

Improving collaborative filtering music recommender systems: A focus on user characterization from behavioral and contextual factors

Doctoral Thesis

Ph.D. PROGRAM IN COMPUTER ENGINEERING

University of Salamanca



**VNiVERSiDAD
D SALAMANCA**

Diego Sánchez Moreno

Salamanca, 2020



UNIVERSITY OF SALAMANCA

DOCTORAL THESIS

**Improving collaborative filtering music recommender systems:
A focus on user characterization from behavioral and contextual
factors**

**Mejora de los sistemas de recomendación de música de filtrado
colaborativo: Un enfoque en la caracterización del usuario a
partir de factores de comportamiento y contextuales**

Author: Diego Sánchez Moreno

Supervisor: María N. Moreno García

Salamanca, 2020

Statement of Authorship

Diego Sánchez Moreno presents the thesis work entitled “Improving collaborative filtering music recommender systems: A focus on user characterization from behavioral and contextual factors” to apply for the Doctorate Degree in Computer Engineering of the University of Salamanca, and states that it has been carried out under the supervision of Dr. María N. Moreno, Full Professor in the Department of Computing and Automation of the University of Salamanca.

Author:

Diego Sánchez Moreno

Supervisor:

María N. Moreno García

Acknowledgements

I want to start by thanking my supervisor at the University of Salamanca Professor María N. Moreno for her hard work throughout this process. This thesis would not have been possible without all her dedication.

I would also like to thank my colleagues in the University of Salamanca and the members of the Data Mining research group (MIDA) for their support along the years that I have been developing my research in this group. My gratitude also to Dr. Yong Zheng from the Illinois Institute of Technology in Chicago for allowing me to work under his supervision and learn from him during my stay in Chicago, in addition to his active collaboration in the publications that are part of this PhD work.

Last but most important, I would like to thank my parents Fernando and María for all their love, guidance and support throughout my entire life, and for providing me with the best possible education both professionally and personally, as well as to my girlfriend María for being an indispensable part in all my life accomplishments.

Abstract

The popularization of digital distribution of multimedia content, known as streaming, allows more and more users to access almost all existing music from anywhere without the limitation of the storage capacity of the devices. This enormous availability, as well as the great variety of providers of these services, makes it very difficult for users to find music that can fit their tastes. Hence, the great current interest in developing recommendation algorithms that help users to filter and discover the music that fits their preferences from the enormous amount of music content available in the digital space. Most platforms have search services, and some of them have recommendation mechanisms and offer personalized playlists, but many improvements are still required.

The methods used in recommender systems are very diverse, although those based on collaborative filtering (CF) are among the most extended. The recommendations they provide are based on the ratings that users assign to items to be recommended, which in the case of music recommender systems are songs or artists. The recommendations for a given user are based on the ratings of other users with similar tastes. The results of this type of technique are quite good, however, the difficulty of obtaining explicit evaluation of the items by users makes the number of ratings insufficient, causing problems of *sparsity*, which prevent or hinder the application of such methods. For this reason, on some occasions implicit ways of obtaining such information are used, which are usually complex and not always effective. Other problems caused by the scarcity of ratings associated with new users or new products in the system are *cold start* and *first-rater*, respectively. In addition, it is difficult to provide reliable recommendations to users with unusual tastes (*gray sheep* users). A related drawback is the popularity bias that results in the most popular items being more likely to be recommended. This is due to the power law distribution of the frequency of plays of musical items (artists or songs) since high frequencies of plays are concentrated in very few items, while the remaining ones are part of the long tail of the curve.

To address the above problems, content-based algorithms have been proposed as an alternative to CF methods. These methods can be used to recommend any item using its characteristics, so that the user receives recommendations of items similar to others for which he or she has shown interest in the past. Most current recommender systems use hybrid techniques to take advantage of the benefits of both approaches and avoid their drawbacks. These methods usually make use of item and user attributes, as well as rating information.

This work focuses on user characterization in order to increase the degree of personalization and thus improve the recommendations provided by collaborative filtering methods. The proposals presented, although they could be extended to other application domains, are focused on the field of music because the way music is consumed differs significantly from other products and, consequently, some aspects related to recommendations are also different.

The approaches proposed to characterize the user have in common the fact that they require only the information available on the streaming music platforms, without the need for any

additional data such as user demographics or item attributes. Besides the fact that there are no explicit ratings of the music items and it is necessary to obtain them implicitly from the plays of artists or songs by each user.

The first proposal addresses the gray sheep problem by characterizing the user according to the popularity of the music he or she listen to, which is closely related to the power law distribution of the item play frequency. This approach is applicable for both artist and song recommendations. In the latter case, recommendations are improved by considering the position of the songs in the user's sessions.

Time is another important factor related to user behavior and habits. The proposal to improve recommendation methods in relation to this factor is addressed from three user-centered perspectives: modeling both the evolution of user preferences and user listening habits over time, and using time as a contextual variable to generate context-aware recommendations. The preference evolution model is involved in the process of obtaining implicit ratings.

Another way to characterize the users is through their social context. Streaming music platforms do not have much information of this kind. However, available data on friendship connections and social tags can be used for this purpose. In particular, this information has been used in this work to model their degree of influence through the properties of trust and homophily, and their level of expertise respectively.

Although the methods presented are not specifically designed to address the cold start drawback, some of them have been tested for that scenario, showing that they also contribute to minimize that problem.

Resumen

La popularización de la distribución digital de contenido multimedia, conocido como *streaming*, permite a cada vez más usuarios el acceso a prácticamente toda la música existente desde cualquier lugar sin la limitación de la capacidad de almacenamiento de los dispositivos. Esa enorme disponibilidad, así como la gran variedad de proveedores de estos servicios hace muy difícil al usuario encontrar música que pueda encajar en sus gustos. De ahí deriva el gran interés actual por el desarrollo de algoritmos de recomendación que ayuden al usuario a filtrar y descubrir la música que se ajusta a sus preferencias a partir de la enorme cantidad de contenido musical disponible en el espacio digital. La mayoría de las plataformas disponen de servicios de búsqueda y algunas de ellas disponen de mecanismos de recomendación y ofrecen listas personalizadas de reproducción (*playlists*), aunque todavía se requieren muchas mejoras.

Los métodos utilizados en los sistemas de recomendación son muy variados, aunque los basados en filtrado colaborativo (FC) se encuentran entre los más extendidos. Las recomendaciones que proporcionan se basan en las valoraciones (*ratings*) que los usuarios hacen de los ítems a recomendar, que en el caso de los sistemas de recomendación de música son canciones o artistas. Las recomendaciones para un usuario dado se basan en las valoraciones realizadas por otros usuarios con gustos similares a él. Los resultados de este tipo de técnicas son bastante buenos, sin embargo, la dificultad de obtener la evaluación explícita de los ítems por parte de los usuarios hace que el número de valoraciones sea insuficiente, causando problemas de dispersión (*sparsity*), que impiden o dificultan la aplicación de tales métodos. Por este motivo, en algunas ocasiones se recurre a formas implícitas de obtener dicha información, las cuales son usualmente complejas y no siempre son efectivas. Otros problemas causados por la incorporación de nuevos usuarios o nuevos productos en el sistema son los de arranque en frío (*cold start*) y primera valoración (*first rater*) respectivamente. A esto hay que sumar la dificultad para ofrecer recomendaciones fiables a usuarios con gustos inusuales (*gray sheep users*).

Para hacer frente a los problemas anteriores se han propuesto algoritmos basados en el contenido como alternativa a los métodos de CF. Estos métodos pueden utilizarse para recomendar cualquier ítem haciendo uso de sus características, de manera que el usuario recibe recomendaciones de ítems similares a otros por los que ha mostrado interés en el pasado. La mayoría de los sistemas de recomendación actuales utilizan técnicas híbridas destinadas a aprovechar las ventajas de ambos enfoques y evitar sus inconvenientes. Estos métodos hacen uso de atributos de ítems y usuarios, además de información de valoraciones.

Este trabajo se centra en la caracterización del usuario con el fin de aumentar el grado de personalización y así mejorar las recomendaciones proporcionadas por los métodos de filtrado colaborativo. Las propuestas que se presentan, aunque pudieran hacerse extensivas a otros dominios de aplicación, se centran en el ámbito de la música debido a que la forma de consumo de la música difiere significativamente de la forma de consumir otros productos y, en consecuencia, algunos aspectos relativos a las recomendaciones también son diferentes.

Los diferentes enfoques propuestos para caracterizar al usuario tienen en común el hecho de requerir únicamente la información disponible en las plataformas de música en *streaming*, sin

necesidad de ningún dato adicional como puede ser información demográfica de los usuarios o atributos de los ítems. Además del hecho de no disponer de valoraciones explícitas de los ítems de música y tener que obtenerlos implícitamente a partir de las reproducciones de artistas o canciones por parte de cada usuario.

La primera propuesta aborda el problema de la oveja negra mediante la caracterización del usuario en función de la popularidad de la música que escucha, lo que está estrechamente relacionado con la distribución de ley de potencia de la frecuencia de reproducción de los ítems. Este enfoque es aplicable tanto para la recomendación de artistas como de canciones, y en este último caso, las recomendaciones se pueden mejorar teniendo en cuenta la posición de las canciones en las sesiones del usuario.

El tiempo es otro factor importante relacionado con el comportamiento y los hábitos del usuario. La propuesta de mejora de los métodos de recomendación en relación con este factor se aborda desde tres perspectivas centradas en el usuario: modelado tanto de la evolución de sus preferencias, como de sus hábitos de escucha en función del tiempo, y uso del tiempo como variable contextual para generar recomendaciones sensibles al contexto. El modelo de evolución de preferencias está incluido en el proceso de obtención de calificaciones implícitas.

Otra forma de caracterizar al usuario es a través de su contexto social. Las plataformas de música en *streaming* no disponen de mucha información de este tipo. Sin embargo, los datos disponibles sobre relaciones de amistad y etiquetado social se pueden utilizar para este propósito. En concreto, esta información se ha utilizado en este trabajo para modelar su grado de influencia, a partir de las propiedades de confianza y homofilia, y su nivel de conocimiento (*expertise*) respectivamente.

Aunque los métodos presentados no están diseñados específicamente para abordar el inconveniente del arranque en frío, algunos de ellos se han probado en este escenario, mostrando que también contribuyen a minimizar ese problema.

Contents

1. MODALITY OF THE THESIS.....	1
1.1. SUPERVISOR AUTHORIZATION	1
1.2. LIST OF CONTRIBUTIONS	2
2. OVERVIEW OF THE CONTRIBUTIONS.....	5
2.1. INTRODUCTION	5
2.2. OBJECTIVES	6
2.3. STATE OF THE ART	7
2.3.1. <i>Gray sheep problem</i>	7
2.3.2. <i>Time-aware recommender systems</i>	8
2.3.3. <i>Using social information in recommender systems</i>	9
2.4. PROPOSALS	10
2.4.1. <i>Dealing with the gray sheep problem</i>	11
2.4.2. <i>Modeling time effects on user preferences</i>	12
2.4.3. <i>Inferring user characteristics from social information</i>	13
2.5. VALIDATION	15
2.6. CONCLUSIONS AND FUTURE WORK	15
2.7. REFERENCES	16
3. RESUMEN EN ESPAÑOL DE LAS CONTRIBUCIONES	23
3.1. CONTRIBUCIÓN 1	23
3.1.1. <i>Referencia</i>	23
3.1.2. <i>Objetivos</i>	23
3.1.3. <i>Método de recomendación</i>	23
3.1.4. <i>Resultados</i>	25
3.1.5. <i>Conclusiones</i>	25
3.2. CONTRIBUCIÓN 2	26
3.2.1. <i>Referencia</i>	26
3.2.2. <i>Objetivos</i>	26
3.2.3. <i>Metodología propuesta</i>	26
3.2.4. <i>Resultados</i>	27
3.2.5. <i>Conclusiones</i>	28
3.3. CONTRIBUCIÓN 3	28
3.3.1. <i>Referencia</i>	28
3.3.2. <i>Objetivos</i>	28
3.3.3. <i>Modelado del efecto del tiempo en las recomendaciones</i>	29
3.3.4. <i>Resultados</i>	30
3.3.5. <i>Conclusiones</i>	31
3.4. CONTRIBUCIÓN 4	31
3.4.1. <i>Referencia</i>	31
3.4.2. <i>Objetivos</i>	31
3.4.3. <i>Uso del contexto social del usuario en los métodos de recomendación</i>	32
3.4.4. <i>Resultados</i>	33

3.4.5. Conclusiones.....	34
3.5. CONTRIBUCIÓN 5	34
3.5.1. Referencia	34
3.5.2. Objetivos	34
3.5.3. Uso de etiquetado social en los métodos de recomendación	35
3.5.4. Resultados.....	36
3.5.5. Conclusiones.....	36



1. Modality of the thesis

The presentation of this doctoral thesis at the University of Salamanca is done in the format of a compendium of previously published articles. The thesis includes five contributions: four articles published in journals and a book chapter.

1.1. Supervisor authorization

María N. Moreno García, Full Professor in the Department of Computer Science and Automation of the University of Salamanca and supervisor of the doctoral thesis of Mr. Diego Sánchez Moreno,

AUTHORIZES

To Mr. Diego Sánchez Moreno to present and defend his doctoral thesis in the form of a compendium of articles.

María N. Moreno García



1.2. List of contributions

The following is a list of published articles with their reference data and quality indexes.

CONTRIBUTION 1

Sánchez-Moreno D., Gil A.B., Muñoz M.D., López V.F. and Moreno M.N. "A collaborative filtering method for music recommendation using playing coefficients for artists and users", *Expert Systems with Applications*, 66 (2016), 234-244, 2016.

DOI: 10.1016/j.eswa.2016.09.019

Authors:

- Diego Sánchez Moreno, University of Salamanca, Spain.
- Ana Belén Gil González, University of Salamanca, Spain.
- María Dolores Muñoz Vicente, University of Salamanca, Spain.
- Vivian F. López Batista, University of Salamanca, Spain.
- María N. Moreno García, University of Salamanca, Spain.

Journal: Expert Systems with Applications (ISSN: 0957-4174)

Quality indexes:

- **WoS JCR¹ Impact Factor 2016:** 3.928. **Rank:** 18/133 (Q1). **Area:** Computer Science-Artificial Intelligence.
- **SCOPUS Cite Score 2016:** 8.2. **Rank:** 15/160 (Q1). **Area:** Artificial Intelligence.

CONTRIBUTION 2

Sánchez-Moreno D., Muñoz M.D., López V.F., Gil A.B., Moreno M.N. "A session-based song recommendation approach involving user characterization along the play power-law distribution", *Complexity*, Article ID 7309453, 2020.

DOI: 10.1155/2020/7309453

Authors:

- Diego Sánchez Moreno, University of Salamanca, Spain.
- María Dolores Muñoz Vicente, University of Salamanca, Spain.
- Vivian F. López Batista, University of Salamanca, Spain.
- Ana Belén Gil González, University of Salamanca, Spain.
- María N. Moreno García, University of Salamanca, Spain.

Journal: Complexity (ISSN: 1099-0526)

Quality indexes:

- **WoS JCR Impact Factor 2019:** 2.462, **Rank:** 28/106 (Q2). **Area:** Mathematics-Interdisciplinary Applications.
- **SCOPUS Cite Score 2019:** 3.2. **Rank:** 54/221 (Q1). **Area:** General Computer Science.

¹ Web of Science Journal Citation Reports



CONTRIBUTION 3

Sánchez-Moreno D., Zheng, Y. Moreno M.N. "Time-aware music recommender systems: Modeling the evolution of implicit user preferences and user listening habits in a collaborative filtering approach", *Applied Sciences*, Special Issue on Recommender Systems and Collaborative Filtering, 10 (15), 5324, 33 pages, 2020.

DOI: 10.3390/app10155324

Authors:

- Diego Sánchez Moreno, University of Salamanca, Spain.
- Yong Zheng, Illinois Institute of Technology, Chicago, USA.
- María N. Moreno García, University of Salamanca, Spain.

Journal: Applied Sciences (ISSN 2076-3417)

Quality indexes:

- **WoS JCR Impact Factor 2019:** 2.474. **Rank:** 32/91 (Q2). **Area:** Engineering-Multidisciplinary.
- **SCOPUS Cite Score 2019:** 2.4. **Rank:** 299/636 (Q2). **Area:** Computer Science Applications.

CONTRIBUTION 4

Sánchez-Moreno D., López V.F., Muñoz M.D., Sánchez, A.L., Moreno M.N. "Exploiting the user social context to address neighborhood bias in collaborative filtering music recommender systems", *Information*, Artificial Intelligence Section, 11(9), 439, 16 pages, 2020.

DOI: 10.3390/info110904392020

Authors:

- Diego Sánchez Moreno, University of Salamanca, Spain.
- Vivian F. López Batista, University of Salamanca, Spain.
- María Dolores Muñoz Vicente, University of Salamanca, Spain.
- Ángel Luis Sánchez Lázaro, University of Salamanca, Spain.
- María N. Moreno García, University of Salamanca, Spain.

Journal: Information (ISSN: 2078-2489)

Quality indexes:

- **SCOPUS Cite Score 2019:** 2.4. **Rank:** 157/300 (Q3). **Area:** Computer Science-Information Systems.
- **Web of Science:** Emerging Sources Citation Index



CONTRIBUTION 5

Sánchez-Moreno, D., Moreno-García M.N., Sonboli, N., Mobasher, B. and Burke, R. “Inferring user expertise from social tagging in music recommender systems for streaming services”, Hybrid Artificial Intelligence Systems, Lecture Notes in Artificial Intelligence, Series: *Lecture Notes in Computer Science*, pp. 39-49, Springer, 2018.

DOI: 10.1007/978-3-319-92639-1_4

Authors:

- Diego Sánchez Moreno, University of Salamanca, Spain.
- María N. Moreno García, University of Salamanca, Spain.
- Nasim Sonboli, University of Colorado Boulder, USA.
- Bamshad Mobasher, DePaul University, Chicago, USA.
- Robin Burke, University of Colorado Boulder, USA.

Book: Hybrid Artificial Intelligence Systems (ISBN: 978-3-319-92639-1)

Series: Lecture Notes in Computer Science

Quality indexes:

- **SCOPUS Cite Score 2018:** 1.6. **Rank:** 95/211 (Q2). **Area:** General Computer Science.



2. Overview of the contributions

2.1. Introduction

Current music platforms make available to users a large amount of songs through their streaming services. This great advantage also carries with it the drawback of the difficulty that users have in finding music that they might like. Thus, the overload problem, which has its origin in the context of information retrieval, have been extended to these systems. To tackle it, many platforms provide users with advanced search functions and some of them are endowed with recommendation mechanisms that are the result of the intensive research that has been carried out for some time in this area.

Recommender systems are present in many domains where their role is to predict users' preferences in order to suggest items that might be of interest to them. The recommendation methods they implement are very varied, although Collaborative Filtering (CF) is one of the most successful approaches. It is based on ratings that users give to items or on the implicit feedback obtained from users' interactions with the items. This kind of methods can be classified in two main groups, user-based CF, where the predictions for the active user are based on the opinions of other users with similar preferences, and item-based CF, which provide recommendations by computing similarities between items. These similarities are also obtained from the ratings given by users. In the first case, a given user receives recommendations for items that similar users have rated positively, while in the second case, recommended items are similar to those that he/she has rated positively.

Collaborative filtering methods are not exempt from problems despite their success. Some of them derive from the difficulty in obtaining enough ratings to produce reliable recommendations. This is known as the *sparsity* problem, which occurs mainly when users are required to explicitly express their preferences about items. Additional drawbacks presented by CF methods are *early-rater (first-rater)* and *cold start* problems. The first one takes place when new products are introduced in the system. These items have never been rated; therefore, they cannot be recommended. Cold start affects new users, who cannot receive recommendations since they have no evaluations about items. *Gray sheep* users also generate a problem in recommender systems. These are users who have unusual preferences, and therefore they do not have enough similar users for computing recommendations. A related drawback is the popularity bias that results in the most popular items being more likely to be recommended. This is due to the power law distribution of the frequency of plays of musical items (artists or songs) since high frequencies of plays are concentrated in very few items, while the remaining ones are part of the long tail of the curve.

Content-based algorithms have been proposed as an alternative to CF methods in order to deal with the shortcomings discussed previously. These methods can be used for recommending any



kind of item by making use of its features. Thus, new items can be recommended according to their similarity to other items for which the user showed interest in the past. However, most current recommender systems use hybrid techniques aiming at taking advantage of the strengths of both approaches and avoiding their drawbacks. These methods take into account the preferences of other users as well as the characteristics of items and users (age, gender, occupation...). Therefore, new and gray sheep users can receive recommendations depending on their characteristics.

There are many proposals in the literature aimed at improving recommendation methods by addressing the problems described above. However, there are other factors that influence recommendations and should be taken into account in order to enhance them. A very important factor is time since it can influence user preferences in different ways. They can evolve over time and can change because of the temporal context (time of day, workday or weekend, season, etc.). Systems considering these temporal effects are commonly known as time-aware recommender systems (TARS).

Other user-related factors such as expertise or influence, that can contribute to improve the recommendations, can be inferred from their social context and their behavior in social networks. Social networks are currently the focus of intensive research, as they are a great source of information that can be exploited in many areas, including recommender systems. The incorporation of some social functionalities to the music streaming platforms, such as tagging or friendship connections, has allowed its use in this specific domain.

There are currently multiple directions of research aimed at improving recommender systems. In this thesis work we have focused on proposals aimed at characterizing the user to increase the degree of personalization of the recommendations. In addition, the proposals are addressed to the music domain, and provide solutions suitable to this specific field, specially to recommender systems for music streaming services. Some aspects of music recommendation are different from other fields; therefore, approaches used in other domains may be difficult to be reused or extended to this area. This is mainly due to differences in the way music is consumed compared to other products, but also in the effect of contextual and other factors on music preferences. For instance, the time of day can have little influence when choosing a book but can have a significant influence when choosing a song.

In addition, the proposed approaches do not need information about items or users as in content-based and hybrid methods, but only use play data and social information available on the streaming platforms. Nor is it required that the users explicitly express their preferences, but rather the ratings are calculated implicitly from the plays.

2.2. Objectives

The goal of this work is to develop user-centered approaches to improving recommender systems by increasing the degree of personalization of recommendations. Since the work is focused on the music field and more specifically on music streaming services, the purpose is to make use of the information available on the music platforms to model user behavior and social relationships in order to improve recommendations.



The specific objectives are as follows:

- Modeling user behavior in terms of the popularity of the music they listen to in order to deal with the gray sheep problem as well as with the long tail or power law distribution bias.
- Inferring ratings from implicit user feedback in which user sessions are considered.
- Modeling the influence of the time factor on user preferences from the point of view of both their evolution and the user's listening habits, also taking into account changes in preferences according to the time context.
- Taking advantage of the user social context to infer characteristics such as expertise, trust or homophily, which somehow reflect the influence of some users on the preferences of others in the social network.

2.3. State of the art

2.3.1. Gray sheep problem

Although there is much work in the literature aimed at addressing the problems of CF recommender systems, there are fewer works specifically focused on improving recommendations for the gray sheep users. However, it has been proven that the existence of many individuals with unusual preferences might have an important impact in the recommendation quality of the entire community [1]. Content-based and hybrid methods, as previously described, can produce some improvement but are not usually very significant. Moreover, they require additional information that may not be available. Semantic Web Mining can also be used to solve gray sheep and other typical problems of recommender systems. Semantic information is added to the existing data in order to formalize and classify product and user features. In this way, content-based models at different abstraction levels can be generated to provide recommendations based on those taxonomies. They can be combined with other approaches in order to improve recommendations [2, 3, 4, 5]. The main drawback of these approaches is their complexity and difficult extensibility since every application domain would involve the time-consuming task of defining a specific ontology.

Clustering is an alternative and simpler procedure to treat users with few neighbors [6]. In [1], an extensive review of recommender systems based on diverse clustering techniques is reported. The work also includes a new proposal involving the application of the k-means algorithm to generate clusters in order to detect the gray sheep users and a recommendation procedure for them based on their profiles. In some works, the clustering technique is used to address the sparsity and gray sheep problems at the same time since some authors consider that both problems are related. In [7] and [8], fuzzy class association rules are induced from previously clustered data in order to assign more than one cluster to each user with different degrees of belonging. A simulated scenario for gray sheep users proved the effectiveness of the method. The process, implemented in a tourist system, is not simple and requires user and items features. The Last.fm dataset is used in [9] to validate a hierarchical agglomerative clustering method for recommending resources in folksonomies, which considers the users' current navigation context



in cluster selection. As far as we know, most of the methods proposed for dealing with the gray sheep problem make use of user and/or item attributes.

2.3.2. Time-aware recommender systems

Time influence on user preferences is other aspect widely studied in the literature. Systems considering these temporal effects are commonly known as time-aware recommender systems (TARS). They have been classified according to the type of influence being considered, although the categories established by different authors are not the same. Most of these systems differentiate between the evolution of preferences over time and their change as a consequence of the temporal context (time of day, workday or weekend, season, etc.).

Within CF approaches, models to address time effects can be classified in three groups [10]: recency-based models, periodic context-based models and models that explicitly use time as an independent variable. Recency-based models give more importance to recent ratings than to older ones in CF. Depending on the way the ratings are managed; this category can be subdivided in window-based and decay-based models. Periodic context-based models are context-aware approaches where the only contextual dimension is time, which is split into intervals corresponding to diverse temporal contexts (morning/evening, weekend/workday, winter/spring/summer/autumn, etc.). The third group includes more advanced methods for modeling user and item trends.

In [11], three categories of time-aware recommender systems are defined: continuous time-aware, categorical time-aware, and time-adaptive methods. The first category is analogous to recency-based models that capture the evolution of preferences. It is usually done by means of functions that increase the weight of the ratings as they approach the target recommendation time. Categorical time-aware systems implement temporal context-aware methods. In time-adaptive systems, the more recent the ratings, the more important they are, which is similar to continuous time-aware methods, but without considering any target recommendation time.

Proposals for addressing preference evolution commonly require explicit ratings and information about the time when they were obtained from users. Timestamped ratings are necessary to adjust their weight [12], using either decay functions or time windows. In the group of recency-based models, decay-based models usually involve exponential functions of time to compute a weight that represents the importance of every rating at the recommendation time [10, 13]. In [14], a recency-based model for recommending songs is proposed to capture the changes in user preferences over time when explicit ratings are not available. In window-based models, a time interval preceding the target time is established. Then, ratings with a timestamp outside the time window are discarded [15, 16]. The size of the window depends on the application domain. There are also some approaches that combine time windows and decay-based models [17,18].

Using time as an independent variable allows us to model preference trends with respect to both users and items. Although time-series methods can be applied, the most extended approaches are temporal factor models [10]. One of the most known methods in this group is time-SVD++ [19], which introduces a modification in the original SVD++ matrix factorization method. Time effects are included in some terms already defined in SVD++ (user biases, item biases and user



factors), but are handled as time-dependent variables. Matrix factorization is also used in other proposals in different application domains [20, 21, 22, 23].

Studies that deal with time as a contextual variable are much more numerous [24, 25]. In many of them time is combined with additional contextual factors, such as place and company for film recommendation [26], or location and weather for recommending activities or tourist points of interest [27,28]. In the field of music, contextual aspects have a particular importance with respect to the type of music listened to by the user. In this regard, more and more methods are emerging that allow recommender systems to adapt to specific contexts [29, 30]. One of these is the Lifetrak system [31], which generates a playlist from the user's music library based on factors related to the user's physical context such as the location, time, day, traffic, level of noise, temperature and weather. In most of the work involving this category, time is only studied from the point of view of the context, without considering time dynamics.

2.3.3. Using social information in recommender systems

The objective generally pursued in studies that exploit social information is to improve recommendations by including social data processing into the rating prediction method so that the predicted value is closer to the actual value. This is mostly done by modifying either neighborhood-based CF techniques [32, 33] or matrix factorization methods [34].

The concept of social trust is the most studied in the literature about recommender systems. It is usually used to give more relevance to the ratings of trusted users against others [35] since it can be considered as a form of social influence that is often obtained from friendship connections, comments, messages, etc. Some systems allow users to explicitly express their trust on opinions, reviews, comments given by other users, but, in most cases, this is not possible, and it is necessary to infer it implicitly [10]. Social trust can be used locally when only opinions of connected friends are taken into account, and globally, when reputed individuals in the entire network are considered [36]. On the other hand, some approaches use social trust without considering similarity between users, but in others, it is used jointly with similarity values [37, 38] or even with additional factors, such as different types of interactions in social networks [36]. There are many works in the literature where diverse factors affecting social influence are addressed, but most of them are focused on social networks like Facebook or Twitter, from which a great variety of social information can be extracted.

Homophily and trust are two related concepts [39]. The effect of homophily can even be used for trust prediction [40], although homophily effects have been less studied and are often included in the general study of social influence without explicitly differentiating. Some recent work analyses the influence of homophily on consumers' purchasing decisions in the context of YouTube and Instagram influencers' popularity [41, 42]. However, in these works, homophily is treated as a complex factor that encompasses aspects like attitude, background, morality and appearance. Therefore, it cannot only be inferred from the structure of social relations. In the area of recommender systems, the study of homophily is much scarcer. In [43], recommendations of tourist attractions are generated by classifying users into several types, depending on factors such as homophily. This factor is determined by the membership of users in social communities.



Although trust and homophily principles have been much less studied in the field of music recommendation, we can highlight the work of Fields et al. [44], where music recommendations are based on the social relevance of musical data obtained through complex network technologies. A different objective is pursued in [45], in which the factors influencing the music listening homophily are analyzed. The analysis includes social information and user demographic attributes. None of these studies have addressed the problem of bias in the recommendations.

Both the use of social information and dealing with bias is the focus of much recent research in recommender systems, but there is little work in the literature in which both topics are addressed together. Regarding the work facing recommendation bias, the main proposals involve data preprocessing as resampling or clustering or postprocessing procedures as reranking, as set out below.

Bias in machine learning models is a widely studied and discussed problem that can be seen from different perspectives. Several types of bias have been studied in the recommender systems area, although most are related to unfair recommendations, from race or gender discrimination [46, 47] to popularity bias [48]. In the former, the problem is usually addressed through recommendation algorithms that are sensitive to this bias and focus on the protection of discriminated groups. Popularity bias is mainly associated with neighborhood-based methods, the most frequently used, and is one of the major concerns of recent research in this field. There are proposals for facing this problem that focus on improving recommendations for gray sheep users [49, 50] while others are focused on increasing the recommendations of the less frequently rated items and improving item diversity [51, 52, 53, 54]. There are studies that address aggregate diversity that refers not only to diversity of individual recommendations but also across recommendations of all users [55, 56]. The aim of these studies is to improve diversity while maintaining accuracy or with a minimum loss of it.

Concerning the topic of user expertise, the main application in the field of recommender systems is modeling the evolution of users' preferences as they gain knowledge. In domains such as CQA (Community Question Answering) and online reviews, most of the works make use of voting scores that users give to posts, answers and so on [57], [58]. In some of them, directed relationships between asker and answerer users are also considered [59]. There are some works in which domain specific ontologies are defined for inducing expertise considering the quantity and diversity of products of different subcategories rated by the users [60]. In our work, expertise is inferred from social tags. In prior research, social tags have been used for obtaining music similarity [61, 62], preferences of users for specific types of music [63] or learning latent features for users and items in matrix factorization approaches [64]. As far as we know, no proposal in the literature derives user expertise from social tags.

2.4. Proposals

All the proposals included in the thesis are aimed at improving collaborative filtering music recommender systems from different perspectives but focusing on characterizing users to achieve a greater degree of personalization in the recommendations provided to them.

Another common aspect is the absence of explicit ratings, thus implicit ratings are calculated from the frequency of plays. Some methods for computing ratings from plays have been



proposed and most of them are based on simple frequency functions [65] [66]. However, these methods are not indicated in the context of artist/song recommendation since play frequencies have a clear power law distribution, also known as the “long tail” distribution. In this scenario, it is more suitable to calculate them with a function of the play frequency percentile for each user [67]. In the case of song recommendation, we have also proposed a calculation procedure based on the position of the songs in the user sessions. This proposal is included in contribution 2. Moreover, a method to compute implicit ratings that involves the evolution of user preferences over time is presented in contribution 3.

Although the approaches presented are not specifically focused on addressing the cold start problem, the fact of using implicit ratings minimizes it. In addition, some of the proposals (contributions 3 and 4) have been validated for the cold start scenario, showing that they contribute to reducing this drawback.

The following is a summary of the contributions of this thesis work organized according to the specific topic addressed by each one.

2.4.1. Dealing with the gray sheep problem

The gray sheep problem, suffered by users with unusual tastes, is addressed in contributions 1 and 2. The first focuses on the recommendation of artists and the second on the recommendation of songs. In the latter, the proposal to obtain implicit ratings considering the user sessions is also presented.

Although there are many proposals in the literature for dealing with the weaknesses of recommender systems, the gray sheep problem has received less attention, and is mainly addressed by means of hybrid approaches that involve some content-based technique. Thus, these methods require information about items and users that often is not available. In the context of music recommender systems, some content-based filtering algorithms use musical content for inducing the models; therefore, a complex extraction task of music features is necessary. In contrast, our proposed recommendation methodology addresses the above-mentioned drawback when little information is available. The recommendation process could be incorporated into any music platform as long as it stores user and artist identification and the number of times the user plays a song in the platform, without the need for collecting rating data.

The key feature of the recommendation algorithm is the characterization of artists and songs depending on their popularity as well as the characterization of users according to the music they listen to. We consider that gray sheep users, who have unusual preferences, are those who mostly play unpopular music (artists or songs). However, we neither consider a division line between the types of users nor create a group for gray sheep users in order to manage them in a different way. The aim of the proposed procedure is to compute a coefficient for all users that characterizes them depending on the popularity of the music they play most. Popularity of the artists or songs is based on a listening coefficient that involves the number of users who play them and the number of plays they have.



A second coefficient related to users, the User Playing Coefficient (UPC), is proposed, aimed at characterizing them according to the popularity of the music they play. The computed coefficient UPC is used as a user attribute in user-based CF methods. That is, when applying neighborhood-based algorithms, two users are considered similar not only if they give similar rates to the same items but also if they have similar playing coefficients.

UPC can be used either as the only attribute if there is no more user information or jointly with other user attributes. This is one of the main advantages of the proposed method since an improvement in the recommendations is achieved even when no user and item information is available.

2.4.2. Modeling time effects on user preferences

In the field of music, time is one of the most important factors influencing user preferences. Its effects are manifested both in the drift of preferences and in the daily listening habits of users. These aspects have been treated extensively in the literature, but, in the music domain, dealing with the effect of time on recommendations is different from other fields, especially in the context of streaming services. Therefore, approaches used in some domains may be difficult to be reused or extended to this area. This is mainly due to the differences in the way music is consumed compared to other products. A music item is usually played many times by the same user, while a book is generally read only once. Another difference is the number of items that can be consumed in a given period, which can be high for music and lower for other items such as books. Likewise, the evolution of preferences may be different depending on the application domain since in some domains, interests evolve quickly, while in others these changes occur more slowly. In addition, when considering time as a contextual variable, its effects on the field of music are quite different from those generated in other areas. For instance, the time of day can have little influence when choosing a book but can have a significant influence when choosing a song.

Additional problems give rise when explicit ratings are not available, and it becomes necessary to resort to implicit feedback. The common methods for obtaining implicit ratings, such as using the number of records purchased or the time spent examining items, are not suitable for the music domain. The reason for this is that music items are not individually purchased on streaming platforms, and the same item can be listened to many times by the same user. This forces music preference evolution, managed from the perspective of implicit feedback, to be different from other areas. Quite often, the number of times a song is played is used to derive implicit ratings about a particular musical item. However, computing ratings over time, as a means to manage the drift in preference, requires establishing time windows to distribute timestamped plays and compute the evolving ratings for each interval. In other domains, the procedure is more straightforward, since only an observation per user and item is needed to apply a decay function or to discard old ratings. The observation can correspond to an action, such as a purchase. Nevertheless, timestamped ratings of music items require several observations in the form of plays for each time window. The same problem arises in context-aware recommender systems when setting time contexts.



Contribution 3 of this thesis addresses those drawbacks inherent to music recommender systems with respect to time dynamics. The purpose is to model both the evolution of user preferences in the form of evolving implicit ratings and user listening behavior.

The recency models described previously assume that user preferences change over time in such a way that the more recent the rating is, the more important it is. Usually, an exponential function of time is used to reduce monotonically ratings' importance, but, as previously stated, it requires ratings to be stamped with a date. In the context of this work, explicit ratings are not available; therefore, implicit ratings must be computed from the frequency of plays. Given that the information about the date and time of the plays is available, it can be used to calculate the degree of importance of the implicit ratings obtained from them. Our proposal can be considered a recency approach, as it involves changes in preferences over time. However, it differs from traditional models in the fact that it does not require timestamped explicit ratings since the evolution of preferences is encompassed in the process for inducing implicit ratings. In this process, unlike other recency methods, the time function is applied to plays instead of ratings.

The proposed approach involves the calculation of decay plays from the timestamped plays. In the same way that the weight of ratings decreases as their age increases, the importance of plays will also decrease. Thus, ratings can be calculated from the frequencies of weighted plays. We propose the use of an exponential function to weight the plays according to the moment in which they occur. The frequency of weighted plays will be used to obtain implicit ratings that incorporate time effects.

In addition, the listening habits of users throughout the day are also analyzed. In this case, we focus on the individual behavior patterns of each user to achieve a greater customization of the recommendations. This data analysis reveals very different behaviors among users. Nevertheless, certain likenesses among those who play songs at similar times of day can be found. Moreover, in order to introduce listening habits in the recommender models, the average time of day representing this behavior for every user is used as an additional attribute for computing the similarity between users, which may ultimately improve the recommendation reliability.

Finally, the two previous proposals were incorporated to a context-aware recommendation scenario where the time of day is used as a contextual factor.

2.4.3. Inferring user characteristics from social information

Social networks are currently being exploited in many domains because of the valuable information they contain regarding relations and user behavior. Recommender systems are one of the areas in which social data can be exploited to improve the reliability of recommendations. The problem of incorporating social information obtained from music streaming systems is addressed in contributions 4 and 5. In the first, the concepts of trust and homophily derived from the social structure are used to deal with the neighborhood bias of some CF recommendation methods which limits the number of items that can be recommended. In the second, these methods are improved by incorporating user expertise inferred from social tags.



Trust and homophily are two factors that influence users when choosing products or services, and, therefore, must be considered when predicting their interests and preferences. Trust refers to individuals that are more likely to adopt recommendations not only from opinion leaders but from their closest social context, while homophily refers to the similarity of connected users in social networks since they usually share tastes and interests. The graph of social connections between users can be the subject of structural measures that capture these two factors and allow their impact on recommendations to be considered.

In the proposed approach, trust is derived from the friendship connections since it is assumed that users trust more those users directly linked to them. To determine the degree of influence between them, a function is used in which trust is directly proportional to the number of friends of the influential user and inversely proportional to the number of friends of the influenced user. Homophily is inferred from structural equivalence. This is a property often used to identify implicit communities in social networks by computing the equivalence similarity between pairs of nodes in the network. Equivalence similarity is based on the overlap between the neighborhood of those nodes. In the context of this work, this metric can be applied to friendship structure whose nodes represent the users. Nodes with high similarity are considered to be part of the same implicit community. This is a way to capture the homophily concept since users belonging to the same community usually share interests and preferences. Therefore, their ratings can be used by the recommendation methods. The similarities between users based on trust and homophily are used to extend the neighborhood of the active user and thus increase the number of potentially recommendable items.

User expertise is another feature that can be exploited in recommender systems. It can be inferred from the social tags that users give to items. Although social tagging is usually used for music characterization, we propose to use it for characterizing users aiming at taking advantage of this information for improving user-based collaborative filtering methods.

To understand the idea in an intuitive way, let us consider tags like *rock* and *pop* that correspond to well-known and easily identifiable genres, thus they could be given by a user with a low level of expertise. However, assigning tags as *darkwave* or *trip hop*, associated with less identifiable or more specialized genres, requires a higher level of music expertise. Therefore, the expertise of users can be determined by analyzing their tags.

Based on the fact that the tags requiring more knowledge are less common than the others, we have resorted to tag frequency analysis to establish the degree of user expertise. We have used TF-IDF (Term Frequency-Inverse Document Frequency), a measure widely used for document retrieval and classification, to characterize users according to the quantity and relevance of the tags they provide to items. In this context, terms are replaced by tags and documents by users.

When TF-IDF is used for classifying documents, the weights of terms that occur very frequently in the document set are decreased while the weights of terms that occur rarely are increased. When finding user expertise profiles from tags, TF-IDF provides an indication of the tags' frequency but giving more relevance to tags that are less frequently used by many users. This metric has been used in some works to identify relevant tags for items (i.e. songs or artists), however, our aim is to identify relevant tags for users. Tags from expert users would have high



values of TF-IDF (high level of specialization) while tags from non-expert users would have low values of this metric.

We introduce user expertise to increase the influence of expert users in the recommendations against that of non-experts. Since we assume that the degree of expertise of a user is given by his TF-IDF average, we use this value as a weight in the similarity measure in order to give more relevance to the ratings of expert users in the rating prediction.

2.5. Validation

The validation of the proposals has been conducted with different public datasets, all of them extracted from Last.fm platform.

In all cases, k-fold cross validation was applied, except in the time-aware approach since here it is necessary to validate the model of preference evolution over time. In this case, the validation was performed by splitting the entire dataset into a training set containing the data corresponding to the first 15 months and a test set including data from the last 2 months. The dataset was divided before the computation of the implicit ratings. The same procedure was used when contextual periods were considered, computing the ratings separately for each training and test set for each contextual period.

Since the validation was performed for both rating prediction and top N recommendations, two types of metrics were used, error rate and rank-based metrics. The metrics used to evaluate rating prediction reliability were MAE (Mean Absolute Error), RMSE (Root-Mean-Square Error) and the normalization of these. For top-N recommendations, the following ranking specific metrics were used: AUC (Area Under the Receiver Operating Characteristics (ROC) Curve), NDCG (Normalized Discounted Cumulative Gain) and MAP (Mean Average Precision).

The baseline methods were selected according to the objective of the validation. However, different CF methods based on neighborhood and matrix factorization have been tested in the comparative studies carried out.

2.6. Conclusions and future work

Music recommendation is a research topic of increasing interest since online music platforms have become popular. However, to provide reliable recommendations it is necessary to deal with the general drawbacks of recommender systems as well as with those specific to this application domain.

This research work includes some proposals to improve collaborative filtering music recommender systems that are especially aimed at modeling user behavior and context in order to infer characteristics that allow an increase in the degree of personalization of the recommendations.

First, the gray sheep problem is addressed by characterizing users according to the popularity of the music they listen to. Recommendations are made to the users according to their degree of gray sheep, without the need to create separate groups or different recommendation procedures for this type of users. Next, a time-aware approach for music recommendation is presented. It



involves both modeling the evolution of user preferences over time and their listening habits throughout the day. Time is also considered as a contextual factor in order to use both models to provide context-aware recommendations. Finally, different ways of exploiting social information to improve recommender systems are proposed. On the one hand, the structure of the social network is used to infer the characteristics of trust and homophily, and on the other hand, social tags are used to obtain the degree of expertise of the user. The first two allow us to expand the neighborhood in CF methods and the third to weight the ratings of the most expert and therefore most influential users.

In all the proposed approaches, implicit ratings have been obtained from the frequency of plays using different procedures. Additionally, some of the proposals have been validated also in the cold start scenario.

The results of the validations performed with public datasets, prove the improvement achieved by the different methods proposed against others in the literature.

In the future we will continue to explore the possibilities offered by social information in the field of recommender systems. Especially in the application of word embedding techniques to social tags, as well as in the area of sentiment analysis. It is well known that intensive work is currently being done in these lines of research, although in the field of recommender systems they have not yet been sufficiently developed.

2.7. References

1. Ghazanfar, M. A., and Prügel-Bennett, A. Leveraging clustering approaches to solve the gray-sheep users problem in recommender systems. *Expert Systems with Applications*, 41 (2014), pp. 3261-3275, 2014.
2. Kim, H. N., Alkhaldi, A., El Saddik, A., and Jo, G. S. Collaborative user modeling with user-generated tags for social recommender systems. *Expert Systems with Applications*, 38 (2011), pp. 8488-8496, 2011.
3. Moreno, M. N., Segrera, S., López, V. F., and Muñoz, M. D. Web Mining based Framework for solving usual problems in recommender systems. A case study for movies' recommendation. *Neurocomputing*, 176 (2016), pp. 72-80, 2016.
4. Cantador, I., Bellogín, A., and Castells, P. A multilayer ontology-based hybrid recommendation model. *AI Communications*, 21 (2008), pp. 203-210, 2008.
5. Wang, F., Hu, L., Sun, R., Hu, J., and Zhao, K. SRMCS: A semantic-aware recommendation framework for mobile crowd sensing. *Information Sciences*, 433-434 (2018), pp. 333-345, 2018.
6. Ghorbani, H., and Novin, A. H. An introduction on separating gray-sheep users in personalized recommender systems using clustering solution. *International Journal of Computer Science and Software Engineering*, 5 (2016), pp. 14-18, 2016.
7. Lucas, J. P., Luz, N., Moreno, M. N., Anacleto, R., Almeida, A., and Martins, C. A hybrid recommendation approach for a tourism system. *Expert Systems with Applications*, 40 (2013), pp. 3532-3550, 2013.



8. Lucas, J. P., Segrera, S., and Moreno, M. N. Making use of associative classifiers in order to alleviate typical drawbacks in recommender systems. *Expert Systems with Applications*, 39 (2012), pp. 1273-1283, 2012.
9. Shepitsen, A., Gemmell, J., Mobasher, B., and Burke, R. Personalized recommendation in social tagging systems using hierarchical clustering, in *Proceedings of the 2008 ACM conference on Recommender systems*, pp. 259–266, ACM, New York, NY, USA, 2008.
10. Aggarwal, C.C. *Recommender Systems. The Testbook*, Springer, Berlin/Heidelberg, Germany, 2016.
11. Campos, P. G., Díez, F., and Cantador, I. Time-aware recommender systems: A comprehensive survey and analysis of existing evaluation protocols. *User Modeling and User-Adapted Interaction*, 24, pp. 67–119, 2014.
12. Tuan, C. C., Hung, C. F., and Wu, Z. H. Collaborative location recommendations with dynamic time Periods. *Pervasive and Mobile Computing*, 35, pp. 1–14, 2017.
13. Ding, Y., and Li, X. Time weight collaborative filtering, in *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*, Bremen, Germany, October 31-November 5, 2005, ACM, New York, NY, USA, pp. 485–492, 2005.
14. Sánchez-Moreno, D., Zheng, Y., and Moreno-García, M.N. Incorporating Time Dynamics and Implicit Feedback into Music Recommender Systems, in *Proceedings of the 2018 IEEE/WIC/ACM International Conference on Web Intelligence*, Santiago, Chile, 3–6 December 2018, pp. 580–585, 2018.
15. Campos, P., Bellogin, A., Díez, F., and Chavarriaga, J. Simple time-biased KNN-based recommendations, in *Proceedings of the Workshop on Context-Aware Movie Recommendation*, Barcelona, Spain, September 30, 2010, pp. 20–23, 2010.
16. Gordea, S., and Zanker, M. Time filtering for better recommendations with small and sparse rating matrices, in *Proceedings WISE*, Springer, Berlin/Heidelberg, Germany, 2007, pp. 171–183, 2007.
17. Zhang, Y., and Liu, Y. A collaborative filtering algorithm based on time period partition, in *Proceedings of the IEEE 2010 Third IITSI*, Jinggangshan, China, 2–4 April 2010, pp. 777–780, 2010.
18. Wang, C., Zhang, M., Ma, W., Liu, Y., and Ma, S. Modeling Item-Specific Temporal Dynamics of Repeat Consumption for Recommender Systems, in *Proceedings of the World Wide Web Conference*, San Francisco, CA, USA, May 13-17, 2019, pp. 1977–1987, 2019.
19. Koren, Y. Collaborative Filtering with Temporal Dynamics, in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Paris, France, June 28 - July 1, 2009, pp. 447–456, 2009.
20. Xiong, L., Chen, X., Huang, T., Schneider, J., and Carbonell, J.G. Temporal collaborative filtering with bayesian probabilistic tensor factorization, in *Proceedings of the 2010 SIAM International Conference on Data Mining*, Columbus, Ohio, USA, April 29 - May 1, 2010, pp. 211–222, 2010.
21. Li, X., Xu, G., Chen, E., and Zong, Y. Learning recency based comparative choice towards point-of-interest recommendation. *Expert Systems with Applications*, 42, pp. 4274–4283, 2015.
22. Zafari, F., Moser, I., and Baarslag, T. Modelling and analysis of temporal preference drifts using a component-based factorised latent approach. *Expert Systems with Applications*. 2019, 116, pp. 186–208, 2019.



23. Tong, C., Qi, J., Lian, Y., Niu, J., and Rodrigues, J.J.P.C. TimeTrustSVD: A collaborative filtering model integrating time, trust and rating information. *Future Generation of Computer Systems*, 93, pp. 933–941, 2019.
24. Panniello, U., Tuzhilin, A., Gorgoglione, M., Palmisano, C., and Pedone, A. Experimental comparison of pre- vs. post-filtering approaches in context-aware recommender systems, in *Proceedings of the Third ACM Conference on Recommender Systems*, New York, NY, USA, October 23-25, 2009, pp. 265–268, 2009.
25. Qi, L., Wang, R., Hu, C., Li, S., He, Q., and Xu, X. Time-aware distributed service recommendation with privacy-preservation. *Information Sciences*, 480, pp. 354–364, 2019.
26. Adomavicius, G., Sankaranarayanan, R., Sen, S., and Tuzhilin, A. Incorporating contextual information in recommender systems using a multidimensional approach. *ACM Transaction on Information Systems*, 23, pp. 103–145, 2005.
27. Baltrunas, L., Ludwig, B., Peer, S., and Ricci, F. Context-aware places of interest recommendations for mobile users, in *Proceedings of the International Conference on Human-Computer Interaction*, Orlando, Florida, USA, July 9–14, 2011, Springer, Berlin/Heidelberg, Germany, pp. 531–540, 2011.
28. Lu, J., and Indeché, M.A. Multi-Context-Aware Location Recommendation Using Tensor Decomposition. *IEEE Access*, 8, pp. 61327–61339, 2020.
29. Adomavicius, G., Mobasher, B., Ricci, F., and Tuzhilin, A. Context-aware recommender systems. *AI Magazine*, 32(3), pp. 67–80, 2011.
30. Baltrunas, L., and Amatriain, X. Towards time-dependant recommendation based on implicit feedback, in *Proceedings of the Workshop on Context-Aware Recommender Systems*, New York, NY, USA, October 25, 2009.
31. Reddy, S., and Mascia, J. Lifetrak: Music in tune with your life, in *Proceedings of the 1st ACM International Workshop on Human-centered Multimedia*, Santa Barbara, California, USA October 23-27, pp. 25–34, 2006.
32. Sánchez-Moreno, D., Moreno-García, M.N., Sonboli, N., Mobasher, B., and Burke, R. Inferring user expertise from social tagging in music recommender systems for streaming services, in F.de Cos et al. (eds.), *Hybrid Artificial Intelligence Systems*, Lecture Notes in Artificial Intelligence, pp. 39-49, 2018.
33. Sánchez-Moreno, D., Pérez-Marcos, J., Gil, A.B., López, V.F., and Moreno-García, M.N. Social influence-based similarity measures for user-user collaborative filtering applied to music recommendation, in Rodríguez S. et al. (eds), *Distributed Computing and Artificial Intelligence*, Special Sessions, 15th International Conference, 2018. *Advances in Intelligent Systems and Computing*, 801, pp. 1-8, 2019.
34. Yadav, P., Rani, K. S., and Kumari, S. Review of Social Collaborative Filtering Recommender System's Methods. *International Journal of Engineering and Computer Science*, 4 (11), pp. 14927-14932, 2015.
35. Massa, P., and Avesani, P. Trust-aware recommender systems, *ACM Conference on Recommender Systems (RecSys)*, Minneapolis, MN, USA, pp. 17–24, 2007.
36. Kalaï, A., Abdelghani, W., Zayani, C.A., and Amous, I., LoTrust: A social trust level model based on time-aware social interactions and interests similarity, 14th IEEE Fourteenth Annual Conference on Privacy, Security and Trust, Auckland, NewZeland, pp. 428-436, 2016.



37. Akcora, C.G., Carminati, B., and Ferrari, E. User similarities on social networks. *Social Network Analysis and Mining*, 3 (3), pp. 475-495, 2013.
38. Ziegler, C., and Golbeck, J. Investigating interactions of trust and interest similarity. *Decision Support Systems*, 43 (2), pp. 460-475, 2006.
39. Zafarani, R., Abbasi, M.A., and Liu, H. *Social Media Mining*, Cambridge University Press, 2014.
40. Zhou, Z., Xu, K., and Zhao, J. Homophily of music listening in online social networks of China. *Social Networks* 55 (2018), pp. 160–169, 2018.
41. Ladhari, R., Massa E., and Skandrani, H. YouTube vloggers' popularity and influence: The roles of homophily, emotional attachment, and expertise. *Journal of Retailing and Consumer Services*, 54, 2020.
42. Sokolova, K., and Kefi, H. Instagram and YouTube bloggers promote it, why should I buy? How credibility and parasocial interaction influence purchase intentions. *Journal of Retailing and Consumer Services*, 53, 2020.
43. Esmaeili, L., Mardani, S., Golpayegani, S.H. Madar, Z.Z. A novel tourism recommender system in the context of social commerce. *Expert Systems with Applications*, 149, Article 113301, 2020.
44. Fields B., Jacobson K., Rhodes C. Inverno M., Sanler M., and Casey M. Analysis and Exploitation of Musician Social Networks for Recommendation and Discovery. *IEEE Transactions on Multimedia*, 13(4), pp. 674–686, 2011.
45. Zhou, Z., Xu, K., and Zhao, J. Homophily of music listening in online social networks of China. *Social Networks*, 55 (2018), pp. 160–169, 2018.
46. Mansoury, M., Abdollahpouri, H., and Smith, J. Dehpanah, A., Pechenizkiy, M., Mobasher, B.: Investigating Potential Factors Associated with Gender Discrimination in Collaborative Recommender Systems, in *Proceedings of 13th FLAIRS Conference*, pp. 193-196, 2020.
47. Burke, R., Sonboli, N., and Ordoñez-Gauger, A. Balanced Neighborhoods for Multi-sided Fairness in Recommendation. *Proceedings of Machine Learning Research*, *Proceedings of the 1st Conference on Fairness, Accountability, and Transparency*, PMLR 81, pp. 202-214, 2018.
48. Abdollahpouri, H., Mansoury, M., Burke, R., and Mobasher, B. The Unfairness of Popularity Bias in Recommendation, *13th ACM Conference on Recommender Systems (RecSys)*, 2019, RMSE workshop, Copenhagen, Denmark, 2019.
49. Sánchez-Moreno, D., Gil, A. B., Muñoz, M. D., López, V. F. and Moreno M. N.: Recommendation of songs in music streaming services. Dealing with sparsity and gray sheep problems. *Trends in Cyber-Physical Multi-Agent Systems. The PAAMS Collection-15th International Conference, PAAMS 2017. Advances in Intelligent Systems and Computing series*, vol. 619, pp. 206-213, 2017.
50. Sánchez-Moreno D., Muñoz M. D., López V. F., Gil A. B., and Moreno M. N.: A session-based song recommendation approach involving user characterization along the play power-law distribution. *Complexity*, Article ID 7309453, 2020.
51. Vargas S., and Castells, P. Rank and relevance in novelty and diversity metrics for recommender systems, in *Proceedings of the 5th ACM Conference on Recommender Systems*, pp. 109-116, 2011.
52. Steck, H. Calibrated recommendations, in *Proceedings of the 12th ACM Conference on Recommender Systems*, pp. 154–162, 2018.



53. Adomavicius, G. and Kwon, Y. Toward More Diverse Recommendations: Item Re-Ranking Methods for Recommender Systems, in *Proceedings of the 19th Workshop on Information Technologies and Systems*, 2009.
54. Kamishima, T., Akaho, S., Asoh, H., and Sakuma, J. Recommendation Independence. *Proceedings of Machine Learning Research. Proceedings of Conference on Fairness, Accountability and Transparency*, PMLR 81, pp. 187–201, 2018.
55. Adomavicius, G., and Kwon, Y. Maximizing aggregate recommendation diversity: A graph-theoretic approach, in *Proceedings of the 1st International Workshop on Novelty and Diversity in Recommender Systems*, pp. 3–10, 2011.
56. Mansoury, M., Abdollahpouri, H., Pechenizkiy, M., Mobasher, B., and Burke, R. FairMatch: A Graph-based Approach for Improving Aggregate Diversity in Recommender Systems. *arXiv:2005.01148*, 2020.
57. Yang, L., Qiu, M., Gottopati, S., Zhu, F. and Jiang, J. CQARank: Jointly model topics and expertise in Community Question Answering, in *Proceedings of the 22nd ACM International Conference on Information and Knowledge Management*, San Francisco, CA, USA, pp. 108, 2013.
58. Yang, B. and Manandhar, S. Tag-Based Expert Recommendation in Community Question Answering, in *Proceedings of the 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 960-963, 2014.
59. Zhou, G., Lai, S., Liu, K. and Zhao, J. Topic-sensitive probabilistic model for expert finding in question answer communities, in *Proceedings of the 21st ACM international conference on Information and knowledge management*, Maui, Hawaii, USA, pp. 1662-1666, 2012.
60. Martín-Vicente, M. I., Gil-Solla, A., Ramos-Cabrer, M., Blanco-Fernández, Y. and López-Nores, M. Semantic inference of user's reputation and expertise to improve collaborative recommendations. *Expert Systems with Applications* 39 (2012), pp. 8248–8258, 2012.
61. Hariri, N., Mobasher, B. and Burke, R. Context-aware music recommendation based on latent topic sequential patterns, in *Proceedings of the sixth ACM conference on Recommender systems*, pp. 131-138, 2012.
62. Su, J. H., Chang, W. Y. and Tseng, V. S. Personalized music recommendation by mining social media tags. *Procedia Computer Science*, 22, pp. 303-312, 2013.
63. Wang, F., Hu, L., Sun, R., Hu, J., and Zhao, K. SRMCS: A semantic-aware recommendation framework for mobile crowd sensing. *Information Sciences*, 433-434 (2018), pp. 333-345, 2018.
64. Liang, N., Zheng, H. T., Chen, J. Y., Sangaiah A. K., and Zhao, C. Z. TRSDL: tag-aware recommender system based on deep learning–intelligent computing systems. *Applied Sciences*, 8 (5), 799, 2018.
65. Vargas, S., and Castells, P. Rank and relevance in novelty and diversity metrics for recommender systems, in *Proceedings of the fifth ACM conference on recommender systems*, ACM, New York, NY, USA, pp. 109-116, 2011.
66. Lee, K., and Lee, K. Escaping your comfort zone: A graph-based recommender system for finding novel recommendations among relevant items. *Expert Systems with Applications*, 42 (2015), pp. 4851-4858, 2015.



67. Pacula, M. A matrix factorization algorithm for music recommendation using implicit user feedback. <http://www.mpacula.com/publications/lastfm.pdf>. Last accessed date: September 10, 2020.



3. Resumen en español de las contribuciones

3.1. Contribución 1

3.1.1. Referencia

Sánchez-Moreno D., Gil A.B., Muñoz M.D., López V.F. and Moreno M.N. “A collaborative filtering method for music recommendation using playing coefficients for artists and users”, *Expert Systems with Applications*, 66 (2016), 234-244, 2016.

3.1.2. Objetivos

En este trabajo se abordan los problemas de dispersión y de la oveja negra (*gray sheep*) mediante un método de recomendación de música basado en coeficientes de reproducción.

Aunque en la literatura hay muchas propuestas para abordar las deficiencias de los sistemas de recomendación de filtrado colaborativo (FC), el problema de la oveja negra relativo a los usuarios con gustos inusuales ha recibido menos atención, y normalmente se aborda mediante enfoques híbridos que implican alguna técnica basada en el contenido. Sus resultados suelen ser buenos, pero estos métodos requieren información sobre los ítems y los usuarios que a menudo no está disponible. En el ámbito de los sistemas de recomendación de música, los algoritmos de filtrado basados en el contenido utilizan el contenido de la música para inducir los modelos; por lo tanto, es necesario realizar una compleja tarea de extracción de características musicales. En este contexto, el objetivo del trabajo es la propuesta de una metodología de recomendación que aborde el inconveniente mencionado anteriormente cuando se dispone de poca información. Otro de los propósitos es hacer frente al problema de la dispersión mediante la obtención de valoraciones implícitas.

3.1.3. Método de recomendación

El procedimiento propuesto tiene por objeto formular recomendaciones fiables cuando no se dispone de información sobre el contenido y valoración de los ítems. La mejora de las recomendaciones se logra principalmente centrándose en los usuarios ovejas negras, dado que su presencia tiene un efecto negativo tanto en sus recomendaciones como en las de los demás usuarios. Sin embargo, a diferencia de la mayoría de los métodos de la literatura, los usuarios considerados ovejas negras no son tratados de manera diferente al resto de los usuarios ni son separados en un grupo diferente, sino que nuestra estrategia implica la determinación para cada usuario de un coeficiente que represente el grado en que son ovejas negras.

La propuesta se basa en el uso de información implícita tanto para obtener las preferencias de los usuarios como para caracterizar a los usuarios y a los ítems que se recomiendan.



El método está centrado en la recomendación de artistas. Dado un conjunto de artistas A y un conjunto de usuarios U donde $a_i \in A, i = 1, \dots, n$ y $u_j \in U, j = 1, \dots, m$ representan un artista y un usuario respectivamente. El número de veces que el usuario j reproduce un artista i se representa como $p_{i,j}$.

El coeficiente de escucha (l_{a_i}) para el artista a_i se utiliza para determinar su popularidad (ecuación 1).

$$l_{a_i} = (TU_{a_i} / \overline{TU}) \frac{\sum_j (p_{i,j} / \overline{p_j})}{\sum_i \sum_j (p_{i,j} / \overline{p_j}) / n} \quad (1)$$

Donde n es el número total de artistas, TU_{a_i} es el número de usuarios que reproducen al artista a_i , \overline{TU} es el número medio de usuarios por artista y $\overline{p_j}$ el número medio de reproducciones por artista del usuario j .

El coeficiente de escucha de un artista determinado tiene en cuenta el número de usuarios que lo escuchan con respecto al promedio de usuarios por artista, así como la relación entre el comportamiento de los usuarios que escuchan al artista y el comportamiento promedio de todos los artistas. El comportamiento del usuario se cuantifica como el número de reproducciones para el artista dado con respecto al número promedio de reproducciones de este usuario. El coeficiente de escucha se normaliza y posteriormente se propone un segundo coeficiente relacionado con los usuarios, UPC (*User Playing Coefficient*), destinado a caracterizarlos según la popularidad de los artistas que reproducen. Valores altos de UPC representan a usuarios que escuchan a artistas populares y los valores bajos corresponden a usuarios ovejías negras.

En primer lugar, se introduce un parámetro $\alpha_{i,j}$ para retener a los artistas reproducidos por cada usuario j . Después se calcula el coeficiente UPC_{u_j} mediante la ecuación 3.

$$\alpha_{i,j} = \begin{cases} 0, & p_{i,j} = 0 \\ 1, & p_{i,j} > 0 \end{cases} \quad (2)$$

$$UPC_{u_j} = \frac{\sum_i \alpha_{i,j} l_{a_i}}{TA_{u_j}} \quad (3)$$

Siendo TA_{u_j} el número de artistas reproducidos por el usuario j .

El coeficiente UPC se utiliza como atributo de usuario en los métodos de recomendación de filtrado colaborativo basados en k-NN (*k-Nearest Neighbors*). Se puede utilizar como único atributo si no se dispone de más información de usuario o conjuntamente con otros atributos de usuario. Esta es una de las principales ventajas del método propuesto, ya que se logra una mejora de las recomendaciones incluso cuando no se dispone de información de usuarios y artistas.

Los métodos de filtrado colaborativo utilizados requieren la obtención de valoraciones implícitas puesto que no se dispone de información de valoraciones de los usuarios sobre los artistas. Esas valoraciones (*ratings*) se obtienen a partir de las reproducciones usando una función lineal del percentil de frecuencia.



3.1.4. Resultados

La validación de la propuesta se realizó mediante un estudio comparativo en el que se analizaron los resultados frente a los enfoques tradicionales de filtrado colaborativo. Se probó el método k-NN basado en usuarios, por una parte, y basado en el ítems, por otra, usando medidas de similitud tanto del coseno como de Pearson. Además, también se probó el algoritmo k-NN basado en el usuario utilizando otros atributos de usuario además de UPC.

El estudio comparativo se llevó a cabo para dos tipos de recomendaciones, predicción de *ratings* y recomendación de listas top-N. Para el primer tipo, las métricas de validación utilizadas fueron RMSE (*Root Mean Square Error*), MAE (*Mean Absolute Error*) y NMAE (*Normalized Mean Absolute Error*), mientras que para el segundo, las métricas fueron MAP (*Mean Average Precision*), NDCG (*Normalized Discounted Cumulative Gain*) y el área bajo la curva ROC (AUC). En la evaluación de todos los algoritmos se realizó una validación cruzada.

Los resultados obtenidos en la predicción de *ratings* demostraron que el algoritmo k-NN basado en el usuario con el único atributo UPC da mejores resultados que el mismo método con los otros atributos. Otra observación positiva es la menor amplitud de los intervalos de confianza en los resultados de todas las métricas para este método, en comparación con el resto de los enfoques probados. En el caso de las recomendaciones top-N, los mejores valores de las medidas de precisión también fueron obtenidos por nuestro enfoque. La diferencia fue especialmente significativa para MAP y NDCG mientras que fue menor para el área bajo la curva ROC. Otra observación a destacar son los malos resultados de los algoritmos k-NN basados en usuarios y en ítems, cuando utilizan solo la similitud basada en las valoraciones del usuario. El uso de otros atributos de usuario mejora las recomendaciones, especialmente el uso de UPC.

3.1.5. Conclusiones

La recomendación de música es un tema de investigación de creciente interés desde que las plataformas de música en línea se han hecho populares. Sin embargo, para proporcionar recomendaciones fiables es necesario abordar los inconvenientes generales de los sistemas de recomendación, así como los específicos de este dominio de aplicación, como la dificultad de extraer información de contenido de la música. En el presente trabajo se abordan algunos de esos inconvenientes mediante una metodología de recomendación especialmente diseñada para tratar los problemas de la oveja negra y la dispersión sin necesidad de utilizar atributos de usuario, datos de contenido o valoraciones explícitas. De esta forma se evita una de las principales dificultades de otras propuestas que a menudo requieren una compleja tarea de extracción de las características de la música.

El procedimiento propuesto en este trabajo utiliza coeficientes de reproducción de artistas y usuarios para determinar el grado en que un usuario es una oveja negra. Los coeficientes necesarios para aplicar los algoritmos de CF se derivan de la frecuencia de reproducciones. La validación de la propuesta se realizó mediante un estudio en el que se probaron varios algoritmos de FC tanto para la predicción de *ratings* como para recomendaciones top-N. Los resultados mostraron que nuestro método supera significativamente los otros enfoques de FC.



3.2. Contribución 2

3.2.1. Referencia

Sánchez-Moreno D., Muñoz M.D., López V.F., Gil A.B., Moreno M.N. “A session-based song recommendation approach involving user characterization along the play power-law distribution”, Complexity, Article ID 7309453, 2020.

3.2.2. Objetivos

El desarrollo de sistemas fiables de recomendación en el ámbito de la música supone abordar muchos problemas, algunos de los cuales son genéricos y ampliamente estudiados en la literatura, mientras que otros son específicos de este dominio de aplicación y, por lo tanto, menos conocidos. Uno de ellos es la obtención de realimentación implícita a partir de la frecuencia en la reproducción de música y otro tiene que ver con un sesgo habitual en los sistemas de recomendación de música conocido como la larga cola (*long tail*). Este último se debe al hecho de que son pocos los artistas o canciones muy reproducidas mientras que hay muchos artistas o canciones que se reproducen mucho. Este problema, relacionado también con el de la oveja negra, produce un sesgo en las recomendaciones hacia los artistas o canciones más populares. En este caso el trabajo está enfocado en la recomendación de canciones.

Nuestro primer objetivo es inferir las valoraciones de los usuarios considerando las sesiones de estos. El segundo se centra en manejar la frecuencia de reproducciones que se ajusta a una distribución de probabilidades de ley de potencia (*long tail*) y que representa el comportamiento de escucha de los usuarios. Se trata de abordar el problema de la oveja negra mediante la caracterización de cada usuario según la frecuencia de reproducción de las canciones que escucha.

3.2.3. Metodología propuesta

La mejora de este enfoque en comparación con los principales métodos de filtrado colaborativo se logra en dos aspectos principales, una nueva forma de obtener valoraciones implícitas a partir de las sesiones de los usuarios y la caracterización de los usuarios según el lugar que ocupan las canciones que escuchan en la curva de frecuencia de reproducciones.

Las formas tradicionales de obtener evaluaciones implícitas de ítems, como a partir de registros de compra, clics o información de tiempo empleado en la consulta de los ítems, no son posibles en el contexto de nuestro estudio, ya que el modo de interacción de los usuarios con las canciones en las plataformas de música es bastante diferente. Por lo general, se utilizan valores binarios o funciones de frecuencia simple de reproducciones para derivar las preferencias a partir de la retroalimentación implícita del usuario. Sin embargo, en este trabajo, proponemos un modelo más complejo para inferir de una manera más fiable los intereses de los usuarios a partir de su comportamiento.



Este enfoque tiene en cuenta las sesiones en las que los usuarios reproducen canciones a través de los servicios de *streaming*, así como una función de percentil de frecuencia de reproducción en el cálculo de los *ratings*. Aunque todas las canciones escuchadas por el usuario pueden ser elegidas por él mismo, el método se basa en el hecho de que la primera canción de una sesión de usuario es más importante ya que tiene una mayor probabilidad de ser una elección directa del usuario que las canciones en otras posiciones.

Se considera que una sesión de usuario es un período en el que el usuario está escuchando canciones sin interrupción. Consiste en canciones que se reproducen en un orden determinado; por lo tanto, puede caracterizarse como una cadena de Markov en la que las probabilidades iniciales son proporcionales al número de veces que se visitó un estado determinado. Así, consideramos el número de veces que cada canción ha sido inicio de sesión ($s_{i,j}$) y el número de veces que no lo ha sido ($ns_{i,j}$). La frecuencia para cada par usuario-canción se calcula de la siguiente forma:

$$Freq_{i,j} = \alpha \frac{s_{i,j}}{\sum_j s_{i,j}} + (1-\alpha) \frac{ns_{i,j}}{\sum_j ns_{i,j}} \quad (4)$$

El parámetro α se utiliza para ajustar la importancia de cada término en la ecuación.

Si se ordenan las canciones según esa frecuencia, el *rating* para una canción en la posición k se calcula mediante la ecuación 5, en la que $k' = 1$ corresponde a la canción con mayor frecuencia:

$$r_{i,j} = 4 \left(1 - \sum_{k'=1}^{k-1} Freq_{k'}(i) \right) \quad (5)$$

Una vez obtenidos las valoraciones, se caracterizan las canciones de acuerdo a su popularidad. Las menos populares serán las situadas en la “larga cola” de la distribución de ley de potencia. Para ello se calcula un coeficiente de escucha L_{g_j} para cada una de las canciones g_j , el cual se usará posteriormente para caracterizar al usuario en función de la popularidad de la música que escucha. Esto se hace mediante el cálculo de *UPC* (*User Playing Coefficient*) utilizando la siguiente ecuación.

$$UPC_i = \frac{\sum_j \beta_{i,j} L_{g_j}}{TG_i} \quad (6)$$

Siendo $\beta_{i,j}$ un parámetro que toma el valor 1 si la canción g_j ha sido reproducida por el usuario u_i y el valor 0 en caso contrario. TG_i es el número total de canciones reproducidas por el usuario u_i . Los usuarios con valores altos de *UPC* tienen preferencias en común con muchos otros, mientras que los que tienen valores bajos serían usuarios ovejas negras.

El coeficiente *UPC* se incorpora como un atributo más en el cálculo de las similitudes entre usuarios al aplicar los métodos de filtrado colaborativo basados en vecindad.

3.2.4. Resultados

El enfoque propuesto se aplicó en un conjunto de datos público procedente de la plataforma Last.fm y sus resultados se compararon con varios métodos de filtrado colaborativo, tanto basados en k-NN como en factorización de matrices. En todos los experimentos se aplicó validación cruzada.



El solo hecho de introducir los *ratings* basados en sesiones en lugar de en el simple percentil de la frecuencia de reproducciones produjo un aumento significativo en la fiabilidad de las recomendaciones con todos los métodos evaluados en el estudio comparativo. Además, nuestro enfoque de incorporación del coeficiente *UPC* resultó ser el que menores tasas de error produjo con los dos tipos de *ratings*. Los porcentajes de mejora entre este método y los demás osciló entre el 5.0% y el 14.3%.

3.2.5. Conclusiones

La forma de obtener una retroalimentación implícita de los usuarios en los servicios de *streaming* es totalmente diferente de otras aplicaciones, debido, entre otras razones, a que no hay registros de compra individualizada de canciones, y la forma de consumir música es diferente del consumo de otros productos. Por otra parte, la mayoría de las propuestas para hacer frente al problema de la oveja negra hacen uso de información de contenido que es difícil de obtener y que, en muchos casos, no conduce a los resultados esperados.

En este trabajo se presenta un enfoque para mejorar la fiabilidad de las recomendaciones en el contexto de los servicios de música en *streaming*. Su principal valor es el de abordar el cálculo de las valoraciones implícitas y la caracterización del usuario sólo conociendo el momento en el que cada usuario reproduce las canciones, información que se recoge regularmente por las plataformas de *streaming*. El procedimiento propuesto para obtener las valoraciones difiere de la mayoría de los métodos, que generalmente se basan solo en número de reproducciones, puesto que utiliza además la posición de las canciones en las sesiones de los usuarios. Además, se presenta una nueva forma de gestionar a los usuarios con gustos inusuales basada en la distribución de ley de potencia. Los resultados muestran una mejora significativa de la fiabilidad de las recomendaciones con respecto a los métodos tradicionales de filtrado colaborativo.

3.3. Contribución 3

3.3.1. Referencia

Sánchez-Moreno D., Zheng, Y., Moreno M.N. "Time-aware music recommender systems: Modeling the evolution of implicit user preferences and user listening habits in a collaborative filtering approach", *Applied Sciences, Special Issue on Recommender Systems and Collaborative Filtering*, 10 (15), 5324, 33 pages, 2020.

3.3.2. Objetivos

Existen muchos motivos por los que algunos sistemas de recomendación actuales no proporcionan al usuario resultados satisfactorios, aunque uno de gran importancia radica en no tener en cuenta los aspectos contextuales o la evolución de los gustos del usuario con el tiempo.

En el campo de la música, el tiempo es uno de los factores más importantes que influyen en las preferencias de los usuarios, por lo que este trabajo está dirigido a modelar sus efectos desde diferentes perspectivas. El propósito es modelar tanto la evolución de las preferencias del usuario en forma de evolución de las valoraciones implícitas como el comportamiento de escucha



del usuario. Se pretende capturar los hábitos de escucha diarios para caracterizar a los usuarios y proporcionarles recomendaciones más fiables.

3.3.3. Modelado del efecto del tiempo en las recomendaciones

Aunque los efectos del tiempo en los sistemas de recomendación han sido el centro de muchos estudios, los que abordan cuestiones particulares relacionadas con el ámbito de la música son muy pocos. En este trabajo, este tema se trata mediante un enfoque de recomendación que incorpora los factores temporales y que está diseñado específicamente para el área de la música cuando no se dispone de valoraciones explícitas. En este contexto, muchos de los métodos definidos para otros dominios no son aplicables.

En el caso de la evolución de las preferencias con el tiempo, o bien se utilizan funciones del tiempo para reducir la importancia de los *ratings* menos recientes o bien se establecen intervalos de tiempo en los que los *ratings* son válidos descartando directamente los anteriores. Para ello es necesario disponer de esos *ratings* con la marca de tiempo correspondiente al momento en el que fueron asignados a los ítems. En el contexto de este trabajo no se dispone de valoraciones explícitas, por lo que se calculan valoraciones implícitas a partir de la frecuencia de las reproducciones. Dado que se dispone de la información sobre la fecha y hora de cada reproducción, esta se usa para calcular el grado de importancia de las valoraciones implícitas obtenidas de ellas. Nuestra propuesta difiere de otros modelos en el hecho de que no requiere *ratings* explícitos, ya que el modelado de la evolución de las preferencias está incluido en el proceso de inducción de valoraciones implícitas

Así, los *ratings* se obtienen a partir de las frecuencias de reproducciones (*plays*) ponderadas. Proponemos el uso de una función exponencial para la ponderación según el momento en que se produzcan. Un *decay play* (*dPlay*) se define como el peso que tiene una reproducción p de una canción s_j por un usuario u_i según su marca de tiempo $t_{i,j,p}$, como se ve en la ecuación 7, siendo 1 el peso máximo, que se asigna a la reproducción más reciente, la correspondiente a t_{max} . T_u se usa para convertir las unidades en las que está medido el tiempo t a las unidades objetivo.

$$dPlay(u_i, s_j, p) = e^{-\lambda(t_{max}-t_{i,j,p})/T_u} \quad (7)$$

La frecuencia de *decay plays* se utilizará para obtener las valoraciones implícitas, que ya incorporan los efectos del tiempo, mediante una función del percentil de frecuencia. Esta frecuencia para un usuario u_i y una canción s_j dentro de una ventana de tiempo determinada se calcula como se indica en ecuación 8.

$$dPlayF(u_i, s_j) = \frac{\sum_{p=1}^{n_p(u_i, s_j)} dPlay(u_i, s_j, p)}{\sum_j \sum_{p=1}^{n_p(u_i, s_j)} dPlay(u_i, s_j, p)} \quad (8)$$

El comportamiento del usuario dependiente del tiempo en el dominio de la música es otro aspecto que puede influir en las recomendaciones. Los usuarios tienen diferentes hábitos dependiendo de la hora del día en que están escuchando música, lo que puede ser tenido en cuenta para mejorar los modelos de recomendación. Los métodos de recomendación



contextuales se basan en estos patrones generales de comportamiento del usuario. Se utilizan para gestionar situaciones en las que el tiempo es una variable contextual, ya que las preferencias del usuario pueden depender de ello. Para poder aplicarlos, el día se divide en intervalos de tiempo y se generan diferentes recomendaciones para cada una de estas ventanas de tiempo. Esto puede causar problemas de dispersión, más aún en el ámbito de la música, donde las valoraciones implícitas se calculan a partir del número de reproducciones debido a la necesidad de intervalos adicionales. En el estudio experimental llevado a cabo en este trabajo, en el que se prueba el método descrito anteriormente, sólo se consideran dos intervalos contextuales de tiempo, mañana y tarde, como una forma de minimizar este problema. Además, se aplica el método de prefiltrado para generar recomendaciones para cada período.

Además de establecer intervalos de tiempo contextuales, nuestra propuesta también contempla los hábitos de escucha de los usuarios a lo largo del día. En este caso, nos centramos en los patrones de comportamiento individual de cada usuario para lograr una mayor personalización de las recomendaciones. Este análisis de datos revela comportamientos muy diferentes entre los usuarios. Sin embargo, se pueden encontrar ciertas similitudes entre los que escuchan canciones en momentos similares del día. Para introducir los hábitos de escucha en los modelos de recomendación, se define una variable ($dTAvg_i$) que representa este comportamiento (ecuación 9).

$$dTAvg_i = \frac{\sum_j \sum_p t_{i,j,p}}{NP_i} \quad (9)$$

Esto nos permite utilizar la variable contextual de la hora del día como un atributo adicional para calcular la similitud entre los usuarios, lo que en última instancia puede mejorar la fiabilidad de la recomendación.

3.3.4. Resultados

En el estudio experimental realizado con datos obtenidos de Last.fm se validó tanto el método de evolución de las preferencias como el modelado de los hábitos de escucha del usuario. Así mismo, ambos fueron probados sin considerar el contexto del tiempo y considerando intervalos contextuales. Adicionalmente, se comprobó la validez de la propuesta para el escenario *cold start*, es decir para usuarios que llevan poco tiempo en el sistema y por tanto han reproducido pocas canciones.

En este caso, debido a que se quería demostrar la validez del modelo de evolución de preferencias, la validación se realizó dividiendo todo el conjunto de datos en un conjunto de entrenamiento que contenía los datos correspondientes a los primeros 15 meses y un conjunto de prueba que incluía los datos de los últimos 2 meses.

Los resultados obtenidos para la predicción de *ratings* fueron significativamente mejores tanto considerando de forma separada la evolución de *ratings* y los hábitos de escucha como aplicando ambos enfoques conjuntamente. Esta mejora se produjo a su vez en los dos escenarios, el de la generación de recomendaciones sensibles al contexto y libres de contexto. En el caso de las recomendaciones top-N las diferencias entre el uso de *ratings* y *decay ratings* no fueron



significativas. Sin embargo, el enfoque de modelado de los hábitos de escucha sí produjo mejoras significativas para el día completo y para los intervalos contextuales de mañana y tarde.

3.3.5. Conclusiones

En este trabajo se propone un enfoque para abordar los efectos del tiempo en las recomendaciones de música. El método está diseñado para tratar problemas característicos de este ámbito que están ausentes en otros dominios de aplicación. El estudio también intenta tratar la retroalimentación implícita, ya que normalmente no se dispone de valoraciones explícitas. Así pues, la evolución de las preferencias de los usuarios a lo largo del tiempo se incluye en el propio método propuesto para obtener las valoraciones implícitas de las canciones a partir de la frecuencia de las reproducciones. Otros métodos comúnmente utilizados para incorporar la dinámica temporal, como las funciones de decaimiento, no pueden utilizarse aquí porque requieren valoraciones explícitas con indicación de tiempo. Asimismo, la forma tradicional de calcular los *ratings* implícitos, como la utilización del tiempo que los usuarios dedican a examinar un ítem o la información de compra, no es aplicable en el ámbito de la música. Esto ocurre porque los usuarios tienden a escuchar una canción más de una vez y las canciones no se compran una por una en las plataformas de *streaming*. Además, el modelo de recomendación propuesto incorpora información del comportamiento de escucha de los usuarios para mejorar las recomendaciones.

Los resultados de la validación realizada con un conjunto de datos de Last.fm demostraron que el método propuesto para recomendar canciones supera a otros enfoques de filtrado colaborativo, tanto cuando se tiene en cuenta el contexto temporal como en escenarios libres de contexto.

3.4. Contribución 4

3.4.1. Referencia

Sánchez-Moreno D., López V.F., Muñoz M.D., Sánchez, A.L., Moreno M.N. “Exploiting the user social context to address neighborhood bias in collaborative filtering music recommender systems”, Information, Artificial Intelligence Section, 11(9), 439, 16 pages, 2020.

3.4.2. Objetivos

Investigaciones recientes en el ámbito de los sistemas de recomendación se están centrando en la incorporación de información social en los métodos de filtrado colaborativo para mejorar la fiabilidad de las recomendaciones. Las redes sociales encierran datos valiosos sobre el comportamiento y las conexiones de los usuarios que pueden ser explotados en esta área para inferir conocimiento sobre las preferencias de los usuarios y la influencia social. El hecho de que las plataformas de música en *streaming* tengan algunas funcionalidades sociales, permite que este tipo de información se utilice para la recomendación de música. Este trabajo tiene como principal objetivo hacer uso de la estructura de la red social para abordar un tipo de sesgo de recomendación derivado de la forma en que los métodos de filtrado colaborativo computan la vecindad. Estos métodos restringen las predicciones de los *ratings* de un usuario a los ítems que



han sido valorados por sus vecinos más cercanos, mientras que dejan fuera otros ítems que podrían ser de su interés. Este problema es diferente al del sesgo de popularidad causado por la distribución de la frecuencia de valoración de los ítems (*long tail*), bien conocido en el ámbito de la música, aunque ambas deficiencias pueden estar relacionadas. Nuestro propósito se focaliza en la ampliación y diversificación de la vecindad mediante la captura de los efectos de los factores de confianza y homofilia entre los usuarios a partir métricas de estructura social.

3.4.3. Uso del contexto social del usuario en los métodos de recomendación

Los métodos de filtrado colaborativo pueden mejorarse mediante la incorporación de información social. La confianza y la homofilia son dos factores que influyen en las recomendaciones que pueden inferirse de la estructura de las relaciones entre los usuarios y otros recursos de las redes sociales. Sin embargo, en la mayoría de los servicios de música en *streaming*, esos recursos son mucho más limitados, y la estructura está restringida a relaciones de amistad bidireccionales, lo que no permite aplicar algunas métricas basadas en grafos. En este trabajo, usamos la estructura de la amistad para derivar esos factores y posteriormente incluirlos en el proceso de recomendación.

La equivalencia estructural es una propiedad aplicable a las comunidades en las redes sociales, a menudo utilizada para identificar comunidades implícitas mediante el cálculo de la similitud de equivalencia entre pares de nodos en la red. La similitud de equivalencia se basa en el solapamiento entre la vecindad de esos nodos. En el contexto de este trabajo, esta métrica puede aplicarse a la estructura de amistad cuyos nodos representan a los usuarios. Los nodos con alta similitud se consideran parte de la misma comunidad implícita. Esta es una forma de capturar el concepto de homofilia, ya que los usuarios que pertenecen a la misma comunidad suelen compartir intereses y preferencias. Por lo tanto, las valoraciones de esos usuarios pueden ser utilizadas por los métodos de recomendación. El cálculo de la similitud de equivalencia entre dos usuarios u_i y u_j con vecindades $N(u_i)$ y $N(u_j)$ respectivamente puede hacerse mediante diferentes métricas, entre las que se encuentra la de Jaccard:

$$socialSim_{Jaccard}(u_i, u_j) = \frac{|N(u_i) \cap N(u_j)|}{|N(u_i) \cup N(u_j)|} \quad (10)$$

De la misma estructura social se puede inferir la confianza. Suponemos que la confianza de un usuario en otro depende de la influencia que éste tenga en el primero. Por un lado, parece razonable que la influencia de los amigos en un usuario determinado sea mayor cuanto menor sea el número de amigos de este último. Por otra parte, los usuarios que tienen más influencia son los que tienen mayor número de amigos. Teniendo en cuenta estas premisas, podemos obtener una función que representa el grado de confianza que un usuario u_a tiene en otro usuario u_f perteneciente a su grupo de amigos. Si F_i es el conjunto de usuarios conectados directamente al usuario u_i , podemos definir la confianza de la forma siguiente:

$$t(u_a, u_f) = \frac{1}{1 + \log F(u_a)} \log F(u_f) \quad (11)$$

Donde $\log F(u_i) = \log(|F_i|)$



Las dos métricas anteriores se utilizarán para encontrar nuevos vecinos del usuario activo, además de los encontrados a partir de la similitud basada en *ratings* $sim_{rat}(u_i, u_j)$. Esos conjuntos adicionales de vecinos del usuario u_a son $k_NN_{social_a}$ y $k_NN_{friends_a}$, definidos a continuación.

$$k_NN_{social_a} = \{u_1, u_2, \dots, u_k\} \quad (12)$$

Donde $sim_c(u_a, u_1) > sim_c(u_a, u_2) > \dots > sim_c(u_a, u_k)$. y sim_c es una similitud obtenida mediante la siguiente ecuación:

$$sim_c(u_i, u_j) = \sqrt{\alpha sim_{rat}(u_i, u_j)^2 + (1 - \alpha) sim_{social}(u_i, u_j)^2} \quad (13)$$

El conjunto de vecinos basado en la confianza se define según la ecuación 14.

$$k_NN_{friends_a} = \{u_1, u_2, \dots, u_k\} \mid t(u_a, u_1) > t(u_a, u_2) > \dots > t(u_a, u_k) \quad (14)$$

La idea del enfoque propuesto es complementar las recomendaciones generadas por los métodos tradicionales de filtrado colaborativo con recomendaciones basadas en la estructura de la red social de los usuarios. En el algoritmo de recomendación intervienen todos los tipos de recomendaciones descritos anteriormente. El objetivo es aumentar el número de predicciones para ampliar el conjunto de ítems potencialmente recomendables y reducir el sesgo de vecindad, al tiempo que se mejora la fiabilidad de las recomendaciones.

3.4.4. Resultados

Para validar el método propuesto, se compararon sus resultados con los de otras propuestas de la literatura. Para ello, se probaron con el mismo conjunto de datos de canciones dos métodos que no utilizan la información social y otros dos métodos que utilizan la información inferida de la estructura de amistad de la red social. Entre los primeros, se eligieron los más representativos, el método k-NN basado en el usuario y un método de factorización de matrices. Entre los que explotan la información social, se probó un enfoque en el que las recomendaciones se restringen al contexto social del usuario y otro que combina las similitudes sociales y las similitudes basadas en *ratings*. Además, la validación se realizó también para el escenario *cold start*.

Los resultados en ambos escenarios demostraron que el método propuesto aumenta la cobertura de las recomendaciones en lo que se refiere a los ítems potencialmente recomendables. Además, este aumento no se produce a expensas de la fiabilidad de las recomendaciones, sino que, por el contrario, da lugar a una disminución de los errores en la predicción de las valoraciones de esos elementos, así como a un aumento de los valores de las métricas utilizadas para evaluar la calidad de las listas top-N recomendadas. En las diferentes propuestas de la literatura relacionadas con nuestro trabajo, hasta donde sabemos, ambas mejoras no se obtienen conjuntamente, es decir, la mejora en la cobertura se logra a expensas de la fiabilidad. Además, la mayoría de estos trabajos abordan el sesgo de popularidad, mientras que nuestro objetivo es aumentar el espectro de elementos potencialmente recomendables, independientemente de si estos elementos son populares o no. Tampoco encontramos ningún



trabajo que utilice las conexiones sociales para expandir la vecindad en los métodos de filtrado colaborativo.

3.4.5. Conclusiones

El creciente uso de los servicios de música en *streaming* y el interés en su personalización es incuestionable hoy en día. Este es uno de los principales motivos por los que el auge de la investigación en el uso de información social en muchos dominios se haya extendido a los sistemas de recomendación de música.

En este trabajo se propone un enfoque centrado en la explotación de la información social disponible en las plataformas de música en *streaming*. Se trata de un esquema de filtrado colaborativo que amplía los métodos clásicos basados en los vecinos más cercanos utilizando métricas estructurales obtenidas de la red de conexiones de amistad de los usuarios. El objetivo es minimizar el sesgo de vecindad, así como aumentar la fiabilidad de las recomendaciones. La propuesta difiere de otras en la literatura en el hecho de que es un enfoque centrado en el usuario en lugar de estar centrado en los ítems. Además, no está dirigida específicamente a aumentar la diversidad o reducir el sesgo de popularidad, sino a ampliar y diversificar la vecindad de los usuarios explotando el contexto social de éstos. Los resultados muestran que el enfoque propuesto supera a otros métodos tanto en la reducción de las tasas de error de las predicciones como en el aumento del número de ítems potencialmente recomendables.

3.5. Contribución 5

3.5.1. Referencia

Sánchez-Moreno, D., Moreno-García M.N., Sonboli, N., Mobasher, B. and Burke, R. "Inferring user expertise from social tagging in music recommender systems for streaming services", Hybrid Artificial Intelligence Systems, Lecture Notes in Artificial Intelligence, Series: *Lecture Notes in Computer Science*, pp. 39-49, Springer, 2018.

3.5.2. Objetivos

Los proveedores de servicios de música en *streaming* están mostrando un creciente interés por proporcionar a los usuarios recomendaciones personalizadas fiables, ya que la oferta prácticamente ilimitada de contenido musical dificulta a los usuarios la búsqueda de artistas o canciones que les gusten. En este trabajo se pretende aprovechar la información que proporcionan las etiquetas sociales que los usuarios asignan a los ítems musicales para mejorar las recomendaciones. La mayoría de los trabajos de la literatura utilizan las etiquetas en el contexto de métodos basados en el contenido para encontrar similitudes entre canciones o entre artistas, sin embargo, nuestra intención es utilizarlas para caracterizar a los usuarios, con el fin de mejorar los algoritmos de filtrado colaborativo basado en el usuario. En concreto, el objetivo del trabajo es inferir el nivel de experiencia (*expertise*) de los usuarios a partir del análisis de la frecuencia de sus etiquetas mediante el uso de TF-IDF (*Term Frequency-Inverse Document Frequency*).



3.5.3. Uso de etiquetado social en los métodos de recomendación

El etiquetado social permite la creación de una descripción en múltiples dimensiones y dirigida por el usuario de los ítems musicales, así como una clasificación dinámica y no preestablecida. Este tipo de indexación obtenida mediante etiquetas sociales es mucho más apropiada para la música que la que se basa en las taxonomías clásicas, ya que sus categorías, como los géneros musicales, tienen límites difusos e inestables, se introducen regularmente nuevos tipos y los existentes cambian.

Aunque el etiquetado social se utiliza habitualmente para la caracterización de la música, nosotros proponemos que se utilice en la caracterización de los usuarios, con el fin de aprovechar esta información para mejorar los métodos de filtrado colaborativo basados en el usuario. Nuestra propuesta implica la inferencia del grado de experiencia del usuario mediante el análisis de las etiquetas que estos les dan a los ítems musicales.

Para entender la idea de manera intuitiva, consideremos etiquetas como *rock* y *pop* que corresponden a géneros conocidos y fácilmente identificables, por lo que podrían ser dadas por un usuario con un bajo nivel de pericia. Sin embargo, la asignación de etiquetas como *darkwave* o *trip hop*, asociadas a géneros menos identificables o más especializados, requiere un mayor nivel de conocimientos musicales. Por lo tanto, la experiencia de los usuarios puede determinarse analizando sus etiquetas.

Basándonos en el hecho de que las etiquetas que requieren más conocimientos son menos comunes que las otras, recurrimos al análisis de la frecuencia de las etiquetas para establecer el grado de experiencia. Hemos utilizado TF-IDF, una medida ampliamente usada para la recuperación y clasificación de documentos, para caracterizar a los usuarios según la cantidad y la relevancia de las etiquetas que proporcionan a los ítems. En este contexto, los términos son sustituidos por etiquetas y los documentos por los usuarios.

Al buscar perfiles de especialización de los usuarios, TF-IDF proporciona una indicación de la frecuencia de las etiquetas, pero da más relevancia a las etiquetas que son utilizadas con menos frecuencia por muchos usuarios. Esta métrica se ha utilizado en algunos trabajos para identificar las etiquetas relevantes para los ítems (es decir, canciones o artistas), sin embargo, nuestro objetivo es identificar las etiquetas relevantes para los usuarios. Las etiquetas de usuarios expertos tendrían valores altos de TF-IDF (alto nivel de especialización), mientras que las etiquetas de usuarios no expertos tendrían valores bajos de esta métrica. Para un usuario $u \in U$ y una etiqueta t , definimos TF-IDF de la forma siguiente:

$$tf-idf(U, u, t) = tf(u, t) \times idf(U, t) \quad (15)$$

$$idf(U, t) = 1 + \log\left(\frac{|U|}{df(U, t)}\right) \quad (16)$$

Where $tf(u, t)$ is the frequency of the tag t for the user u .

Al aplicar los métodos de filtrado colaborativo, en la predicción de las recomendaciones se introduce la experiencia de los usuarios como un peso, de manera que se aumenta la influencia



de los usuarios expertos en las recomendaciones frente a la de los no expertos. La similitud ponderada se calcula de la siguiente manera:

$$w\text{-sim}(u_a, u_i) = \text{sim}(u_a, u_i) \times \overline{tf-idf}(u_i) \quad (17)$$

Donde $\overline{tf-idf}(u_i)$ es el promedio de $tf-idf$ para el usuario u_i .

3.5.4. Resultados

Para comprobar si la influencia de los usuarios expertos, obtenida a partir de las etiquetas sociales, mejora los métodos de filtrado colaborativo basados en k-NN, se han comparado los resultados de este método obtenidos utilizando la similitud ponderada con TF-IDF con los obtenidos sin introducir la ponderación. Para el cálculo de la similitud entre usuarios se han utilizado diferentes métricas: Similitudes del coseno y de Jaccard por una parte, y similitudes basadas en las distancias de Chebyshev y euclídea por otra.

El estudio se llevó a cabo utilizando un conjunto de datos que contiene información sobre relaciones sociales, etiquetas y datos de reproducción de artistas del sistema de música online Last.fm. Las pruebas se realizaron aplicando validación cruzada.

La introducción del peso correspondiente al grado de experiencia produjo mejoras significativas respecto al método básico con todas las métricas de similitud utilizadas.

3.5.5. Conclusiones

El etiquetado social se suele utilizar en el contexto de los sistemas de recomendación en los métodos de filtrado colaborativo basados en contenido para encontrar similitudes entre los ítems. Sin embargo, en este trabajo se hace uso las etiquetas sociales para inducir el grado de experiencia de los usuarios. Esta información se utiliza para mejorar los métodos basados en el usuario, dando más relevancia en el cálculo de la similitud a las opiniones de los usuarios más expertos.

Para la validación se realizó un estudio con un conjunto de datos que contiene información obtenida del sistema de música Last.fm. Los resultados demostraron que la introducción de un peso proporcional a la experiencia del usuario produce predicciones significativamente mejores, independientemente de la métrica de similitud aplicada.