Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Automatic Control Laboratory

Semester Project

Analysis of the Closed-Loop between Recommender Systems and Opinion Dynamics

Jules Authier

July 3, 2022

Advisors
Prof. Dr. Florian Dörfler
Dr. Nicolò Pagan
Nicolas Lanzetti

Chapter 1

Abstract

In online platforms, recommender systems (RSs) are responsible for suggesting user-personalized content with the objective of maximizing user's engagement, i.e., the time spent on the platform, likes, shares, etc. RSs adapt their suggestions to the reaction of the users, who are in turn affected by the recommended content. Practically, a RS tries to estimate the opinion of the users by looking at their reactions to previously sent suggestions and uses this estimate to suggest personalized content. Looked through the lenses of control theory, this process describes a feedback loop in which a RS acts as a controller of the users' opinion. Here, we statistically study the effect of the described closed-loop system on the opinion of a population under different modeling assumptions and evaluate the importance of prior knowledge about the users and the suggestions on the closed-loop system. Interestingly, our experiments show large deviations of the users' opinion in this set-up as well as the importance of prior knowledge for the RS performance.

Contents

List of Figures

Chapter 2

Introduction

Opinion Dynamics is a well-established field of research that models the individual opinion formation under the influence of their peers. To do so, it considers a set of agents with individual opinion x_i that evolves according to a linear dynamic $x(t+1) = A \cdot x(t)$, where A is the adjacency matrix of the graph that connects the agents and A_{ij} denotes the strength of the influence of user j on user i. However, in a real-world setting, agents are influenced by exogenous factors such as advertising and influencers. Nowadays, most of the above-mentioned interactions take place online, on platforms such as Facebook, Instagram, Twitter, or YouTube. On those platforms, a RS tries to maximize the user's engagement and chooses the content of the user's feed accordingly. Unlike old-generation sources of information, the content provided by RS of online platforms is personalized. Such recommendations are generated in two steps: first, the RS tries to track the opinion of the user, then it chooses the next recommendation to maximize the user's engagement. However, the position of the recommendations is not always known, in this case, the RS needs to evaluate the recommendations' position and to track the users' opinion.

On online platforms, personalization improves the user experience, but it may lead to so-called "filter bubbles" that isolate users into their own ideologies, leading to polarization and an increase of misinformation [1–3].

In a prior paper [1], the feedback loop between one user's opinion and the RS of an online platform has been analytically studied. In their approach, a user with a scalar opinion between -1 and 1 receives opposite binary news items from a RS which influence the user's opinion. The user's engagement is measured by its click-through rate and the aim of the RS is to maximize it. They showed that the RS produces a shift of opinions towards more extreme positions, inducing a polarization effect, and that mitigating the impact of the RS on the opinions reduces the user engagement. Thus, they highlight the existence of a trade-off between the engagement of the user and the polarization of its opinion.

The interaction between users on social platforms has been studied in [4], the users adopt opinions as function of the viewpoints they are exposed to on the platform. An algorithmic filtering mechanism selects the opinions of the friends of the user who are presented in the user timeline. They found out that in a balanced bipartisan set-up, the algorithmic filtering is not able to change the status quo, however, the formation of echo chambers of like-minded individuals is observed. Alternatively, they showed that if the algorithmic filtering is biased, the population's opinion moves towards the bias, showing that the opinion of a population can be manipulated by algorithmic personalization methods.

The natural extension of these two works is: a) to study the closed-loop between the opinions of a large set of users and a RS and b) to extend the range of opinions and recommendations from a binary set to the real numbers.

Contributions

The main contributions of the present report are:

- A new closed-loop model describing the interaction between multiple users with homogeneous opinion dynamics and a recommender system,
- The development of two recommender systems: a data-based and a model-based,
- Extensive simulations of the new set-up with a specific study of the variation of the dynamics and of the prior knowledge available.

We showed that the closed-loop interaction between users and a RS may lead to large deviation of the users' opinion and we highlighted the importance of the prior knowledge about the users and the recommendations for the RS.

Organization

The remainder of this paper is organized as follows: Chapter 3 describes the problem set-up and the models used. The main results are presented and discussed in Chapter 4, followed by a conclusion in Chapter 5.

Chapter 3

Formulation

We are interested in the interaction between users of an online platform, with a given opinion dynamic and initial opinion, and the RS of the online platform. The users receive individually tailored news articles from the RS and react to these articles with rewards, as illustrated in Figure 3.1. The rewards represent the liking of the users for the news article they received from the RS.

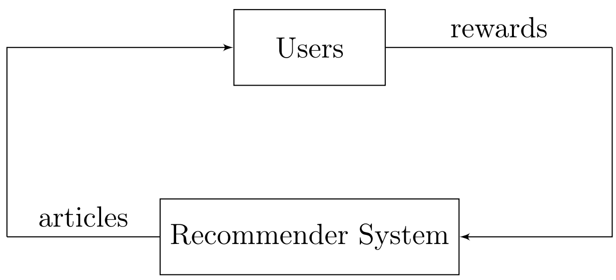


Figure 3.1: Closed-loop interaction between users and a RS.

We consider a set of user $\mathcal{U} = \{1, ..., n\}$ and a set of articles $\mathcal{A} = \{1, ..., m\}$, with $m, n \in \mathbb{N}_*$.

3.1 User Model

3.1.1 Opinion Dynamics

The model described here has been introduced in [1].
Each user $i \in \mathcal{U}$ has a scalar opinion that evolves in discrete time, $o_i \in \mathbb{R}$, as well as a time independent prejudice, $o_i^0 \in \mathbb{R}$, that encodes influences external to the RS, such as family or friends. The prejudice determines the initial opinion of the user, i.e. $o_i(0) = o_i^0$. Moreover, the user receives at each time-step an article $j \in \mathcal{A}$ from a recommender system, this article has position $p_j \in \mathbb{R}$ and influences the future opinion of the user.
The evolution of the opinion of the user i is described by:

$$o_i(t+1) = \alpha o_i^0(t) + \beta o_i(t) + \gamma p_j(t), \tag{3.1}$$

where $\alpha, \beta, \gamma \in \mathbb{R}^+$ and $\alpha + \beta + \gamma = 1$.
The weights α, β, and γ describe the importance of external influences, the user's previous opinion, and the articles received from an online source, respectively, in the evolution of the user's opinion. We assume the weights α, β, and γ are identical for all users but the model can readily be extended to non-identical weights in order to have a more realistic model of the users.

Note that the user's opinion is only influenced by the articles received and not by the other users' opinion as traditionally the case in opinion dynamics models. This allows us to focus on the influence of the RS on the users' opinion.

In this Section, we extended the model introduced in [1] from a bounded interval of possible opinions between -1 and 1 and binary position of the articles to opinions and positions belonging to the real numbers.

3.1.2 Reward

When receiving an article $j \in \mathcal{A}$, the user $i \in \mathcal{U}$ reacts with an output signal $r_i \in \mathbb{R}$ that represents the liking of the user for the article received, i.e., likes, clicks or share. This output signal is viewed as a reward by the RS and is defined by:

$$r_i(t) = \exp\left(-4(o_i(t) - p_j(t))^2\right) + \epsilon, \qquad \epsilon \sim \mathcal{N}(0, 0.2) \tag{3.2}$$

where $\mathcal{N}(0, 0.2)$ is a Gaussian distribution with zero-mean and a standard deviation of 0.2.

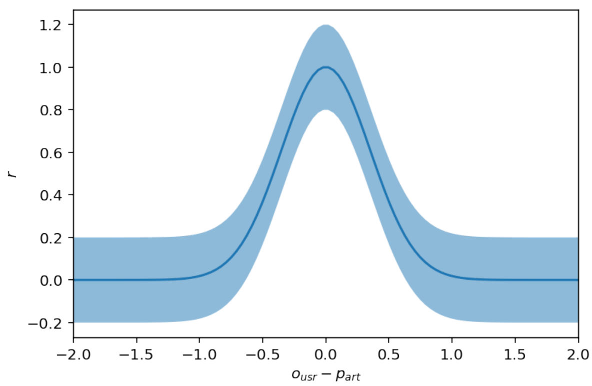


Figure 3.2: Reward function.

The reaction of the user is a function of the distance between the current opinion of the user and the article's position, the smaller the distance the higher the expected value, this can be observed in Figure 3.2. Such behavior is described as confirmation bias [5, 6], i.e. the general tendency for people to interpret evidence in ways that confirm their existing beliefs. Note that Equation 3.2 is not deterministic, the randomness represents the inconsistency of the users' reaction to articles sent by the RS and is used to make the reward more realistic. The non-deterministic nature of the reward is illustrated in Figure 3.2 by the shaded area representing the standard deviation of the reward function. Furthermore, the -4 factor in Equation 3.2 determines the width of the peak of the reward function, i.e., the liking of users for articles with positions far from their opinion. A bigger factor (but still negative) would model more tolerant users and a smaller factor users more reluctant to articles with positions different from their opinion.

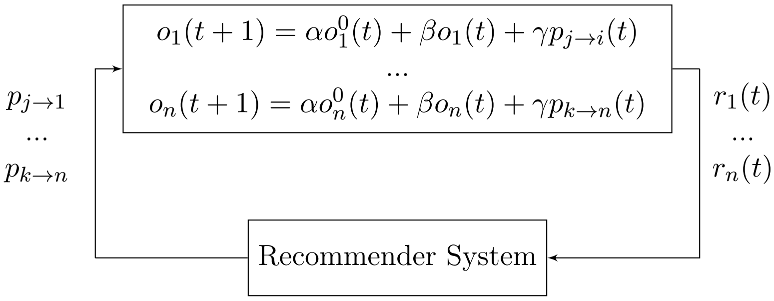


Figure 3.3: Multi-user closed-loop model.

The rewards and opinion dynamics described above allow us to describe the closed-loop of Figure 3.1 more precisely. The RS sends an article $j \in \mathcal{A}$ with position p_j to an user $i \in \mathcal{U}$, written $p_{j \to i}$, the user responds with a reward $r_i(t)$ and its opinion is influenced by $p_{j \to i}$. This happens for all the users, as illustrated in Figure 3.3.

3.2 Recommender Systems

Here, we assume that a Recommender System has the purpose to maximize the users' engagement, represented by the rewards. This is one simple, but yet realistic, way of modeling the objective of a RS. To maximize the rewards, the RS adapts the articles sent to each user according to the rewards collected in the past. In the case of a scenario with the set of users $\mathcal{U}$ over N time-steps, the objective of the RS is:

$$\max . \sum_{h=0}^{n} \sum_{i=0}^{N} r_h(i) \tag{3.3}$$

We analyzed two different RSs:

- the utility matrix [7], which is a data-driven RS relying on the rewards received from the users,
- the learning approach, which is a model-based approach relying on previous knowledge about the users and the articles.

These two RSs observe individual users' reaction to articles to choose the next recommendations. Additionally, one could think about using collaborative filtering techniques in which the suggestions to a given user depends also on the preferences of a similar user.

3.2.1 Utility Matrix

The utility matrix (UM), represented by a matrix $U \in \mathbb{R}^{n \times m}$, stores a discounted average of the rewards of each user in $\mathcal{U}$ for each article in $\mathcal{A}$. At initialization, the UM is a matrix of zeros, representing the absence of knowledge about the preferences of the users. The UM is used and populated in the following way:

1. At time $t \geq 0$, the RS sends an article j to user i,
2. The user i reacts with a reward $r_i(t)$,

3. $U_{i,j}$ is updated according to:

$$U_{i,j} = \begin{cases} r_i(t), & \text{if } U_{i,j} = 0, \\ \delta \cdot U_{i,j} + (1-\delta) \cdot r_i(t), & \text{otherwise.} \end{cases} \qquad 0 \leq \delta \leq 1 \text{ is the discount factor;}$$

   $U_{i,j}$ stores a discounted average of the rewards of user i when receiving article j,

4. the next recommendation for user i is:

$$\begin{cases} \text{argmax}_{1 \leq j \leq m} U_{i,j} & \text{with probability } 1-\zeta \\ \text{a random article in } \mathcal{A} & \text{with probability } \zeta \end{cases} \qquad 0 \leq \zeta \leq 1 \text{ is the exploration rate;}$$

5. go back to 2.

This is made for all users in $\mathcal{U}$. In our set-up $\delta = 0.7$ and $\zeta = 0.05$.
Notice that the first article is chosen randomly and that the exploration rate ζ is used to populate the matrix, i.e., explore unknown areas of the UM. When not exploring, the UM sends the article that, to the best of its knowledge, maximizes the rewards the RS will receive from the users. The UM is purely data-based: no assumption is made about the users, the rewards, or the articles. However, it may take time to find the users' favorite articles, and populating the matrix may be costly in terms of maximization of the user's engagement. Moreover, in a more realistic scenario, if the UM takes too long to find the favorite articles of some users, they may simply stop using the online platform.

3.2.2 Learning Approach

In the learning approach (LA), the RS uses prior knowledge about the users' opinion and the position of the articles to optimize the users' engagement. To do so, the RS stores a belief of the users' opinion $\tilde{\mathbf{o}} \in \mathbb{R}^n$ and a belief of the articles' position $\tilde{\mathbf{p}} \in \mathbb{R}^m$. These two beliefs are updated at each time step in order to track the users' opinion and to improve the estimate of the articles' position. At each time step, the following steps are followed:

1. The user i receives an article j,

$$j = \underset{1 \leq j \leq m}{\text{argmin}} \, \|\tilde{\mathbf{o}}_i - \tilde{\mathbf{p}}_j\|_2$$

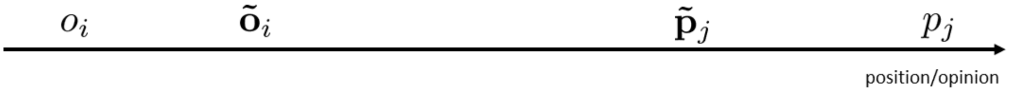
2. The RS estimates the reward it would receive from the user i,

$$\tilde{\mathbf{r}}_{\mathbf{i}} = \exp\left(-4(\tilde{\mathbf{o}}_i - \tilde{\mathbf{p}}_j)^2\right)$$

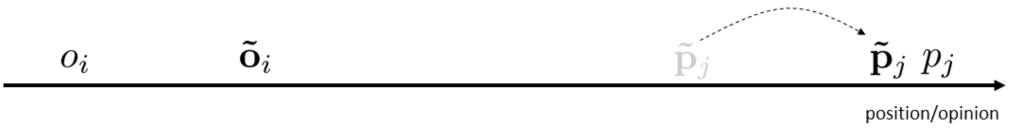
3. if $r_i > \tilde{\mathbf{r}}_{\mathbf{i}}$:
      bring $\tilde{\mathbf{o}}_i$ and $\tilde{\mathbf{p}}_j$ closer together
   else $r_i \leq \tilde{\mathbf{r}}_{\mathbf{i}}$:
      split $\tilde{\mathbf{o}}_i$ and $\tilde{\mathbf{p}}_j$ further away

4. go back to 1.

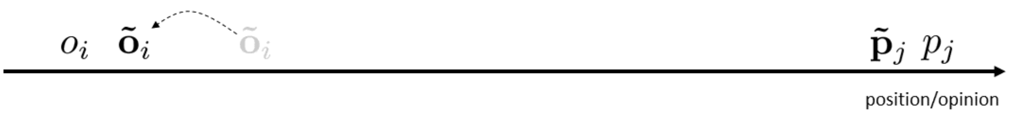
This is made for all users $i \in \mathcal{U}$.
In step 3., the beliefs of the opinions and positions are not updated synchronously. At one time-step, the opinion beliefs are updated and at the next, the position beliefs are updated. Moreover, the change of beliefs are made by step-size proportional to the distance between $\tilde{\mathbf{o}}_i$ and $\tilde{\mathbf{p}}_j$. In our set-up, we choose a step size of 10% of the distance.



(a) Initial situation.



(b) Update of the position of the article.



(c) Update of the opinion of the user.

Figure 3.4: Example of the Learning Approach.

In Figure 3.4, two iterations of the learning approach are displayed. The initial situation is illustrated in Figure 3.4a, the update of the believed position of article j is shown in Figure 3.4b and the update of the believed opinion of user i occurring in the next iteration of the LA is shown in Figure 3.4c. We see that both beliefs get closer to their real values.
However, note that the algorithm described in this Section does not always lead to the convergence of the beliefs to their correct values. If a believed opinion and a believed position are at the correct distance from each other but shifted towards the left or the right of the opinion/position space, the algorithm will not be able to detect it. This is a fundamental limitation of the model and it comes from the fact that the reward is a function of the distance between a user's opinion and an article's position and does not take their absolute position or opinion into account.

3.3 Performance Measurement

Now that the opinion dynamics of the users, as well as the models of the RS, have been defined, we need metrics to measure the performance of the different RS. We are interested in the performance of an RS in two ways, the first one is how well the RS maximizes the rewards of the users and the second is the impact of the RS on the society, i.e. the importance of the shift between the initial and the final opinion of the users. These two metrics allow us to compare different RS and opinion dynamics parameters objectively and to highlight some trade-off between the societal impact and the reward maximization of an RS.

3.3.1 Rewards Maximization

At each time-step, the highest expected reward of each user is 1.

$$\max_{j \in \mathcal{A}} \ \mathbb{E}[r_i(t)] = 1 \qquad \forall i \in \mathcal{U},\ 0 \leq t \leq N$$

The reward maximization (RM) performance is calculated as the percentage of the highest possible total expected reward, i.e.:

$$RM = \frac{\sum_{i=0}^{n}\sum_{t=0}^{N} r_i(t)}{\max\sum_{h=0}^{n}\sum_{i=0}^{N}\max_{j\in\mathcal{A}}\mathbb{E}_\epsilon[r_i(t)]} = \frac{\sum_{i=0}^{n}\sum_{t=0}^{N} r_i(t)}{n\cdot N}\ [\%]$$

3.3.2 Societal Performance

In order to measure the similarity between the initial and final opinion of the users, we use the Pearson correlation coefficient [8] defined as:

$$p_{\text{corr}} = \frac{\text{Cov}(o_{\text{init}}, o_{\text{final}})}{\sigma_{o_{\text{init}}}\sigma_{o_{\text{final}}}} \qquad -1 \leq p_{\text{corr}} \leq 1$$

Where Cov denotes the cross-correlation of two vectors.
The Pearson correlation coefficient is a normalized measurement of the linear correlation between the variables. However, deviations of the mean are not considered in this measurement. In order to have a more complete picture of the shift of the users' opinion, we also consider the percentage of users whose deviation from their initial opinion to final opinion is smaller than a given threshold.

Chapter 4

Results

In the following sections, we analyze the performance of the two different RSs, described in Section 3.2, in closed-loop with users with two different opinion dynamics. We looked at "slow" opinion dynamics, described by $o_{\mathrm{usr}}(t+1) = 0.2 \cdot o^0_{\mathrm{usr}} + 0.7 \cdot o_{\mathrm{usr}}(t) + 0.1 \cdot p_{\mathrm{art}}(t)$ and "fast" opinion dynamics described by $o_{\mathrm{usr}}(t+1) = 0.7 \cdot o_{\mathrm{usr}}(t) + 0.3 \cdot p_{\mathrm{art}}(t)$ where the influence of the prejudice is ignored. The slow and the fast opinion dynamics describe two different types of users; the formers are little influenced by online platforms and rely more on their friends' opinions, leading to more stable opinions trajectories, the latters rely solely on the online platforms to form their opinions, leading to opinions more sensitive to the articles recommended by the RS.

We studied two relatively extreme behaviors, but this allows us to stretch the effect of each component in order to get an idea of the whole spectrum. In a more realistic set-up, we observe heterogeneous individual behavior, so different conclusions may apply to different individuals.

4.1 Simulation Set-Up

In the following Sections, we show the results of simulations of the closed-loop described in Chapter 3 with different RS models and opinion dynamics parameters. In all of our simulations we have:

- 10'000 users,
- 400 articles,
- 1000 time-steps.

The initial opinions of the users are drawn from the standard normal distribution and the positions of the articles are evenly spaced in the "logspace" between -2 and 2, such that approximately 95% of the users' opinion is covered by the articles' position.

The threshold used to calculate the percentage of users whose opinion deviated little during the interaction with the RS is set to 0.2. This represents 10% of the most extreme positions the users can receive from the RS.

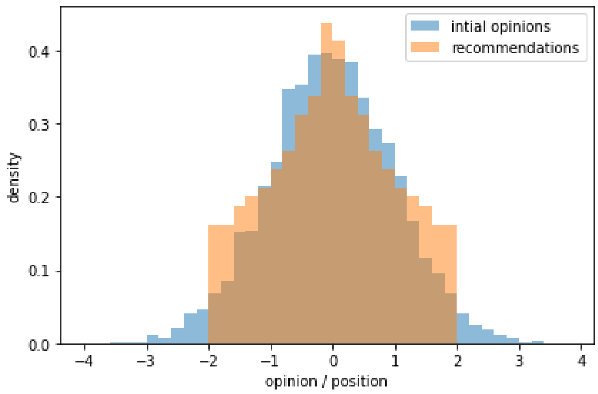


Figure 4.1: Distribution of the initial opinions and the position of the articles.

The distribution of the position of the articles as well as the distribution of the opinion of the users are illustrated in Figure 4.1. The mismatch between the two distributions highlights the fact that not all opinions of the society are equally represented in the media and online.

The plots in the following Sections are the results of a single simulation with same initial conditions for the different set-ups. On the other hand, the numerical results are the mean of 10 different simulations, thus producing more representative results.

4.2 Perfect Knowledge

First, we analyze the ideal scenario: all the users' opinion and articles' position are known to the RS at the beginning of the simulation. We compare the societal performance and the reward maximization of the learning approach and the utility matrix.

4.2.1 Slow Opinion Dynamics

Figure 4.2 shows a histogram of the RM performance of the RS for each of the users. The mean RM, as defined in 3.3.1, is denoted by a dashed line. The RM performance for the LA is 98.82% and is 86.18% for the UM. This difference can be very well explained as in the LA, the RS knows the opinion of all the users and the position of all the articles from the beginning, therefore the RS sends the best possible article at the first time-step and keeps doing so, leading to a RM of almost 100%. On the other hand, the UM has no information to start with and needs exploration to find the articles leading to high rewards, thus inducing a relative loss of RM. However the RM of the UM is surprisingly high, this may be due to the large number of time steps giving the UM enough time to compensate for its absence of initial knowledge.

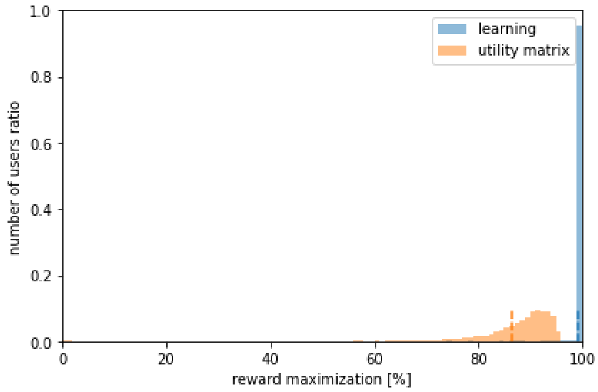
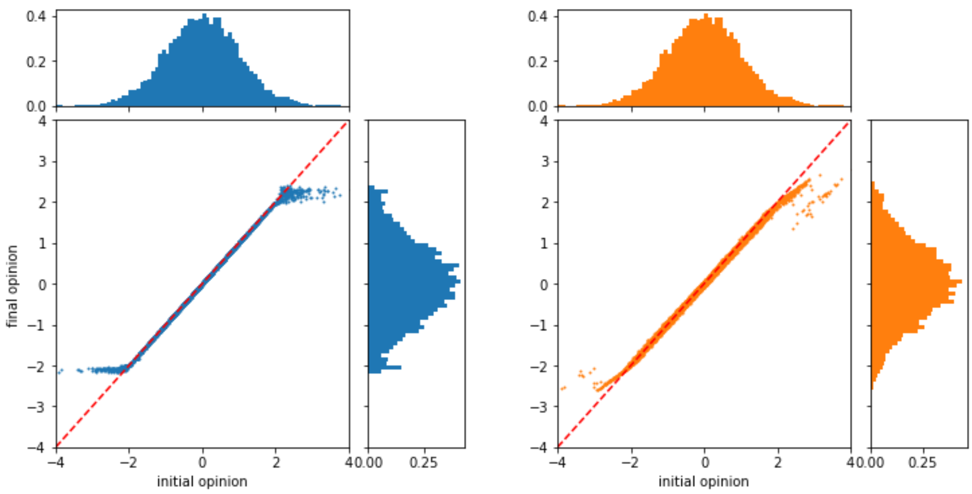


Figure 4.2: Histogram of the reward maximization (RM) performance of the RS.

Figure 4.3 shows a scatter plot of the initial and final opinions of the users as well as a histogram showing the initial distribution of the opinions (on the top) and a histogram showing the final distributions of the opinions (on the right) for the LA and UM RS. The red diagonal line in the scatter plots marks the absence of shift of opinions. We can observe that the points of the scatter plots are mostly concentrated around the red diagonal lines except for the most extreme opinions that shifted towards the most extreme articles' position, particularly in the case of the LA RS. This is confirmed by the SP numbers:

- The LA has a Pearson correlation coefficient of 0.994 and 97.88% of the users shifted their opinion less than 0.2 during the simulation,
- The UM has a Pearson correlation coefficient of 0.997 and 98.53% of the users shifted their opinion less than 0.2 during the simulation.

Overall, both the LA and the UM RS lead to very small shifts in the users' opinion, it is interesting to notice that the UM achieved to learn the users' preferences without drastically changing the users' opinions.



(a) Learning Approach.

(b) Utility Matrix.

Figure 4.3: Scatter plots and histograms of the initial and final opinions of the users.

4.2.2 Fast Opinion Dynamics

We now analyze the same configuration as in Section 4.1 but with faster dynamics, here the influence of the prejudice on the formation of the opinions is neglected and more importance is given to the articles coming from the RS. In Figure 4.4, we observe that both the LA and UM RS achieve very high reward maximization. The RM performance is 99.89% for the LA and 91.31% for the UM. As in the case of slower dynamics, we notice that the lack of initial knowledge of the UM leads to a lower reward maximization than the LA. However, the fast opinion dynamics leads to higher reward maximization for both RS. This is due to the fact that the users' opinion will converge to the position of the articles sent relatively fast, thus increasing the RM performance. This is particularly relevant for the most extreme users whose opinions converge faster to the range of the articles' position, thus increasing the RM performance.

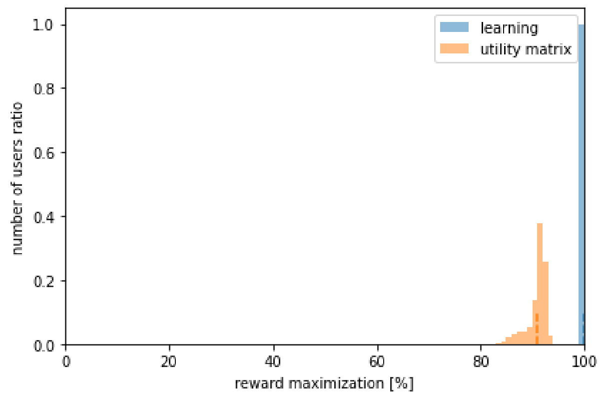


Figure 4.4: Histogram of the reward maximization (RM) performance of the RS.

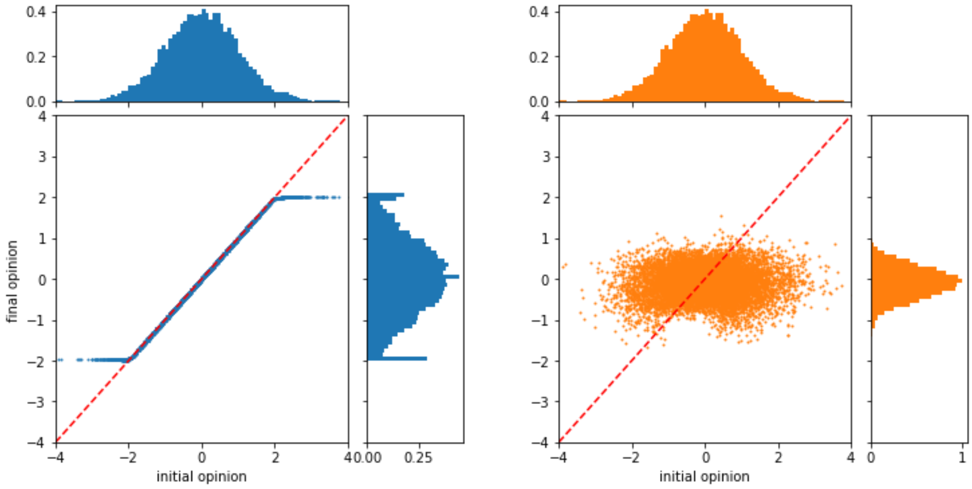
.

In Figure 4.5, we observe two very different situations. On one hand, the scatter plot of the LA RS matches almost perfectly the red diagonal between -2 and 2, meaning no shift of the opinions, and the most extreme users have been shifted to the most extreme articles' positions, the two peaks in the histogram of the final opinions illustrate this well. On the other hand, in Figure 4.5b almost no correlation between the initial opinions and the final opinions of the UM RS can be observed. This is confirmed by the SP numbers:

- The LA has a Pearson correlation coefficient of 0.995 and 97.06% of the users shifted their opinion less than 0.2 during the simulation,

- The UM has a Pearson correlation coefficient of 0.155 and 17.42% of the users shifted their opinion less than 0.2 during the simulation.

Looking at the histogram of the final opinions of the UM RS, we can observe that the opinions converged to where most of the articles are positioned. This means that while trying to learn the users' preferences, the UM shifted those preferences thus leading to the final opinions observed.



(a) Learning Approach. (b) Utility Matrix.

Figure 4.5: Scatter plots and histograms of the initial and final opinions of the users.

4.3 Noisy Knowledge

In this section, we analyze the performance of the learning approach when the beliefs of the users' opinion and articles' position are noisy. More precisely, a Gaussian noise blurs the users' opinion belief, $\tilde{\mathbf{o}} + \epsilon,\ \epsilon \sim \mathcal{N}(0, 0.5)$ and two types of noise are tested and compared on the on the articles' position belief, a zero and a non-zero mean Gaussian noise $\tilde{\mathbf{p}} + \zeta,\ \zeta \sim \mathcal{N}(m, 0.5),\ m = 0, 2$

4.3.1 Slow Opinion Dynamics

We observe in Figure 4.6 that the RM performance is far from being as high as in Section 4.2, moreover, it seems that the non-zero mean noise on the articles' position belief further decreases the RM performance as a large part of users have a RM performance smaller than 20%. Indeed the RM performance of the LA with $\epsilon \sim \mathcal{N}(0, 0.5)$ is 62.29% and only 26.82% for the LA with $\epsilon \sim \mathcal{N}(2, 0.5)$. It can be concluded that for slow opinion dynamics, the LA algorithm is not able to recover the users' opinion and articles' position needed to deliver a great RM performance.

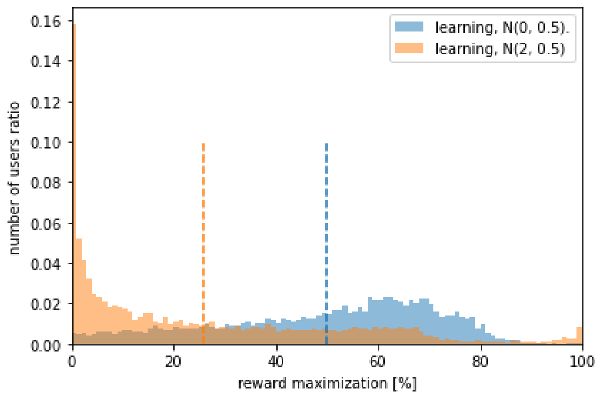
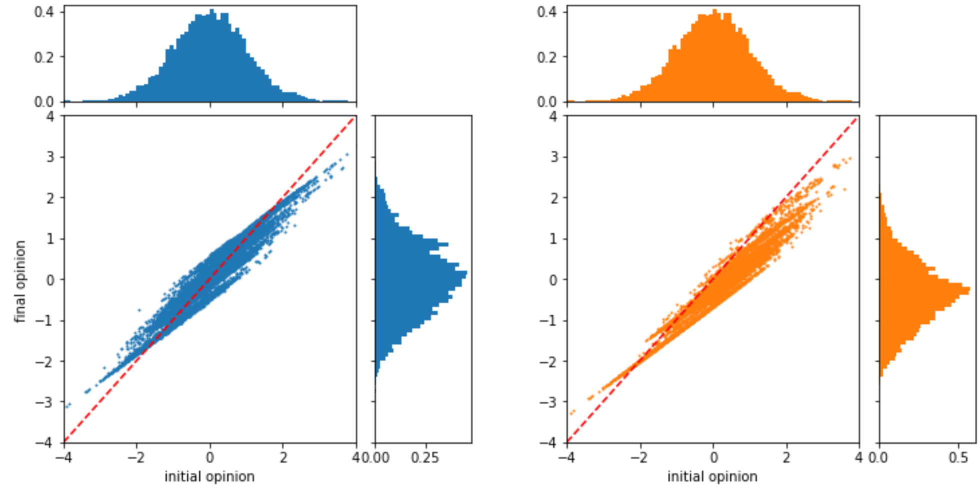


Figure 4.6: Histogram of the reward maximization (RM) performance of the RS.

In Figure 4.7, we see that the final opinions keep a shape relatively close to the diagonal line, however, in the case of the LA with the non-zero mean noise, most of the final opinions are below the diagonal red line, which is not the case in the zero-mean case, thus highlighting the effect of the non-zero mean noise on the final opinions. These two observations are confirmed by the SP performances numbers:

- The LA with $\epsilon \sim \mathcal{N}(0, 0.5)$ has a Pearson correlation coefficient of 0.986 and 76.76% of the users shifted their opinion less than 0.2 during the simulation,

- The LA with $\epsilon \sim \mathcal{N}(2, 0.5)$ has a Pearson correlation coefficient of 0.982 and 24.48% of the users shifted their opinion less than 0.2 during the simulation.



(a) LA with $\epsilon \sim \mathcal{N}(0, 0.5)$.

(b) LA with $\epsilon \sim \mathcal{N}(2, 0.5)$.

Figure 4.7: Scatter plots and histograms of the initial and final opinions of the users.

4.3.2 Fast Opinion Dynamics

We now analyze the same configuration as in Section 4.3.1 but with faster dynamics. Here, the influence of the prejudice on the formation of the opinions is neglected and more importance is given to the articles coming from the RS. In Figure 4.8 ,we see that the RM performance is very good for almost all the users for the two kinds of noise applied to the articles' position belief, this is due to the fact that the users' opinions converge relatively fast to the received articles' positions leading to a good RM performance but it could lead to large deviations of the users' opinion. Here the RM performance of the LA with $\epsilon \sim \mathcal{N}(0, 0.5)$ is 93.36% and 96.19% for the LA with $\epsilon \sim \mathcal{N}(2, 0.5)$.

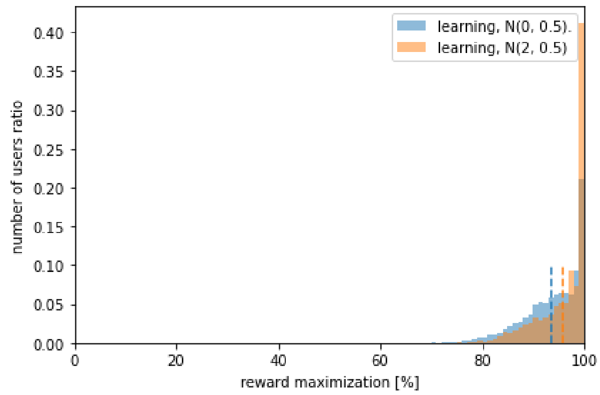
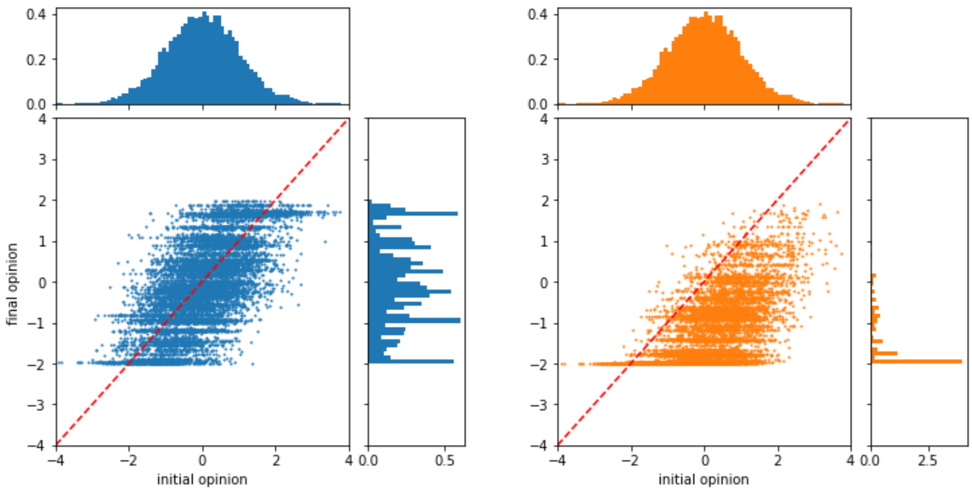


Figure 4.8: Histogram of the reward maximization (RM) performance of the RS.

In Figure 4.9, we observe that the gain of RM performance observed above leads to large deviations of the users' opinions. Moreover, the effect of the non-zero mean noise can be very well seen in the histogram of the final opinions of the LA with $\epsilon \sim \mathcal{N}(2, 0.5)$ as a lot of users are concentrated around the -2 opinion. The deviation, as well as the dispersion of the users' opinion, can be seen in the SP performance numbers:

- The LA with $\epsilon \sim \mathcal{N}(0, 0.5)$ has a Pearson correlation coefficient of 0.842 and 28.48% of the users shifted their opinion less than 0.2 during the simulation,
- The LA with $\epsilon \sim \mathcal{N}(2, 0.5)$ has a Pearson correlation coefficient of 0.730 and 3.38% of the users shifted their opinion less than 0.2 during the simulation.



(a) LA with $\epsilon \sim \mathcal{N}(0, 0.5)$). (b) LA with $\epsilon \sim \mathcal{N}(2, 0.5)$.

Figure 4.9: Scatter plots and histograms of the initial and final opinions of the users.

4.4 No knowledge

In Section 4.3, we observed a trade-off between RM and SP, the learning approach is either inducing small shifts of opinions or achieving large RM, for slow or fast dynamics respectively.

Now we study the behavior of the LA when the RS has no knowledge and compare it with the UM. To do that we will use the following opinion dynamics:

$$o_{\text{usr}}(t+1) = \beta \cdot o_{\text{usr}}(t) + \gamma \cdot p_{\text{art}}(t),\ \beta + \gamma = 1$$

and make the two coefficients vary between 0 and 1 in order to observe the importance of opinion dynamics in the RS performances.
Figure 4.10 shows the performance of the LA RS for ten different opinions dynamics, we clearly see a trade-off between the SP given by the Pearson correlation coefficient on the y-axis and the RM on the x-axis. The faster the dynamics, the higher the RM performance but the lower the SP. On the other hand, slow dynamics lead to an improved SP but a decreased RM.

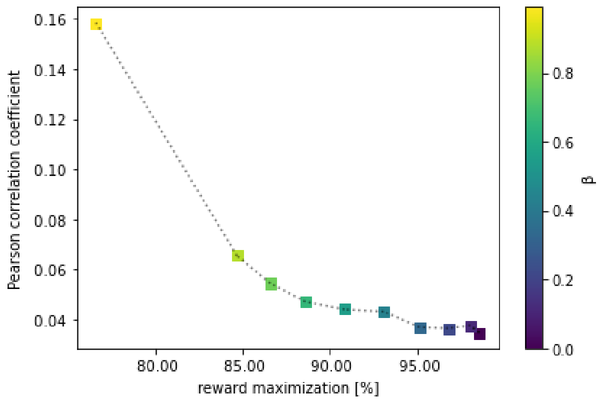


Figure 4.10: Performances of the LA with variations of the opinion dynamics.

In Figure 4.11, we see the Pearson correlation coefficient as well as the RM performance for ten different parameters settings of opinion dynamics for the UM scenario. Interestingly, no clear conclusion can be made on the dependence of the RM on the opinion dynamics as the RM vary between 91.5% and 93.5%, which is almost negligible. However, there is a clear trend showing that faster dynamics lead to a decrease in the Pearson correlation coefficient, hence the SP. As observed in Section 4.2.2, the UM changes the users' opinion during the learning process when the opinion dynamics are fast.

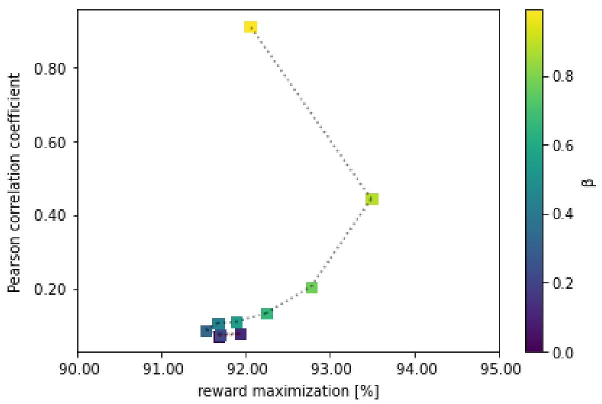


Figure 4.11: Performances of the UM with variations of the opinion dynamics.

Figure 4.12 combines the plots of Figure 4.10 and 4.11 in one plot in order to compare their performances. We clearly see that the LA reaches almost perfect performances for very fast opinion dynamics but with very low Pearson correlation coefficient values whereas the opinion dynamics influences mostly the SP of the UM and have little impact on the RM.

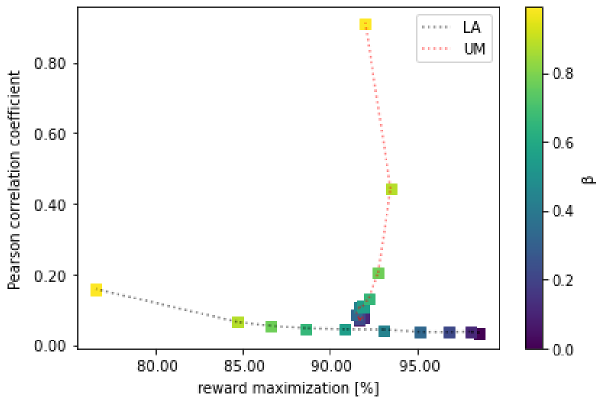


Figure 4.12: Performances of the LA and UM with variations of the opinion dynamics.

Chapter 5

Conclusion

In the present project, we developed a new closed-loop model of the interaction between RSs of online platforms and their users. To do so, we extended an opinion dynamics model and we defined an equation determining the user engagement and two new RS models, a data-driven and a model-based relying on prior knowledge of the opinions of the users and the position of the articles. We evaluated the impact of each of the RS on the users with three different metrics: the first measures how well the users' engagement is maximized, the second the correlation between the initial and final opinions, and the third the deviation of the users' opinion above a given threshold. These two last measurements try to evaluate the impact of the RS on the society, i.e., on the opinion of the users while the first only focuses on the users' engagement.
We observed that the absence of prior knowledge about the users and the articles is detrimental to the societal performance, mostly when the users have fast opinion dynamics. However, a biased prior knowledge leads to a shift of the opinion of the users. The data-based approach used with the utility matrix achieves high reward maximization, but in case of fast opinion dynamics the distribution of the position of the articles determines the final opinion of the users. In particular, this implies that a bias in the articles' position would lead to a bias in the opinion of the users. In all the cases, the fast opinion dynamic case is critical: the online platform does not notice that the opinion of the users shifts, as it is not observed in the RM performance that stays above 90%.
In future works, users with heterogeneous opinion dynamics as well as other distributions of the initial opinion of the users and of the position of the articles, such as bipartisan users' opinion distribution, should be investigated. A theoretical understanding of the impact of the distribution of the articles' position on the users' opinion would also be an interesting investigation.

Bibliography

[1] Wilbert Samuel Rossi, Jan Willem Polderman, and Paolo Frasca. "The closed loop between opinion formation and personalised recommendations". In: *IEEE Transactions on Control of Network Systems* (2021), pp. 1–1.

[2] Michela Del Vicario et al. "Echo Chambers: Emotional Contagion and Group Polarization on Facebook". In: *Scientific Reports* 6 (2016).

[3] Eli Pariser. *The filter bubble: What the Internet is hiding from you.* Penguin, UK, 2012.

[4] Perra Nicolas. and Rocha Luis E.C. "Modelling opinion dynamics in the age of algorithmic personalisation". In: Sci Rep 9, 7261 (2019).

[5] Raymond Nickerson. "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises". In: *Review of General Psychology* 2 (June 1998), pp. 175–220.

[6] Joshua Klayman. "Varieties of Confirmation Bias". In: ed. by Jerome Busemeyer, Reid Hastie, and Douglas L. Medin. Vol. 32. Psychology of Learning and Motivation. Academic Press, 1995, pp. 385–418.

[7] Jure Leskovec, Anand Rajaraman, and Jeffrey David Ullman. "Recommendation Systems". In: *Mining of Massive Datasets.* 2nd ed. Cambridge University Press, 2014.

[8] Jacob Benesty et al. "Pearson correlation coefficient". In: *Noise reduction in speech processing.* Springer, 2009, pp. 1–4.

Appendix A

Average Gradient Ascent

The Average Gradient Ascent (AGA) is a recommender system. It tries to maximizes the users' engagement by using an average gradient ascent of the reward function. In order to perform the average gradient ascent, the RS stores the value of the position of the article it would ideally send to the users. The users receive the articles with positions the closest to the ideal positions. The AGA RS stores the ideal position of the last m articles sent to each user along with the corresponding rewards. The AGA RS has no knowledge about the users but stores a belief of the position of the articles. At each time step, the following steps are followed:

- Calculation of the discounted average of the ideal position of the articles sent to user i in the last m time steps,

$$p^{-}(t) = \frac{p_{\text{ideal}}(t-1) + d \cdot p_{\text{ideal}}(t-2) + ... + d^{m-1} \cdot p_{\text{ideal}}(t-m)}{1 + d + ... + d^{m-1}}$$

- Calculation of the discounted average of the rewards received from user i in the last m time-steps,

$$r_i^{-}(t) = \frac{r_i(t-1) + d \cdot r_i(t-2) + ... + d^{m-1} \cdot r(t-m)}{1 + d + ... + d^{m-1}}$$

- Choice of the next ideal article position using the following logic:

if $r_i(t) \geq r_i^{-}(t)$:
    if $p_{\text{ideal}}(t) \geq p^{-}(t)$:
        $p_{\text{ideal}}(t+1) = p_{\text{ideal}}(t) + \Delta$
    else:
        $p_{\text{ideal}}(t+1) = p_{\text{ideal}}(t) - \Delta$
else :
    $p_{\text{ideal}}(t+1) = p^{-}(t)$:

- The next article for the user i is:

$$j = \underset{1 \leq j \leq m}{\operatorname{argmin}} \|\tilde{\mathbf{p}}_j - p_{\text{ideal}}(t+1)\|_2$$

The first article j is chosen randomly and $p_{\text{ideal}}(0) = \tilde{\mathbf{p}}_j$.
This is done for all users $i \in \mathcal{U}$.
In our set-up, $m = 10, d = 0.5$ and $\Delta = 0.1$.

A.1 Results

Here, we look at the impact of a noise ϵ on the position of the recommendations belief as well as the impact of the dynamics on the AGA RS. The same set-up is used as in the Chapter 4, except that all the results come from one simulation only.

A.1.1 Perfect knowledge

The reward maximization performance of the AGA RS with slow dynamics is 84.11% and 87.165% with the fast dynamics.

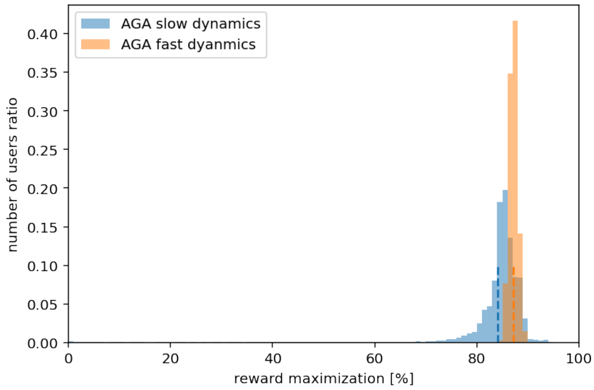
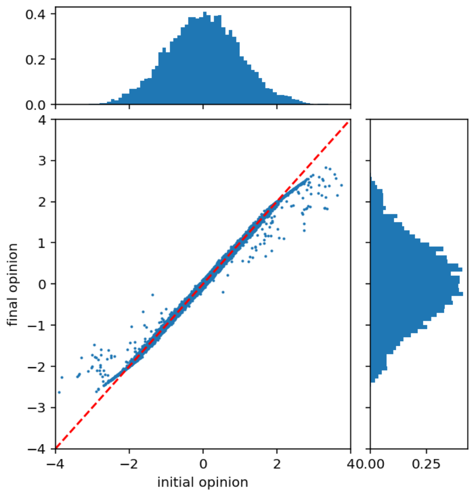
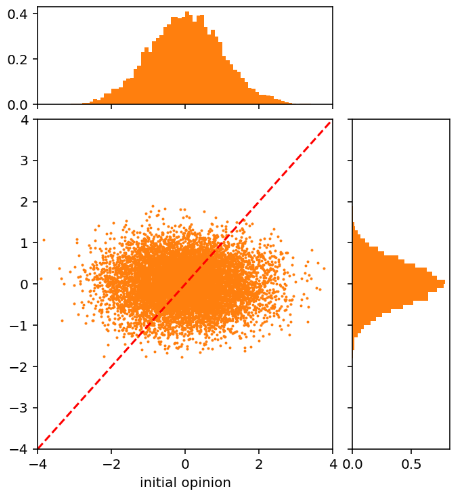


Figure A.1: Histogram of the reward maximization (RM) performance of the RS.

With slow dynamics, the AGA RS has a Pearson correlation coefficient of 0.995 and 94.84% of the users shifted their opinion less than 0.2 during the simulation.
With fast dynamics, the AGA RS has a Pearson correlation coefficient of 0.006 and 7.45% of the users shifted their opinion less than 0.2 during the simulation.



(a) Slow dynamics.



(b) Fast dynamics.

Figure A.2: Scatter plots and histograms of the initial and final opinions of the users.

We clearly see that the fast dynamics decrease the SP of the AGA RS while the RM performance increases a little compared to the slow dynamics.

A.1.2 Noisy knowledge

Here, a Gaussian noise with a standard deviation of 0.5 and zero-mean is applied on the belief of the position of the articles.
The reward maximization performance of the AGA RS with slow dynamics is 41.07% and 64.80% with the fast dynamics.

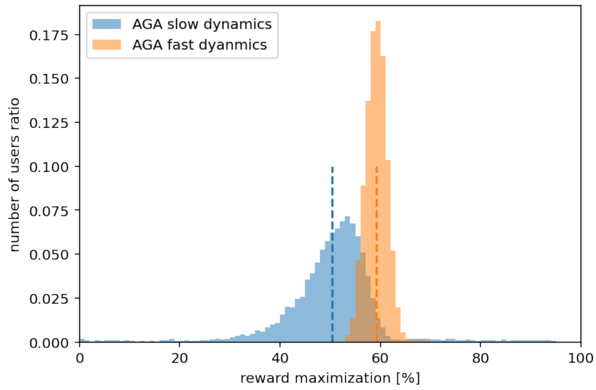
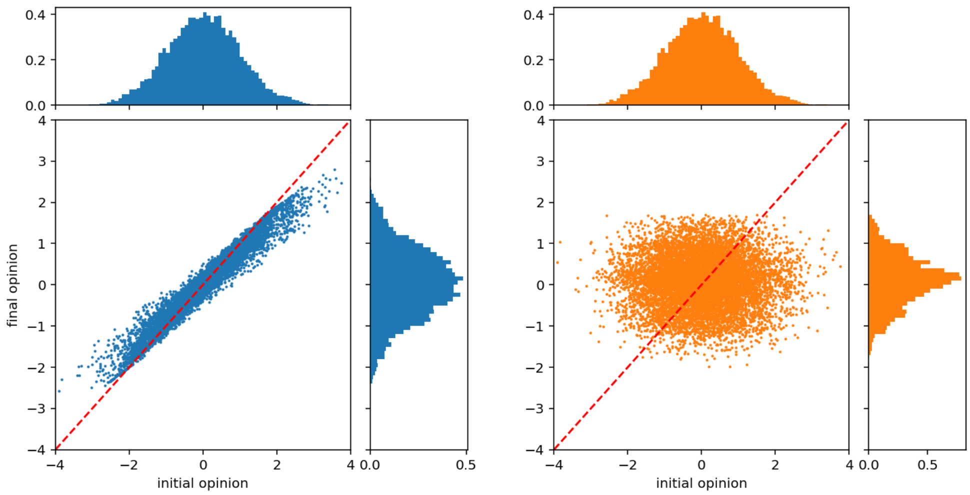


Figure A.3: Histogram of the reward maximization (RM) performance of the RS.

With slow dynamics, the AGA RS has a Pearson correlation coefficient of 0.975 and 41.07% of the users shifted their opinion less than 0.2 during the simulation.
With fast dynamics, the AGA RS has a Pearson correlation coefficient of 0.001 and 6.48% of the users shifted their opinion less than 0.2 during the simulation.



(a) Slow dynamics. (b) Fast dynamics.

Figure A.4: Scatter plots and histograms of the initial and final opinions of the users.

In the slow dynamics case, we clearly the that the noise on the belief of the recommendation decreases the SP. The RM performance decreases for both the slow and the fast dynamics case.

A.1.3 Mean deviation

Here the belief of the position of the articles is shifted by 2 in order to test the impact of a bias on the closed-loop system.

The reward maximization performance of the AGA RS with slow dynamics is 83.76% and 86.99% with the fast dynamics.

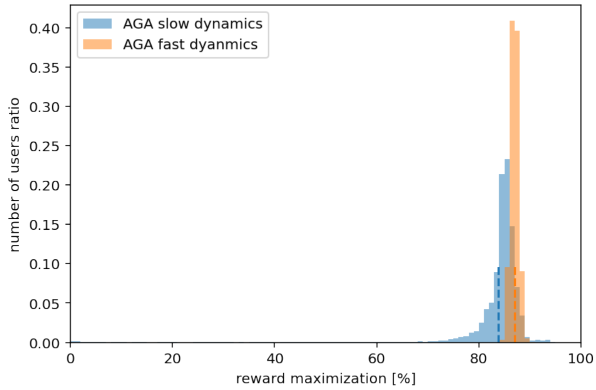
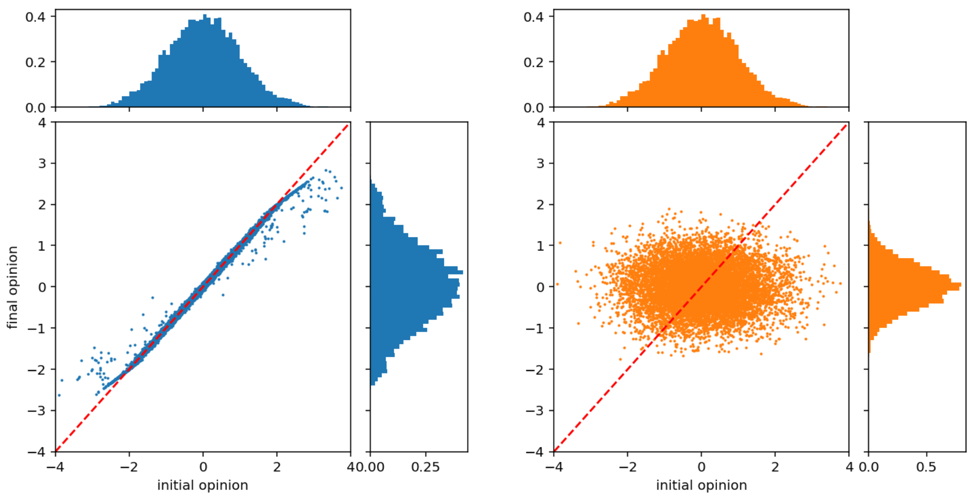


Figure A.5: Histogram of the reward maximization (RM) performance of the RS.

With slow dynamics, the AGA RS has a Pearson correlation coefficient of 0.995 and 94.48% of the users shifted their opinion less than 0.2 during the simulation.
With fast dynamics, the AGA RS has a Pearson correlation coefficient of 0.005 and 7.29% of the users shifted their opinion less than 0.2 during the simulation.



(a) Slow dynamics.

(b) Fast dynamics.

Figure A.6: Scatter plots and histograms of the initial and final opinions of the users.

We see that the mean deviation of the belief of the recommendations has no impact on the closed-loop system, the results are the same as in Section A.1.1.