

Just Rate It!

Gamification as Part of Recommendation

Angelina de C.A. Ziesemer, Luana Müller, and Milene S. Silveira

Faculdade de Informática — PUCRS
Avenida Ipiranga 6681, 90619-900, Porto Alegre - RS, Brazil
{angelina.ziesemer,luana.muller}@acad.pucrs.br,
milene.silveira@pucrs.br

Abstract. In attempt to help users in filtering available products, recommender systems are being used by e-commerce systems to try to predict users' preferences and suggest them new products. Some recommender systems are based in previous ratings and evaluations provided by users to purchased items. When new users or new items join in recommender systems they can suffer by the so called cold-start problem. However, do you rate the products that you bought? This question and other ones were made to 367 participants by an online survey that aims to identify customer profiles and motivations. Also, we investigated user engagement in gamified systems and the effects of tangible and intangible rewards in their behavior. This work presents a theoretical framework that provides basis for defining how gamification can be used to encourage ratings and improve user engagement in tasks that benefit user reputation, item reliability and to overcome cold-start problem.

Keywords: gamification, recommendation, e-commerce.

1 Introduction

E-commerce systems have growth in the quantity of users and content available for searching and due to this, Recommender Systems (RS) become essential to help users find relevant products. The purpose of recommendation algorithms on e-commerce is to filter products according to users' preferences. There are distinct approaches [3, 4, 16, 17, 22] for recommending content and one of the most popular technique is the Collaborative Filtering (CF) [19] that uses ratings/reviews from users on items.

Ratings and reviews are known as explicit feedback [21] and using this technique the user is the one who informs the content relevance. Amazon¹ has used ratings and reviews for years and a well known problem caused by this is that CF can face the cold-start problem. The cold-start problem occurs when recommendations are required for users that have no activity (ratings, reviews) or there are items in the dataset that no one has yet rated [18]. To understand this

¹ <https://www.amazon.com/>

problem we performed a survey with 367 participants. We asked users about their consumer routine, their behavior related to rating/reviews on e-commerce systems and whether users are familiar with gamification [10]. Based on the results we propose a theoretical framework to overcome the cold-star problem using gamification approaches.

The term gamification consists in the use of game design elements in a non-game context to motivate and increase user activity [9], and many services like Foursquare², LinkedIn³, Stackoverflow⁴ have adopted this technique.

We did not find any research related to the use of recommendation with gamification, so this motivated us to develop directions to use the best of the two approaches to engage users in the rating and review tasks using game elements on e-commerce systems.

In the next section we present an overview about Recommender Systems and Gamification. Next, we present the results from the survey performed with 367 participants and after a theoretical framework proposing the use of gamification in order to engage users to rate products on e-commerce systems. Finally, we present the conclusion about the survey and future research prospects.

2 Background

2.1 Recommender Systems

Due to the amount of information available on the Internet, recommender systems have become an efficient alternative for helping users find relevant content and are being widely deployed on the web in many domains including large e-commerce powerhouses such as Amazon [15]. Regardless of the approach, recommender algorithms use information about the users preferences to try recommend items that people with similar tastes and preferences liked in the past [1].

A Likert' five-point response is a popular technique of explicit feedback and it is widely used in CF approaches. This kind of feedback is used internally to measure the similarity among items/products and increase the quality of recommendations [11]. However, explicit feedback requires additional effort from users and sometimes they might not be willing to provide it. Thus, the number of available ratings explicitly, might be too small and can result in a poor recommendations quality [3]. Further, new products in an e-commerce dataset do not have evaluations, which exclude them from the recommendation item-based algorithm and the same occurs with users that are new in the system.

One of the most challenging key questions in recommender systems is: "how to provide recommendations at the initial stage of the system when available data about users, items and ratings is extremely sparse?" [15]. Recently, users are willing to contribute to their community knowledge (a direct reflection from the Web 2.0 and the role of online communities). Still, recent researches are focusing

² <http://foursquare.com>

³ <http://www.linkedin.com/>

⁴ <http://stackoverflow.com/>

on developing techniques that can help to persuade the users to provide more ratings [11].

However, when it is necessary to deliver content to new users or deliver recommendation items that were not rated in the system, one of the recommender systems' problems can appear: the cold-start problem.

Beyond the Cold-Start Problem. Pure CF methods base their recommendations on users' preferences ignoring user and item attributes as demographics and product descriptions. The cold-start problem can occur in the following situations: recommendations for existing items for new users, recommendations for new items for existing users and recommendations for new items for new users [15]. To try overcome these problems, recommender systems try to get users to rate items at the beginning of their profile creation, asking them to answer questions about their preferences or using stereotypes (e.g. elderly people usually enjoy classical music) [14].

Another approach is to use content information to infer similarities from existing items compared to new items. Content recommendations can be made for new items that seem similar to others. However, the cold-start problem is only one of the problems related to new users and new items. In [13] users reported that customer reviews played a role in deciding whether or not to buy a product, but the reviews are not the only factors. Users are also concerned about reviewers' expertise and reputation, and products that have five stars might be attractive, but if it is based only on one or few reviews it is possibly not very interesting. Some systems can rank reviews and products based on the author's reputation, but if a new product is rated by new users some recommendation will possibly fail in quality.

Therefore, user engagement in the rating task is needed to improve recommendation. In the next section we present the gamification definition and the elements that are involved to motivate users.

2.2 Gamification

The term gamification refers to the use of game elements and concepts within systems and applications. It has recently increased and triggered a buzz in the academy, mainly for those trying to better define, understand and explore its effects. One of the major reasons to use gamification approaches is to motivate people to become more involved in an activity, environment or any task that requires user engagement.

According to the Self Determination Theory [6] (SDT), people are intrinsically or extrinsically motivated. Intrinsic Motivation (IM) is related to activities that people do because they have an internal aspiration, such as personal development, or because it is the right thing to do or just because it is enjoyable. On the other hand, Extrinsic Motivation (EM) is related to activities that people do because they have a reward or status and it is all about its value. SDT examines people life goals or aspirations, showing the variations regarding intrinsic versus extrinsic life motivation to perform [8] a task, activities etc.

The goal of implementing game elements in the way that users can interact with a system is to motivate them through rewards. Rewards in a gamified system are generally related to badges, leaderboards, status and reputation, whereas their structure is related to progress, cooperation, feedback, etc. Moreover, Deci and Ryan [7] conceptualize the Cognitive Evaluation Theory (CET) to categorize the type of rewards that can influence behavior [20]. Rewards in CET are categorized as tangible or intangible, expected or unexpected, or contingent rewards. We are particularly interested in tangible vs. intangible rewards and contingent rewards to understand users motivation regarding products evaluation and their behavior:

- Tangible rewards are related to prizes/rewards that are material, real or physical. They are related to money or prizes that people can have.
- Intangible rewards are related to prizes/rewards that are not material. They are game elements that can be used in a gamified system, like users reputation, digital badges or leaderboard position etc.

In the other hand contingent rewards are related to the rewards structures:

- Task non-contingent: users will win this reward every time e.g. they logon.
- Engagement-contingent: users will win when they start a new task.
- Completion-contingent: users will win this reward when they complete a task.
- Performance-contingent: reward based on the quality of the task performed.

3 Case Study and Data Analysis

To analyze users of e-commerce websites, their product evaluation behavior and their use of gamified applications or systems, data was gathered through an online survey. A total of 367 individuals completed the survey, 221 (60%) male and 146 (40%) female, with an average age of 30. From this sample, we used 356 (97%) responses, that represent only those participants who usually buy on e-commerce systems.

The focus of the survey was to identify and compare users behavior on e-commerce systems and understand users motivation for rating products.

The data obtained from the survey is composed by categorical and qualitative variables. We use the Chi-square test and Z-Test of Proportion to gauge the comparative frequency of variables obtained from this survey and also a qualitative analysis from the participants answers.

Next, we present an analysis of the qualitative responses obtained during the survey. The answers show that although users tend to express more interest in tangible rewards, they have other motivations to evaluate products.

3.1 Users' Behavior and the Cold-Start Problem

In a previous section we explained the cold-start problem and its causes. According to [13], those users that intend to buy e.g. a book in the Amazon website,

claim for expertise reviewers. It is possible for Amazon costumers say whether a rating and a review are helpful or not, and this action can increase the reviewers reputation and also improving the quality of recommendation.

However, what are the kind of ratings expected by websites and recommender systems? All kind of rating play a role in recommendation, even those ratings that do not represent high or low level of satisfaction because the item was even the way that the consumer expected.

For the 356 (97%) participants that are used to buy online, 210 (58%) said that they do not usually return to the website for rating products. Here we have the first point: the longer the products wait for ratings, the longer the cold-start problem will persist.

We asked participants an open question with the purpose of understanding why they evaluate products. 119 participants explained when and why they would return to the website to rate a product. 104 respondents said that they usually rate their purchased products when an item does not meet their expectations. However, 64 respondents also said that they rate products for the opposite reason, when it positively exceeds their expectations. This action can difficult the recommender system to recognize the real user feelings about new products that they could possibly purchase. Also, future buyers can be influenced by these opposite reviews.

Only 48 participants reported that they always rate a product regardless of the situation, even when the products are consistent with the description available on the website and other ratings. Also 30 participants stated that they only go to the website to rate products when the website or the seller (on websites like eBay⁵) send to them an email asking for a review. Another reasons that participants pointed for evaluating products were:

- When he/she has being a website user since a long time.
- When his/her opinion is different from other buyers.
- When the system provides to him/her some points or advantages.
- When the website is a C2C⁶.
- When the product presents problem.
- When the customer service was not satisfactory.
- When he/she intends to return to the website to make more purchases.

Product ratings are important to e-commerce and users need to feel that ratings are necessary and indispensable. When this is not established in users minds, it is necessary to develop and improve tools that encourage them to perform this task. A good product rating (not only in the extremists cases) can improve recommendations, and consequently increase sales, users loyalty and their online shopping experience.

⁵ <http://www.ebay.com/>

⁶ Customer-to-customer

3.2 Rewards and Motivations

To try understand users motivation regarding the type of rewards discussed in this paper, we asked participants two similar questions related to tangible and intangible rewards.

First we asked them: “Would you feel motivated to return to an e-commerce website and evaluate a purchased product if this task would help you to improve your reputation or points in the system?”. We analyzed the proportion of participants that care about intangible rewards (IR). The results show that the proportion of users that use gamification and those who not, do not differ significantly ($Z = -0.55, p = 0.57$) as described in Table 1. The participants motivation to rate products if the system give them a reputation or points is not related to the previous use of gamification.

Table 1. Individuals behavior related to tangible and intangible rewards and those who use and not use gamified systems

Motivation	Not Use Gamification	Use Gamification	Overall	Proportion
IR	41 (54%)	161 (58%)	202 (56%)	$Z = -0.55$
TR	71 (93%)	252 (90%)	323 (90%)	$Z = 0.91$

Furthermore, we observed the same behavior for the second question. Participants were asked about their motivation regarding tangible rewards (TR) and the rating/review task. The proportion of users that would return to evaluate products from those that use and do not use gamification are similar ($Z = 0.91, p = 0.36$).

However from the overall participants, 90% said that they could evaluate products if it would give them tangible rewards and only 56% would do this if the rewards were intangible. Then, the motivation for users keep doing a gamified task could not be just related to the rewards it offers, users can have a intrinsic motivation to keep them engaged. In the next section we investigate why users rating products and their motivation for keeping rating items.

3.3 Rating and Motivation

In this section we intend to analyze the motivations for users to rate products. The results shown here are related to qualitative and a quantitative analysis and helped us to evaluate issues related to intrinsic and extrinsic motivation.

Regarding the question “Why would you feel motivated to return in a website and evaluate a purchased product if you received some type of reputation or points in the system?”, 172 participants answered this question. Each answer was read, analyzed and categorized according to the motivations reported by them. Some respondents had more than one motivation, but the answer was categorized according to the strongest one.

Fifty-four users stated that their motivation is related to virtual economy. For them, reputation or points need to provide a discount, benefit or reward. One user reported: *"I would return to the website if it generates points or reputation that could be useful for new acquisitions, also discounts or improvement in payment methods."* Another user stated, *"Indirectly I would be contributing with the seller. Thus, I would like to receive some benefit. Even if the benefit takes a long time to bring me a credit I like the simple fact of knowing that the benefit is mutual."* These are examples of extrinsic motivation for rating products and it is all about virtual currency or virtual economy.

Another 34 respondents said that their mainly motivation is to have their loyalty recognized by the website. These users would like to be recognized as "good buyers" and consequently improve their reputation in the system.

For 28 participants, simply to be encouraged to do something is a motivation. One of them stated: *"Beyond winning points, you are helping other people to know about the product reputation. I believe it would be a motivation"*. Moreover, 11 users reported that they usually rate their purchased products because it is the right to do. For them, gamification will be an additional motivation to keep them doing something that they already do. This behavior is related to intrinsic motivation, in other words, people complete tasks because they think it is important and the right to do.

Moreover, regarding intrinsic motivation, the game feeling and the possibility to compete with friends were one of the motivations reported by 7 respondents. One of them reported: *"the simple fact that it seems like a competition is the necessary motivation to keep me busy performing some task"*.

For three participants the motivation is the loyalty promoted when the user feels stimulated to always return to the same website to do their purchasing. According to one of them, *"the company gives to you a reason to return and buy again"*. The main motivation to use a gamified system is to keep users engaged with the business objectives through the tasks provided by the system. Also, other 35 remaining participants did not know how to explain their motivations.

Moreover, we asked participants if they usually rate products on the internet. Overall, less than a half (41%) participants responded that they rate products in the internet. We used the Pearson Chi-square to gauge the relationship between the use of gamification and the rating task on e-commerce systems. Results show that those who use gamified systems and those who not, tend to rate products with equal frequency ($X^2 = 1.32$, $p = 0.24$). However, 98% of participants answered that they read product reviews online before purchasing an item and more than a half (55%) evaluate reviews left by other users in products that they intend to buy.

According to the results, users have mainly intrinsic motivation for rating products, although they think that tangible rewards could lead them to more frequently rate and review items.

4 Theoretical Framework

Explicit feedback is a common approach for creating recommender systems and there are several works trying to improve recommendation by algorithms to treat rating accuracy [2, 5, 12]. However, ratings and reviews are feedbacks given by humans and it is also necessary to motivate users engagement to complete these tasks and to try to overcome the cold-start problem.

Moreover, individuals are also concerned about the quality and reliability of recommendations that they have received. It was reported by [13] that those users who receive recommendations also analyze the reviews and ratings and they take into account not only the rating but the quantity of users that evaluate the product and their expertise about the subject.

Figure 1 shows the effect of new ratings on new items from new users. Those users that are new in the system have low/no reputation because they do not have too much interaction/ratings in the systems and consequently the recommended item may fail in reliability because it has received a new evaluation from a new user. The same occurs with new items and new users, users can have weak or no recommendations until items have a x number of reviews by “trusted users”. Also, new items reviewed only few times could be considered with low reliability.

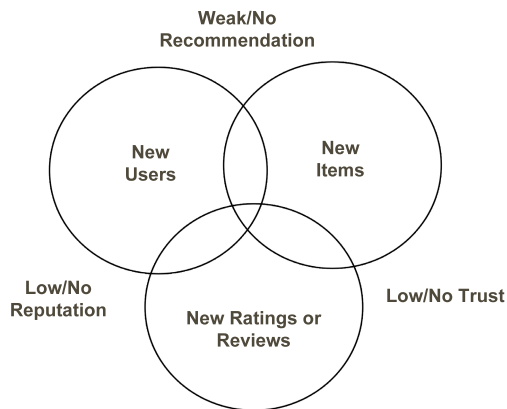


Fig. 1. Beyond the cold-start problem, new items may fail in trust if new users with low reputation rate it

In order to gather as much information as necessary to improve recommendation, recommender systems typically ask users to rate a diverse set of items, especially those users that are new in the system [2]. We propose the use of gamification to improve recommendation, explicit feedback, items and users reputation. Gamification is the way to design systems that motivate people to do things and using the right approaches it could succeed as many success stories recently reported by [20]. The main purpose of using gamification on e-commerce systems is to promote user engagement in those tasks that result in better recommendation for themselves.

To better recommend content to new users and also to improve their reputation, an e-commerce system may be designed to call the user's attention by intrinsic motivation and offering rewards for each task completed by users.

Users that are new in the system can be motivated to complete a task by intrinsic reasons. For example, Netflix⁷ encourages users to evaluate movies that were already seen by them to improve recommendation. However, Netflix has no game elements involved to keep users doing the rating task.

The use of intangible and engagement-contingent rewards can keep users involved with the results as they receive a feedback about their journey through the tasks, using badges and leaderboard to motivate them. Also, tasks that have a Fixed Ratio [20] (every x number of times a user performs a task he/she will receive a reward) can engage users in tasks such as reviews evaluation.

For those users that have high reputation, gamified system can encourage them to evaluate (new) products if they have already used or bought it. For this task, it is interesting to stimulate the users by intrinsic and extrinsic motivation by the use of performance-contingent rewards given users tangible and intangible rewards when they perform good reviews. In addition, other users can evaluate the review and it can also be gamified. Users can be encourage to evaluate reviews from other users to ensure the quality of recommended items and users reputation.

5 Conclusion

This work began by thinking about how gamification could help to fix the cold-start problem present in recommender systems. Through an online survey, we have gathered information in order to understand how users behave on e-commerce websites, when it is required for them to evaluate purchased products.

We asked participants questions about their purchase routine and their motivations to evaluate products. We investigated and presented the results from two perspectives. First, we investigated the reasons why users rate products, results show that most users usually rate their purchased products when it is not according to their expectations (positively and negatively). This action can result in a poor recommendation because the recommender system needs to work only with extreme ratings.

Second, we investigated the motivations for users to return to the website to rate products if the e-commerce website provides them points or reputation. Most part of the participants showed that they are moved by tangible rewards, however, they have intrinsic motivations to perform this action.

We developed a theoretical framework that supports the use of game elements on e-commerce websites to improve product ratings. The framework was developed based on the results of the survey presented here.

We intend to implement this framework and compare the user engagement before and after the use of game elements. We have also observed that users

⁷ <http://www.netflix.com/>

are using gamification without noticing it and this could be an interesting point for future research. This paper provides the initial elements to start a deeper investigation about the subject. We showed how users are motivated, why users usually rate products, how gamification can improve the ratings, and mainly, which problems are related to the lack of (good) ratings.

References

1. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Trans. on Knowl. and Data Eng.* 17(6), 734–749 (2005)
2. Amatriain, X., Pujol, J.M., Tintarev, N., Oliver, N.: Rate it again: increasing recommendation accuracy by user re-rating. In: *Proceedings of the third ACM Conference on Recommender Systems, RecSys 2009*, pp. 173–180. ACM, New York (2009)
3. Balabanović, M., Shoham, Y.: Fab: content-based, collaborative recommendation. *Commun. ACM* 40(3), 66–72 (1997)
4. Bostandjiev, S., O’Donovan, J., Höllerer, T.: Tasteweights: a visual interactive hybrid recommender system. In: *Proceedings of the Sixth ACM Conference on Recommender Systems, RecSys 2012*, pp. 35–42. ACM, New York (2012)
5. Choi, S.M., Han, Y.S.: Identifying representative ratings for a new item in recommendation system. In: *Proceedings of the 7th International Conference on Ubiquitous Information Management and Communication, ICUIMC 2013*, pp. 64:1–64:6. ACM, New York (2013)
6. Deci, E.L., Ryan, R.M.: Self-Determination Theory: A Macrotheory of Human Motivation, Development, and Health. *Canadian Psychology* 49(10), 182–185 (2008)
7. Deci, E.L., Cascio, W.F., Krusell, J.: Cognitive Evaluation Theory and Some Comments on the Calder and Staw Critique. *Journal of Personality and Social Psychology* 31(1), 81–85 (1975)
8. Deci, E.L., Ryan, R.M.: Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology* (25), 54–67 (2000)
9. Deterding, S., Dixon, D., Khaled, R., Nacke, L.: From game design elements to gamefulness: defining “gamification”. In: *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, MindTrek 2011*, pp. 9–15. ACM, New York (2011)
10. Deterding, S., Khaled, R., Nacke, L., Dixon, D.: Gamification: Toward a definition. In: *CHI 2011 Gamification Workshop Proceedings, Vancouver, BC, Canada* (2011)
11. Jannach, D., Zanker, M., Felfernig, A., Friedrich, G.: *Recommender Systems: An Introduction*, 1st edn. Cambridge University Press (September 2010)
12. Kwon, Y.: Improving top-n recommendation techniques using rating variance. In: *Proceedings of the 2008 ACM Conference on Recommender Systems, RecSys 2008*, pp. 307–310. ACM, New York (2008)
13. Leino, J., Rähkä, K.J.: Case amazon: ratings and reviews as part of recommendations. In: *Proceedings of the 2007 ACM Conference on Recommender Systems, RecSys 2007*, pp. 137–140. ACM, New York (2007)
14. Masthoff, J.: Group recommender systems: Combining individual models. In: *Recommender Systems Handbook*, pp. 677–702 (2011)
15. Park, S.T., Chu, W.: Pairwise preference regression for cold-start recommendation. In: *Proceedings of the Third ACM Conference on Recommender Systems, RecSys 2009*, pp. 21–28. ACM, New York (2009)

16. Rashid, A.M., Albert, I., Cosley, D., Lam, S.K., McNee, S.M., Konstan, J.A., Riedl, J.: Getting to know you: learning new user preferences in recommender systems. In: Proceedings of the 7th International Conference on Intelligent User Interfaces, IUI 2002, pp. 127–134. ACM, New York (2002)
17. Said, A., Fields, B., Jain, B.J., Albayrak, S.: User-centric evaluation of a k-furthest neighbor collaborative filtering recommender algorithm. In: Proceedings of the 2013 Conference on Computer Supported Cooperative Work, CSCW 2013, pp. 1399–1408. ACM, New York (2013)
18. Schein, A.I., Popescul, A., Ungar, L.H., Pennock, D.M.: Methods and metrics for cold-start recommendations. In: Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2002, pp. 253–260. ACM, New York (2002)
19. Su, X., Khoshgoftaar, T.M.: A survey of collaborative filtering techniques. *Adv. in Artif. Intell.*, 4:2–4:2 (January 2009)
20. Werbach, K., Hunter, D.: *For the Win: How Game Thinking Can Revolutionize Your Business*. Wharton Digital Press (2012)
21. Ziesemer, A., Oliveira, J.: How to know what do you want? a survey of recommender systems and the next generation. In: Proceedings of the Eighth Brazilian Symposium on Collaborative Systems, SBSC 2011, pp. 104–111 (October 2011)
22. de C.A. Ziesemer, A., de Oliveira, J.B.S.: Keep querying and tag on: Collaborative folksonomy using model-based recommendation. In: Antunes, P., Gerosa, M.A., Sylvester, A., Vassileva, J., de Vreede, G.-J. (eds.) CRIWG 2013. LNCS, vol. 8224, pp. 10–17. Springer, Heidelberg (2013)