

A Novel Short-term and Long-term User Modelling Technique for a Research Paper Recommender System

Modhi Al Alshaikh, Gulden Uchyigit and Roger Evans

School of Computing, Engineering and Mathematics, University of Brighton, Brighton, U.K.

Keywords: Recommender System, Personalization, User Profile, Research Papers, Short-term, Long-term.

Abstract: Modelling users' interests accurately is an important aspect of recommender systems. However, this is a challenge as users' behaviour can vary in different domains. For example, users' reading behaviour of research papers follows a different pattern to users' reading of online news articles. In the case of research papers, our analysis of users' reading behaviour shows that there are breaks in reading whereas the reading of news articles is assumed to be more continuous. In this paper, we present a novel user modelling method for representing short-term and long-term user's interests in recommending research papers. The short-term interests are modelled using a personalised dynamic sliding window which is able to adapt its size according to the ratio of concepts per paper read by the user rather than purely time-based methods. Our long-term model is based on selecting papers that represent user's longer term interests to build his/her profile. Existing methods for modelling user's short-term and long-term interests do not adequately take into consideration erratic reading behaviours over time that are exhibited in the research paper domain. We conducted evaluations of our short-term and long-term models and compared them with the performance of three existing methods. The evaluation results show that our models significantly outperform the existing short-term and long-term methods.

1 INTRODUCTION

A major challenge in recommender systems is the modelling of dynamically evolving short-term and long-term user's interests. The short-term interests represent the user's most recent interests which are more erratic, whereas the long-term interests are more stable in comparison (Challam et al., 2007). Recommender systems for research papers suffer from a number of limitations; for example, fast deviations in short-term interests may remain undetected and stable long-term interests may not be appropriately updated to reflect the user's evolving short-term and long-term interests. The importance of this stems from the need to design automatically adaptable user profiling techniques that should keep track of multiple information that is needed by the user. It is important to recommend right papers at the right time. Therefore, there is a need for user profiling models and techniques that automatically adapt to the diverse and frequently changing users' short-term and long-term interests.

Existing short-term and long-term user modelling techniques have been developed for domains such as

recommending web pages (Gao et al., 2013; Hawalah and Fasli, 2015; Li et al., 2007) and news articles (Zeb and Fasli, 2011; Agarwal and Singhal, 2014; Zeb and Fasli, 2012), where a user reading behaviour is different from the research paper domain. These models depend on continuous time-based user behaviour measured in days for the web pages domain and in hours in the news domain. These models also assume that users are continuously active in their reading with no significant breaks.

In this paper, we present analysis of users' reading behaviour of research papers using the BibSonomy dataset (Knowledge & Data Engineering Group, 2017). The BibSonomy dataset contains actual records of users' interests as posts for research papers. We consider these posts as users' reading records of research papers. Our analysis shows that users are actively reading during some days and inactive on other days. Moreover, they may also be inactive for several months. Furthermore, the users have different reading behaviours from each other, and reading behaviour for a user may change during a year. Therefore, utilizing continuous time-based models for building a user's profile based on continuous

timing algorithms (such as Hawalah and Fasli, 2015) or time-based window (such as Gao et al., 2013) are not appropriate. In this paper, we propose a novel user modelling method for short-term and long-term interests as follows:

- a. *Short-term model:* this model is based on a novel personalized dynamic sliding window (PDSW) technique where the window length is adapted according to the ratio between the number of concepts/interests and number of papers recently read by the user. The content of these papers are then used to build the user's short-term profile.
- b. *Long-term model:* this model determines the user's long-term concepts/interests and then selects papers that represent those concepts/interests. The user's long-term profile is built from the selected papers.

The rest of this paper is organized as follows. Section 2 analyses users' reading behaviour of research papers using the BibSonomy dataset. Section 3 presents our short-term and long-term models. Section 4 presents evaluation and results produced by our models. Finally, the conclusions are presented in section 5.

2 ANALYZING USERS' READING BEHAVIOUR OF RESEARCH PAPERS USING THE BIBSONOMY DATASET

The BibSonomy dataset contains actual records of users' interests as posts for research papers over approximately a ten-year period. Each post contains: metadata for a research paper, date and time of the post. We consider these posts as users' reading records of research papers. For our analysis, we used records of users' reading behaviour over the last two years 2015 and 2016 for users in computing area. This included analysis of 1,642 user records and 43,140 research papers. Our analysis involved automatically searching for patterns of users' reading behaviour. Firstly, we analysed the periods of days and months that a user was inactive (an inactive day/month is a day/month that the user did not read any papers). Secondly, we analysed the users' reading behaviour during active months.

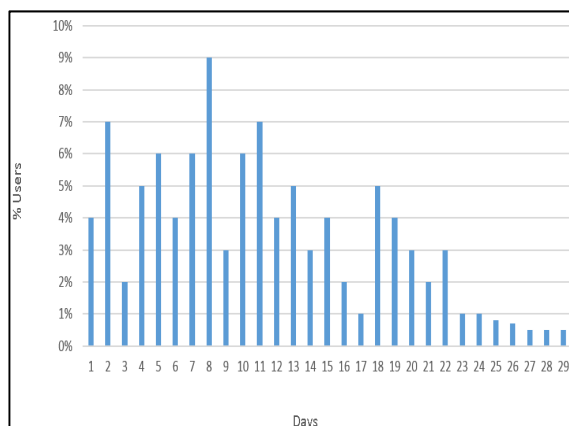


Figure 1: Average inactive days in one active month.

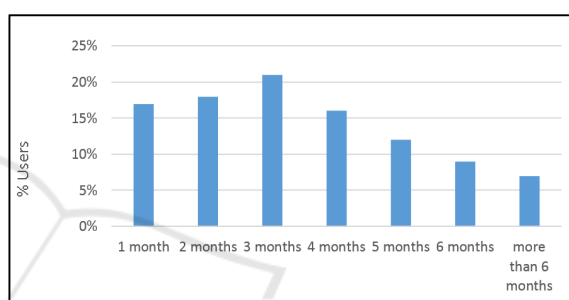


Figure 2: Average inactive months.

We analysed the periods of days and months that a user was inactive as follows:

- a. Average number of consecutive inactive days during one active month. (An inactive day is a day that the user did not read any papers.)
- b. Average consecutive inactive months. (An inactive month is a month that the user did not read any papers.)

Figure 1 shows the average number of consecutive inactive days in one active month. It can be seen that users are not active every day; they do not read papers continuously. Also, users have different patterns of this short-term inactivity. For example, 9% of users are inactive for eight days per active reading month. Therefore, using a fixed duration in time-based models for short-term user profiling is not suitable in this domain. This is because the users can be inactive for several days, which will lead to inaccuracies if modelled based on fixed time periods.

Figure 2 presents the average consecutive inactive months. Our results show that users may not read for several months and may have long inactive periods. For example, our results show that 21% of users are

inactive in reading papers for three continuous months.

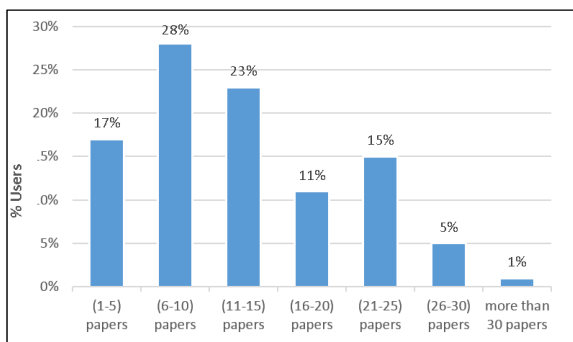


Figure 3: Average number of papers per active month.

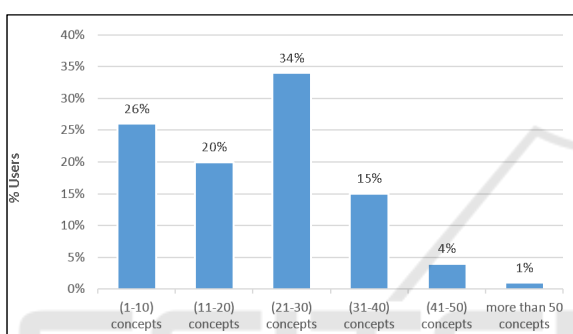


Figure 4: Average number of concepts per active month.

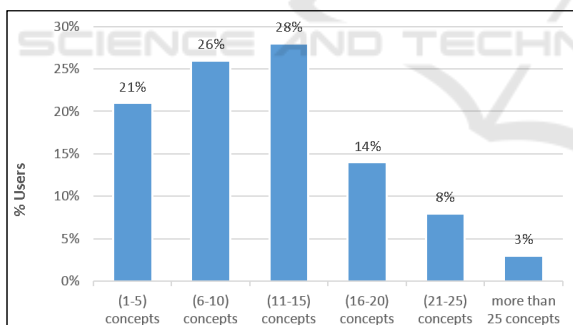


Figure 5: Number of long-term concepts.

Our analysis for the users' behaviour during active months includes the following:

- Average number of papers that are read by a user per active month.
- Average number of concepts/interests encountered in a user's reading per active month.
- Number of long-term concepts that stay in a user's record more than one active month.

Figure 3 shows the average number of papers read by a user per active month. There is significant variability in the number of papers read by users in one active month. For example, 28% of the users read 6-10 papers and 23% of the users read 11-15 papers per one active month.

We analyse average number of concepts per one active month as follows. From the BibSonomy metadata we extracted papers' title, abstract and keywords. Then, each paper is entered to the classifier in our earlier work (Al Alshaikh et al., 2017) to classify it to the three most closely related concepts in 2012 ACM Computing Classification System (CCS) ontology (ACM, 2012).

Figure 4 shows the average number of concepts that are encountered by a user per active month. Figure 5 presents number of long-term concepts that remain in a user's record for more than one active month. It can be seen that the number of long-term concepts in Figure 5 are fewer than the number of concepts in Figure 4. For example, the largest group of users in Figure 4 (34%) encounters 21-30 concepts per month, whereas the largest group of users in Figure 5 (28%) have 11-15 concepts remaining for more than one active month. This is because some of the concepts represented in Figure 4 can be short-term interests. Not all the short-term concepts can be considered as being long-term concepts. The current recommender systems for research papers do not involve short-term and long-term models; they mostly use the whole user reading history. Hence, they are not efficient in recommending the right papers at the right time for evolving users' interests. Therefore, it is important to develop short-term and long-term models for a research paper recommender system. The next section presents our novel short-term and long-term models.

3 SHORT-TERM AND LONG-TERM USER MODELS

In this section, we present our novel short-term and long-term models which automatically adapt to different users' reading behaviour.

3.1 Short-term Model

The short-term model uses novel personalized dynamic sliding window (PDSW) technique. The PDSW length is the number of latest papers that are read by a user. These papers are then used to build a short-term user's profile, represented as Dynamic

Normalized Tree of Concepts (DNTC) as in our earlier work (Al Alshaikh et al., 2017). Figure 6 presents the basic idea of our short-term model. In Figure 6 the PDSW length is four papers. P_1 is the first paper read by the user, P_2 is the second paper and so on, the current time is T and the short-term user's DNTC tree is U_T .

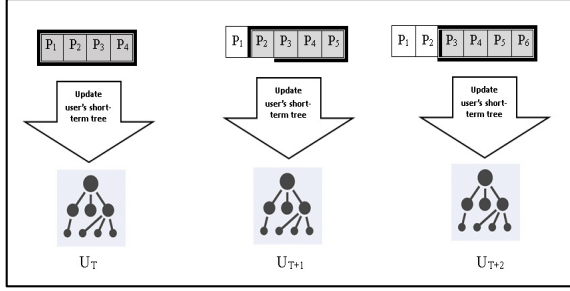


Figure 6: Building DNTC using our short-term dynamic window.

The PDSW length is modified according to the ratio between number of concepts and number of papers that are read by the user. The ratio is calculated for the previous active reading days for a user and results in the length of the sliding window to extend or shrink according to the user's behaviour. The ratio R on time T is calculated as follows:

$$R_T = \frac{\sum_{i=1}^{PAD_T} \frac{nCi}{nP_i}}{PAD_T} \quad (1)$$

where PAD_T is the number of previous active days on time T , nCi is the number of concepts in active day i and nPi is the number of papers in active day i . Each time a new paper is read by a user, the new ratio R_{T+1} is compared with the previous ratio R_T . If R_{T+1} is larger than R_T , then the previous PDSW length has a greater distribution of concepts. Hence, we shrink the PDSW length to focus on the latest papers and concepts to discover the new short-term interests. If R_{T+1} is smaller than R_T , then we extend the PDSW length. If R_{T+1} is equal to the R_T then the window length remains unchanged. To shrink or extend the length (L) of PDSW, Signum function¹ (sgn) is used as follows:

$$L_{T+1} = L_T + \beta * sgn(R_T - R_{T+1}) * R_{T+1} \quad (2)$$

Where L_{T+1} is the new window length on time $T+1$, L_T is the previous window length on time T , β is decay

factor and sgn function as follows:

$$sgn(R_T - R_{T+1}) = \begin{cases} -1 & \text{if } R_T - R_{T+1} < 0 \\ 1 & \text{if } R_T - R_{T+1} > 0 \\ 0 & \text{if } R_T - R_{T+1} = 0 \end{cases}$$

After calculating the new PDSW length, the latest papers that are read by the user are selected to represent the user's short-term profile. The number of selected papers is an integer equal to the PDSW length. Then, the short-term user's profile is represented as DNTC profile as in (Al Alshaikh et al., 2017). Dynamic Tree Edit Distance technique as in (Al Alshaikh et al., 2017) is then used to recommend a set of papers to the user that match his/her short-term interests.

3.2 Long-term Model

The long-term model is updated at the end of each active month for a user. Long-term concepts are the concepts that remain for more than one active month in a user's record. The long-term model selects the papers that represent long-term concepts, then these papers represent a user's long-term profile. The set of long-term concepts is defined as $LC = \{LC_1, LC_2, \dots, LC_n\}$, where n is the total number of long-term concepts. After selecting the long-term concepts, the papers that are related to at least one of the long-term concepts are selected to represent a user's long-term profile. The set of long-term papers is defined as $LP = \{Lp_1, Lp_2, \dots, Lp_m\}$, where m is the total number of long-term papers and Lp_i is related to at least to one of LC concepts. Then the set of papers LP is used to build a user's long-term DNTC as in (Al Alshaikh et al., 2017). Then, the Dynamic Tree Edit Distance technique (Al Alshaikh et al., 2017) is used to recommend a set of papers to the user that match his/her long-term interests.

4 EVALUATIONS

4.1 Evaluation of Short-term Model

We evaluated the performance of our short-term model using the BibSonomy dataset. The BibSonomy dataset in section 2 was pruned to remove users with fewer than 60 active days (an active day is a day that the user reads at least one paper). The remaining dataset consists of 1,074 users in the year 2015 and

¹ https://calculus.subwiki.org/wiki/Signum_function

2016. Every day in the 60 active days for each user is evaluated. The **training set for an active day i** is the papers in the user's record for previous active days before the active day i . The **testing set for an active day i** is the papers that exist in day i and the next 29 calendar days in the user's record (we assume that the duration for short-term interests is 30 calendar days). At every active day i , if a recommended paper exists in its testing set, then it is relevant to his/her short-term interests. The measurement that is used for evaluation is precision at top k papers of an active day i for a user a as follows:

$$P_k(d_i, a) = \frac{NP_{i,a}}{k} \quad (3)$$

where $NP_{i,a}$ is the number of recommended papers that match the testing set for active day i for user a . Then, the average precision is calculated for all users U for an active day i as follows:

$$AVG P_i = \frac{\sum_{a=1}^U P_k(d_i, a)}{U} \quad (4)$$

The mean average precision for all active days is calculated for all active days (AD) as follows:

$$MAP = \frac{\sum_{i=1}^{AD} AVG P_i}{AD} \quad (5)$$

4.1.1 Evaluating B Parameter

In this section we evaluated different values of β (the decay factor in equation 2) parameter to find the optimal value that provide the best overall performance for our short-term model. The optimal value of the decay parameter β was determined by measuring the precision of the model for different values of β . The measurement that is used for evaluation is precision at top 10 papers ($k=10$). Figure 7 presents the MAP for all users using different values of β in the range of [0.1 to 1]. When $\beta = 0.1$, the PDSW length is very small to detect the short-term interests. The results increase when the β value increases until $\beta = 0.6$, where MAP is 0.76. Then, the PDSW length becomes very large and may include some of the old short-term interests that do not belong anymore to the user's current short-term interests. The value of β used in our model was therefore $\beta = 0.6$.

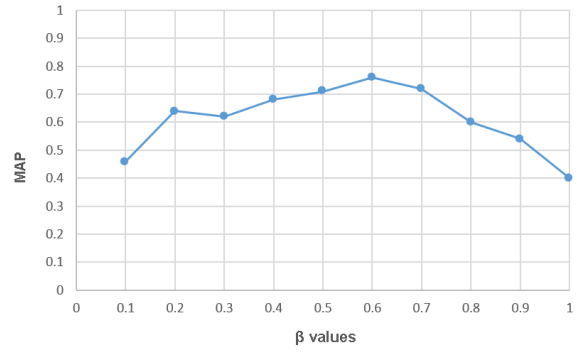


Figure 7: MAP results using different β values for PDSW.

4.1.2 Comparing Our Short-Term Model against Baselines

We compared our PDSW short-term model against three baselines:

1. DNTC system (Al Alshaikh et al., 2017).
2. Static window time-based model in (Gao et al., 2013).
3. Dynamic time-based model for short-term model in (Hawalah and Fasli, 2015).

Our PDSW short-term model and the three systems are run for each day during the 60 active days. Figure 8 shows the overall comparison for our short-term model against the three systems over 60 active days. Table 1 shows the MAP that reflect the results of those of Figure 8. It can be seen that the DNTC system achieves the lowest precision performance with MAP over the 60 active days of 0.47. The DNTC system does not consider short-term behaviour but includes all the papers read by a user. Considering all previous papers in a user's record give the previous existing concepts high weights in a user's profile, hence they are considered as short-term interests. However, new concepts receive lower weights in a user's profile, which can cause sharp drops in the precision in some active days. When it comes to the Static window time-based system, the performance is slightly better than the DNTC system with MAP of 0.49. This is because this system considers only the latest papers during the static window time-based. The low performance of this system because it assumes a user's reading behaviour is static, whereas in reality the user behaviour changes over time. Moreover, each user has different personalized behaviour. When it comes to the Dynamic time-based system, there is improvement in the performance with MAP of 0.55. This system is better than the previous two systems because it can

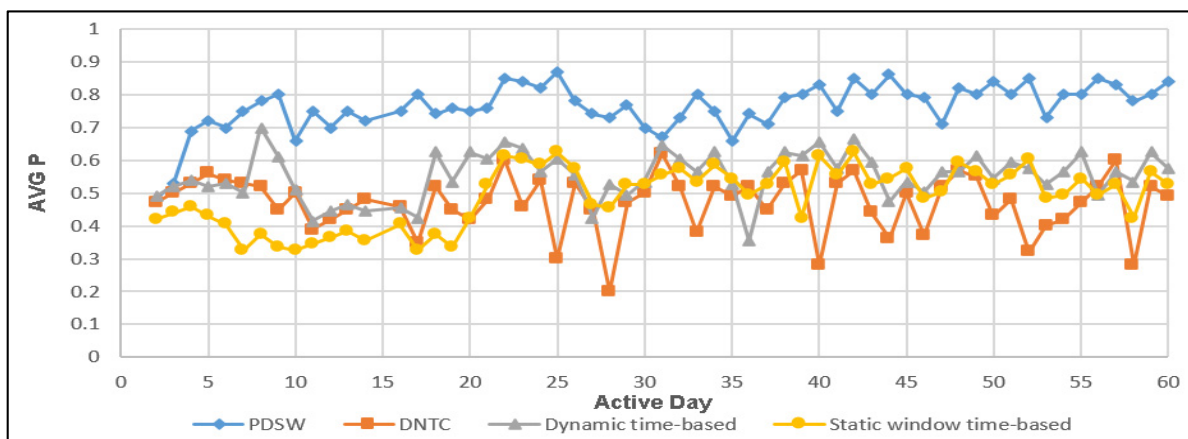


Figure 8: Comparing average precision for our short-term model against baselines.

handle the situation when new short-term concepts arise in a user's profile, and it does not depend on static time-based behaviour. However, it has a limitation that it cannot handle the problem of different inactive days for different users' behaviour. Our PDSW system achieves MAP of 0.76 which is an improvement on each of the previous three systems. These results show that our short-term model can effectively learn different users' reading behaviour even if there are different patterns of inactive days. Moreover, it dynamically adapts with the changes in a user's reading behaviour over time.

Table 1: MAP results for the four short-term systems.

System	MAP
DNTC	0.47
Static window time-based	0.49
Dynamic time-based	0.55
PDSW	0.76

4.2 Evaluation of the Long-term Model

We evaluated the performance of our long-term model using the BibSonomy dataset. The BibSonomy dataset in section 2 was pruned to remove users with fewer than 12 active months during the years 2015 and 2016 (an active month is a month that the user reads at least one paper). The remaining dataset consists of 261 users. Every month in the 12 active month for each user is evaluated. The **training set for an active month i** is the papers in the user's record for previous active months before the month i . The **testing set for an active month i** is the papers that exist in the rest of the user's record and one of its concepts is long-term concept 'LC'. At every active month i , if a recommended paper exists in its testing set, then it is relevant to his/her long-term interests.

The measurement that is used for evaluation is precision at top k papers of an active month i for a user a as follows:

$$P_k(m_i, a) = \frac{MP_{i,a}}{k} \quad (6)$$

Where $MP_{i,a}$ is the number of recommended papers that are exist in the testing set for active month i for user a .

Then, average precision is calculated for all users U for active month i as follows:

$$AVG P_i = \frac{\sum_{a=1}^U P_k(m_i, a)}{U} \quad (7)$$

The mean average precision for all active months is calculate for all active months (AM) as follows:

$$MAP = \frac{\sum_{i=1}^{AM} AVG P_i}{AM} \quad (8)$$

We compared our long-term model against three baselines:

1. DNTC system (Al Alshaikh et al., 2017).
2. Time-based forgetting factor model in (Gao et al., 2013).
3. Dynamic time-based for long-term interests in (Hawalal and Fasli, 2015).

Our long-term model and the three systems are run at the end of each active month for each user. The top 10 recommended papers ($k=10$) are evaluated. Figure 9 shows the overall comparison for our long-term model against the three systems over 12 months.

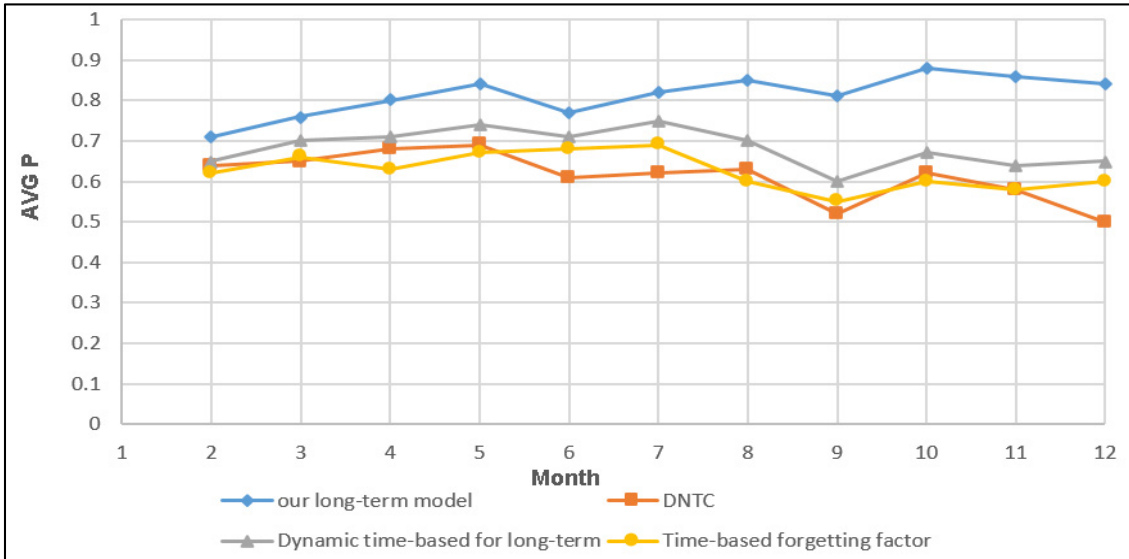


Figure 9: Comparing average precision for our long-term model against baselines.

Table 2 shows the MAP that reflect the results of those of Figure 9. It can be seen from Figure 9 and table 2 that the DNTC achieves the lower precision performance with MAP of 0.61. After the fifth month DNTC performance declined dramatically because of cumulative calculations for all the papers that are read by the user. This low performance is because DNTC includes all the papers in a user’s record even the papers for short-term interests. When it comes to the time-based forgetting factor model, the performance is slightly better than the DNTC with MAP of 0.63. This is because this model has a forgetting factor. However, this forgetting factor is fixed for all users and does not consider different users’ behaviour. When it comes to the Dynamic time-based model for long-term interests, there is improvement in the performance with MAP of 0.68. This model is better than the previous two models because it can handle the situation when there is short-term concepts and long-term concepts, and it does not depend on static time-based technique. However, it has a limitation that it does not handle well long inactive periods in users’ behaviour. Therefore, after the seventh month its performance declined significantly. Our long-term model achieves MAP of 0.81 which is better than each of the previous three models. This is because our model can effectively learn different users’ reading behaviour even if there are different long inactive periods. Our long-term model significantly outperforms the other three baselines after the seventh month as shown in Figure 9.

Table 2: MAP results for the four long-term systems.

System	MAP
DNTC	0.61
Time-based forgetting factor	0.63
Dynamic time-based	0.68
Our long-term model	0.81

5 CONCLUSIONS

In this paper, we presented our novel short-term and long-term models for a research paper recommender system. First, we analysed users’ reading behaviour in the BibSonomy dataset. Our analysis shows that the users’ reading of research papers is different to that of reading web pages and news articles. Therefore, we developed our short-term and long-term models based on our analysis of users’ reading behaviour for the research paper domain. Our evaluations of performance demonstrate that our models significantly outperforms the other baseline systems. Our short-term PDSW model achieves MAP of 0.76 and our long-term model achieves MAP of 0.81. The performance advantage is because our models can effectively learn different users’ reading behaviour. Moreover, they dynamically adapt to the changes in users’ reading behaviour over time. In future work, we will combined our short-term and long-term models and add collaborative model to develop a hybrid system for the research paper domain.

REFERENCES

- ACM Computing Classification System, 2012, URL: <https://www.acm.org/about/class/2012>.
- Agarwal, N., Haque, E., Liu, H. and Parsons, L., 2005. Research paper recommender systems: A subspace clustering approach. *In Advances in Web-Age Information Management*, Springer, pp.475-491.
- Agarwal, S. and Singhal, A., 2014. Handling skewed results in news recommendations by focused analysis of semantic user profiles. *In IEEE International Conference on Optimization, Reliability, and Information Technology (IEEE ICROIT)*, pp. 74-79.
- Al Alshaikh, M., Uchyigit G. and Evans, R., 2017. A Research Paper Recommender System Using a Dynamic Normalized Tree of Concepts Model for User Modelling. *In IEEE Eleventh International Conference on Research Challenges in Information Science (IEEE RCIS 2017)*, pp.200-210.
- Challam, V., Gauch, S. and Chandramouli, A., 2007. Contextual search using ontology-based user profiles. *In Large Scale Semantic Access to Content (Text, Image, Video, and Sound)*, pp. 612-617.
- Gao, Q., Xi, S.M. and Im Cho, Y., 2013. A multi-agent personalized ontology profile based user preference profile construction method. *In IEEE 44th International Symposium on Robotics (ISR)*, pp. 1-4.
- Hawalah, A. and Fasli, M., 2015. Dynamic user profiles for web personalisation. *Expert Systems with Applications*, 42(5), pp.2547-2569.
- Knowledge & Data Engineering Group, University of Kassel: Benchmark Folksonomy Data from BibSonomy, version of January 1st, 2017.
- Li, L., Yang, Z., Wang, B. and Kitsuregawa, M., 2007. Dynamic adaptation strategies for long-term and short-term user profile to personalize search. *In Advances in Data and Web Management*. Springer. pp. 228-240.
- Zeb, M. A. and Fasli, M., 2011. Adaptive user profiling for deviating user interests. *In Computer Science and Electronic Engineering IEEE Conference (CEEC)*, pp. 65-70.
- Zeb, M. A. and Fasli, M., 2012. Dynamically Adaptive User Profiling for Personalized Recommendations. *In IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, pp. 604-611.