# Applying the HYRIWYG incentive mechanism in a Recommender System

*Ana Cristina Bicharra Garcia, Leandro Neumann Ciuffo*
*Instituto de Computação - Universidade Federal Fluminense (UFF)*
*Rua Passo da Pátria, 156 Bloco E Sala 350 - 24210-240 - Niterói – RJ - Brazil.*
*bicharra@dcc.uff.br,lciuffo@ic.uff.br*

## Abstract

*The use of Recommender Systems (RS) has become common in several e-commerce websites. Collaborative Filtering is one of the most popular RS techniques. It infers an user's predilections based on similarity with other users. However, to work correctly, the RS depends on the evaluator's participation with truly opinions. The HYRIWYG mechanism proposes an incentive model to motivate users to contribute with a high number of opinions and, more important, guaranteeing the truthfulness of the information. In this paper, we present HYRIWYG mechanism and discuss it using an empirical study in the movie evaluation domain. Initial results highlight the great potential benefits of using this kind of incentive mechanism.*

## 1. Introduction

In several daily situations we need to make one choice among a lot of available options, like to decide which DVD to rent in a video store, which city to travel to your vacations or which website to visit among the thousands listed in a Google search.

Frequently we lose a lot of time filtering and selecting which better information, product or service will attend our expectations. For example, there are hundreds of restaurants in a city, each one with dozens of plates. Guides and specialized magazines are just some examples of attempts to filter information, helping us to save some time.

Besides, each person have an individual preference. A certain product can please most of its buyers, but it doesn't satisfy all of them. And when we analyzed the people's taste for movies or songs, the individual preferences are much more heterogeneous. A same movie can be loved by a person and hated by another.

Recommender Systems (RS) are softwares that can suggest tailored items (products or services) to consumers in an e-commerce website, for example.

One of the most popular techniques used to develop a RS is the Collaborative Filtering approach [2;6]. This technique essentially automates the process of "word-of-mouth" recommendations. Items are recommended to an user based on the evaluations supplied by other users with similar taste.

This technique depends on the participation of a great number of users to tuning the RS. This way, is essential to motivate the users to contribute supplying its evaluations on the available items in the System.

Thus, two potential problems appears: (i) how to motivate the users to spend part of its time to supply its opinions and evaluations for the System? There is no apparent reason for the user to contribute with the tuning of RS, once that it will benefit just other users; (ii) how to guarantee that the supplied opinions will be true? In order to promote a certain product, a group of users can lie about the quality of it. Or a hurried user can supply random evaluations for the RS.

One proposal for the inhibition of the occurrence of those problems is the use of the incentive mechanism HYRIWYG (How You Rate Influences What You Get) [3;4], that encourages the evaluators to volunteer their true opinion, by compensating users with gifts and rewards in agreement with each user's profile.

We propose in this paper a case study to verify the effectiveness of the incentive mechanism. So, we implemented a simple Movie Recommender System to apply the rules suggested by HYRIWYG and published it in a website specialized in movies.

## 2. Recommender Systems

Recommender Systems are computational agents capable to manipulate a set of data from preferences and behaviors of users when they interact with the System. These data are used to recommend items to each registered user, and it can be obtained, for example, through the forms filling in websites. Thus, the user can declare its preferences and supply evaluations about items that he or she had already tried. The gathering of those data is used to configure the user's profile and are vital for tuning the RS. In general, the larger the number of evaluations supplied, the more correct recommendations will be.

To manipulate a great amount of data, we use one of the earliest and most successful recommender technologies: Collaborative Filtering, also called Social Filtering [6;10].

Collaborative Filtering considers the following hypothesis: users who agreed in the past will probably agree again in the future. For example, imagine that an active user want to receive a product recommendation. The RS will look for the nearest neighbors, i.e., the users whose taste (profile) is most similar to the taste of active user. Then, the RS will recommend to active user products that other users with the same taste had already enjoyed.

The success of this technique depends strongly on the users' interaction with the System. If the users not motivated to supply honest evaluations about the items, it doesn't matter which technique was used in the RS development.

In the next section, we present a proposal to motivate users to contribute with truthful evaluations, essential for the good operation of RS.

## 3. How You Rate Influences What You Get

Usually people want to use the Recommender System to receive good recommendations, but they don't like to waste time supplying their evaluations to adjust the RS. This kind of behavior is harmful to RS, making the System give either bad or doubtful recommendations.

Nowadays, there are several RS in the world wide web that offer different kind of recommendations, like songs, restaurants, news, movies or even joke recommendations, but they are most commonly used in e-commerce websites. In that kind of application, an RS can become an important marketing tool, used to increase the consumers' loyalty and, consequently, the store profit through good recommendations of products or services. On the other hand, an RS that supplies bad recommendations can make the costumers lose the credibility on the website, decreasing the sales.

The HYRIWYG mechanism is a proposal to motivate users to contribute with truthful evaluations. Consequently, the RS effectiveness increases and its inference mechanism is better adjusted [3]. It is important to stand out that HYRIWYG can be used in any RS that generates personalized recommendations, but it is not applicable to RS that generate general recommendations.

### 3.1. Incentive mechanism operation

The user can interact with RS by two different ways and in different times: first supplying its evaluations on the available items and later receiving the recommendations from RS.

In general, users don't have incentives to supply evaluations to the RS. A particular individual evaluation benefits others, but not directly the user who provides the evaluation.

A good RS tuning depends on the collection of each user's individual evaluations. Thus, people need to be motivated to increase the number of evaluations on the RS, and consequently, improve the quality of the recommendations. So, a user should be rewarded by supplying information that will benefit all the participants of the RS.

The HYRIWYG propose a social function that provides incentives for each individual to tune the RS for his or her profile [3]. The incentive can be points, gifts, coupons or any other reward according to the company's marketing strategy. The larger the user's amount of evaluations, the more prizes he can receive.

However, the incentives don't guarantee that the evaluations are true. The RS can reward a user that rated a lot of items, but it was not honest in its evaluations. To avoid this problem, the solution proposal is to use the RS to choose the prize that will be offered to each user.

Analyzing each user's profile, the System can infer which reward can be recommended for each person. Thus, if a user lie in order to amass incentives, he or she can receive a non-desired prize.

This is a simple rewarding system that stimulates people to say the truth, according to each individual perspective, since the reward will be as good as the quality of the RS [3].

## 4. The developed RS

The RS developed for this work uses the memory-based collaborative filtering algorithm [6;7]. This algorithm compares the active user profile, i.e., the profile of user who want to receive a movie recommendation with all users registered in the System. Once identified the most similar user (the nearest neighbor), the RS looks for the movies well rated (and supposedly watched) by the neighbor. The movies found and that haven’t been watched by the active user yet will be recommended.

To illustrate the functioning of RS, consider table 1. The table represents the matrix of rating of five different movies. The user's profile is formed by each line in the matrix. The cells marked with "N" indicate that the movie has not been watched by the user yet.

**Table 1. A simple matrix of rating**

| | Aliens | Matrix | Jaws | Patton | Troy |
|---|---|---|---|---|---|
| **John** | N | 2 | 2 | 5 | N |
| **Rick** | 6 | 4 | 4 | 2 | N |
| **Jully** | 5 | 4 | 3 | 2 | 5 |
| **Maria** | N | N | 5 | 4 | 4 |
| **Paul** | 4 | 5 | 6 | 3 | 5 |

To recommend a movie for Rick, the system has to seek which user have the most similar profile, i.e., which

user liked the same movies as Rick. If we just take a look in the table, we can notice that Jully is the most similar user. To compute this, we use the Pearson coefficient to measure the similarity among users [8]. The correlation between user **a** and another user **b** is calculated by:

$$r(a,b) = \frac{\sum (N_{a,f} - \overline{N}_a)(N_{b,f} - \overline{N}_b)}{\sqrt{\sum (N_{a,f} - \overline{N}_a)^2 \sum (N_{b,f} - \overline{N}_b)^2}} \quad (1)$$

where: $f \in$ Movies watched by both users: ***a*** and ***b***;

$N_{i,f}$ is the rating of user ***i*** on movie ***f***

$\overline{N}_i$ is the mean value of ratings by user ***i***

Applying the correlation calculus for each pair of users, the following matrix is generated.

**Table 2. Matrix of similarity**

| | John | Rick | Jully | Maria | Paul |
|---|---|---|---|---|---|
| **John** | **1** | - 0,816 | - 0,618 | - 0,8 | - 0.943 |
| **Rick** | - 0,816 | **1** | + 0,916 | + 0,447 | + 0,314 |
| **Jully** | - 0,618 | + 0,916 | **1** | - 0,176 | + 0,269 |
| **Maria** | - 0,8 | + 0,447 | - 0,176 | **1** | + 0,754 |
| **Paul** | - 0,943 | + 0,314 | + 0,269 | + 0,754 | **1** |

The value of correlation degree varies from +1 (identical profiles) to -1 (opposed profiles). The closer calculated value is to those limits, the more accurate will be the inference generated by RS.

Analyzing the matrix once again, we can notice that John and Paul have opposite tastes, i.e., John like the movies that Paul doesn't like. The RS also can make use of these information to recommend movies to both. But the System cannot generate good recommendations if the correlation degree is close to zero, as between Maria and Jully. After calculating the matrix, to determine which is the profile more similar to Rick's profile, the algorithm just seeks the highest number in its respective line or column in the matrix. In this example, Troy will be recommended to Rick because it is the movie that he hasn't watched and it was well rated by Jully.

## 5. Case study

In order to test empirically the benefits of the application of the HYRIWIG in Recommender Systems, we developed an RS to recommend movies to Cinefilia visitors. Cinefilia is a non-professional website specialized in movies [5] from Juiz de Fora, a medium-sized Brazilian city.

The site visitors are invited to try the movie recommender system by a full-banner in the home page. When it is clicked, a page with basic information about the RS is loaded. To receive the recommendations, the user needs to register by filling out a small electronic form. In every RS user logon, a list of movies is showed to user's evaluation. To all, information about 300 movies is stored in our database, from classics to the new releases.

As well as in the MovieLens site [9], the user needs to rate at least 15 movies so that the system will be able to generate recommendations. The rating varies from 0 to 6 for each movie watched by user. Also there is the option "haven't seen" to indicate the movies that he or she hasn't watched yet.

Our dataset consists of 201 users from Juiz de Fora and 170 users from other cities. The total number of ratings (from 0 to 6) is 34031. All those data are automatically stored in a MySQL database in a web server. Analyzing the database queries and the access statistics, we present some data obtained during the follow phases of the experiment:

Phases 1a and 1b: the site visitors were invited to try the RS altruistically. No incentives were granted for the user's participation.

Phase 2: gifts were offered to the users that evaluated the largest number of movies. Every 12 days, the top 10 users that rated the highest number of movies received gifts.

Phase 3: free movie tickets were given for each 50 movies rated by users. For example, if a user rates 100 movies, it would win 2 tickets. To accomplish this experiment, we made a partnership with a small movie theater in Juiz de Fora. The movie theater has two screening rooms and it shows an average of four different movies a week.

Phase 4: free movie rents were given for each 40 movies rated. This time, we made a partnership with five movie rental store in the same city.

Phase 5: this phase is very similar to the previous one, the difference is that in this phase it is not allowed to the user to choose the movie that he or she wants to watch. The RS infers which films can be recommended to the user. This time, when the user has to return the video, he or she is invited to rate the movie recommended by RS.

When the user is motivated to contribute with RS by receiving gifts in exchange for his/her participation, the number of evaluations increases and, consequently, and so does the number of lies. We hope that the occurrence of those lies shall be reduced in the last phase, when the mechanism HYRIWYG is fully applied in the RS.

### 5.1. Analysis of the results

Comparing the evaluations average supplied by new users during the first four phases of the experiment, we observed a considerable increase of that number especially in the third and fourth phases, in consequence of the granted incentive (figure 1).

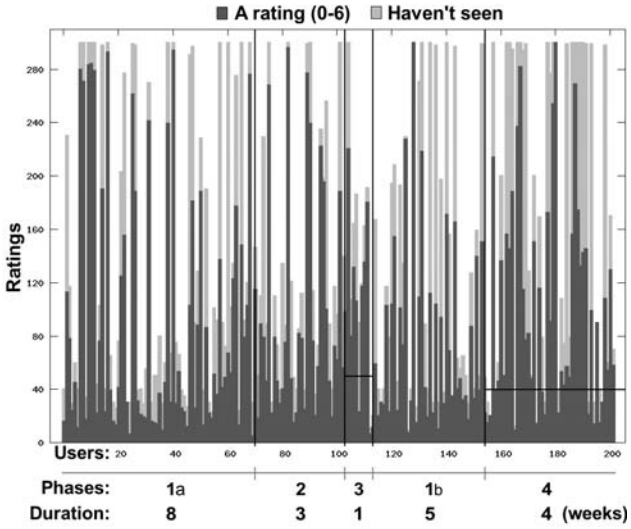


**Figure 1. Users's Profile from Juiz de Fora**

The figure shows only the reviews of Juiz de Fora users, as users from other cities don't get prizes. The horizontal lines show the minimum number of evaluations necessary to win one prize. We can notice a rating increase during phases 3 and 4.

The number of evaluations received during phase 2 in relation to phase 1a is almost the same, but a curious behavior was observed: After the first gifts distribution, the users that didn't get any prizes supplied its evaluations to the system with less frequency than the winning users. These, motivated by the received prize, startd to access regularly the RS.

In the three occasions of gifts distribution during that phase, when selecting the 15 users that rated more movies, were just listed the users contemplated recently in the previous gift distribution or the new registered users, i.e., just winners continued contributing with the RS.

Analyzing the users' profiles, some behaviors that suggest false evaluations were detected, as for example, users that supply the same rating for all available movies, or when the evaluation of the movie is accomplished in a short time interval.

As we expected, the incentive granted in exchange of evaluations increased considerably the number of information stored in the RS. However, we don't know if the supplied information is true. At the end of the phase 1a, when we asked the users' opinion about the recommendations, 67,47% declared to have liked the recommendations given by the system. But during the phase 3 period, just 34,92% declared satisfaction with the recommendation received.

Until the deadline for submission of this paper, we had not begun the phase 5 data collection yet. In the next section we describe the next steps and the expected results.

## 6. Expected results and conclusions

The interest in the development of Recommender Systems has been growing, mainly in e-commerce websites. Many works are being developed in the academy, most of them presents new algorithms or new information filtering techniques. However, to motivate users to contribute with truthful evaluations is beyond the technology used in the RS development.

After the conclusion of the five proposed phases, we expected a reduction of behaviors that suggest false evaluations, maintaining the high number of supplied evaluations. An increase of the number of users satisfied with the recommendations received is also expected.

The analysis of the all obtained data will be presented in future papers. But observing the results obtained so far, we believe that we are heading to attest empirically the efficiency of HYRIWYG in to reduce the number of dishonest ratings and to increase the volume of users' participation in Recommender Systems, vital for the good operation of this kind of application.

## 7. References

[1] P. RESNICK, R. ZECKHAUSER, E. FRIEDMAN and K. KUWABARA, "Reputation Systems: Facilitating Trust in Internet Interactions", Communications of the ACM, 2000.

[2] P. RESNICK, N. IACOVOU, M. SUCHAK, P. BERGSTROM and J. RIEDL, "GroupLens: an open architecture for collaborative filtering of netnews", Proceedings of the 1994 Conference on Computer Supported Cooperative Work (CSCW'94), Chapel Hill, NC, EUA. ACM Press, P. 175 - 186.

[3] M. EKSTRÖM, A.C.B. GARCIA and H. BJÖRNSSON, "Rewarding honest ratings thought personalized recommendations in eletronic commerce", International Journal of Eletronic Business (IJEB), 2004.

[4] M. EKSTRÖM, A.C.B. GARCIA and H. BJÖRNSSON, "HYRIWYG: Leveraging Personalization to Elicit Honest Recommendations", Proceedings of the 5th ACM conference on Electronic commerce, New York, NY, EUA. ACM Press, 2004.

[5] L.N. CIUFFO, "Cinefilia", *http://www.canalcinefilia.com.br*

[6] R. CÖSTER, "The architecture and implementation of a system for collaborative and content-based filtering", Technical Report, Stockholm University, 2002.

[7] K. YU, X. XU, M. ESTER and H. Kriegel, "Selecting relevant instances for efficient and accurate collaborative filtering", Proceedings of the 10th International Conference on Information and knowledge Management (CIKM'01), Atlanta, Georgia, USA. ACM Press, P. 239 - 246.

[8] M.J. PAZZANI, "A Framework for Collaborative, Content-based and Demographic Filtering", University of California, Irvine, CA, USA, 1999.

[9] GROUPLENS RESEARCH. "MovieLens", University of Minnesota, USA. *http://movielens.umn.edu*