

An Evaluation on User Profiling Methodologies and the Challenges Associated with it in Recommender Systems

Priyanka Muniyandi¹ & Dr. Ananthi Sheshasaayee²

¹ Research Scholar, Quaid-E-Millath Government College for Women

² Head & Associate Professor, Quaid-E-Millath Government College for Women.

Abstract: *One of the important tasks of building a recommendation system is analyzing the characteristics of users' interests. This is often referred to as user profiling. User profiling refers to the extraction of user labels on different attributes such as age, gender, occupation, income, and interests. Complete and accurate attribute labels will effectively reveal the inherent characteristics of users, thus greatly facilitating accurate personalized recommendations.*

1. INTRODUCTION

User profiling is an important part of recommender systems. These profiles model users' interests and preferences and are used to assess an item's relevance to a particular user. In any domain it is difficult to extract explicit signals from the users about their interests, and user profiling depends on in-depth analyses of users' habits in that particular domain.

In recent years, the concept of personal assistants helping individuals to do their work has arisen. Interface agents are computer programs that use artificial intelligence techniques to provide a user with computer-based tasks with active assistance.

Agents need some knowledge of the activities they have to perform in order to support users, and they have to be aware of the user's needs, behaviors and preferences. A requirement for the creation of systems that provide customized services is to focus on user profiles, i.e. a representation of individual user interests. User profiles differ from one agent to another in content, in the acquisition process and in its use. The type of information that forms a profile is typically based on the application and much of this data is simply the collection of interests of a user. Some agents often consider a user's dislikes, personal user data, and some other type of information specifically required by the role in which the agent is engaged. To build profiles, several methods have been developed. Most profiles are either created directly by users who supply items of interest or by automated methods through which an agent can learn user preferences. [1]

In general, user profiling began with the mere extraction of information and compilation of user data. Older systems were more worried about receiving information directly from users than the system was specifically informing users about the appropriate data. But this approach is not considered effective since the user is never involved in supplying the feedback explicitly, so today's research focuses more on profiling user data indirectly based on those user behavior.

2. USER PROFILE

A profile is a summary of someone that includes the most significant or interesting information about him or her. A user profile or user model provides important information about an individual user in the context of users of software applications. The explanation for creating user profiles is that when using software apps, users vary in their tastes, interests, history and goals. To provide consumers with customized services, it is important to discover these variations.[2]

User profiling is a virtual depiction of individual knowledge connected to a single user in a personalized desktop configuration. User profiling was defined by *Kanoje et al.*[3] as a means of determining the interest data of the user that is based on the user information and the retrieval of the exact device the happiness of users

2.1. Types of User Profiles

User profile can be broadly grouped into two main types: Inactive profile and Active profile [4]. These categories are explained in details below.

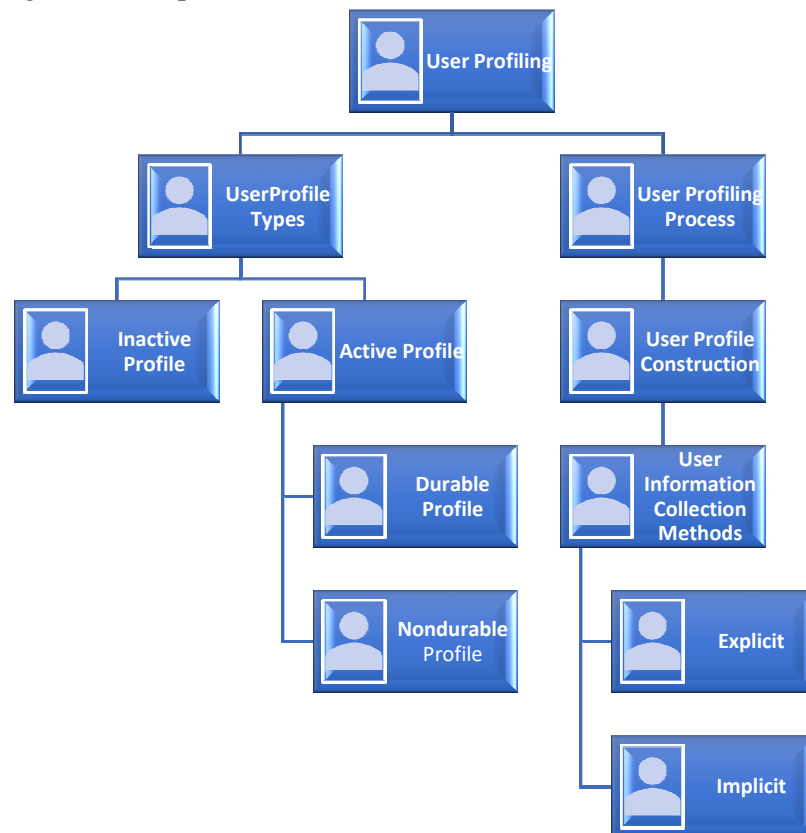


Fig 2.1 User Profiling Categorization

2.1.1 Inactive Profile

This kind of profile consists of all the details that remains largely unchanged, Such as name, email ID, gender, etc. are therefore considered static in nature. According to Poo et al. [5] “Static profiling approach is a process of analysing user’s predictable and static characteristics”

2.1.2 Active Profile

In a situation where there is a high data delivery velocity, the dynamic or Active profile is often correct. This is also known as a behavioral or adaptive profile. In addition, the current ontology of the user is used to guide the extraction of the profile, identify the set of relationships in question and provide the dictionary of the person.[6]. Active profile is further classified into Durable and Nondurable profiles.

3. USER PROFILING CHALLENGES

There are two main challenges in user profiling process. These are the generation of an initial user profile for a new user and the continuous update of the profile information to adapt user's changing preferences, interests and needs. In literature two fundamental user profiling methods have been proposed to tackle these challenges. These are the content-based and the collaborative methods. Both of these methods have limitations and the hybrid user profiling has been proposed to overcome these limitations by combining these two. The user profiles that are created based on the traditional user profiling methods were not adequate to personalize different services. For this reason, various clustering and classification algorithms have been utilized to create more comprehensive user profiles. However, these profiles are lack in representing the multi-dimensionality of the user profiles and still not adequate to personalize different services. [7]

Established methods of user profiling are typically based on simple linear regression or classification models, so they do not learn the abstracted high-level features from user data automatically or model the interaction between features. Furthermore, current user profiling approaches are mostly based on homogeneous data from one source, which is not sufficiently rich for users to be accurately represented. In fact, many user data comes from various sources, which can help create higher-quality user profiling.

To create a single model of user representation for user profiling. For each user attribute, current user profiling methods also train an individual model. The number of user attributes, however, can be huge in reality. Therefore, current methods of user profiling also require a lot of model training and storage. Furthermore, the fundamental similarities between various user attributes have yet to be thoroughly explored. Is it possible to find a way to create a single user representation from heterogeneous data such that the model can collect a user's details from different dimensions in a detailed manner? For such a dilemma, multi-task learning-based learning algorithms and user-embedding technologies provide promising solutions.

4. USER PROFILING METHODOLOGIES

The content and amount of the information within a user profile can vary depending on the application area. However, regardless of the information, the accuracy of the user profile is based on how the user information is gathered and organized, and how accurately this information reflects the user. In other words, it depends on the user profiling process in which the information is gathered, organized and interpreted to create the summarization and the description of the user [8].

4.1 User Profiling Process

This section describes the user profiling process. It provides the entire processes that are involved in user profiling such as profile construction and user information collection methods. The detailed description of each process is given in subsections below.

4.1.1 User Profile Construction

For a person, the user profile can be constructed by collecting user information through direct contact with the user or automatically by the system monitoring the user's activities. In user profile development, various learning algorithms/information retrieval systems are employed based on the choice of representation. Profile construction can be divided into the profile of the semantic network, keyword and definition. A user profile may be manually created by users or experts. For most consumers, however, this strategy is difficult and time consuming, which hinders the expansion of personalized adoption of services. The technique that automatically employs user input for profile formation is most common in contracts. Other methods are commonly used, such as neural networks/genetic algorithms that rely on probabilities or model of vector space, which have been found to be more successful in many areas.

4.1.2 User Information Collection Methods

The starting point for user profiling techniques is data collection regarding a single user. However, the system is required to classify users exclusively, which serves as an integral requirement of the system. Information about users can be accessed in the form of feedback from the user or collected automatically by an intelligent agent. Five common methods, such as registration, software agents, improved proxy servers, cookies and session ids, are used to gain user identification. Despite these challenges, due to its transparency for users and the provision of cross-section tracking capabilities, cookies are more effective and commonly used among the techniques. In addition, login approach is favored in terms of continuity and improvement of accuracy as it tracks the behavior of the user over a session between computers if users are willing to register and login at any visit. In order to create a user profile, user information can be obtained explicitly or implicitly [9].

4.1.2.1 Explicit Method: The explicit method of collecting user information, also known as acquiring user input, depends on collecting personal information from the user. In this case, the user profile can be accessed directly through direct user interference. While filling out the form, a consumer can be asked to express their opinion. The explicit data is given by users through a method of survey and registration process [10]. Demographic attributes such as the name of the user, the address of the user, his telephone number, marital status, work status, birthday, personal interest, and hobbies can be the data collected. Other pieces of user data, such as the online transaction or web operation of the user, may also be categorized as explicit information. For example, certain pieces of user evidence which consist of the average amount consumed by the user on the purchased item, the most purchased item categories, and the user's most frequent web visit. [11]. This information collection method is inefficient and has a low usability rate since users are often reluctant to make their profile

information public as their privacy is considered more essential. Filling out the forms, on the other hand, is burdensome, and users often prefer to avoid it. This method of data collection has an inherent time consumption, in addition to the aforementioned drawbacks, and is subject to the user's willingness to participate. Thus, the user's inability to provide personal details would halt the profile's creation.

4.1.2.2 Implicit method This profile selection method relies on implicit user feedback for its creation. Implicit data can be obtained in this process through intelligent agent or datamining techniques that evaluate user behavior, such as obtaining the user rule [12]. According to Alaoui et al. [13] the implicit method of profile collection has been described as a means of collecting information in order to create a profile that tracks the behavior of the user. This approach, however, has an advantage over the explicit one as it does not require any extra task to be performed during the construction process by the users. During their analysis, Kelly and Teevan [14] demonstrated this strategy by providing the most effective technique used to gain implicit input and the type of knowledge that can be derived from the actions of the user. Furthermore, automated updates benefit from the implicit profile process, as it mostly depends on the application of machine learning. Nevertheless, in the initial stage before the development of an accurate user profile, a significant amount of interaction between the user and the content is needed. Implicit profiling is synonymous with user profiling based upon ontology.

5. CONCLUSION

User profiles represent users and represents the desires, needs, habits and interests of each user. These profiles are the result of the process of user profiling and are important for the personalization of the service. This paper provides a study of user profiling, including its associated principles, approaches, strategies, and the literature's current solutions. As a future work, by concentrating on other function weighing algorithms, it is expected to continue working on this ideal approach to achieve more detailed multi-dimensional user profiles that could be used for the personalization of different services.

References

[1] Schiaffino, Silvia N., and Analía Amandi. "User profiling with Case-Based Reasoning and Bayesian Networks." *IBERAMIA-SBIA 2000 Open Discussion Track*. 2000.

[2] Schiaffino, Silvia, and Analía Amandi. "Intelligent user profiling." *Artificial Intelligence An International Perspective*. Springer, Berlin, Heidelberg, 2009. 193-216.

[3] Kanoje, Sumitkumar, Sheetal Girase, and Debajyoti Mukhopadhyay. "User profiling trends, techniques and applications." *arXiv preprint arXiv:1503.07474* (2015).

[4] Farseev, Aleksandr, et al. "" 360° user profiling: past, future, and applications" by Aleksandr Farseev, Mohammad Akbari, Ivan Samborskii and Tat-Seng Chua with Martin Vesely as coordinator." *ACM SIGWEB Newsletter Summer* (2016): 1-11.

- [5]Poo, Danny, Brian Chng, and Jie-Mein Goh. "A hybrid approach for user profiling." *36th Annual Hawaii International Conference on System Sciences, 2003. Proceedings of the. IEEE, 2003.*
- [6]Krishnan, Gokul S., and S. Sowmya Kamath. "Dynamic and temporal user profiling for personalized recommenders using heterogeneous data sources." *2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE, 2017.*
- [7]Cufoglu, Ayse. "User profiling-a short review." *International Journal of Computer Applications 108.3 (2014).*
- [8]Henczel, Sue. "Creating user profiles to improve information quality." *Online (Weston, CT) 28.3 (2004): 30-33.*
- [9] Gauch, Susan, et al. "User profiles for personalized information access." *The adaptive web. Springer, Berlin, Heidelberg, 2007. 54-89.*
- [10]Raghu, T. S., et al. "Dynamic profiling of consumers for customized offerings over the Internet: A model and analysis." *Decision Support Systems 32.2 (2001): 117-134.*
- [11]Yang, Yinghui Catherine. "Web user behavioral profiling for user identification." *Decision Support Systems 49.3 (2010): 261-271.*
- [12]Adomavicius, Gediminas, and Alexander Tuzhilin. "Expert-driven validation of rule-based user models in personalization applications." *Data Mining and Knowledge Discovery 5.1-2 (2001): 33-58.*
- [13]Alaoui, Sara, Younès EL Bouzekri EL Idrissi, and Rachida Ajhoun. "Building rich user profile based on intentional perspective." *Procedia Computer Science 73 (2015): 342-349.*
- [14]Kelly, Diane, and Jaime Teevan. "Implicit feedback for inferring user preference: a bibliography." *Acm Sigir Forum. Vol. 37. No. 2. New York, NY, USA: ACM, 2003.*