C. R. Physique 20 (2019) 371–379

Contents lists available at ScienceDirect

Comptes Rendus Physique

www.sciencedirect.com

From statistical physics to social sciences/*De la physique statistique aux sciences sociales*

Towards novelty-driven recommender systems

Vers des systèmes de recommandation axés sur la nouveauté

Pietro Gravino [a,*], Bernardo Monechi [a], Vittorio Loreto [a,b,c]

[a] *Sony Computer Science Laboratories, Paris, 6, rue Amyot, 75005 Paris, France*
[b] *Sapienza University of Rome, Physics Department, Piazzale Aldo Moro 2, 00185 Roma, Italy*
[c] *Complexity Science Hub, Josefstädter Strasse 39, A-1080 Wien, Austria*

ARTICLE INFO

Article history:
Available online 26 June 2019

Keywords:
Recommender systems
Novelties
Adjacent possible
Comfort zone

Mots-clés :
Systèmes de recommandation
Nouveautés
Adjacente possible
Zone de confort

ABSTRACT

We get recommendations about everything and in a pervasive way. Recommender systems act like compasses for our journey in complex conceptual spaces and we more and more rely on recommendations to ground most of our decisions. Despite their extraordinary efficiency and reliability, recommender systems are far from being flawless. They display instead serious drawbacks that might seriously reduce our open-mindedness and our capacity of experiencing diversity and possibly conflicting views. In this paper, we carefully investigate the very foundations of recommendation algorithms in order to identify the determinants of what could be the next generation of recommender systems. We postulate that it is possible to overcome the limitations of current recommender systems, by getting inspiration from the way in which people seek for novelties and give value to new experiences. From this perspective, the notion of adjacent possible seems a relevant one to redesign recommender systems in a way that better aligns with the natural inclination of human beings towards new and pleasant experiences. We claim that this new generation of recommenders could help in overcoming the pitfalls of current technologies, namely the tendency towards a lack of diversity, polarization, the emergence of echo-chambers and misinformation.

RÉSUMÉ

Nous recevons des recommandations à propos de tout et d'une manière omniprésente. Les systèmes de recommandation agissent comme des boussoles pour notre voyage dans des espaces conceptuels complexes, et nous comptons de plus en plus sur les recommandations pour fonder la plupart de nos décisions. Malgré leur efficacité et leur fiabilité extraordinaires, les systèmes de recommandation sont loin d'être parfaits. Ils présentent plutôt de sérieux inconvénients qui pourraient sérieusement réduire notre ouverture d'esprit et notre capacité à faire l'expérience de la diversité et à avoir des opinions contradictoires. Dans cet article, nous examinons attentivement les fondements mêmes des algorithmes de recommandation afin d'identifier les déterminants de ce que pourrait être la prochaine génération de systèmes de recommandation. Nous postulons qu'il est possible de surmonter les limites des systèmes de recommandation actuels

* Corresponding author.
E-mail address: pietro.gravino@gmail.com (P. Gravino).

https://doi.org/10.1016/j.crhy.2019.05.014

en s'inspirant de la façon dont les gens recherchent les nouveautés et valorisent les nouvelles expériences. De ce point de vue, la notion de possible adjacent semble pertinente pour repenser les systèmes de recommandation d'une manière qui s'aligne mieux avec l'inclination naturelle des êtres humains vers des expériences nouvelles et agréables. Nous affirmons que cette nouvelle génération de prescripteurs pourrait aider à surmonter les écueils des technologies actuelles, à savoir la tendance au manque de diversité, à la polarisation, à l'émergence d'écho-chambres et à la désinformation.

1. Introduction

Recommender systems are ubiquitous in our everyday experience. We get books recommended, music recommended, food recommended, items to buy, hotels, trips. Even opinions and ideas. Recommendations are so entangled in our experience that perhaps we cannot even conceive our life without them. Still we base most of our decisions on the suggestions or tips we receive or seek for.

Recommender systems act like compasses for our journey in complex conceptual spaces [25,1,2,9,13,22,14,13,22,14]. They exploit the knowledge about user behaviours and about the structure of the conceptual space itself, to suggest new directions to take, new experiences. Recommendations often concern something we may like because it is similar to something liked by someone else with a personal history similar to ours. This idea is brilliantly in place in almost all recommender systems with little variations taking into account personal histories, historical dependences, etc. And it turns out to be also very effective in many sectors thanks to the abundance of data about people and their choices. But all that glitters is not gold!

Recently, a lot has been written about potential drawbacks of recommender systems in shaping our approach to information [25,1,2,9]. For instance, a controversy emerged about the impact of those systems on the diversity of contents experienced by users. While in some cases an enhancement of diversity has been reported [13,24,32], it is often claimed that personalization results in a loss of diversity [22,25,5]. This reduction could bring to a dangerous amplification of the human natural tendencies to homophily, to polarization, to the emergence of echo chambers and misinformation [25,1,2,9,10,7,29,11,8].

Is it possible to conceive recommender systems able to overcome these difficulties? This is a very pressing question that calls for a global rethinking of the way in which recommenders should accompany us in our everyday experience. Here we address this question by postulating the possibility of a new generation of recommender systems able to nudge us out of our "comfort zone," experiencing instead novel, diverse, though still pleasant, experiences.

New recommenders can be inspired by the way in which people experience novelties. The experience of the new is something very common in our lives, either at a personal level (e.g., reading a of a new book) or at the global level (e.g., writing a new book). In both cases, the experience of the new can be pictured as a path in a very special space, the space of the possible, the space of what could be. Francois Jacob highlighted the dichotomy between "actual" and "possible." The "actual" is the set of things we experienced already. The "possible" is the set of things we might possibly experience in the future.

The question of the exploration of the new is a fascinating one that puzzled generations of scholars. How can we conceive and define the space where novelties can be experienced? How can one estimate the probability for a new event to occur? A very interesting hypothesis in this area has been proposed by Stuart Kauffman, who introduced the notion of Adjacent possible [15,31,12,21]. Adjacent possible is loosely defined as the set of things – could be molecules, organisms, technological products, ideas, etc. – that are one step away from what actually exists and that could be reached by the incremental recombination of the existing elements. The key feature of the adjacent possible space is that its structure is not predefined – Kauffman would say unprestatable –; it gets instead constantly reshaped by the exploration of the space itself. Somehow, each of us carves his own adjacent possible space while experiencing the new.

The exploration of the space of possibilities is paralleled by the learning process that makes users progressively more familiar with specific entities [23,26]. To extract value from each new piece of information, humans need to get more familiar with it. The learning process is precisely what makes new things more familiar to us while restructuring our evaluation of the whole past experience.

Here we postulate that the notion of adjacent possible and the learning dynamics can be adopted as cornerstones to design the new generation of recommender systems. Those new recommenders need to be able to nudge us in a path of discovery of the new that progressively dislodges us from our comfort zone, but without pushing us too far.

These paths will touch progressively new elements within the adjacent possible, enlarging in this way the comfort zone and creating the preconditions for further expansions.

The outline of the paper is as follows. Section 2 will be devoted to reviewing the assumptions behind a typical recommender system and decompose the problems in two parts: modelling the space of items to be recommended and modelling user behaviour. In Section 3, we shall focus on a preliminary assessment of the impact of recommendations. We shall exhibit hints supporting that a reconstruction of the space of items solely based on user behavioural data brings to a poor exploration of the space itself. Already this evidence points to the need for a new generation of recommenders that could allow for a better exploration of the space of items. In Section 4 we further explore how present recommenders can be improved. To this end, we disprove the common assumption that user behaviours are static. Users display instead memory effects in their paths, presumably guided by saturation phenomena and the quest for something new, like in learning processes. We conclude the paper by the more speculative Section 5, which proposes a roadmap that, through the definition of a novel theoretical framework, could lead to the emergence of a new generation of recommender systems, able to enrich the user experiences while overcoming the drawbacks of current systems.

2. Recommender systems: space and user modeling

Each recommender system makes assumptions, more or less implicitly, in order to model the space of the items to be recommended. Modeling a space of contents in an accurate way does not really mean to assess the value of each item, which could be subjective and time-dependent. It means to infer the relationships among those items. In other words, what we could call semantics. Usually, this semantics is not only used for items, but also to localize users into this space, i.e. profiling the tastes of the users.

The most common techniques adopted to model the space of contents are the so-called Content-Based Filtering (CBF) [19] and the Collaborative Filtering (CF) [28]. In the CBF, the space structure is inferred by the meta-information of the items (e.g., for a song the album it belongs to, the authors, the year of release, the genre, etc.), usually provided by creators, editors, experts or simple users. Collaborative Filtering techniques, instead, aim at reconstructing the semantics of the space of items based on user histories and behaviours. The underlying assumption is that people with a similar past will have a similar future. Both methods spawned a several variations and hybridizations. From the point of view of the accuracy performance in rating forecast, which is the standard validation method for recommender systems, both models have been constantly improving. In particular, CF algorithms raised to a major popularity since the Netflix Prize [16], won by the combined BellKor's Pragmatic Chaos team. Since then, matrix factorization methods have been very popular for their performances, scalability, and ease of implementation, and also because they do not rely on other external data source (which could be hard to find and validate), as CBF methods.

Despite recommender systems have been almost omnipresent for decades, only in recent years an interest emerged in the literature on their impact. There is no clear consensus on the effect of recommender systems. Some studies claim negative effects [22,25,5], while others state the opposite [13,24,32], the debate being perhaps affected by an ill-posed formulation of the problem. Under the denomination of recommender, fall, in fact, many different techniques applied in very different scenarios. It is thus very hard to make a universal point. Still, if we focus on the comparison between the main space reconstruction technique, CBF and CF, several works confirm a difference in perceived diversity. CF algorithms seem to offer a representation of the space biased by the popularity, often leading to a decrease in the diversity of the recommendations [13]. CBF is instead observed to be less prone to this kind of issues, and cases have been reported of increase of diversity [22]. This is somehow unsurprising if we think about the data adopted within the different approaches. If one adopts data on user behaviour to infer the semantic relationships of items, i.e. if one says that two items are trivially related because they have been jointly chosen by many pairs of users, it is very likely that the estimate will be strongly biased by the popularity of items and the homophily linking different users [20]. The adoption of external information, focused on the actual semantics of items, seems to be a better choice to model the space of items.

In the development of a recommender system, a lot of attention has been given to the modelling of the space of items. A similar attention has not been devoted to modelling user behaviour. Only more recently, the rising tide of polarization and misinformation online contributed to shift the focus on the effect of recommendations on user behaviours. Still little or no consideration has been devoted to revisit the main and critical assumption that considers the user as a static agent who would behave in a similar way under similar circumstances. This assumption has several implications. For example, many basic recommender systems consider user history as flat and equivalent [18,27], e.g., the last song listened to has the same importance as the first song listened to. Also, the impact of the recommender itself, i.e. how the suggestions of the recommender affect the perception and push the development of future tastes, is rarely considered. There are instead evidences of a temporal evolution [17] and that even the very fact of showing ratings has an influence on the opinions of the users [6]. Users evolve while they "consume" information, and they do in a non-trivial way. Things that were favoured in the past can be considered boring in the future. Also, if we consider a new recommended to a polarized user, the content will be elaborated in a biased way, thus the impact will be affected by the same bias, and amplified. If a recommendation system does not take into the account the possibility of evolution, its proposal will present an inertia pushing back users toward old tastes, toward the so-called "comfort zone". Moreover, this inertia, combined with cognitive biases, might result in a further push towards extremization and polarization. Some systems have been proposed to consider memory effects to overcome the limitations mentioned above, usually just weighing more the most recent past of the history of the users [17]. In other cases, diversity was suggested as a new metric of optimization, together with accuracy.

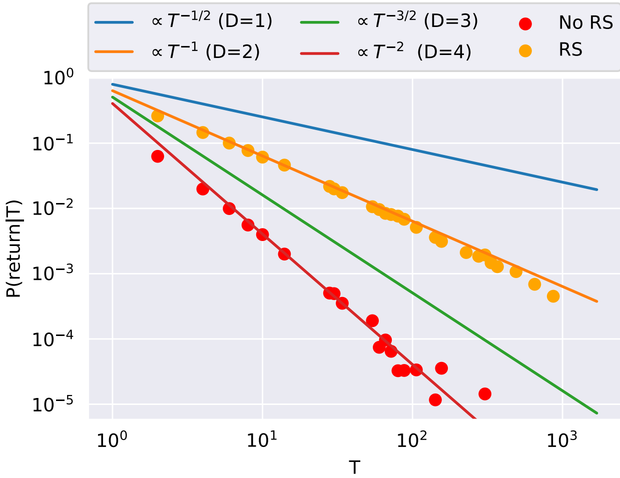

Fig. 1. A comparison of return probabilities as a function of the return time for the random walk in $D = 4$ dimension with and without the recommender system. Without (red dots), the curve is close to what is theoretically expected (red line). Theoretical predictions for various dimensionalities are reported as a guide for the eye (blue, orange, green and red line). If the recommender system influences the dynamics with a probability of 50% (users do not always follow recommendations), the curve (orange dots) is closer to the theoretical curve foreseen for $D = 2$. The system adopted is a CF using a simple matrix factorization technique, the Singular Value Decomposition. The simulation is the result of 100 random walker traveling for 1000 steps.

3. An agnostic assessment of recommender systems

In this section, we focus on the assessment of the impact of recommender systems on the choices made by the users. This operation is far from being trivial because typically one cannot compare what happens with and without the recommendations. What is typically done is to assess the level of diversity of the recommendations by looking at how different are the items recommended. Here differences are quantified in terms of the semantic distances as defined in the space of items. In order to overcome the above-mentioned difficulty, here we first consider a simplified version of the problem, where the presence of the recommender can be switched on and off. In this version, we can estimate the impact of the recommender by looking at the topological properties of the space sampled by each user with and without the recommender in action. Let us assume we can represent the space of items as a graph, where nodes are items and links mark the transition from one node to another. Let us take, as an example, a regular D-dimensional lattice, with symmetric and equally weighted links. On this topology, the dynamics of the user in the absence of any recommendations will be represented by a D-dimensional random walk. Let us further assume that, at a given point in time, we turn on the recommender system, namely a Collaborative Filtering one. How can we now assess the impact of the recommender without relying on the knowledge of the space structure, i.e. without knowing that the dynamics lives on a D-dimensional lattice? The request of an assessment agnostic with respect to the structure of the underlying space of items comes from the need to avoid a circular loop. In Collaborative Filtering systems, in fact, the structure of the space of items is deduced by the behaviours of the users. But the latter are themselves affected by the recommender system, whose impact is what one is looking for in the first place. This conundrum is not present in Content-Based Filtering recommenders, which rely instead on external information about the structure of the space. One possible way out to the above-mentioned problem consists in performing the analysis of return times in the dynamics of single users. Let us consider a large set of virtual users in a given position $x(t) = X$ at time t. The fraction of users who will be brought back in X by the dynamics at a given time $t' > t$ is a good estimate of the return probability in position X after $T = t' - t$ steps. By performing the averages on all different starting positions and all different times t and t', with $T = t' - t$ constant, one obtains, for random walks on a 1-dimensional lattice, a return probability, as a function of the return time T, given by:

$$P(\text{return}|T) = \binom{T}{T/2}(1/2)^T \propto T^{-1/2}$$

In D dimensions, one has:

$$P(\text{return}|T) \propto T^{-D/2} \tag{1}$$

This is the situation without the recommender system in action. We are now interested in measuring the dimensionality of the space sampled by the users once a recommender system is switched on. We consider in particular a standard recommender system working through a collaborative filtering process. We then reproduce the same procedure as before and we compute the return probability as a function of the return time. The compared results of the two cases, with and without recommender systems are given in Fig. 1.

Though the functional form for the return probability is not affected by the presence of the recommender system, a strong effect is observed in the exponent. In particular the exponent of the return probability under the effect of the

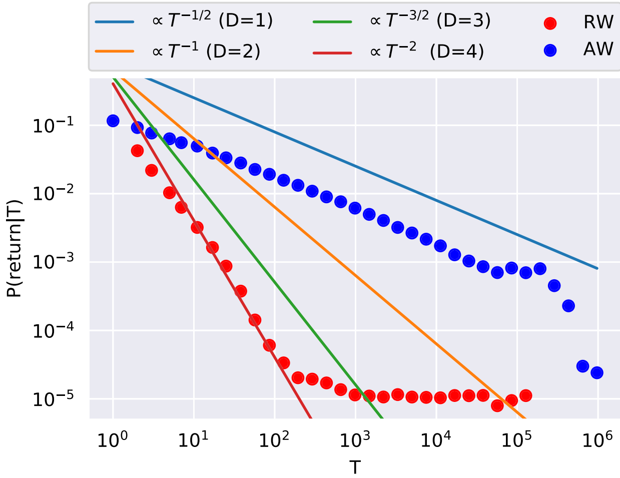

Fig. 2. A comparison of return probabilities as a function of return time for the random walk (RW, red dots, ∼500 simulated users) on the Last.fm weighted graph and for "actual walks" (AW, blue dots, ∼8000 real users). Theoretical predictions for grids of various dimensionalities are reported as a guide for the eye (blue, orange, green, and red line). The length of the path for simulated users are randomly extracted from the distribution of real users' path lengths. The shape of the curve does not allow us to define a precise dimensionality for the simulated users. It seems an overposition of a $D = 4$ grid ($T < 10^2$) and a complete graph ($T > 10^2$), for which the return probability as function of the return time is constant. This is unsurprising in a complex graph. Still, $P(\text{return}|T)$ for actual users shows a completely different behaviour, closer to a RW on a $D = 1$ grid.

recommender system turns out to be reduced with respect to the case where the recommender is not active. In other words, the effect of Collaborative Filtering is that of reducing the dimensionality of the space explored by the user. This implies a less diverse experience.

In order to go beyond this simple proof of concept, let us repeat the same analysis applied to a real case. To this end, we considered a dataset downloaded from the Last.fm music streaming platform. The dataset features the listening to a sample of about 8000 users who listened to about 14 millions of songs for about 51 milions of times over a period of about 9 years. During their listening experience, users were exposed to a Collaborative Filtering recommender [4]. The availability of the listening histories of all the users allowed us to reconstruct the weighted graphs of songs of Last.fm. In this graph, two songs are connected if they have been listened to in sequence by at least one user. Each link is weighted with the number of co-occurrences of the two songs. The resulting graph includes 14 million songs and features a small-world property with a scale-free degree distribution with an average degree of ∼14. By exploiting the individual listening histories, we measured the return probability as a function of the return times. In order to simulate the absence of a recommender system, we simulated a set of random walks on the weighted graph of Last.fm. The compared results are reported in Fig. 2.

Also in this case, the comparison highlights a decrease of the dimensionality of the space sampled by users. The exponent of the return probability vs. the return time turns out to be lower for real users, affected by the recommender, than for users performing random walks. It is important to observe that this is not a direct proof of the impact of the recommender system embedded in Last.fm. Still, the results seem to suggest that something similar to what we observed in the example of a D-dimensional lattice is in place. The result shows that the exploration performed by Last.fm users seems to take place in a space much narrower than is actually available. This is the same kind of effect as that we observed in the simple in silico experiment on D-dimensional lattices.

A first important conclusion we can draw is that, perhaps unsurprisingly, a poor reconstruction of the space of items seems to lead to a poor exploration of the space itself. In particular, if recommendations are based on the behaviour of previous users, truly creative suggestions are less likely to appear.

4. A temporal non-trivial dynamics

In this section, we focus on the behaviour of users to test whether it features a non-trivial temporal dynamics. To this end, we consider the patterns of recurrence of songs, i.e. how recurrent listenings to the same song are distributed over time. A non-uniform temporal distribution of listenings to the same song may reveal changes in the preferences of users over time.

Let us consider a specific song S listened M times by a user during his/her whole listening experience. Here, time is the so-called intrinsic time, which is the simple count of the overall number of listenings in the whole history of the user. Be n_1 the intrinsic time of the first listening to S (i.e. if S is listened to for the first time as the 10th song of the user's history, $n_1 = 10$). Equivalently, n_2 is the intrinsic time of the second occurrence of S and so on. If we define $\Delta n_1 = n_2 - n_1$, we can use it to give a rough estimate of the probability for the user to listen to S in the time interval $[n_1, n_2]$ as:

$$P(S, [n_1, n_2]) \simeq 1/\Delta n_1$$

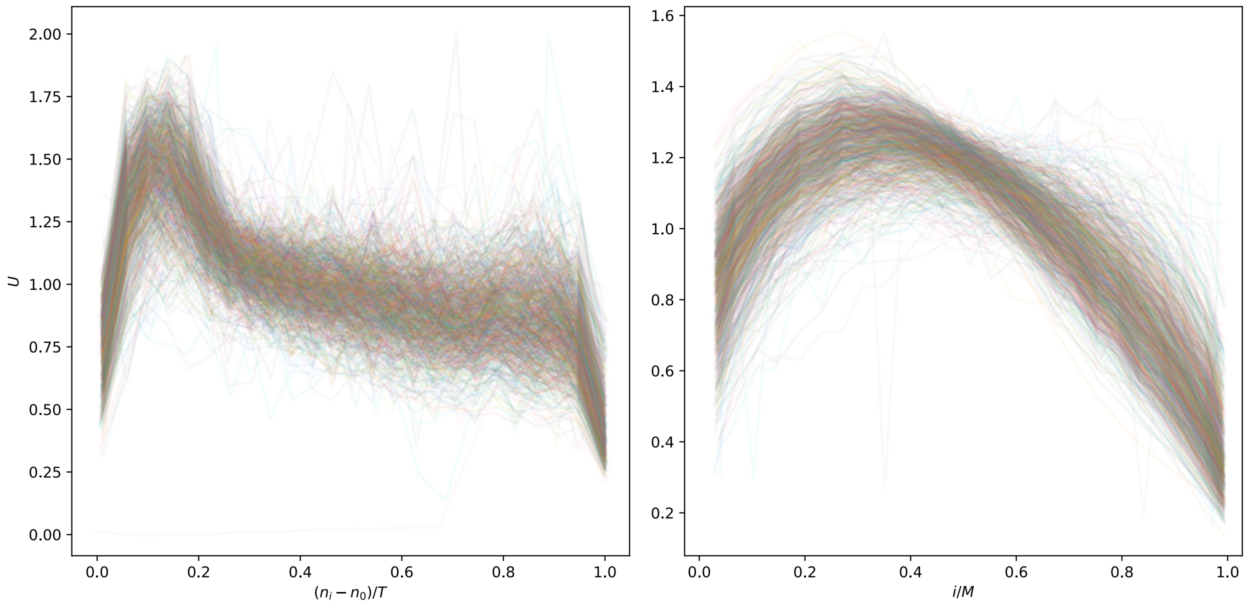

Fig. 3. The temporal probability unbalance $U(r_i)$ as function of the fraction of the time span elapsed $r_i = (n_i - n_1)/T_S$ (left) and as a function of the fraction of repetition elapsed i/M (right). Each line represents a user. The graphs represent a random sample of 1000 users. For each user, the curve is calculated by averaging the temporal behaviour of all the repetitions of the songs listened to by that user at least 10 times. The peak observed in the left panel for a value of $r_i \simeq 0.1$ indicates that, for a generic user, the maximum return probability occurs for around 10% of the whole listening experience. For instance, a user who listened to a given song $M = 100$ times in a time span of $T_S = 10\,000$ songs has the maximum return probability when he/she listened to around 1000 songs. An equivalent interpretation holds for the right panel.

and more generally:

$$P(S, n_i) \simeq 1/\Delta n_i \tag{2}$$

The above-defined quantity is subject to huge fluctuations from song to song and, to average out the fluctuations, we average over all songs listened to more than once by the same user. To do so, we have to keep in mind that each song is characterized by two quantities: the total number of repetitions M and the overall intrinsic time span $T = n_M - n_1$. We can thus rewrite $P(S, n_i)$ as a function of the fraction of the elapsed time span $r_i = (n_i - n_1)/T$. Further, in order to make the probability of repetition of S comparable among different songs listened to by the same user, we normalize $P(S, r_i)$ with the average probability during its time span $< P(S) >$. In this way, we define the unbalance of the return probability as:

$$U(S, r_i) = P(S, r_i)/ < P(S) > \tag{3}$$

Since $U(S, r_i)$ is normalized according to M and T, it allows for safely averaging over all the different songs listened to by the same user. By performing this average, we obtain the average of unbalance of the return probability for a generic user $U(r_i)$ as a function of the fraction of time, r_i, elapsed at the moment of *thei*-th listening to a generic song. For instance, if a song has been listened to $M = 50$ times in a time span of $T = 10\,000$, and its 10-th recurrence happens 100 songs after the first listening, $r_i = 100/10\,000$ of the overall timespan T.

The quantity $U(r_i)$ allows us to quantify how much the listening patterns of a given user deviates from uniformity. A flat shape of $U(r_i)$ would signal that the user features a constant repetition probability of listenings. This would confirm the hypothesis of staticity of user behaviour. On the other hand, any deviation from a pure flat shape of $U(r_i)$ would signal a temporal dynamic of user preferences. The left panel of Fig. 3 report the behaviour of $U(r_i)$ as a function of r_i for a subset of about 1000 users of the Last.fm dataset mentioned in the previous section.

The measurement shows a strikingly similar behaviour for all users: a non-trivial pattern of listening habits. We observe in particular a pronounced peak around $r_i \simeq 0.1$, i.e. around 10% of the listening experience for each song. In addition, the raise and fall around the peak are strongly asymmetric. This pattern reveals the existence of a non-trivial temporal dynamics of tastes. Songs appear to enter and exit the user's radar of attention at different speeds. A specific song does not actually exit the sequence of listenings. Their repetition rate becomes smaller and smaller over time after the peak, i.e. after the moment of maximum attention from the user. All these evidences point to a sort of "boredom effect": after the first moment of enthusiasm, a song, even though still attractive, becomes gradually less interesting according to a well-defined law.

In order to get a more precise picture, we can also look at the unbalance of the return probability U as a function of the fraction of repetitions, i/M, instead of r_i. The right panel of Fig. 3 reports the resulting curves. The result is reminiscent

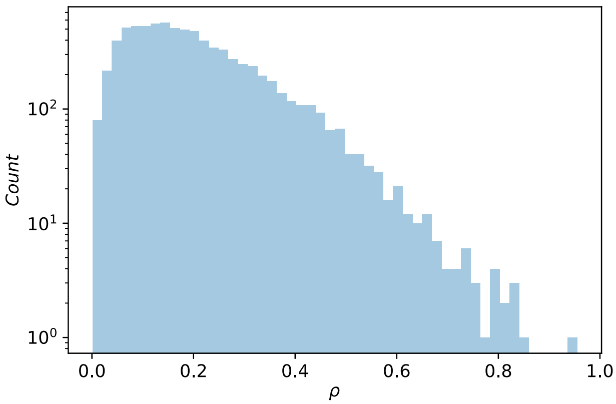

Fig. 4. The histogram of the ratio between the total number of distinct songs listened to by a user and the total number of his/her listenings in the Last.fm dataset. The values ρ are bounded in the interval $[0, 1]$ and its distributions display a strong heterogeneity.

of a learning dynamics [26]. The user needs some repetition to "learn" the song and get to a moment when he/she can really enjoy it (as signalled by very frequently repeated listenings around the maximum of U). After that phase, his/her attention fades out, possibly replaced by some other song, in a universal way for all users and all songs, and after the need for recalling it becomes less and less important.

This kind of dynamics reveals a highly non-trivial user behaviour that mixes many cognitive and emotional factors. Very often, these kinds of features are totally neglected when conceiving recommender systems. Even in the few cases of Collaborative Filtering algorithms that take into account memory effects, the memory is taken as uniform for the user, in striking contrast with the finding of Fig. 3. This implies that the natural tendency of users to explore new things and move on with respect to what they just experienced can be frustrated by recommender systems that take the past history of users as equally important to determine what will be next. Acknowledging instead that a learning dynamics is in place could be the key to conceive the next generation of recommender systems.

We conclude this section with a last remark. Though Fig. 3 points out some sort of universality in the learning-like dynamics, it does not imply that all users behave in the same way. On the contrary, each user has a peculiar inclination towards exploring and learning new things. To quantity this heterogeneity at the population level, we measured the explore/exploit ratio, which is simply defined as follows. For each user, one computes the ratio ρ between the total number of distinct songs listened to in the whole listening history and the total number of listenings. Fig. 4 shows the result of this measurement. ρ displays a wide distribution pointing out to a strong heterogeneity in the behaviours of the users. This heterogeneity suggests that the customization of recommendations should not be just revolving around new contents to be suggested. It should also take care of the overall personal balance between the exploration of new items and the exploitation of already experienced ones.

5. Conclusion and perspectives: a new theoretical framework

In this final section, we present some considerations that might be useful to redesign recommender systems in a way that better aligns with the natural inclination of human beings towards new and pleasant experiences. We claim that this new generation of recommenders could help in overcoming the pitfalls of current technologies, namely the tendency towards a lack of diversity, polarization, the emergence of echo-chambers, and misinformation.

While designing the next generation of recommender systems, one should take into account what we learned so far. Here is a non-exhaustive list.

Semantic spaces. One first point concerns a better description of the space of items. Semantic relations between items need to be grasped in a non-biased way, e.g., by removing the *potential biases* of popularity or homophily, as it happens for instance in Collaborative Filtering-based systems. In the literature, contrasting results about the diversity of recommendations might be related to different reconstructions of the semantic relationships. In Section 3, we provided hints of the possible effects of Collaborative Filtering, which might result in a poorer exploration and perception of the space of items from the point of view of the user. It is important to remark that our approach does not rely on any predefined notion of distance in the space of items. Rather we exploit the distribution of return probabilities in the sequence of listenings to each song, as witnessed in the Last.fm dataset.

User modeling. A second important point concerns the need to take into account in a more realistic way how humans acquire and react to information. This kind of aspects is widely studied, though hardly implemented in current recommender systems. In Section 4 we gave evidences of a learning-like dynamics in the way people navigate the space of items. We highlighted in particular a non-trivial temporal dynamics through which items trigger the interest of users, eventually leaving the radar of attention. We could describe this phenomenon as a continuous reshape of individual Comfort Zones. Everything goes as if a user "learns" a song for a while, progressively increasing the rate of listening to that song, to get

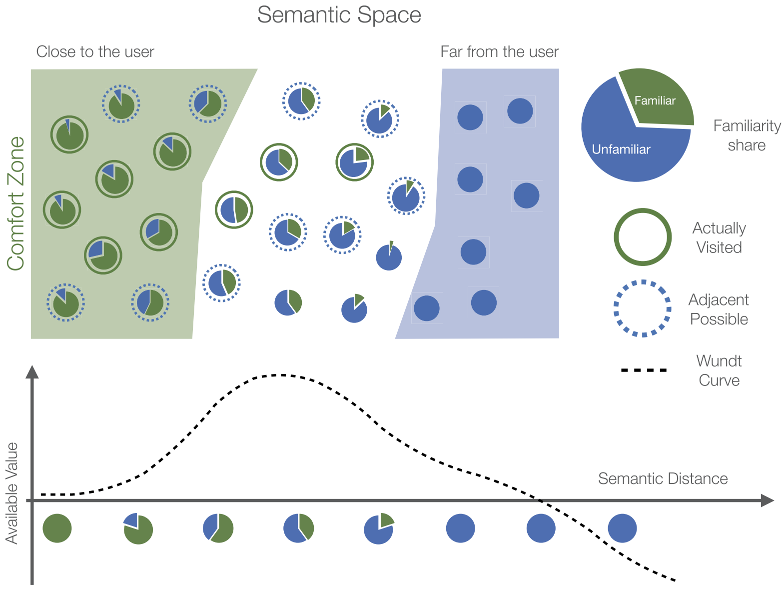

Fig. 5. A schematic representation of the conceptual framework for a new generation of recommender systems. In the upper part, we depict the space of items. Each item is depicted as a circle, whose content can be more or less familiar to the user. The green part quantifies the level of familiarity, while the blue one quantifies the level of unfamiliarity. Items close in space are also semantically similar. Items circled in green are those already experienced by the user at least once. Those circles in dotted blue belong to the *adjacent possible* space, i.e. the set of items the user can possibly experience given his/her past history. Items belonging to the *adjacent possible* space can still be familiar to the user because they could be semantically similar to something already experienced in the past. From this perspective, they could lie in the *comfort zone* of each user, depicted in green in figure, i.e. in the area of items very familiar to him/her. The curve in the lower panel is inspired by the famous Wundt curve and quantifies the value a user can associate with each item, as a function of how familiar the item is to the user. Items very close to the past experience of the user can be of little value for him/her. On the other hand, items very far from the user's experience are depicted in the blue area at the top right. Also those items are of little value to the user (as mirrored by the underlying curve). In between (white area), there are items that the user is not too unfamiliar with. Those items are potentially of great value for the user, as mirrored by the underlying curve.

to a climax moment where the listening rate to that song is maximum. After that, the song starts fading out in the user's interest. It is remarkable as this process appears to be universal, independently of a specific user or a specific song.

With these two ingredients in mind, we can make a first attempt to design what the new conceptual framework for recommendation could look like. Fig. 5 displays in a pictorial way a possible scheme. The experience of a generic user can be seen as the process of extracting value while exploring the space of items. Based on the his/her history, one can define how much a specific item is familiar or unfamiliar to the user, independently of whether the item has been experienced or not. When a user experiences a given item, his/her familiarity with it increases, but also that of similar items. Familiar items are somehow unsurprising and are located in the comfort zone of each user. The possibility to extract value from the experience of a given item depends on the familiarity of the user with that item. In order to give a rough quantification of the relationship between the familiarity of a given item and its perceived value, we refer to the famous Wundt curve [3]. In psychology, the Wundt curve is related to theories of human motivation and novelty seeking. The Wundt curve is bell-shaped and describes as the pleasantness of a given stimulus increases from low to moderate as the stimulus intensity increases. When the stimulus intensity increases too much, its effect is unpleasant and even painful. If we follow this line of reasoning, we can imagine that: (i) if the item is too far from the past user experience (i.e. outside his/her comfort zone) or if it is too close to the past experience (well within the comfort zone), little or no new value can be extracted. On the other hand, items neither too close nor too far away can trigger the curiosity of the user and allow for the extraction of a possibly significant value. The shape of the curve of the available value (i.e. the Wundt curve) as a function of the level of familiarity of the item is reported in the lower part of Fig. 5. This figure mirrors the empirical results obtained in 3, though with the X-axis reversed.

Using the same formalism as in Section 4, familiar means that the ratio between the number of encounters i with a given item is very close to its asymptotic value M is close to one, $i/M \sim 1$. On the other hand, for a still unfamiliar item $i/M \sim 0$.

The above-proposed picture implies that, in order to trigger an optimal response from users, recommenders should take into account the personal histories of the users and how new information may affect their inner state. To this end, a better reconstruction of the space of items is key to recommend items that could bring value to users both on the short and on the long term. This approach is reminiscent of what Reinforcement Learning does [30] in machine learning approaches.

As already mentioned, the adjacent possible space is dynamically reshaped while it is explored. So the experience of an item belonging to the adjacent possible produces a reorganization of the space of items, both in terms of the level of familiarity of the different items, but also in terms of new elements becoming possible at some later stage. The reconstruction of the space of items performed by the new generation of recommender systems should also take into account this dynamical reshaping, in order to find paths into the adjacent possible optimizing the amount of exploited value. The selection of the steps into the adjacent possible should aim not only at the optimization of the next step but also at unlocking new areas of unexploited value, that are likely to be semantically different from users previous experiences. Of course, such areas must be approached taking into the account the diversity acceptable by the users, but the general idea is to create long-lasting paths of value that might lead the user to experience and appreciate novelties and diversity.

In this way, the comfort zone of each user can be progressively enlarged, allowing him/her to nurture his/her tendency to novelty seeking, while keeping the level of pleasantness of the whole experience high. The interplay between the mechanism of learning and the restructuring of the adjacent possible space seems to be very promising ingredients for the next generation of recommender systems.

References

[1] E. Bakshy, S. Messing, L.A. Adamic, Exposure to ideologically diverse news and opinion on facebook, Science 348 (6239) (2015) 1130–1132.
[2] P. Barberá, J.T. Jost, J. Nagler, J.A. Tucker, R. Bonneau, Tweeting from left to right: is online political communication more than an echo chamber? Psychol. Sci. 26 (10) (2015) 1531–1542. PMID: 26297377.
[3] D.E. Berlyne, Novelty, complexity, and hedonic value, Percept. Psychophys. 8 (5) (1970) 279–286.
[4] blogs.cornell.edu. Last.fm – Music Recommendation incorporating social network ties and collaborative filtering, https://blogs.cornell.edu/info2040/2012/09/20/last-fm-music-reccomendation-incorporating-social-network-ties-and-collaborative-filtering/.
[5] Ò. Celma, P. Cano, From hits to niches?: Or how popular artists can bias music recommendation and discovery, in: Proceedings of the 2nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition, ACM, 2008, p. 5.
[6] D. Cosley, S.K. Lam, I. Albert, J.A. Konstan, J. Riedl, Is seeing believing?: How recommender system interfaces affect users' opinions, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Ft. Lauderdale, FL, USA, 5–10 April 2003, ACM, New York, 2003, pp. 585–592.
[7] United Nations Human Rights Council, Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, https://www.ohchr.org/EN/Issues/FreedomOpinion/Pages/ContentRegulation.aspx.
[8] M. Del Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H.E. Stanley, W. Quattrociocchi, The spreading of misinformation online, Proc. Natl. Acad. Sci. 113 (3) (2016) 554–559.
[9] S. Flaxman, S. Goel, J.M. Rao, Filter bubbles, echo chambers, and online news consumption, Public Opin. Q. 80 (S1) (03 2016) 298–320.
[10] World Wide Web Foundation, For the web, https://fortheweb.webfoundation.org/.
[11] P. Gravino, S. Caminiti, A. Sîrbu, F. Tria, V.D.P. Servedio, V. Loreto, Unveiling political opinion structures with a web-experiment, in: 1st International Conference on Complex Information Systems, COMPLEXIS 2016, Rome, Italy, 22–24 April, 2016, Science and Technology Publications, Lda, Setúbal, Portugal, 2016, pp. 39–47.
[12] P. Gravino, B. Monechi, V.D.P. Servedio, F. Tria, V. Loreto, Crossing the horizon: exploring the adjacent possible in a cultural system, in: Proc. of the Seventh Int. Conf. on Computational Creativity, Paris, France, 27 June–1 July 2016, 2016.
[13] K. Hosanagar, D. Fleder, D. Lee, A. Buja, Will the global village fracture into tribes? Recommender systems and their effects on consumer fragmentation, Manag. Sci. 60 (4) (2013) 805–823.
[14] D. Jannach, L. Lerche, F. Gedikli, G. Bonnin, What recommenders recommend–an analysis of accuracy, popularity, and sales diversity effects, in: S. Carberry, S. Weibelzahl, A. Micarelli, G. Semeraro (Eds.), Proc. 21st International Conference on User Modeling, Adaptation, and Personalization, UMAP 2013, Rome, Italy, 10–14 June 2013, Springer, 2013, pp. 25–37.
[15] S.A. Kauffman, Investigations, Oxford University Press, 2000.
[16] Y. Koren, The Bellkor solution to the Netflix grand prize, Netflix Prize Doc. 81 (2009) 1–10.
[17] Y. Koren, Collaborative filtering with temporal dynamics, in: Proc. 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '09, Paris, France, 28 June, 28–1 July 2009, ACM, New York, 2009, pp. 447–456.
[18] G. Linden, B. Smith, J. York, Amazon.com recommendations: item-to-item collaborative filtering, IEEE Internet Comput. 1 (1) (2003) 76–80.
[19] P. Lops, M. De Gemmis, G. Semeraro, Content-based recommender systems: state of the art and trends, in: Recommender Systems Handbook, Springer, 2011, pp. 73–105.
[20] M. McPherson, L. Smith-Lovin, J.M. Cook, Birds of a feather: homophily in social networks, Annu. Rev. Sociol. 27 (1) (2001) 415–444.
[21] B. Monechi, Ã. Ruiz-Serrano, F. Tria, V. Loreto, Waves of novelties in the expansion into the adjacent possible, PLoS ONE 12 (6) (2017) e0179303.
[22] T.T. Nguyen, P.-M. Hui, F. Maxwell Harper, L. Terveen, J.A. Konstan, Exploring the filter bubble: the effect of using recommender systems on content diversity, in: Proceedings of the 23rd International Conference on World Wide Web, WWW '14, Seoul, Korea, 7–11 April 2014, ACM, New York, 2014, pp. 677–686.
[23] T.P. Novikoff, J.M. Kleinberg, S.H. Strogatz, Education of a model student, Proc. Natl. Acad. Sci. USA 109 (6) (2012) 1868–1873.
[24] B. Pan, X.R. Li, The long tail of destination image and online marketing, Ann. Tour. Res. 38 (1) (2011) 132–152.
[25] E. Pariser, The Filter Bubble: What the Internet Is Hiding From You, Penguin, UK, 2011.
[26] G.C. Rodi, V. Loreto, V.D.P. Servedio, F. Tria, Optimal learning paths in information networks, Sci. Rep. 5 (2015) 10286.
[27] B. Sarwar, G. Karypis, J. Konstan, J. Riedl, Item-based collaborative filtering recommendation algorithms, in: Proceedings of the 10th International Conference on World Wide Web, WWW'01, Hong Kong, 1–5 May 2001, ACM, New York, 2001, pp. 285–295.
[28] J. Ben Schafer, D. Frankowski, J. Herlocker, S. Sen, Collaborative filtering recommender systems, in: The Adaptive Web, Springer, 2007, pp. 291–324.
[29] Federal News Service, President Obama's farewell address, https://www.nytimes.com/2017/01/10/us/politics/obama-farewell-address-speech.html.
[30] R.S. Sutton, A.G. Barto, Reinforcement Learning: An Introduction, MIT Press, 2018.
[31] F. Tria, V. Loreto, V.D.P. Servedio, S.H. Strogatz, The dynamics of correlated novelties, Sci. Rep. 4 (2014) 5890.
[32] R. Zhou, S. Khemmarat, L. Gao, The impact of youtube recommendation system on video views, in: Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement, IMC '10, 1–3 November 2010, ACM, New York, 2010, pp. 404–410.