TagRec: Towards a Toolkit for Reproducible Evaluation and Development of Tag-Based Recommender Algorithms

Christoph Trattner
Norwegian University of Science and Technology, Trondheim
and
Dominik Kowald AND Emanuel Lacic
Know-Center & Graz University of Technology, Austria

This article presents *TagRec*, a framework to foster reproducible evaluation and development of recommender algorithms based on folksonomy data. The purpose of *TagRec* is to provide the research community with a standardised framework that supports all steps of the development process and the evaluation of tag-based recommendation algorithms in a reproducible way, including methods for data pre-processing, data modeling and recommender evaluation. *TagRec* currently contains 32 state-of-the-art algorithms for tag and item prediction, including a set of novel and very efficient algorithms based on the human cognition theories ACT-R and MINERVA2. The framework should be relevant for researchers, teachers, students and developers working on recommender systems and predictive modeling in general and those interested in tag-based recommender algorithms in particular.

1. INTRODUCTION

In recent years social tagging has become an important instrument of Web 2.0, allowing users to collaboratively annotate and search content [Trattner et al. 2012; Helic et al. 2012; Helic et al. 2011]. In order to facilitate this process, current research has attempted to improve the performance and quality of tag recommendations. Furthermore, a number of algorithms has been created to recommend items to users based on folksonomy data. However, although various tag-based recommender approaches and studies exist showing the performance of the developed methods, most of them use different data pre-processing methods or evaluation protocols, making it difficult for researchers to understand the real value of the methods developed. To tackle this issue, we invented *TagRec* [Kowald et al. 2014], a framework providing researchers with a toolkit for developing and compareing algorithms in a standardized way. The purpose of *TagRec* is not only (a) to increase the transparency in tag-based recommender research (see also [Said and Bellogín 2014]) but also (b) to decrease the workload associated with developing novel recommender algorithms by providing researchers with an easy to use and easy to extend framework.

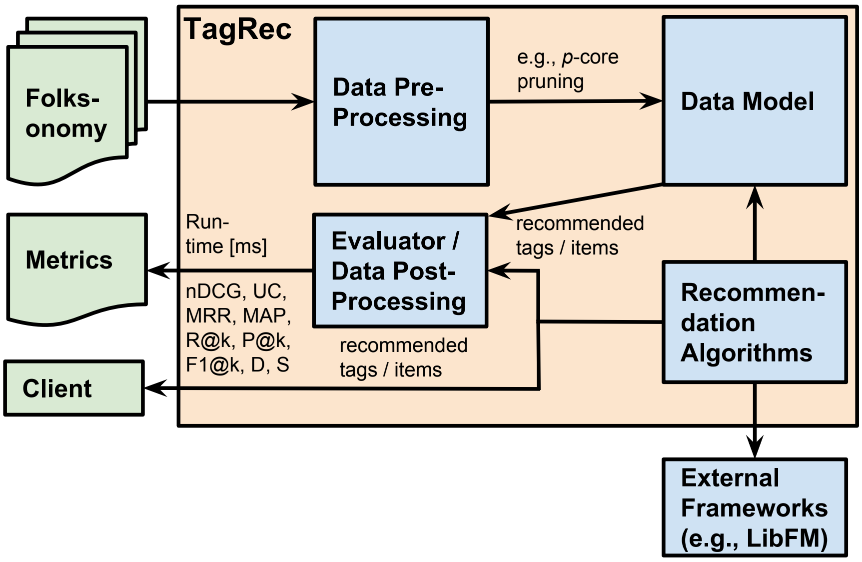


Fig. 1. Current system architecture of *TagRec*.

2. SYSTEM OVERVIEW

Fully implemented in the Java programming language, *TagRec* is open-source software that can be downloaded free of charge from our Github[1] repository. Figure 1 shows the system architecture of *TagRec*, which consists of four main components briefly described below:

Data Pre-Processing. *TagRec* offers various methods for data pre-processing: (1) parsing and processing of social tagging datasets, such as CiteULike, BibSonomy, Delicious, LastFm, MovieLens and Flickr, into the system's data format; (2) *p*-core pruning; (3) training/test set splitting (e.g., leave-one-out, time-based or 80/20 splits) [Jäschke et al. 2008] and (4) creating Latent Dirichlet Allocation [Krestel and Fankhauser 2010] topics for category-based algorithms, such as 3Layers [Kowald et al. 2014a; Seitlinger et al. 2013].

Data Model. The data model of *TagRec* is generated from a folksonomy that represents the bookmarks (i.e., the combination of user-id, resource-id, timestamp and assigned tags) in a dataset. Furthermore, the data model is fully object-oriented and provides distinct classes and powerful methods for modeling and analyzing the relationship and interactions between users, resources and tags (e.g., the number of times a specific tag has been assigned to a target resource or the time since the last usage of a specific tag in the tag assignments of a target user).

Recommendation Algorithms. Along with the state-of-the-art approaches for tag-based recommendations (e.g., Collaborative Filtering or FolkRank) [Marinho et al. 2011], the engine contains a set of novel and recently published algorithms based on models derived

[1] https://github.com/learning-layers/TagRec/

Algorithm	Name	Authors
	Tag Recommender Algorithms	
MP	Most popular tags	[Jäschke et al. 2008]
MP_u	Most popular tags by user	[Jäschke et al. 2008]
MP_r	Most popular tags by resource	[Jäschke et al. 2008]
$MP_{u,r}$	Mixture of MP_u and MP_r	[Jäschke et al. 2008]
CF_u	User-based Collaborative Filtering	[Marinho and Schmidt-Thieme 2008]
CF_r	Resource-based Collaborative Filtering	[Marinho and Schmidt-Thieme 2008]
$CF_{u,r}$	Mixture of CF_u and CF_r	[Marinho and Schmidt-Thieme 2008]
APR	Adapted PageRank	[Jäschke et al. 2008]
FR	FolkRank	[Jäschke et al. 2008]
FM	Factorization Machines	[Rendle and Schmidt-Thieme 2010]
PITF	Pairwise Interaction Tensor Factorization	[Rendle and Schmidt-Thieme 2010]
LDA	Latent Dirichlet Allocation	[Krestel and Fankhauser 2010]
LDA+LM	Mixture of LDA and $MP_{u,r}$	[Krestel and Fankhauser 2010]
3L	3Layers	[Seitlinger et al. 2013]
3L+MP_r	Mixture of 3L and MP_r	[Kowald et al. 2014a]
$3LT_{topic}$	Time-dependent 3L on the level of topics	[Kowald et al. 2014a]
$3LT_{topic}$+MP_r	Mixture of $3LT_{topic}$ and MP_r	[Kowald et al. 2014a]
$3LT_{tag}$	Time-dependent 3L on the level of tags	[Kowald et al. 2014a]
$3LT_{tag}$+MP_r	Mixture of $3LT_{tag}$ and MP_r	[Kowald et al. 2014a]
GIRP	Temporal Tag Usage Patterns	[Zhang et al. 2012]
GIRPTM	Mixture of GIRP and MP_r	[Zhang et al. 2012]
BLL	Base Level Learning Equation	[Kowald et al. 2014b]
BLL+C	Mixture of BLL and MP_r	[Kowald et al. 2014b]
BLL_{AC}	BLL with Associative Component	[Trattner et al. 2014]
BLL_{AC}+C	Mixture of BLL_{AC} and MP_r	[Trattner et al. 2014]
	Item Recommender Algorithms	
MP	Most popular items	[Schafer et al. 2007]
$CF_{u,t}$	User-based CF based on tags	[Schafer et al. 2007]
$CF_{u,b}$	User-based CF based on bookmarks	[Zheng and Li 2011]
$CF_{r,t}$	Item-based CF	[Sarwar et al. 2001]
Z	User-based CF based on bookmarks and time	[Zheng and Li 2011]
H	User-based CF based on tags and time	[Huang et al. 2014]
CIRTT	User-based CF based on bookmarks, tags and time	[Lacic et al. 2014]

Table I. Current tag-based recommender algorithms implemented in *TagRec*.

from human cognition to predict tags or items in folksonomies. All algorithms implement a common interface which makes it easy to develop and integrate new approaches. Moreover, we offer also connectors to external libraries such as LibFM to make use of frameworks in other programming languages such as C++. The predicted items and tags generated by the different algorithms can be forwarded either to the evaluation engine or directly to a client application.

Evaluator / Data Post-Processing. This component evaluates the algorithms based on training/test set splits of a dataset with respect to standard Information Retrieval (IR) metrics, such as Recall (R@k), Precision (P@), F1-score (F1@k), Mean Reciprocal Rank (MRR), Mean Average Precision (MAP), Normalized Discounted Cumulative Gain (nDCG) and User Coverage (UC) [Jäschke et al. 2009]. As for the item recommender algorithms, we also integrated other evaluation metrics such as Diversity (D) and Serendipity (S). Moreover, the evaluation engine offers data post-processing functionality that can, for example, limit the evaluation to users with a given minimum or maximum number of bookmarks or to users with certain tagging behavior (e.g, categorizer vs. describer [Körner et al.

Dataset	$\lvert B\rvert$	$\lvert U\rvert$	$\lvert R\rvert$	$\lvert T\rvert$	$\lvert TAS\rvert$
BibSonomy	400,983	5,488	346,444	103,503	1,479,970
CiteULike	379,068	8,322	352,343	138,091	1,751,347
Delicious	1,416,151	15,980	931,993	180,084	4,107,107
Flickr	864,679	9,590	864,679	127,599	3,552,540

Table II. Properties of the datasets used in our first evaluation experiment (i.e., recommending tags), where $|B|$ is the number of bookmarks, $|U|$ the number of users, $|R|$ the number of resources, $|T|$ the number of tags and $|TAS|$ the number of tag assignments.

2010]).

3. RESULTS

In the following section we report the results of two evaluation experiments to (i) demonstrate the functionalities implemented in *TagRec* and (ii) to reveal the differences of the algorithms in terms of the predictive power and runtime. Both experiments were conducted in memory on a IBM System x3550 M4 Server, with one Intel(R) Xeon(R) CPU E5-2640 v2 @ 2.00GHz and 256GB RAM.

3.1 Evaluation Experiment 1: Recommending Tags

In the first experiment we evaluated and compared several tag recommendation approaches implemented within *TagRec* using four real-world folksonomy datasets gathered from the online bookmarking and sharing platforms BibSonomy[2] (2013-07-01), CiteULike [3] (2013-03-10), Delicious and Flickr [4] (2010-01-07). The properties of the datasets are highlighted in Table II. For the benchmarking experiment we split the datasets into one training and one test set using the time-based leave-one-out pre-processing method implemented in *TagRec* (i.e., the latest bookmark for each user was used for testing and the rest for training). To determine the predictive power of the approaches, we utilized a set of well-known Information Retrieval metrics available in *TagRec* (R@k, P@k, F1@k, MRR, MAP, nDCG). Parameters of the algorithms were set to the default values in *TagRec*, i.e., for CF the neighborhood size was set to $k = 20$, for MP$_u, r$ and BLL+C β was set to .5, the number of factors for PITF and FM were set to 256 based on 50 iterations, LDA as well as 3L and 3LT were calculated based on 1000 latent topics, and for FR and APR the parameter d was set to .7 and the number of iterations was set to 10.

Table III and Figure 2 presents the results of this experiment. As highlighted, 3LT$_{tag}$+MP$_r$ reaches the highest level of estimates in all four datasets followed by the other combinations of 3LT and MP$_r$ and BLL+C. It is interesting to observe, that all these approaches are simple algorithms based on human cognition theories and that they all outperform more complex models such as APR, FR or popular tensor factorization approaches such as FM or PITFM. However, it has to be noted that those approaches also incorporate the vari-

[2]http://www.kde.cs.uni-kassel.de/bibsonomy/dumps
[3]http://www.citeulike.org/faq/data.adp
[4]https://www.uni-koblenz.de/FB4/Institutes/IFI/AGStaab/Research/DataSets/PINTSExperimentsDataSets/

	Measure	MP	LDA	MP_r	$MP_{u,r}$	CF	APR	FR	FM	PITF	GIRPTM	BLL+C	3L+MP_r	$3LT_{topic}$+MP_r	$3LT_{tag}$+MP_r
Bib.	F_1@5	.013	.097	.074	.192	.166	.175	.171	.122	.138	.197	.201	.206	.207	**.211**
	MRR@10	.008	.084	.054	.148	.133	.149	.148	.099	.119	.152	.158	.157	.158	**.162**
	MAP@10	.009	.101	.070	.194	.173	.193	.194	.122	.150	.200	.207	.207	.208	**.214**
	nDCG@10	.019	.142	.089	.240	.214	.244	.242	.172	.198	.248	.254	.254	.256	**.261**
Cite.	F_1@5	.007	.068	.033	.199	.157	.162	.161	.112	.130	.207	.215	.232	.233	**.238**
	MRR@10	.005	.066	.024	.179	.175	.181	.181	.116	.148	.196	.205	.199	.200	**.212**
	MAP@10	.005	.074	.029	.210	.203	.212	.212	.132	.168	.229	.241	.235	.236	**.250**
	nDCG@10	.011	.107	.035	.252	.234	.257	.255	.177	.210	.270	.280	.275	.277	**.289**
Del.	F_1@5	.033	.168	.140	.236	.228	.211	.229	.157	.185	.253	.270	.268	.271	**.278**
	MRR@10	.025	.157	.113	.215	.214	.206	.221	.141	.178	.236	.262	.243	.246	**.263**
	MAP@10	.026	.187	.146	.257	.262	.246	.270	.168	.211	.286	.320	.301	.304	**.323**
	nDCG@10	.049	.249	.179	.329	.316	.315	.341	.233	.278	.358	.388	.370	.374	**.390**
Flickr	F_1@5	.023	.160	-	.435	.417	.328	.334	.298	.318	.509	.523	.568	.571	**.585**
	MRR@10	.023	.165	-	.360	.436	.352	.355	.298	.337	.445	.466	.450	.454	**.477**
	MAP@10	.023	.193	-	.468	.581	.453	.459	.381	.431	.590	.619	.599	.604	**.636**
	nDCG@10	.038	.252	-	.520	.612	.507	.513	.447	.491	.626	.648	.634	.639	**.663**

Table III. F_1@5, MRR@10, MAP@10 and nDCG@10 of the tag recommender approaches in BibSonomy (Bib.), CiteULike (Cite.), Delicious (Del.) and Flickr.

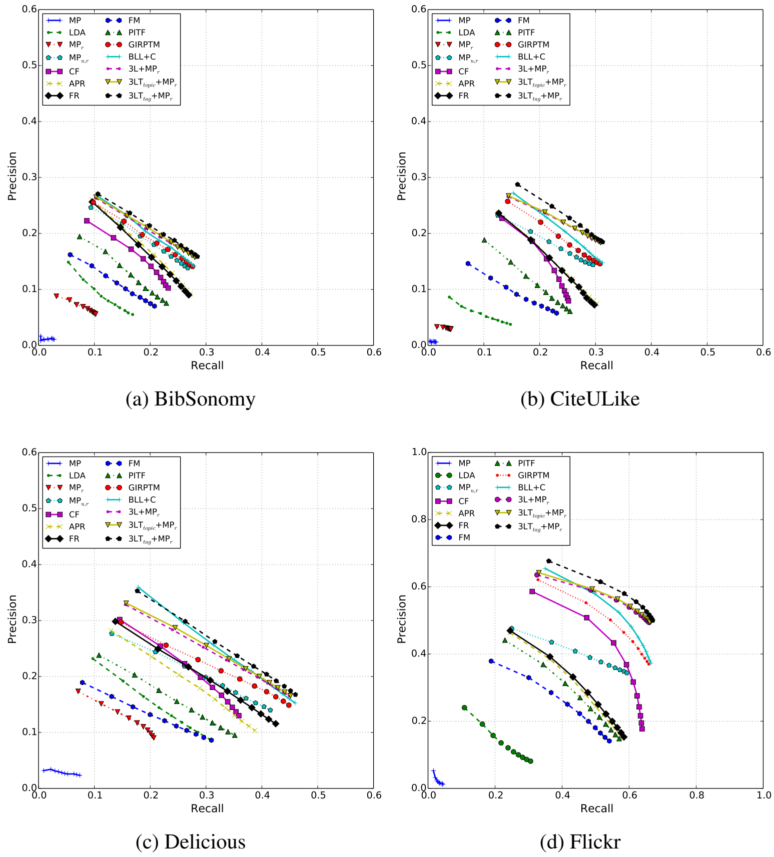
able of time in their models which is not the case for all the other models except GIRPTM [Zhang et al. 2012] which was known to be the best time-based tag recommender approach until the BLL+C [Kowald et al. 2014b] and $3LT_{tag}$+MP_r [Kowald et al. 2014a] algorithms were published.

Figure 5 shows the runtime comparison of the algorithms in the four datasets for the full time required to provide tag recommendations for all user-resource pairs in the test sets. As highlighted, the frequency-based methods such as MP, MP_u, MP_r, MP_u, r or GRPTM and BLL+C perform best in that respect. All other methods perform significantly worse. Summed up, the best algorithm providing both good runtime and high predictive power is BLL+C [Kowald et al. 2014b].

3.2 Evaluation Experiment 2: Recommending Items

In the second experiment we benchmarked a selection of item recommendation approaches implemented in *TagRec* against each other. As in the first evaluation experiment, we utilized folksonomy datasets from the social bookmarking sites BibSonomy and CiteULike (same datasets). Furthermore, we used a tagging dataset from the MovieLens platform[5] (2011-01-01). The properties of the datasets are outlined in Table IV. To determine the predictive power of the approaches we used the time-based 80/20 splitting method within *TagRec*, i.e., for each user we sorted her bookmarks in chronological order and used the 20% most recent bookmarks for testing and the rest for training. As benchmark evaluation metrics R@k, P@k, F1@k, MAP, nDCG, UC (= User Coverage) and D (= Diversity) were utilized. The results of this experiment are presented in Table V and Figure 4. As shown, the CIRTT [Lacic et al. 2014] algorithm, that re-ranks the recommended list with respect to similar items recently bookmarked by the users with important tags, reaches the best results in all settings. Moreover, CIRTT not only provides the highest levels of accuracy but also shows good performance in terms of the diversity measure. Another interesting finding from that experiment is the fact that CF_B performs significantly better than CF_T, although one might assume that semantic information present in the form of tags contains

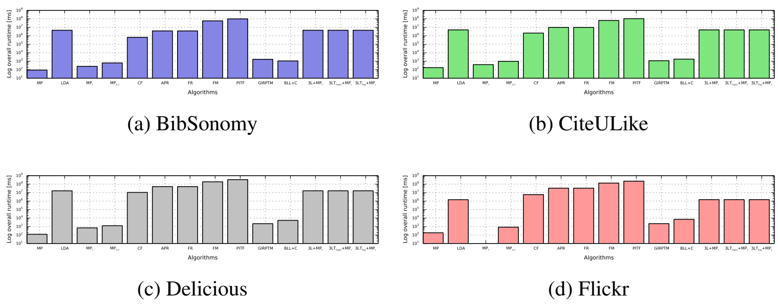
[5] `http://grouplens.org/datasets/movielens/`



(a) BibSonomy (b) CiteULike

(c) Delicious (d) Flickr

Fig. 2. Recall/Precision plots of the tag recommender approaches in BibSonomy, CiteULike, Delicious and Flickr.



(a) BibSonomy (b) CiteULike

(c) Delicious (d) Flickr

Fig. 3. Overall runtime of the tag recommender approaches in milliseconds [ms] for BibSonomy, CiteULike, Delicious and Flickr.

Dataset	$\|B\|$	$\|U\|$	$\|R\|$	$\|T\|$	$\|TAS\|$
BibSonomy	82,539	2,437	28,000	30,919	339,337
CiteULike	36,471	3,202	15,400	20,937	99,635
MovieLens	53,607	3,983	5,724	14,883	92,387

Table IV. Properties of the datasets used in our second evaluation experiment (i.e., recommending items), where $|B|$ is the number of bookmarks, $|U|$ the number of users, $|R|$ the number of resources, $|T|$ the number of tags and $|TAS|$ the number of tag assignments.

	Metric	MP	CF_T	CF_B	Z	H	$CIRTT$
BibSonomy	nDCG@20	.0143	.0448	.0610	.0621	.0564	**.0638**
	MAP@20	.0057	.0319	.0440	.0447	.0394	**.0464**
	R@20	.0204	.0618	.0820	.0834	.0816	**.0907**
	D	.8307	.8275	.8852	.8528	.6209	.8811
	UC	100%	99.76%	99.52%	99.52%	99.76%	99.76%
CiteULike	nDCG@20	.0062	.0407	.0717	.0762	.0706	**.0912**
	MAP@20	.0036	.0241	.0453	.0484	.0459	**.0629**
	R@20	.0077	.0630	.1033	.1077	.0928	**.1225**
	D	.8936	.7969	.8642	.8145	.6318	.8640
	UC	100%	98.38%	96.44%	97.32%	98.38%	97.61%
MovieLens	nDCG@20	.0198	.0361	.0602	.0614	.0484	**.0650**
	MAP@20	.0075	.0201	.0347	.0367	.0263	**.0413**
	R@20	.0366	.0561	.1031	.1013	.0763	**.1058**
	D	.9326	.8861	.9267	.9119	.7789	.9176
	UC	100%	97.82%	95.90%	98.43%	97.82%	95.90%

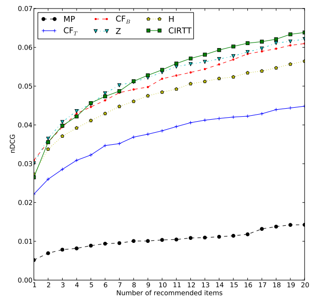
Table V. nDCG@20, MAP@20, R@20, D and UC values of the item recommender approaches for BibSonomy, CiteULike and MovieLens.

more complete and precise information for finding similarities between users than a simple binary representation of the data as it is available in the case of the CF_B approach.

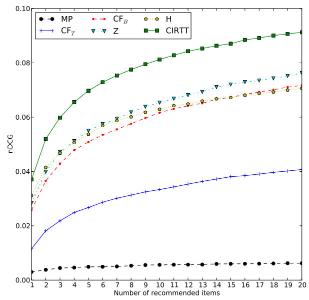
As in the first experiment, we also conducted an evaluation of the recommender runtime to determine which of the approaches is capable of not only generating good recommendations but also in a time efficient manner. The results of this experiment are presented in Figure 5. As highlighted the best results are achieved via the MostPopular (MP) approach, followed by CF_B and CF_T. CIRTT showing the highest estimates of accuracy places 4th just slightly behind the two simple Collaborative Filtering approaches and outperforming the more complex methods Z [Zheng and Li 2011] and H [Huang et al. 2014]. This suggests that CIRTT [Lacic et al. 2014] provides an excellent alternative to simple user-based Collaborative Filtering approaches showing good runtime qualities and high predictive power.

4. CONCLUSIONS & FUTURE WORK

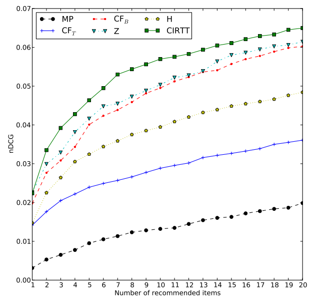
In this work we presented *TagRec*, a toolbox for transparent tag-based recommender benchmarking and construction. *TagRec* is fully implemented in Java and contains a rich set of state-of-the-art recommender algorithms for folksonomy data along with newly developed and published recommendation mechanisms based on models derived from the human cognition theories ACT-R [Anderson et al. 2004] and MINERVA2 [Hintzman 1984]. Al-



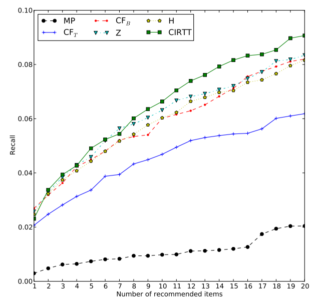
(a) nDCG BibSonomy



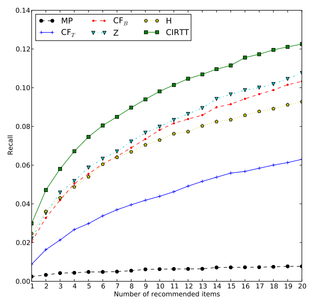
(b) nDCG CiteULike



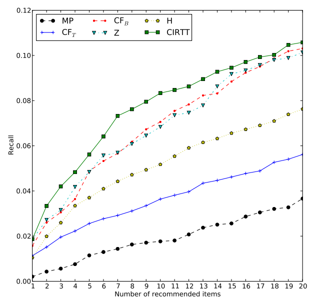
(c) nDCG MovieLens



(d) Recall BibSonomy

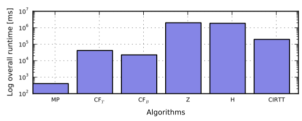


(e) Recall CiteULike

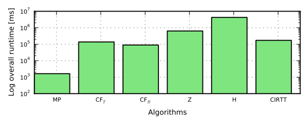


(f) Recall MovieLens

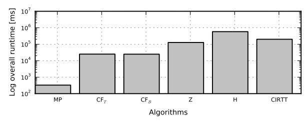
Fig. 4. nDCG, MAP and Recall plots of the item recommender approaches in BibSonomy, CiteULike and MovieLens.



(a) BibSonomy



(b) CiteULike



(c) MovieLens

Fig. 5. Overall runtime of the item recommender approaches in milliseconds [ms] in BibSonomy, CiteULike and MovieLens.

though, *TagRec* contains already a rich set of algorithms for item and tag predictions in folksonomies, there is still a lot of room for improvements. For instance, what is currently not available within our framework are content-based recommender methods that take also page-text information into account. Also recent work on matrix-factorization are not yet included. Another significant contribution would be to enhance the evaluation pipeline with more evaluation metrics and protocols. To that end, we encourage the research community to consider *TagRec*, test it out and, if possible, make a contribution to our research. It is our believe that good research practice should also include - among other things - sharing of code to make experiments reproducible!

ACKNOWLEDGMENTS

This work was carried out during the first authors tenure at Norwegian University of Science and Technology, Trondheim, sponsored by the ERCIM Alain Bensoussan Fellowship Program. The research is supported by the Know-Center and the EU funded project Learning Layers (7th Framework Program, Grant Agreement 318209). The Know-Center is funded within the Austrian COMET Program - Competence Centers for Excellent Technologies - under the auspices of the Austrian Ministry of Transport, Innovation and Technology, the Austrian Ministry of Economics and Labor and by the State of Styria. COMET is managed by the Austrian Research Promotion Agency (FFG).

REFERENCES

ANDERSON, J. R., BYRNE, M. D., DOUGLASS, S., LEBIERE, C., AND QIN, Y. 2004. An integrated theory of the mind. *Psychological Review 111,* 4, 1036–1050.

HELIC, D., KÖRNER, C., GRANITZER, M., STROHMAIER, M., AND TRATTNER, C. 2012. Navigational efficiency of broad vs. narrow folksonomies. In *Proc.* HT '12. ACM, New York, NY, USA, 63–72.

HELIC, D., TRATTNER, C., STROHMAIER, M., AND ANDREWS, K. 2011. Are tag clouds useful for navigation? a network-theoretic analysis. *International Journal of Social Computing and Cyber-Physical Systems 1,* 1, 33–55.

HINTZMAN, D. L. 1984. Minerva 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers 16,* 2, 96–101.

HUANG, C.-L., YEH, P.-H., LIN, C.-W., AND WU, D.-C. 2014. Utilizing user tag-based interests in recommender systems for social resource sharing websites. *Knowledge-Based Systems*.

JÄSCHKE, R., EISTERLEHNER, F., HOTHO, A., AND STUMME, G. 2009. Testing and evaluating tag recommenders in a live system. In *Proc.* RecSys'09. ACM, 369–372.

JÄSCHKE, R., MARINHO, L., HOTHO, A., SCHMIDT-THIEME, L., AND STUMME, G. 2008. Tag recommendations in social bookmarking systems. *Ai Communications 21,* 4, 231–247.

KÖRNER, C., KERN, R., GRAHSL, H.-P., AND STROHMAIER, M. 2010. Of categorizers and describers: an evaluation of quantitative measures for tagging motivation. In *Proceedings of the 21st ACM conference on Hypertext and hypermedia*. HT '10. ACM, New York, NY, USA, 157–166.

KOWALD, D., LACIC, E., AND TRATTNER, C. 2014. Tagrec: Towards a standardized tag recommender benchmarking framework. In *Proc.* HT '14. ACM, New York, NY, USA, 305–307.

KOWALD, D., SEITLINGER, P., TRATTNER, C., AND LEY, T. 2014a. Forgetting the words but remembering the meaning: Modeling forgetting in a verbal and semantic tag recommender. *arXiv preprint arXiv:1402.0728*.

KOWALD, D., SEITLINGER, P., TRATTNER, C., AND LEY, T. 2014b. Long time no see: The probability of reusing tags as a function of frequency and recency. In *Proc.* WWW '14. ACM.

KRESTEL, R. AND FANKHAUSER, P. 2010. Language models and topic models for personalizing tag recommendation. In *Proc.* WI'10. IEEE, 82–89.

LACIC, E., KOWALD, D., SEITLINGER, P., TRATTNER, C., AND PARRA, D. 2014. Recommending items in social tagging systems using tag and time information. *arXiv preprint arXiv:1406.7727*.

MARINHO, L., NANOPOULOS, A., SCHMIDT-THIEME, L., JÄSCHKE, R., HOTHO, A., STUMME, G., AND SYMEONIDIS, P. 2011. Social tagging recommender systems. In *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds. Springer US, 615–644.

MARINHO, L. B. AND SCHMIDT-THIEME, L. 2008. Collaborative tag recommendations. In *Data Analysis, Machine Learning and Applications*. Springer, 533–540.

RENDLE, S. AND SCHMIDT-THIEME, L. 2010. Pairwise interaction tensor factorization for personalized tag recommendation. In *Proc.* WSDM '10. ACM.

SAID, A. AND BELLOGÍN, A. 2014. Rival: A toolkit to foster reproducibility in recommender system evaluation. In *Proc.* RecSys '14. ACM, New York, NY, USA, 371–372.

SARWAR, B., KARYPIS, G., KONSTAN, J., AND RIEDL, J. 2001. Item-based collaborative filtering recommendation algorithms. In *Proc.* WWW '01. ACM, New York, NY, USA, 285–295.

SCHAFER, J. B., FRANKOWSKI, D., HERLOCKER, J., AND SEN, S. 2007. Collaborative filtering recommender systems. In *The adaptive web*. Springer, 291–324.

SEITLINGER, P., KOWALD, D., TRATTNER, C., AND LEY, T. 2013. Recommending tags with a model of human categorization. In *Proc.* CIKM '13. ACM, 2381–2386.

TRATTNER, C., KOWALD, D., SEITLINGER, P., KOPEINIK, S., AND LEY, T. 2014. Modeling activation processes in human memory to predict the reuse of tags. *The Journal of Web Science (under review)*.

TRATTNER, C., LIN, Y.-L., PARRA, D., YUE, Z., REAL, W., AND BRUSILOVSKY, P. 2012. Evaluating tag-based information access in image collections. In *Proc.* HT '12. ACM, New York, NY, USA, 113–122.

ZHANG, L., TANG, J., AND ZHANG, M. 2012. Integrating temporal usage pattern into personalized tag prediction. In *Web Technologies and Applications*. Springer, 354–365.

ZHENG, N. AND LI, Q. 2011. A recommender system based on tag and time information for social tagging systems. *Expert Syst. Appl.*.

Christoph Trattner is currently working as ERCIM Alain Bensoussan Fellow at the Norwegian University of Science and Technology. Furthermore, he is affiliated with the Know-Center, Austria's leading research competence for BigData analytics and data driven business where he leads the SocialComputing research lab. He holds a PhD (with hons), a MSc (with hons) and BSc in Computer Science and Telematics from Graz University of Technology (Austria). His research interests include Social and Semantic Computing, Web-Science, HCI, Recommender Systems, Networks, Data Mining and Machine Learning.

Dominik Kowald and Emanuel Lacic work as university assistants and software developers at Knowledge Technologies Institute at Graz University of Technology (Austria) and the SocialComputing research lab at the Know-Center, Austria's leading research competence for BigData analytics and data driven business. Both hold a MSc (with hons) and BSc in Computer Science and are interested in social tagging and recommender systems. They are currently pursuing a PhD in this context at Graz University of Technology under the supervision of Dr. Christoph Trattner.