

Gamified Incentives: A Badge Recommendation Model to Improve User Engagement in Social Networking Websites

Reza Gharibi

Department of Computer Science & Engineering
Shiraz University
Shiraz, Iran

Mohammad Malekzadeh

School of Electronic Engineering and Computer Science
Queen Mary University of London
London, United Kingdom

Abstract—The online social communities employ several techniques to attract more users to their services. One of the essential demand of these communities is to find efficient ways to attract more users and improve their engagement. For this reason, social media sites typically take advantage of gamification systems to improve users' participation. Among all the gamification services, badges are the most popular feature in online communities which are massively used as a reward system for users. Therefore, the recommendation of relevant unachieved badges to users will have a significant impact on their engagement level; instead of leaving them in the ocean of different actions and badges. In this paper, we develop a badge recommendation model based on item-based collaborative filtering which recommends the next achievable badges to users. The model calculates the correlation between unachieved badges and users' previously awarded badges. We evaluate our model with the data from Stack Overflow question-answering website to examine if the recommendation model can recommend proper badges in an existing real community. Experimental results show that the model has about 70 per cent true recommendation by just recommending one badge and it has about 80 per cent correct recommendation if it recommends two badges for each user.

Keywords—Social Media; Data Mining; Gamification Algorithms; User Engagement; Recommendation Systems

I. INTRODUCTION

The user experience of video games as well as models, methods, and heuristics which had been developed by researchers for the usability or playability of games [1] has become a notable topic of discussions on social networking websites [2]. An obvious matter of interest in this field is the idea of using the game design elements in non-game contexts to improve user experience and user engagement. Since the video games primarily designed to entertain people and motivate people to engage with them, we should be able to use game elements in other non-game products and services to make them more enjoyable and engaging as well [3]. This idea is a new phenomenon which is called Gamification.

Following the success of the location-based services, such as Foursquare and Nike+, gamification has rapidly gained attention in design and digital marketing [4], [5]. Numerous empirical studies have shown that gamification has positive effects in a wide range of contexts [6]. The gamification can

also help finding an effective avenue to attract customers and retain an interest in today's digital world. After all, gamification is really just about getting more people to do more stuff more often. That's why most of the time hours fly by when playing video games without noticing.

These days several vendors offer gamification as a service layer of reward and reputation systems with points, badges, levels and leader boards (e.g. Badgeville) [7]. At the same time, gamification is increasingly catching the interest of social media researchers [4]. In this paper, among all the gamified elements, we focus on badges. Badge-based achievement systems are being used increasingly to drive user participation and engagement across a variety of platforms and contexts [8]. Powerful examples of large-scale successful implementation of badges are Valve's Steam and Microsoft's Xbox Live platform where all games released must have some sort of achievements [9]. Since then, badges have widely implemented on different gaming platforms and have been extremely successful. They also employed across a wide range of domains, from news sites like Huffington Post, where users are recognised for contributing valuable comments and being well-connected; to education sites like Khan Academy and Codecademy, where users are awarded badges for correctly answering questions; to knowledge creation sites like Wikipedia and Stack Overflow, where users are awarded for their contributions to the online communities.

In fact, some social networking websites use the badges as separate milestones for each user, some use them to exhibit each user's skills, and some award badges to users for doing certain actions [10]. This latter use of badges is the main functionality we have worked on in this paper, because as an incentive by recommending certain badges which have associated with certain actions, websites can control users' behaviour and make them eager to participate more in the society. Badges are simpler than other incentives and in practice, many social websites have positioned them as an important part of their incentive system. However, despite their simplicity, they include some aspects of complex users' behaviours. The most fundamental way in which badges can influence users' behaviour is by encouraging them to expand their general level of interest and participation.

1) **Problem Definition.** Social media websites can provide a wide range of activities that users can do for most of their

The work was done while the authors were working at the Persian Gulf University of Bushehr.

parts, and by creating badges for one or a set of these activities they can control users' behaviour and steer them to the direction that they seek. As an example, let's suppose there is a question and answer (Q&A) website with a set of activities like asking and answering questions, up-voting and down-voting questions and answers, editing questions and answers and so on. In addition, we assume that the users of this site are the people who just ask questions and none of them likes answering questions. In this situation, we can create a new badge for answering questions and reward it to users who answer the question. In this way, we provide enough enthusiasm for users not only to ask questions but also try to answer questions to gain the new answer badge. By considering a Q&A website scenario, we assume that there is a complete set of predefined badges for different activities on the site. All badges are threshold badges and are awarded once a user has taken a specified number of actions of certain types. Between all these badges the challenging problem is finding the relevant ones. We want to develop a mechanism for the recommendation of new unachieved badges to each user.

2) **Contribution.** In this paper, we propose an efficient framework based on collaborative filtering—and in specific, item-based collaborative filtering [11]—which is an appropriate solution to our badge recommendation problem. The bottleneck in conventional collaborative filtering or user-based collaborative filtering is the search for neighbours among a large user population of neighbours. This computation increases as the number of users increases and therefore it can simply lose the scalability. However, Item-based techniques avoid this bottleneck by first analysing the user-item matrix to identify relationships between different items, and then using these relationships to compute recommendations for users indirectly. We try to show how well the proposed recommendation model can predict user's future behaviour. Therefore, as the main contribution of this paper, we first develop an item-based collaborative filtering recommendation model for badges which are awarded to users and record their behaviour according to this recommendation model. After that, we evaluate the proposed model on the data from the popular question-answering website Stack Overflow to see how much it's accurate when recommending unachieved badges on an existing environment. We also analyse the dataset to extract some useful statistics from it and develop a baseline algorithm based on this information for more experimental purposes.

The rest of the paper is structured as follows. In Section II, we summarise how this work relates to other research. In Section III, the recommendation model is described. The description of the dataset and the analysis we did are shown in Sections IV and V, respectively. This is followed by Section VI by the empirical evaluation setting of the model and results. Finally, in Section VII, we draw conclusions and describe future work.

II. RELATED WORK

As mentioned in the introduction, the utilisation of badges is a growing pattern in different fields but our research is

principally related to two lines of research: improving user engagement with gamification incentives and designing recommendation algorithms for social media sites.

The first topic is about the study of badge effects in user behaviour and their use to steer and control user actions; leading both to increase participation and changes in activities a user seeks on the website [12], [13]. In this subject, Denny in [8] has worked on the effect of badges on students' engagement in an online learning tool. He discovered a highly significant positive effect on the quantity of students' contributions, without a corresponding reduction in their quality. Students also enjoyed being rewarded by badges for their contributions.

There are also some papers on badge design and placement problem which works on the idea of how to ideally place badges in our system and how should we design them to induce particular user behaviours the way we want [12], [14]. For this matter, Antin and Churchill in [10] presented a conceptual organisation for different types of badges with concentrating on social psychological perspective. Easley and Ghosh in [15] studied the question of how gamification via badges can be most effectively used for incentivising participation in online systems. They analysed various design choices and offered guidance about how, and for what, a website might choose to award badges.

Aside from these topics, the use of badges can also be viewed as part of the growing phenomenon of gamification and more general incentives for contribution in social media [4], [7]. Lounis et al. [16] have worked on the role of incentives and community collaboration. In their study, they investigated the impact of game elements on user's experienced fun during participation in a gamified service.

The second topic is about the field of recommendation systems; our work relates to collaborative filtering recommendation algorithms and in specific item-based collaborative filtering with implicit feedback. In this field, [11] and [17] analysed different item-based recommendation generation algorithms. They looked into different techniques for computing item-item similarities like item-item correlation and cosine similarity. They also compared their results to the basic k-nearest neighbour approach. Results and experiments suggest that item-based algorithms provide better performance than user-based algorithms, while at the same time also providing a better quality recommendation.

For item recommendation based on users' implicit feedback, Hu et al. [18] have identified unique properties of implicit feedback datasets. They transformed user observations into positive and negative preferences with difference confidence levels. Their algorithm is successfully implemented and tested as part of a large-scale TV show recommender system. Liu et al. [19] also proposed a boosting algorithm for item recommendation with implicit feedback. Boosting is a general technique that can improve the accuracy of a given learning algorithm. Herlocker et al. [20] overviewed the factors that have been considered in recommendation systems evaluation and also introduced new factors which should be considered. They described empirical results on accuracy

metrics and showed how results from different accuracy metrics might vary.

In the last few years, some researcher tried to analyse the role of badges [21], reputation systems [22], [23], and question's tags [24] in social medias, especially for Q&A websites. However, on the best of the authors' knowledge, this paper is the first one in examining and analysing a badge recommendation model to improve user engagement through recommending related badges to online social media users.

III. RECOMMENDATION MODEL

In this section, we describe the recommendation model. We are going to develop a badge recommendation model and use collaborative filtering for this purpose. According to the nature of our problem, item-based collaborative filtering [11] best suits this work. Therefore, before anything, it is necessary to find the most similar badges and then combine these similarities with user's badges to generate a recommendation.

The critical step in this item-based collaborative filtering is the computation of the similarity between different badges and then finding the most similar badges to each of the badges. The fundamental idea of similarity computation for two different badges is to find users who have gained both of these badges and then applying a similarity computation technique to determine the similarity scores. There are different similarity scores to use in order to find similarities between badges. We can find similarities by computing distance using measures like Manhattan or Euclidean distance in the n -dimensional space where n is the number of users. We can calculate these distances between two badges as

$$d(x, y) = (\sum_{i=1}^n |x_i - y_i|^r)^{\frac{1}{r}} \quad (1)$$

where x and y are zero-one vectors that show the badge availability of each badge column for each user. For $r = 1$ this formula is the Manhattan distance and for $r = 2$ it's the Euclidean distance. We have tested different similarity measures and among them, we use cosine similarity. A good point of cosine similarity is that it is suitable for sparse data and it doesn't depend on the shared-zero values, so it ignores 0-0 matches of each evaluation of the badge vector. It is defined as

$$\cos(x, y) = \frac{x \cdot y}{||x|| \cdot ||y||} \quad (2)$$

where \cdot indicates the dot product and $||x||$ indicates the length of the vector x which contains a bunch of zeros and ones that show the badge availability of each badge column for our users. We'll show an example of this vector when we start building our model for empirical evaluation. The length of this vector is

$$||x|| = \sqrt{\sum_{i=1}^n x_i^2} \quad (3)$$

and by substituting (3) in (2) the final formula for cosine similarity sums up as

$$\cos(x, y) = \frac{\sum_{i=1}^n x_i \times y_i}{\sqrt{\sum_{i=1}^n x_i^2} \times \sqrt{\sum_{i=1}^n y_i^2}} \quad (4)$$

With this similarity measurement formula, we can calculate the similarity of each badge against other badges and find the most similar badges to each one. The cosine similarity rating ranges from +1 indicating perfect similarity to -1 indicating perfect dissimilarity. By identifying the set of most similar badges with cosine similarity measure, we then develop a technique for badge recommendation which proposes relevant unachieved badges to target users based on the history of their achieving badges (*history* of users' badges).

For each user, we check all the available badges, if the user already has that badge then we won't recommend it, but if he doesn't have that badge we go through our similarity table and extract the most similar badges to this badge and call them *similar badges*. Then using (5), we calculate the similarity between the "history of users' badges" and extracted "similar badges" to get a measure for recommending a new badge according to the users' history.

$$recommendation\ measure = \frac{\sum_{i=1}^s history_i \times similarity_i}{\sum_{i=1}^s history_i + similarity_i} \quad (5)$$

where s is the number of similar badges which can be selected by the model. In (5), *history* is a zero-one vector (of s elements) and *similarity* is a vector that contains s cosine similarity values. For each badge of the user, *similarity* is the top similar badges to that user's badge and *history* is the existence of those top similar badges in the user's badges profile. We calculate this measurement for each badge that the user doesn't already own and finally recommend badges with the highest score from this formula. We can recommend as many badges as we want based on the descending order of scores we get for each badge. We talk more about this part of the recommendation system in Section VI.

We want our model to also cover the cold-start problem for new users who doesn't have any badges. The cold-start problem occurs when it is not possible to make reliable recommendations due to an initial lack of items. As we use threshold badges which are awarded once a user has taken a specified number of actions of certain types, we can simply recommend some common badges with the threshold of one to new users. These are common badges based on the timestamp of achieving them from our existing users and are owned with just one common action. This subject will be discussed more in the analysis section (Section V).

Finally, in this model not only we recommend badges to improve user engagement but also we are considering people's behaviour in our recommendations. We are recommending specific badges according to users' behaviour and history even if they are far from that badge's threshold but their behaviour shows they can go toward it. Figure 1, depicts a big picture of the proposed badge recommendation model in brief.

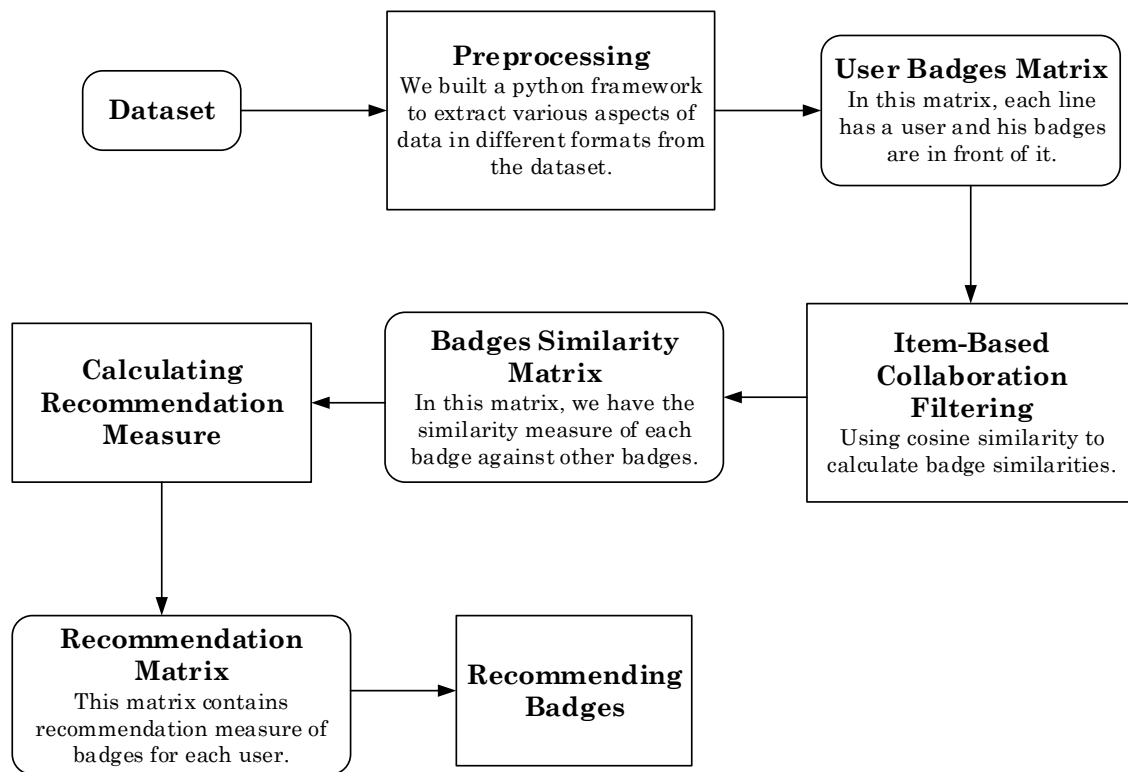


Fig. 1. The process of badge recommendation in the recommendation model

IV. DATASET

We use the dataset of the question-answering site Stack Overflow to test our recommendation model. Stack Overflow is part of the Stack Exchange network and makes extensive use of badges. It is one of the first sites to use badges on a large scale. The anonymised data dump of Stack Overflow is freely available from Internet Archive and it includes an archive file for Posts, Users, Votes, Comments, Badges and so on in the XML format. We use Badges XML dataset which contains all the badges that are awarded to users. In this dataset, each row element shows a single user badge; description of attributes from each row element has shown in Table 1. The data is also available through “Stack Exchange Data Explorer” which lets you run SQL queries directly against a copy of the data.

TABLE I. SUMMARY OF FEATURES FROM THE BADGES DATASET

Feature	Description
UserId	UserId of the badge owner
Name	Badge name
Date	Timestamp of when the user had achieved the badge

There are over 100 different badges on Stack Overflow, which vary greatly in how difficult they are to achieve. For example, there are badges for encouraging new users that nearly everyone obtains, such as the “Autobiographer” badge for filling your profile description which is categorised under bronze badges and also there are complex ones like “Legendary” badge which has a more complex threshold to achieve and is categorised under the gold ones.

In the Badges dataset, each individual badge given to a user is time-stamped. For our work, we use badges that were given from years 2008 to 2010. We turn this period into three separate partitions, one for badges from the year 2008 only, one for years 2008 to 2009 and one for years 2008 to 2010 which covers the whole dataset. We did this so to run experiments on various dataset sizes. The data and source code from our experiments are also available online¹.

V. ANALYSIS

Before running the empirical evaluation, we have done some analysis on the extracted dataset. Table 2, shows the number of users and distinct badges in each partition of the dataset. We can see the number of different badges grew over the years.

TABLE II. NUMBER OF USERS AND DISTINCT BADGES IN EACH PARTITION OF OUR DATASET

Dataset Partition	Number of Users	Number of Badges
2008 only	18,255	88
2008 - 2009	75,182	292
2008 - 2010	210,743	592

Figure 2, shows 20 most frequent badges that were awarded to users in each dataset partition, and Figure 3, shows the frequency of users with one badge to users with 20 badges which makes a nice heavy-tailed like distribution; similar to other observations in social networking websites.

¹ <https://github.com/h4iku/stack-badges>

TABLE III. FIVE MOST FAVOURITE BADGES AS USERS' FIRST BADGE

Badge	Frequency as user's first badge			When to award
	2008	2008 - 2009	2008 - 2010	
Student	4,624	21,859	74,921	First question with score of 1 or more
Teacher	9,041	26,117	54,604	Answer a question with score of 1 or more
Editor	1,628	9,451	30,179	First edit
Scholar	973	8,467	25,065	Ask a question and accept an answer
Autobiographer	1,873	6,931	12,930	Complete "About Me" section of user profile

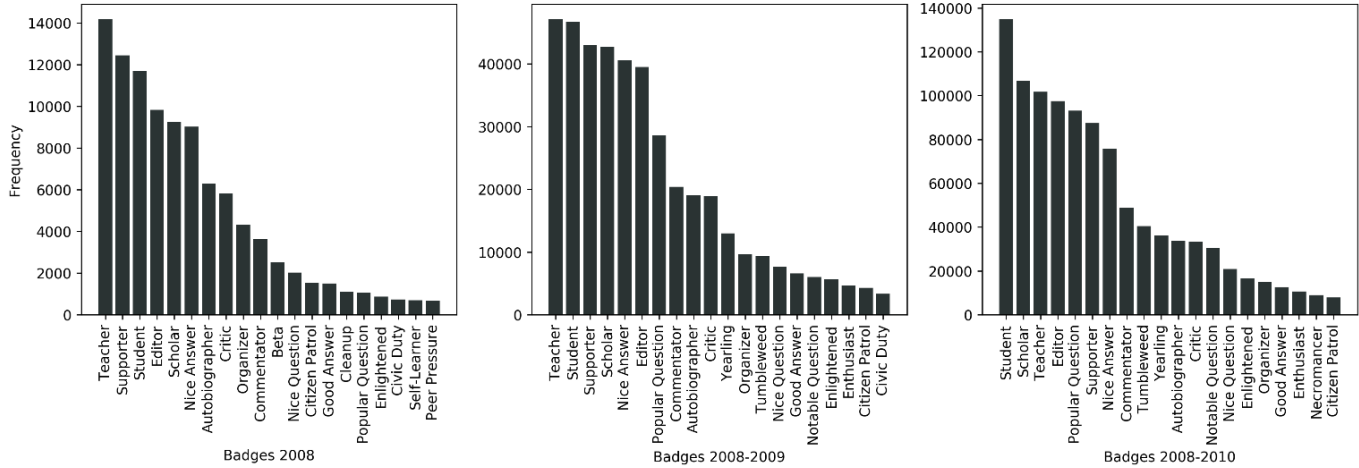


Fig. 2. Twenty most frequent badges in different dataset partitions.

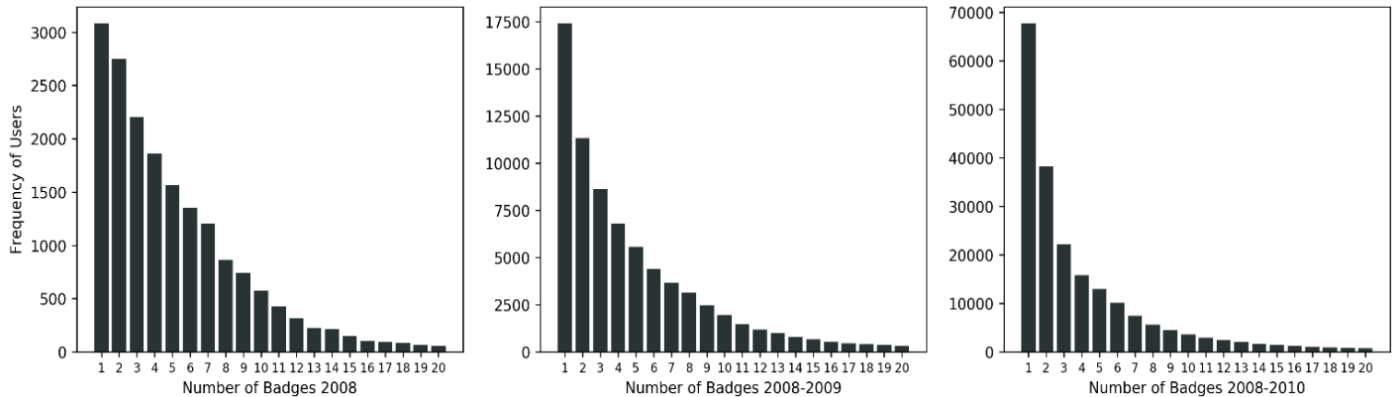


Fig. 3. The frequency of users with one badge to users with 20 badges.

In order to recommend badges to newcomers and cover the cold-start problem, we analysed the first five favourite and frequent badges that users have achieved according to the badges timestamp; the result and description of these badges are shown in Table 3. All these five badges are threshold badges with the threshold of one and they will be awarded by doing just one action, so they are good for recommending to newcomers.

VI. EMPIRICAL EVALUATION

After developing the model, we want to investigate whether the predictions match the users' behaviour we see in the dataset or not. In fact, the training and test setup are designed to evaluate how well the model can predict future user behaviour. As mentioned, our dataset is from the question-answering site Stack Overflow that makes extensive use of badges. We have

built a python framework to extract various aspects of data in different formats from this dataset.

A. Building the Model

The raw data from the dataset is in the XML format. First, we create a matrix file from the XML file, which it has a user in each line and the users' badges are in front of it. So, the row index of our data consists of users and the column index contains all the badges. If a user has a badge, we put 1 under that badge column and if not we put 0. A simple example of this matrix file is shown in Figure 4. Then the dataset is divided into a training set and a test set. We use the training set to build our model and the test set to test it with.

To start building the model for item-based collaborative filtering, we have to determine the similarity between columns which are the badges. We apply cosine similarity between

columns and as a result, we have similarity measure between all the badges. We can also sort this similarity numbers in the descending order to have the most similar badges to each badge. As we are going to test this, we limit the similarities to ten similar badges for each badge.

We also developed a simple baseline algorithm using results of our dataset analysis. For the baseline method, we take five most popular badges in each dataset partition and recommend them to users who doesn't already own them.

B. Test Setting

We made a specific test set from the data to test the model. To build the test set, in each dataset partition we randomly select enough users who have more than five badges. Then for each user, we randomly select one badge and remove it from his badges (leave-one-out evaluation) in the train set [20]. We put this selected badge on the user's test set.

UserId	Autobiographer	Citizen Patrol	Civic Duty	Cleanup
3718	1	1	0	0
994	1	1	1	0
3893	1	1	1	1
4591	1	1	0	1
5196	1	0	1	0
2635	1	1	1	1
1113	0	0	1	0

Fig. 4. An example for the user badges matrix that we extract from the main dataset.

TABLE IV. EVALUATION RESULTS

Title	Dataset Partition		
	2008	2008 - 2009	2008 - 2010
Our Model's Precision	0.70	0.63	0.61
Baseline	0.52	0.43	0.42
Number of Training Users	18,256	75,182	210,743
Number of Test Users	1,344	5,098	10,666

We give the train badges of a user to the recommendation model and the model will recommend one or more badges to that user. Then the recommended badge is checked against the test set to see if the test set contains the recommended badge for that user or not. If the recommended badge is in the test set then we have a true positive because the recommended badge was in the user's test set and if not then we have a false positive because the user didn't have that recommended badge in his test set.

In this phase, we check every badge for each user. If the user has that badge then the model is not going to recommend that badge to that user but if the user doesn't have that badge, we calculate the score of user history badges and badges similar to this badge and get a value for it. After doing this for all the badges, we sort the values in the descending order and recommend badges.

C. Results

In this subsection, we present the experimental results of our empirical evaluation of the Stack Overflow badges dataset. Results are shown in Table 4. As said previously, we divided the dataset into three partitions to run our model on. One

smaller part which contains badges from the year 2008, the second part which contains badges from years 2008 to 2009, and the third part which is our complete dataset and contains badges from years 2008 to 2010.

We also run the model with two badge recommendations and compared the results of top one recommendation with top two recommendations in Table 5.

VII. CONCLUSION AND FUTURE WORK

Although the use of badge incentives is a new trend in online social websites, it has a huge effect on user engagement and participation. Aside from this, recommendation systems are also impressive technologies that help users find their way and are now an important part of online E-commerce systems. Combining these two approaches will give us a nice model to steer user behaviour in online communities.

In this paper, we have built a badge recommendation model using item-based collaborative filtering. We evaluated the model with the Stack Overflow badges dataset to see how well it can predict future user behaviour. This model tried to recommend badges to each user along the user's behavioural activities so that he can find the direction he wants to go in the community. The results show that the model has about 70 per cent true recommendation by just recommending one badge and it has about 80 per cent correct recommendation if it recommends two badges for each user.

TABLE V. RESULTS FOR TOP ONE AND TOP TWO RECOMMENDATION

Title	Dataset Partition		
	2008	2008 - 2009	2008 - 2010
Top 1	0.70	0.63	0.61
Top 2	0.82	0.75	0.74

In future work, we can examine other state-of-the-art algorithms for the recommendation with implicit feedback. We can also use content-based recommendation and combine user's posts, comments, and other features with collaborative filtering recommendations. The fair rate of correct recommendations in this paper shows that this area of work can get better, and really help online social sites toward their goal.

REFERENCES

- [1] P. Sweetser and P. Wyeth, "GameFlow: a model for evaluating player enjoyment in games," *Comput. Entertain. CIE*, vol. 3, no. 3, pp. 3-3, 2005.
- [2] Y. Chou, *Actionable gamification: Beyond points, badges, and leaderboards*. Octalysis Media Fremont, CA, USA, 2015.
- [3] D. R. Flatla, C. Gutwin, L. E. Nacke, S. Bateman, and R. L. Mandryk, "Calibration games: making calibration tasks enjoyable by adding motivating game elements," in *Proceedings of the 24th annual ACM symposium on User interface software and technology*, 2011, pp. 403-412.
- [4] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: defining gamification," in *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*, 2011, pp. 9-15.
- [5] K. Huotari and J. Hamari, "Defining gamification: a service marketing perspective," in *Proceeding of the 16th International Academic MindTrek Conference*, 2012, pp. 17-22.

- [6] J. Hamari, J. Koivisto, and H. Sarsa, "Does gamification work?—a literature review of empirical studies on gamification," in *System Sciences (HICSS)*, 2014 47th Hawaii International Conference on, 2014, pp. 3025–3034.
- [7] S. Deterding, M. Sicart, L. Nacke, K. O'Hara, and D. Dixon, "Gamification. using game-design elements in non-gaming contexts," in *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, 2011, pp. 2425–2428.
- [8] P. Denny, "The effect of virtual achievements on student engagement," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2013, pp. 763–772.
- [9] M. Jakobsson, "The achievement machine: Understanding Xbox 360 achievements in gaming practices," *Game Stud.*, vol. 11, no. 1, pp. 1–22, 2011.
- [10] J. Antin and E. F. Churchill, "Badges in social media: A social psychological perspective," in *CHI 2011 Gamification Workshop Proceedings*, 2011, pp. 1–4.
- [11] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the 10th international conference on World Wide Web*, 2001, pp. 285–295.
- [12] A. Anderson, D. Huttenlocher, J. Kleinberg, and J. Leskovec, "Steering user behavior with badges," in *Proceedings of the 22nd international conference on World Wide Web*, 2013, pp. 95–106.
- [13] M. Montola, T. Nummenmaa, A. Lucero, M. Boberg, and H. Korhonen, "Applying game achievement systems to enhance user experience in a photo sharing service," in *Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era*, 2009, pp. 94–97.
- [14] J. Hamari and V. Eranti, "Framework for designing and evaluating game achievements," *Proc DiGRA 2011 Think Des. Play*, vol. 115, no. 115, pp. 122–134, 2011.
- [15] D. Easley and A. Ghosh, "Incentives, gamification, and game theory: an economic approach to badge design," *ACM Trans. Econ. Comput.*, vol. 4, no. 3, p. 16, 2016.
- [16] S. Lounis, K. Pramatari, and A. Theotokis, "Gamification is all about fun: The role of incentive type and community collaboration," 2014.
- [17] M. Deshpande and G. Karypis, "Item-based top-n recommendation algorithms," *ACM Trans. Inf. Syst. TOIS*, vol. 22, no. 1, pp. 143–177, 2004.
- [18] Y. Hu, Y. Koren, and C. Volinsky, "Collaborative filtering for implicit feedback datasets," in *Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on*, 2008, pp. 263–272.
- [19] Y. Liu, P. Zhao, A. Sun, and C. Miao, "A Boosting Algorithm for Item Recommendation with Implicit Feedback," in *IJCAI*, 2015, vol. 15, pp. 1792–1798.
- [20] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, "Evaluating collaborative filtering recommender systems," *ACM Trans. Inf. Syst. TOIS*, vol. 22, no. 1, pp. 5–53, 2004.
- [21] H. Cavusoglu, Z. Li, and K.-W. Huang, "Can gamification motivate voluntary contributions?: the case of stackoverflow Q&A community," in *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing*, 2015, pp. 171–174.
- [22] D. Movshovitz-Attias, Y. Movshovitz-Attias, P. Steenkiste, and C. Faloutsos, "Analysis of the reputation system and user contributions on a question answering website: Stackoverflow," in *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 2013, pp. 886–893.
- [23] A. Bosu, C. S. Corley, D. Heaton, D. Chatterji, J. C. Carver, and N. A. Kraft, "Building reputation in stackoverflow: an empirical investigation," in *Proceedings of the 10th Working Conference on Mining Software Repositories*, 2013, pp. 89–92.
- [24] C. Stanley and M. D. Byrne, "Predicting tags for stackoverflow posts," in *Proceedings of ICCM*, 2013, vol. 2013.