

The Impact of Time on Hashtag Reuse in Twitter: A Cognitive-Inspired Hashtag Recommendation Approach

Presentation of work originally published in the Proc. of the 26th Intl. Conf. on WWW

Elisabeth Lex¹, Dominik Kowald²

Abstract: In our work [KPL17], we study temporal usage patterns of Twitter hashtags, and we use the Base-Level Learning (BLL) equation from the cognitive architecture ACT-R [An04] to model how a person reuses her own, individual hashtags as well as hashtags from her social network. The BLL equation accounts for the time-dependent decay of item exposure in human memory. According to BLL, the usefulness of a piece of information (e.g., a hashtag) is defined by how frequently and how recently it was used in the past, following a time-dependent decay that is best modeled with a power-law distribution. We used the BLL equation in our previous work to recommend tags in social bookmarking systems [KL16]. Here [KPL17], we adopt the BLL equation to model temporal reuse patterns of individual (i.e., reusing own hashtags) and social hashtags (i.e., reusing hashtags, which has been previously used by a followee) and to build a cognitive-inspired hashtag recommendation algorithm. We demonstrate the efficacy of our approach in two empirical social networks crawled from Twitter, i.e., *CompSci* and *Random* (for details about the datasets, see [KPL17]). Our results show that our approach can outperform current state-of-the-art hashtag recommendation approaches.

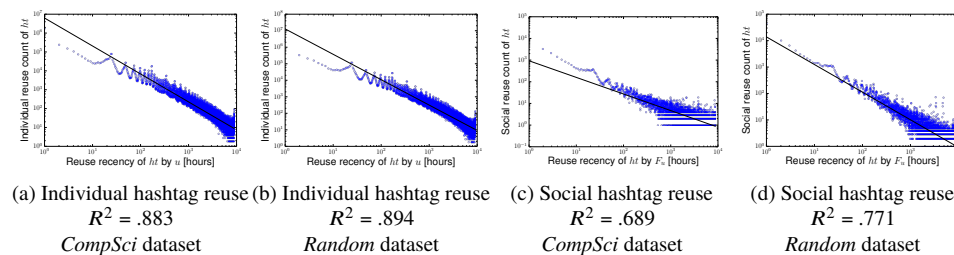


Abb. 1: The effect of time on individual and social hashtag reuse for the *CompSci* and *Random* datasets (plots are in log-log scale). We also report the R^2 estimates for the linear fits of the data.

Temporal Effects of Hashtag Reuse. To determine the plausibility of our approach, we study hashtag use in our two empirical Twitter datasets, i.e., *CompSci* and *Random*. For each hashtag assignment, we investigate whether the hashtag has either been used by the same user before (“individual”), by some of her followees (“social”), by both (“individual/social”), by anyone else in the dataset (“network”) or by neither of them (“external”). We find that depending on the dataset, individual or social hashtag reuse can explain approximately two-third of hashtag assignments. Also, both reuse types follow a time-dependent decay that follows a power-law distribution, as shown in Figure 1. This motivates our

¹ Graz University of Technology, ISDS, Inffeldgasse 13/V, 8010 Graz, Austria elisabeth.lex@tugraz.at

² Know-Center, Social Computing Group, Inffeldgasse 13/VI, 8010 Graz, Austria dkowald@know-center.at

idea to model both individual as well as social hashtag reuse and to recommend hashtags for new tweets with the BLL equation.

Experiments and Results. We implement the BLL equation in two variants, where the first one (i.e., $BLL_{I,S}$) predicts the hashtags of a user solely based on past hashtag usage, and the second one (i.e., $BLL_{I,S,C}$) combines $BLL_{I,S}$ with a content-based tweet analysis to also incorporate the text of the currently proposed tweet of a user. We evaluate our approach using standard evaluation protocols and metrics, and we find that our approach provides significantly higher prediction accuracy and ranking estimates than current state-of-the-art hashtag recommendation algorithms in both scenarios (for more details about the baselines, refer to [KPL17]), as shown in Figure 2.

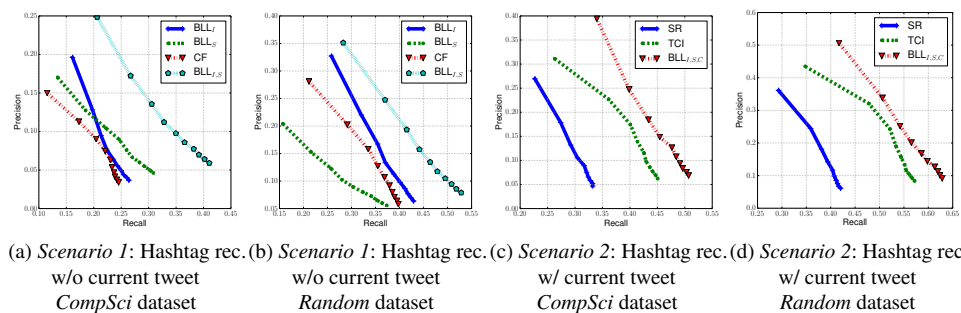


Abb. 2: Precision / Recall plots of our two evaluation scenarios showing the accuracy of BLL_I , BLL_S , and our baselines CF, $BLL_{I,S}$, SR, TCI and $BLL_{I,S,C}$ for $k = 1 - 10$ recommended hashtags.

Conclusion and Reproducibility. We find that temporal effects play an important role in hashtag reuse on Twitter. We propose our cognitive-inspired hashtag recommendation approaches $BLL_{I,S}$ and $BLL_{I,S,C}$ to account for such effects. We compare both algorithms to state-of-the-art hashtag recommendation algorithms and find that our cognitive-inspired approaches outperform these algorithms in terms of prediction accuracy and ranking. With our work, we aim to contribute to the rich line of research on improving the use of hashtags in social networks. We also hope to spark future work to utilize models from human memory theory to model and explain digital traces and user behavior online. For the sake of reproducibility, we implement and evaluate our approach by extending our open-source tag recommender benchmarking framework *TagRec*. The source code and framework are freely accessible for scientific purposes on the Web³.

Literaturverzeichnis

- [An04] Anderson, John R; Bothell, Daniel; Byrne, Michael D; Douglass, Scott; Lebiere, Christian; Qin, Yulin: An integrated theory of the mind. *Psychological review*, 111(4):1036, 2004.
- [KL16] Kowald, Dominik; Lex, Elisabeth: The Influence of Frequency, Recency and Semantic Context on the Reuse of Tags in Social Tagging Systems. In: *Proceeding of Hypertext'16*. ACM, S. 237–242, 2016.
- [KPL17] Kowald, Dominik; Pujari, Subhash Chandra; Lex, Elisabeth: Temporal Effects on Hashtag Reuse in Twitter: A Cognitive-Inspired Hashtag Recommendation Approach. In: *Proceedings of the 26th International Conference on World Wide Web*. 2017.

³ <https://github.com/learning-layers/TagRec>