Privacy issues in holistic recommendations

Privacy Issues in Holistic Recommendations

Federica Cena
Dept. of Computer Science,
University of Turin
Torino, Italy
federica.cena@unito.it

Ruggero G. Pensa
Dept. of Computer Science,
University of Turin
Torino, Italy
ruggero.pensa@unito.it

Amon Rapp
Dept. of Computer Science,
University of Turin
Torino, Italy
amon.rapp@gmail.com

ABSTRACT

In this paper we point out some relevant issues in relation to privacy when providing holistic recommendations. We emphasize that a holistic recommender should be fair, explainable and privacy-preserving to ensure the ethicality of the recommendation process. Further, we point out relevant research questions that should be addressed in the future, as well as propose some preliminary suggestions to face the emergent issues with reference to privacy in the recommendation domain.

CCS CONCEPTS

• **Human-centered computing** → *HCI theory, concepts and models*; • **Security and privacy** → **Privacy protections**.

KEYWORDS

Privacy, Ubiquitous Context, Holistic User Model, Recommender Systems, Context-aware recommender systems

ACM Reference Format:
Federica Cena, Ruggero G. Pensa, and Amon Rapp. 2019. Privacy Issues in Holistic Recommendations. In *27th Conference on User Modeling, Adaptation and Personalization Adjunct (UMAP'19 Adjunct), June 9–12, 2019, Larnaca, Cyprus.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3314183.3323461

1 INTRODUCTION

In recent years, due to an increased use of ubiquitous and wearable technologies as well as social networks, the everyday-life of the individuals is now tightly bounded with the digital life: an increasingly large fraction of what we say and do, from taking a picture to buying a good, from visiting a place to meeting a friend, leaves a digital trace that tells something about our lives [13]. Unprecedented capabilities in collecting user and context data, together with technical and theoretical advancements in computer science give novel opportunities for User Modeling: a User Model (UM) could now exploit information related to many aspects of the user (from her medical records to her food behavior, from her physiological parameters to her psychological states, etc.), creating a sort of total, holistic representation of an individual [4, 17]. The increased complexity of UM would then enable novel forms of personalized services, which may be delivered anywhere at any time, potentially impacting every domain of people' daily life [5]. In particular, new forms of recommendations might be able to give suggestions on an aspect in a specific domain starting from data coming from multiple, and maybe apparently unrelated, contexts. This would enable a sort of holistic recommendation, that is to say, a recommendation built on the ground of a holistic representation of the user's needs, interests, knowledge, preferences [5]. Such a representation is defined through the collection of data from diverse data sources and by reasoning over them in order to populate the different facets describing the person. As an example, the interests of the users may be inferred on the ground of the places she visits, the topics she discusses or the opinions she expresses on social networks or through the reviews she writes. The notion of holistic recommendation can be seen as an evolution of*context-aware (recommender) systems* [1], which have a long tradition in creating algorithms considering different variables about the user in order to provide (possibly just-in-time) recommendations. For instance, they can mine social networks to automatically infer context features [18, 21]. Other works focused on providing real-time dynamic recommendations [23]. Probably the field where real-time recommendations have been most used is tourism [2]. Holistic recommendations are based on a "complete" representation of the user. Such systems could capture every aspect of the user pertaining to the different spheres of her life, even over very long periods of time, and handle their changes, according to a lifelong user modeling vision [12]. In principle, the user could be allowed to explore her holistic user model, which could further scaffold processes of self-reflection. Moreover, holistic recommenders could be exploited to make forecasts on the user's goals, behavior and preferences on the basis of past and current trends in her data. For example, we can think of smart adaptive systems able to predict what would be useful for users, by simulating the future evolution of their data and setting the right goals to be reached based on such predictions. Then, they could provide recommendations triggered by the user's current condition, suggesting which kinds of actions and changes the user should put in place to meet the set goals. Finally, holistic recommenders can provide suggestions in every domain and context of the user's everyday life, becoming a sort of pervasive personal advisor. However, even if extremely accurate, how and when such suggestions are delivered may arise potential issues, e.g. interrupting the user's activities, or being out of context and socially inappropriate (e.g. the user does not want to have recommendations about the diet to follow, which could be seen by her friends, when she is out with them). In this perspective, privacy surely becomes a fundamental concern. How can we define a holistic user model, provide holistic recommendations and ensure, at the same time, the privacy issues that may arise from the

UMAP'19 Adjunct, June 9–12, 2019, Larnaca, Cyprus

ACM ISBN 978-1-4503-6711-0/19/06…$15.00
https://doi.org/10.1145/3314183.3323461

collection of data potentially coming from the whole individual's life?

2 PRIVACY ISSUES IN HOLISTIC RECOMMENDATION

Although holistic recommendation may provide undeniable advantages to end users, they use large amounts of potentially sensitive/private data (e.g., health data, location data, sexual preferences, and so on). Moreover, models and algorithms trained by using such behavioural data may leverage discriminating patterns, e.g., gender- or ethnicity-based decisions that results in recommending two completely different products to two different persons (e.g., with different gender or ethnicity), sharing exactly the same behaviour. In addition, it is important to trace provenance and proof of the data in the user and context model, in the perspective to make the models explainable and scrutable by the user [24] and give her the possibility to know how the data in the models are gathered and inferred. Consequently, an ethical holistic recommender system algorithm or model must be fair, explainable and privacy-preserving.

Fairness. With the recent advances in artificial intelligence and machine learning and the resulting concerns for their consequences on human freedom and rights, in the last decade, many research groups have addressed fairness issues [9, 11]. The topic, however, has been only superficially addressed in the recommender systems community [9], even though this is a crucial aspect of holistic recommendation, due to the heterogeneity and complexity of data sources. Consequently, many research questions are still far from being solved or even discussed, in some cases. First, how widespread is the problem of algorithmic bias in recommender systems? How to address the complexity of holistic recommender systems in an efficient way without affecting the accuracy of the recommendations too much? How to detect discrimination in the underlying algorithms? How to regularize them in order to dismiss potentially discriminative patterns and decisions?

Explainability. Fairness is strictly related to explainability, i.e., the ability of an algorithm to be interpretable and scrutable, a popular issue in machine learning due to the opacity of recent (and popular) nonlinear techniques, such as deep learning algorithms [16]. In holistic recommendation, a way to achieve explainability consists in enabling the user to visualize her models in order to add the meaning associated to some contexts, people, memories, etc. following a *scrutability* vision [24]. This can provide valuable source of information otherwise not easily collectible. But how to make such a complex User Model scrutable? In principle, the user could be allowed to explore her holistic user model, which could further scaffold processes of self-reflection. However, it is not feasible to present all the data to the user, since it could cause information overload. This problem is strictly related to the granularity of the collected data. For example, not all the data about the user's blood pressure should be provided to her. Data to be visualized could change format according to the specific application that is using them, and/or the user's features (e.g., goals or expertise), and/or the specific context. Also the interaction modalities should be adapted to the user's features and context. How should the data be presented? At which level of aggregation? What are the best representations (graphical, textual, etc.)?

Privacy preservation. The problem of privacy has been addressed by the data mining and machine learning communities for twenty years [22], but the recent advances brought by differential privacy [6] have opened new research opportunities in recommender systems as well [3, 14]. Holistic recommendation adds new complex dimensions to the problem of computing privacy-preserving models: the interplay between lifestyle aspects, purchasing behaviour, contextual properties and so on could reveal very private habits and preferences. Hence, the most intriguing research questions concern the definition of leakage or attack models (and the related countermeasures) leveraging the uniqueness of the relations between multiple dimensions, and how to improve user's awareness about privacy, by enhancing her perception of the trade-off between the accuracy of the recommendations and the amount of required disclosure of private sensitive information.

Attitude to privacy. Although fairness and privacy are universally recognized challenges, users may exhibit different attitudes towards them, as shown in several studies [15, 19, 25]. Hence, what is considered private, sensitive by some users, is deemed safe or public by other people. The same can be applied to discrimination: subjective discrimination, in particular, can be perceived differently, according to diverse sensibility degrees. An issue here is how to gather such preferences: gamification approach can be used for the scope [8, 20]. Moreover, if, from one hand, holistic recommendation should take into account all the above-mentioned ethical issues, from another hand, the attitude of the users towards these issues could be used to adjust recommendation models and patterns. Learning the attitude towards privacy or discrimination issues is then crucial, and a holistic recommender system should be able to automatically adjust its models according to explicit or inferred user preferences. Explainability could be then used to enable the user to decide to what extent recommendation should take privacy and fairness into account. An even more challenging achievement would make such mechanism dynamic and able to adapt to different contexts and situations.

Ethical issues. In the light of the characteristics of Holistic-based recommendations, a lot of ethical issues arise, especially in relation to health. Considering problem complexity and data structure as well as security and autonomy aspects, the user empowerment and engagement is of paramount importance, fostering reflection of the recommendations and of integrating the user into the loop. In a preliminary empirical analysis, [10] show that presenting uncertainty to the user might help the user to reflect the recommendation and integrate him into the loop. Moreover, it might increase trust, perceived transparency, system responsibility, and overall user satisfaction. Moreover, it is crucial to provide accurate recommendation, since wrong recommendation, especially in health domain, can be dangerous for the user [7].

3 CONCLUSION

In this paper, we tried to surface some relevant themes in relation with privacy when building holistic user models in order to provide holistic recommendations. We pointed out research questions that should be explored in the future, as well as proposed some tentative suggestions to preliminary address some of the key privacy issues emerging in the user model domain.

REFERENCES

[1] Gediminas Adomavicius and Alexander Tuzhilin. 2005. Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Trans. on Knowl. and Data Eng.* 17, 6 (June 2005), 734–749. https://doi.org/10.1109/TKDE.2005.99

[2] Claudio Biancalana, Fabio Gasparetti, Alessandro Micarelli, and Giuseppe Sansonetti. 2013. An Approach to Social Recommendation for Context-aware Mobile Services. *ACM Trans. Intell. Syst. Technol.* 4, 1, Article 10 (Feb. 2013), 31 pages. https://doi.org/10.1145/2414425.2414435

[3] Antoine Boutet, Davide Frey, Rachid Guerraoui, Arnaud Jégou, and Anne-Marie Kermarrec. 2016. Privacy-preserving distributed collaborative filtering. *Computing* 98, 8 (2016), 827–846.

[4] Federica Cena, Silvia Likavec, and Amon Rapp. 2018. Real World User Model: Evolution of User Modeling Triggered by Advances in Wearable and Ubiquitous Computing: State of the Art and Future Directions. *Information Systems Frontiers* (2018), 1–26. https://doi.org/10.1007/s10796-017-9818-3 cited By 0; Article in Press.

[5] Federica Cena, Amon Rapp, Cataldo Musto, and Pasquale Lops. 2018. Towards a Conceptual Model for Holistic Recommendations. In *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization, UMAP 2018, Singapore, July 08-11, 2018*, Tanja Mitrovic, Jie Zhang, Li Chen, and David Chin (Eds.). ACM, 207–210. https://doi.org/10.1145/3213586.3225248

[6] Cynthia Dwork and Aaron Roth. 2014. The Algorithmic Foundations of Differential Privacy. *Foundations and Trends in Theoretical Computer Science* 9, 3-4 (2014), 211–407.

[7] Jennifer D. Ekstrand and Michael D. Ekstrand. 2016. First Do No Harm: Considering and Minimizing Harm in Recommender Systems Designed for Engendering Health. In *Proceedings of the Workshop on Recommender Systems for Health at RecSys '16*. 14.

[8] Severin Hacker and Luis Von Ahn. 2009. Matchin: eliciting user preferences with an online game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1207–1216.

[9] Sara Hajian, Francesco Bonchi, and Carlos Castillo. 2016. Algorithmic Bias: From Discrimination Discovery to Fairness-aware Data Mining. In *Proceedings of ACM SIGKDD 2016, San Francisco, CA, USA, August 13-17, 2016*. 2125–2126.

[10] Katja Herrmanny and Ayseguel Doganguen. 2017. The Impact of Prediction Uncertainty in Recommendations for Health-Related Behavior. In *Second International Workshop on Health Recommender Systems*. 14.

[11] Toshihiro Kamishima, Shotaro Akaho, Hideki Asoh, and Jun Sakuma. 2012. Fairness-Aware Classifier with Prejudice Remover Regularizer. In *Proceedings of ECML PKDD 2012 (Part II), Bristol, UK, September 24-28, 2012*. 35–50.

[12] J. Kay. 1995. The um Toolkit for Reusable, Long Term User Models. In *User Model. User-Adap.*, Vol. 4. 149–196.

[13] David Lazer, D Brewer, N Christakis, J Fowler, and G King. 2009. Life in the network: the coming age of computational social. *Science* 323, 5915 (2009), 721–723.

[14] Jianqiang Li, Ji-Jiang Yang, Yu Zhao, Bo Liu, Mengchu Zhou, Jing Bi, and Qing Wang. 2017. Enforcing Differential Privacy for Shared Collaborative Filtering. *IEEE Access* 5 (2017), 35–49.

[15] Kun Liu and Evimaria Terzi. 2010. A Framework for Computing the Privacy Scores of Users in Online Social Networks. *TKDD* 5, 1 (2010), 6:1–6:30.

[16] Grégoire Montavon, Sebastian Lapuschkin, Alexander Binder, Wojciech Samek, and Klaus-Robert Müller. 2017. Explaining nonlinear classification decisions with deep Taylor decomposition. *Pattern Recognition* 65 (2017), 211–222.

[17] Cataldo Musto, Giovanni Semeraro, Cosimo Lovascio, Marco de Gemmis, and Pasquale Lops. 2018. A Framework for Holistic User Modeling Merging Heterogeneous Digital Footprints. In *UMAP (Adjunct Publication)*. ACM, 97–101.

[18] Ante Odić, Marko Tkalčič, Jurij F Tasič, and Andrej Košir. 2013. Predicting and detecting the relevant contextual information in a movie-recommender system. *Interacting with Computers* 25, 1 (2013), 74–90.

[19] Ruggero G. Pensa and Gianpiero di Blasi. 2017. A privacy self-assessment framework for online social networks. *Expert Syst. Appl.* 86 (2017), 18–31. https://doi.org/10.1016/j.eswa.2017.05.054

[20] Amon Rapp, Federica Cena, Cristina Gena, Alessandro Marcengo, and Luca Console. 2016. Using game mechanics for field evaluation of prototype social applications: a novel methodology. *Behaviour & IT* 35, 3 (2016), 184–195. https://doi.org/10.1080/0144929X.2015.1046931

[21] Mina Razghandi and Seyyed Alireza Hashemi Golpaygani. 2017. A Context-Aware and User Behavior-Based Recommender System with Regarding Social Network Analysis. In *e-Business Engineering (ICEBE), 2017 IEEE 14th International Conference on*. IEEE, 208–213.

[22] Pierangela Samarati and Latanya Sweeney. 1998. Generalizing Data to Provide Anonymity when Disclosing Information (Abstract). In *Proceedings of ACM PODS 1998, June 1-3, 1998, Seattle, Washington, USA*. 188.

[23] Norma Saiph Savage, Maciej Baranski, Norma Elva Chavez, and Tobias Höllerer. 2012. I'm feeling loco: A location based context aware recommendation system. In *Advances in Location-Based Services*. Springer, 37–54.

[24] Rainer Wasinger, James Wallbank, Luiz Augusto Sangoi Pizzato, Judy Kay, Bob Kummerfeld, Matthias Böhmer, and Antonio Krüger. 2013. Scrutable User Models and Personalised Item Recommendation in Mobile Lifestyle Applications. In *Proceedings of UMAP 2013, Rome, Italy, June 10-14, 2013, Proceedings*. 77–88.

[25] Allison Woodruff, Sarah E. Fox, Steven Rousso-Schindler, and Jeffrey Warshaw. 2018. A Qualitative Exploration of Perceptions of Algorithmic Fairness. In *Proceedings of CHI 2018, Montreal, QC, Canada, April 21-26, 2018*. 656.