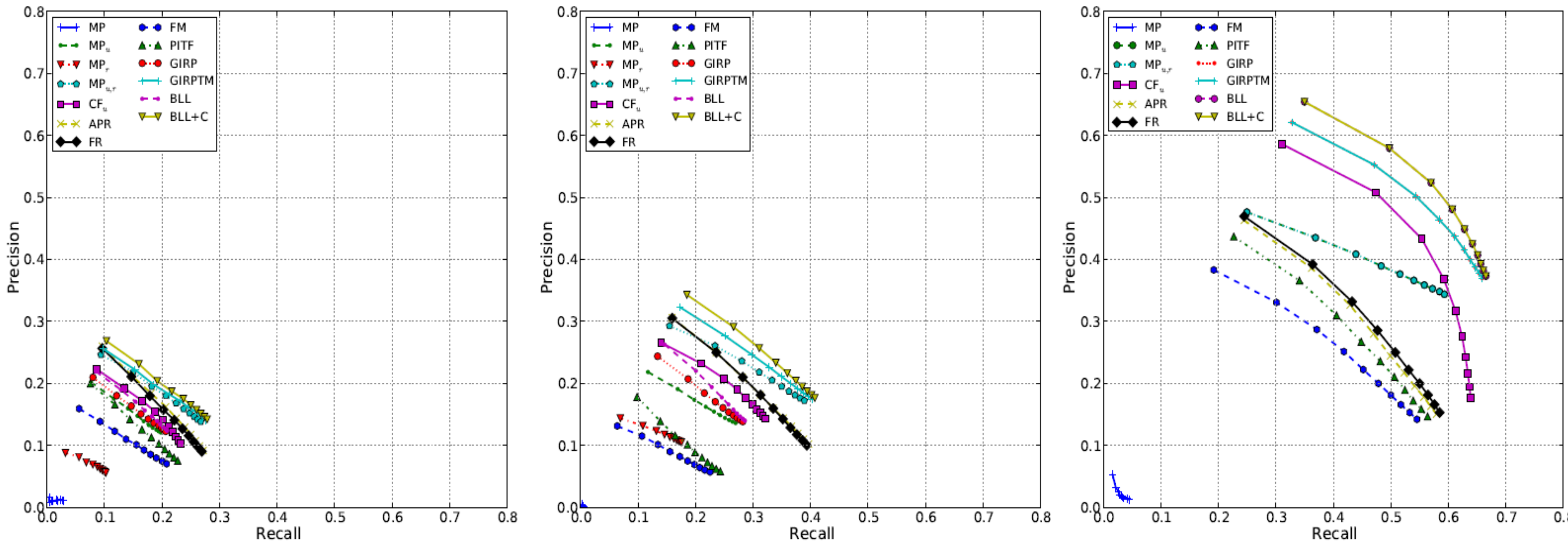
- 3 freely-available folksonomy datasets
  - **BibSonomy** (1.5 Million tag assignments)
  - **CiteULike** (16.7 million tag assignments)
  - **Flickr** (3.5 million tag assignments)

- **Original datasets** and *p*-core pruned datasets (core 3)

- **Leave-one-out evaluation** (for each user latest bookmark/post in test-set, rest in training-set)

- **IR metrics:** Precision, Recall, F1-score, MRR, MAP

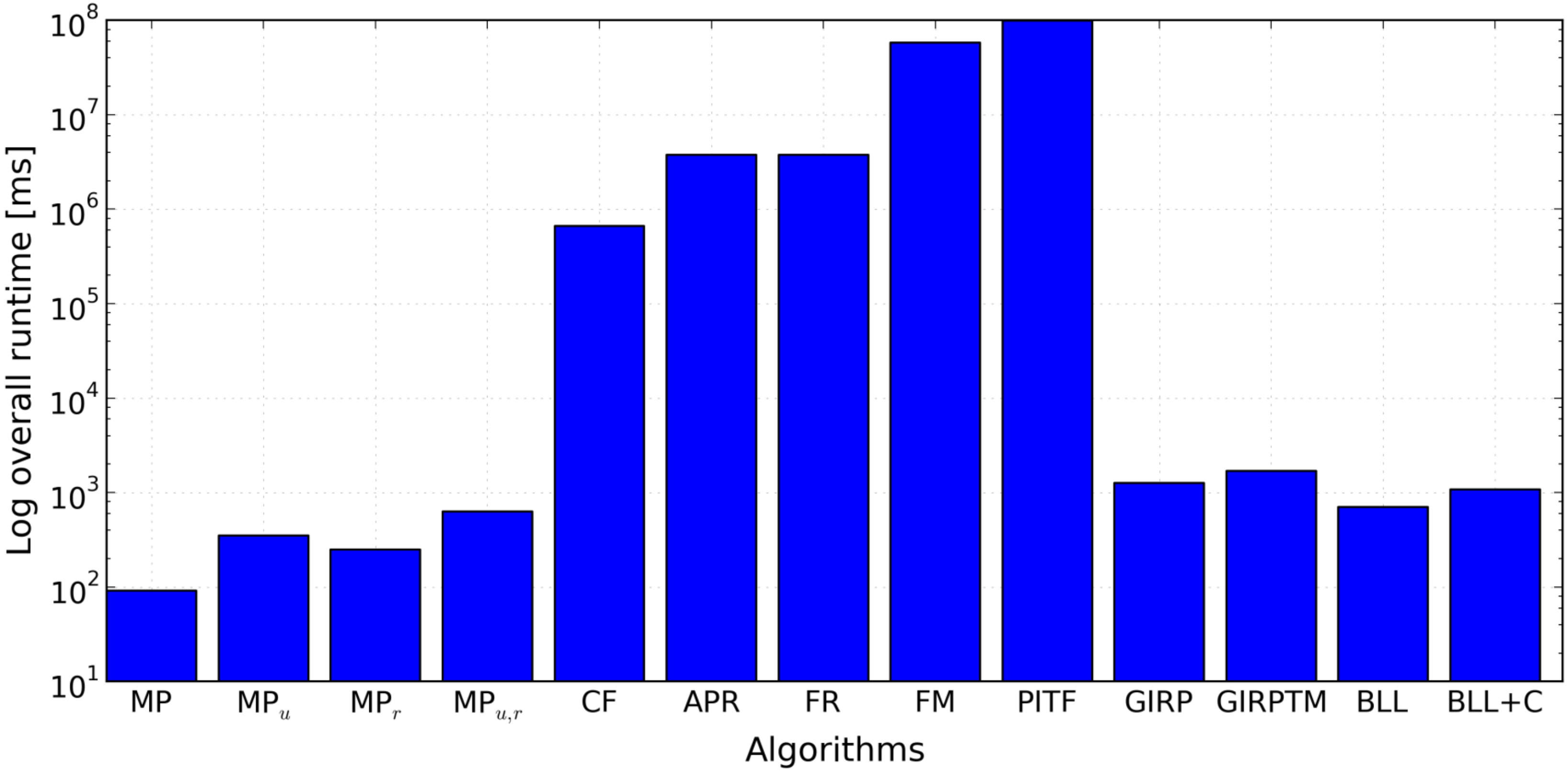
Results: Precision-Recall plots



(a) BibSonomy (no core) (b) CiteULike (no core) (c) Flickr (no core)

- The **time-depended** approaches outperform the state-of-the-art
- **BLL+C** reaches the highest level of accuracy

Runtime: BibSonomy



- **BLL+C** needs only around 1 second to provide accurate tag-recommendations for 5,500 user-resource pairs in the test set

What we have shown

1) The **time component** is an important factor for tag-recommendations

2) The BLL-equation can be used to implement an effective tag recommender
   - Models the time component with a **power function** rather than an exponential function
   - Outperforms current state-of-the art algorithms despite its **simplicity**
   - Computationally efficient: **linear runtime**

3) Effective principles of recommenders in social tagging can be implemented if human memory processes are taken into account

What are we currently doing?

- In previous work we presented a tag recommender based on human categorization (3Layers) [Seitlinger et al., 2013]
  - Combine this recommender with BLL to model the time component on a lexical and semantic layer

- Better modelling of the (resource) context (MP_r)
  - Spreading activation
  - Content-based approaches
- Adapt BLL+C also for the recommendation of resources
- Conduct online evaluation (BibSonomy)

Thank you for your attention!

Code and framework:

https://github.com/learning-layers/TagRec/

Questions?

Dominik Kowald

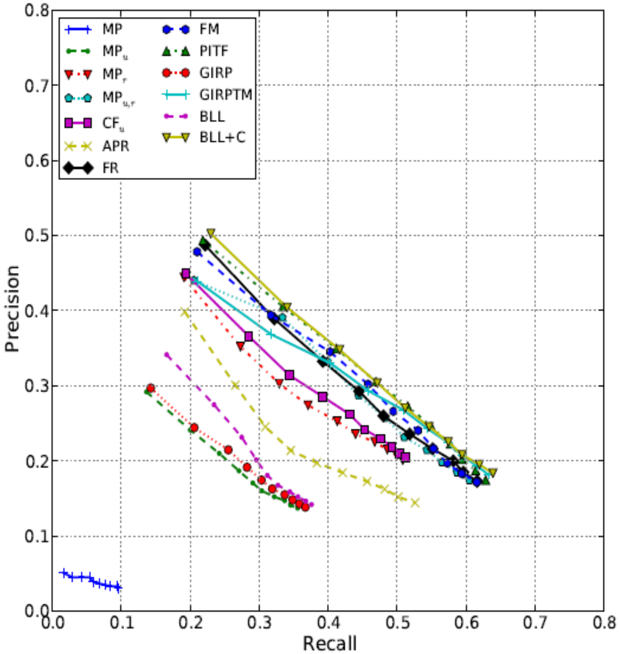
dkowald@know-center.at

Know-Center

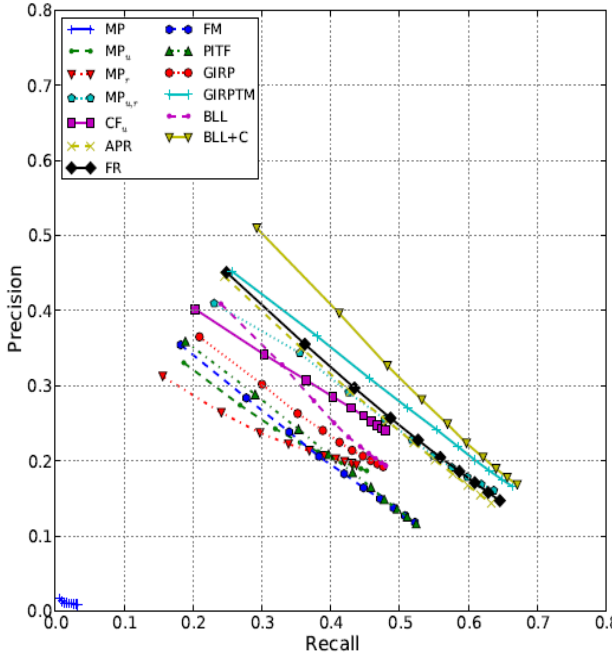
Graz University of Technology (Austria)

Backup

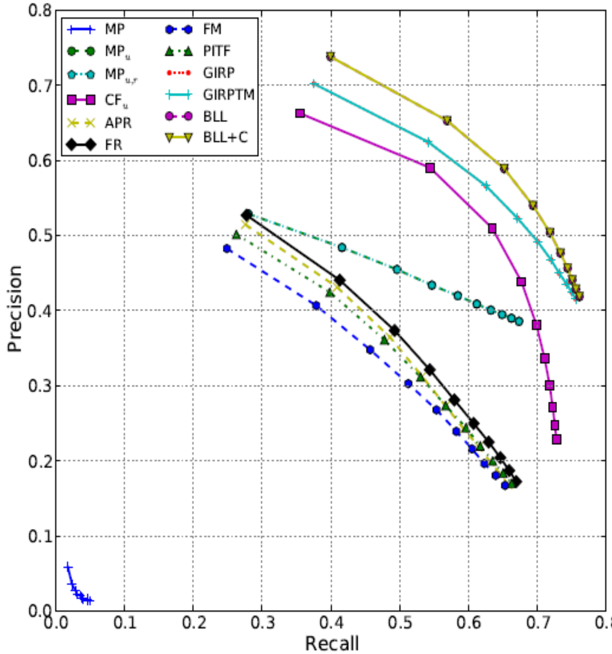
Results: Core 3



(d) BibSonomy (core 3)



(e) CiteULike (core 3)



(f) Flickr (core 3)

Results: F1@5, MRR, MAP

Dataset	Core	Measure	MP	MP_r	$MP_{u,r}$	CF	APR	FR	FM	PITF	GIRPTM	BLL+C
BibSonomy	-	$F_1@5$	.013	.074	.192	.166	.175	.171	.122	.139	.197	**.201**
		MRR	.008	.054	.148	.133	.149	.148	.097	.120	.152	**.158**
		MAP	.009	.070	.194	.173	.193	.194	.120	.150	.200	**.207**
	3	$F_1@5$	.047	.313	.335	.325	.260	.337	.345	.356	.350	**.353**
		MRR	.035	.283	.327	.289	.279	.333	.329	.341	.334	**.349**
		MAP	.038	.345	.403	.356	.329	.414	.408	.421	.416	**.435**
CiteULike	-	$F_1@5$	.002	.131	.253	.218	.195	.194	.111	.122	.263	**.270**
		MRR	.001	.104	.229	.201	.233	.233	.110	.141	.246	**.258**
		MAP	.001	.134	.280	.247	.284	.284	.125	.158	.301	**.315**
	3	$F_1@5$	.013	.270	.316	.332	.313	.318	.254	.258	.336	**.346**
		MRR	.012	.243	.353	.295	.361	.366	.282	.290	.380	**.409**
		MAP	.012	.294	.420	.363	.429	.436	.326	.334	.455	**.489**
Flickr	-	$F_1@5$	.023	-	.435	.417	.328	.334	.297	.316	.509	**.523**
		MRR	.023	-	.360	.436	.352	.355	.300	.333	.445	**.466**
		MAP	.023	-	.468	.581	.453	.459	.384	.426	.590	**.619**
	3	$F_1@5$	.026	-	.488	.493	.368	.378	.361	.369	.577	**.592**
		MRR	.026	-	.407	.498	.398	.404	.375	.390	.511	**.533**
		MAP	.026	-	.527	.663	.513	.523	.481	.502	.676	**.707**

Runtime Complexities

Algorithm	Complexity	Authors
MP	$\mathcal{O}(\|Y_t\|)$	Jäschke et al. [21]
MP_u	$\mathcal{O}(\|Y_{t,u}\|)$	Jäschke et al. [21]
GIRP	$\mathcal{O}(\|Y_{t,u}\|)$	Zhang et al. [54]
BLL	$\mathcal{O}(\|Y_{t,u}\|)$	Kowald et al. [23]
MP_r	$\mathcal{O}(\|Y_{t,r}\|)$	Jäschke et al. [21]
$\text{MP}_{u,r}$	$\mathcal{O}(\|Y_{t,u}\| + \|Y_{t,r}\|)$	Jäschke et al. [21]
GIRPTM	$\mathcal{O}(\|Y_{t,u}\| + \|Y_{t,r}\|)$	Zhang et al. [54]
BLL+C	$\mathcal{O}(\|Y_{t,u}\| + \|Y_{t,r}\|)$	Kowald et al. [23]
CF	$\mathcal{O}(\|V_r\|\|Y_{t,u}\|)$	Marinho et al. [32]
APR	$\mathcal{O}(l \cdot (\|Y_t\| + s))$	Hotho et al. [19]
FR	$\mathcal{O}(l \cdot (\|Y_t\| + s))$	Hotho et al. [19]
FM	$\mathcal{O}(l \cdot \|B_s\| \cdot (k_T \cdot \|T\|^2 + k_U \cdot k_R \cdot k_T))$	Rendle et al. [43]
PITF	$\mathcal{O}(l \cdot \|B_s\| \cdot (k_T \cdot \|T\|^2 + k_U \cdot k_R \cdot k_T))$	Rendle et al. [43]

Runtimes for BibSonomy

Core	Type	MP	MP_u	MP_r	$MP_{u,r}$	CF	APR/FR	FM	PITF	GIRP	GIRPTM	BLL	BLL+C
-	Train $[\|t\|]$	.091	.177	.217	.520	.164	.919	58,182	99,753	.955	1409	.502	**.943**
	Test $[t]$	.000	.001	.001	.001	.120	.683	.000	.000	.001	.001	.001	**.001**
	All $[\|t\|]$	.091	.349	.250	.631	662.724	3,751	58,182	99,753	1.270	1.685	.705	**1.082**
3	Train $[\|t\|]$	.028	.052	.059	.059	.037	.165	4,318	7,437	.127	.170	.095	**.122**
	Test $[t]$	.000	.001	.001	.001	.002	.062	.000	.000	.001	.001	.001	**.001**
	All $[\|t\|]$	.028	.111	.080	.119	2.006	49.354	4,318	7,437	.183	.232	.151	**.168**