

PEN recsys: a Personalized News Recommender Systems Framework

Florent Garcin

Boi Faltings

Artificial Intelligence Lab
Ecole Polytechnique Fédérale de Lausanne
Switzerland
firstname.lastname@epfl.ch

ABSTRACT

We present the Personalized News (PEN) recommender systems framework¹, currently in use by a newspaper website to evaluate various algorithms for news recommendations. We briefly describe its system architecture and related components. We show how a researcher can easily evaluate different algorithms thanks to a web-based interface.

Categories and Subject Descriptors

H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval—*Information filtering*

Keywords

recommender system, news, online evaluation

1. INTRODUCTION

Researchers in the recommender systems community have developed open-source platforms which try to bring a growing number of recommender algorithms under one roof. Most of these platforms are designed for research purposes to conduct offline evaluations and only a few target online evaluations on production websites.

In this paper, we are interested in *online* evaluation of state-of-the-art algorithms for *news* recommendations. Unfortunately, it is not possible to use current open-source platforms because they are not tailored to the specific needs of news recommendations and thus are difficult to adapt to the news domain [1].

To this end, we present the PEN recsys framework for online evaluation of news recommender systems. PEN recsys is designed with 4 criteria in mind. First, it has to be *fast*. The framework must provide real-time recommendations as soon as possible, without making the users wait. Second, it must be *reliable*. It is not acceptable for a newspaper website to suffer from crashes. Third, a *flexible* design is

¹Note that this paper is an extended abstract of [2].

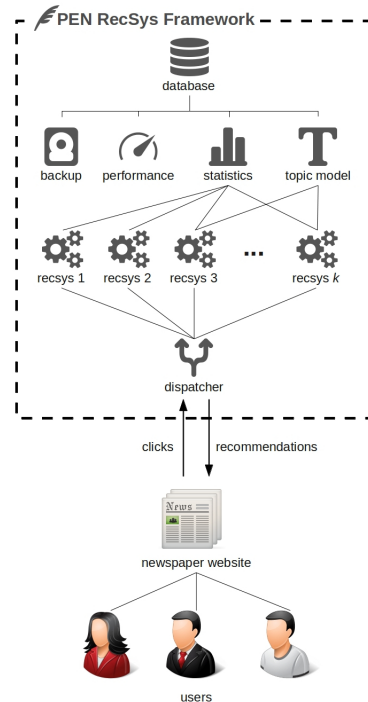


Figure 1: System architecture and components

important. It should be easy to add new components or extend recommender systems. Finally, it must be *scalable*. News websites are subject to unpredictable visit peaks and the framework must be able to handle them by delivering recommendations on time and without problems.

2. SYSTEM ARCHITECTURE

Figure 1 gives a brief overview of the main components in the PEN recsys framework. When a user clicks on a news story, she is assigned to a recommender system which will provide recommendations.

The *dispatcher* randomly assigns a recommender system to a user, performing A/B or multivariate testing. The *recsys* 1, 2, 3, ..., *k* are the different algorithms to evaluate (See Section 3 for more details). Some algorithms relies on click statistics. The component *statistics* gathers click statistics about the stories. Other recommender systems needs the content of news articles, or more specifically to its topic

distribution. The component *topic model* is in charge of keeping the topic model up-to-date.

The *performance* component generates performance reports of the algorithms under evaluation. The *database* stores the clicks, statistics and performance reports for of-line analysis.

Finally, the *backup* component periodically triggers backup to the hard disk of the various states of the system such as the current topic model and click statistics. This is useful if we want to roll back to a previous set of parameters.

It is important to deliver recommendations to the user as soon as possible. With this in mind, the platform is designed to reduce this latency to the minimum. Hence tasks that are not essential for generating recommendations are run in the background. For instance, the database is known to be a bottleneck. Thus statistics are first cached in memory and later stored in the database when resources are available.

The PEN recsys framework follows the software-as-a-service paradigm and is implemented using Java EE technologies. It scales very well since each component can be physically located on different sites.

3. RECOMMENDER SYSTEMS

The framework contains various recommender systems such as the 4 versions of context-tree recommender systems [1], a simple collaborative filtering, a content-based approach, most popular articles, and random articles. We plan to implement more algorithms in the future.

To add a new algorithm, we just need to implement a single method `getRecommendations`. If required, the algorithm can have access to the click statistics, fresh news stories, or topic model via the specific components.

4. INTERFACES

The PEN recsys framework has a web-based control panel (Fig. 2). The researcher can configure the general behaviour of the framework, enable/disable an algorithm, fine tune its parameters (see Fig. 3).

The researcher can also check the performance of the enabled algorithms. Figure 4 shows the performance panel with 3 metrics: success@5, mean average precision and the average clicks per visit.

5. CONCLUSION

We presented the PEN recsys framework which aims at helping researchers and practitioners to evaluate algorithms for news recommendations. The PEN recsys framework is fast, reliable, flexible and scalable. With the help of a simple control interface, it is possible to fine tune each recommender system and have a direct feedback of their performance.

In the future, we plan to increase the set of recommender systems available on this platform.

6. REFERENCES

[1] F. Garcin, C. Dimitrakakis, and B. Faltings. Personalized news recommendation with context trees. In *Int. Conf. on Recommender Systems*, 2013.

[2] F. Garcin and B. Faltings. Pen recsys: a personalized news recommender systems framework. In *News Recommender Systems Workshop*, 2013.

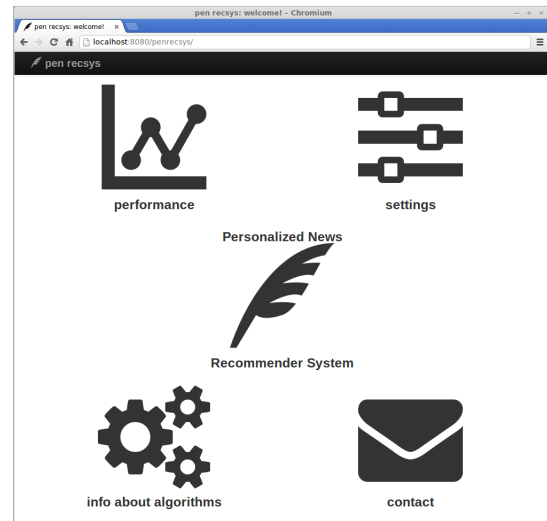


Figure 2: screenshot of the main panel

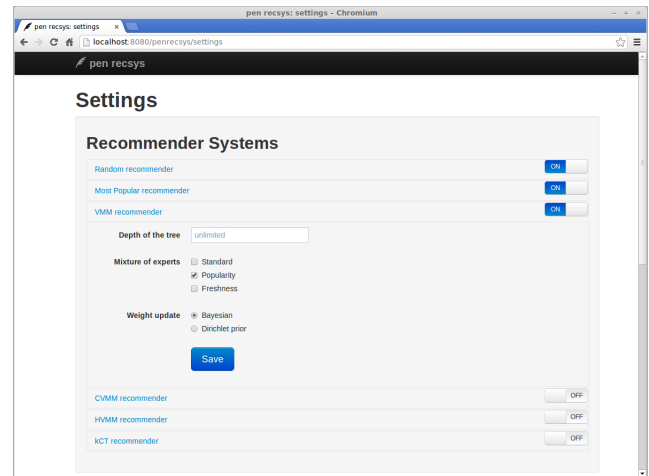


Figure 3: screenshot of the setting panel (partial)

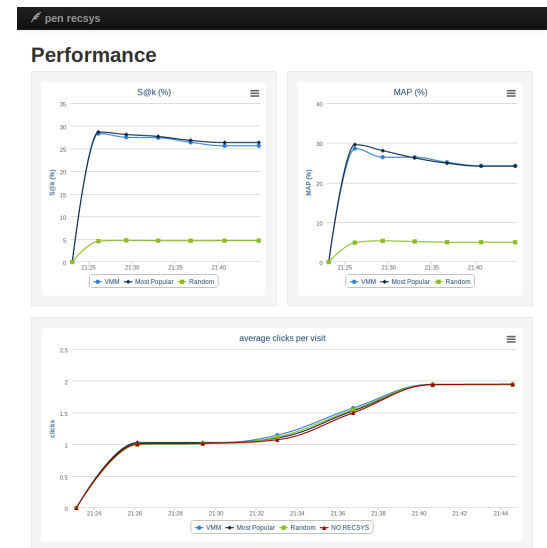


Figure 4: screenshot of the performance panel