Electronic Commerce Research and Applications 41 (2020) 100962

Contents lists available at ScienceDirect

Electronic Commerce Research and Applications

journal homepage: www.elsevier.com/locate/elerap

A 2020 perspective on "A graph-oriented model for hierarchical user interest in precision social marketing"

Zhiguo Zhu[a,*], Liping Kong[c], Xiaoyi Deng[b], Bo Tan[a]

[a] School of Management Science and Engineering, Dongbei University of Finance and Economics, Dalian, Liaoning, China
[b] Business School, Huaqiao University, Fengze, Quanzhou, Fujian, China
[c] Dongbei University of Finance and Economics, Dalian, Liaoning, China

ARTICLE INFO

Keywords:
Consumer interest inference
Interest model evaluation
Personalization
Precision social marketing
User interest profiles
User interest representation

ABSTRACT

Currently, with the rapid development and popularity of online social networks, social marketing is in the ascendant, while the lack of precision and personalization has become a prominent issue in current social marketing practices. A fundamental solution to this issue is how to accurately and comprehensively infer user interests. In our earlier research (Zhu et al., 2019), we constructed a *user interest graph* (UIG) model represented by a hierarchical tree structure with an exponential interest decay scheme. For this study, from a 2020 perspective, this research commentary provides some comments on this line of research in terms of four dimensions for inferring user interests: data acquisition, representation of user interests, construction and enhancement of a user interest profile, and evaluation of the constructed model. Finally, some challenges and opportunities in this domain are discussed for future work.

1. Introduction

With the growing popularity of diverse *online social network* (OSN) services in recent years, how to fully utilize the abundant information has become one of the most important issues in electronic commerce. This involves *user-generated content* (UGC), relationships and behaviors that are identified and used to carry out effective marketing activities in the practice of social marketing (Han et al., 2018; Zhu et al., 2016).

The simple pattern of "pushing, forwarding, and diffusing" marketing messages by leveraging online social relations has been widely adopted in current social marketing practices. Yet, because of the lack of deeper consideration of users' interests and preferences, this pattern can easily result in an uninterested user's antipathy in the course of marketing message diffusion in OSNs. Thus, how to accurately and comprehensively infer user interests in an OSN has become a fundamental solution for precision social marketing.

2. Our prior research

In our earlier work (Zhu et al., 2019), an inverted tree-shaped structure was constructed based on the page categories of Dianping.com and Yelp.com. It contains a three-level interest node representation for the *user interest graph* (UIG). After extracting the feature terms from multidimensional UGC and interaction records in the Dianping user profile, semantic similarity is calculated between the feature terms and the interest nodes in the UIG tree. Furthermore, taking the characteristic of user's interest decaying over time into consideration, we also design a scheme of exponential interest decay. Finally, the calculated scores on all nodes in the tree are able to reflect the distribution and extent of the user's explicit and implicit interests.

In comparison to predicted interests with user's real interests obtained from investigations, the proposed algorithm outperforms three benchmarks and two similar models on various standard metrics at all levels in the tree for users in the collected dataset. The achievements of this study will provide important basic technologies and decision supports for social marketing practices.

3. A 2020 perspective on user interest profiles

From a 2020 perspective, some opportunities worthy of focus for inferring user interests to construct a profile for precision social marketing are provided concerning four dimensions: (1) data acquisition, (2) representation of user interests, (3) construction and enhancement of user interest profiles, and (4) evaluation of the constructed profiles.

3.1. Data acquisition

A straightforward way of increasing a user's interests involves

* Corresponding author.
E-mail addresses: zhu_zg@dufe.edu.cn (Z. Zhu), pltdeng@dlut.edu.cn (X. Deng).

https://doi.org/10.1016/j.elerap.2020.100962
Received 7 February 2020; Received in revised form 18 February 2020; Accepted 24 February 2020
Available online 26 February 2020
1567-4223/

leveraging information from the user's activities and online records in OSNs, such as posting, liking or replying to a post, and commenting. However, inferring user interests from their activities requires users to be active, which is not always the case. For example, Gong et al. (2015) revealed that a significant portion of Twitter users is passive: they only follow other users but do not generate any content. Therefore, it is difficult to infer user interest profiles for passive users in OSNs.

To address this problem, leveraging information from the social relations of a user can be useful to infer their interests, especially for passive users. The social relations information of users, such as topical followees, or list memberships can be leveraged (Besel et al., 2016). However, sometimes it is still difficult to distinguish the activities of a user's followees that are relevant to the interests of that user. In addition, combining external information from other OSNs of users can support better interest inference compared to a single OSN. Nevertheless, how to effectively combine different information sources, and whether external data also have more irrelevant contents for inferring user interests and introduce noise that requires more study.

3.2. Representation of user interests

Concerning this process, using groups of keywords for representing user interests is popular in OSNs (Bhattacharya et al., 2014). As an alternative approach, other special types of keywords such as tags and hashtags have been widely used. In contrast to the words mined from short UGC text, keywords from tags and hashtags may be more informative and categorical in nature. Topics distilled from topic modelling approaches such as *latent Dirichlet allocation* (LDA) are also popular for representing user interest profiles (Trikha et al., 2018). To sum up, one of the shortcomings of the keyword-based user profiles is *polysemy*. This issue occurs when a word may have multiple meanings which cannot be distinguished. In addition, these keyword-based approaches lack semantic information and cannot capture relationships among these words.

Concept-based user interest is represented as conceptual nodes and their relationships, and their formats usually come from a pre-existing *knowledge base* (KB). The KBs leveraged for different purposes of user modelling are varied, such as cross-domain KBs or domain-specific KBs (Nguyen et al., 2015). This type of representation aims to address the polysemy problem of keyword-based ones. Currently, there are two mainstream approaches to concept-based representation: entity-based and category-based. The entity-based approach involves extracting entities from information sources such as a user's posts, and uses these entities to represent user interests (Perera et al., 2016). The category-based approach is another concept-based representation which uses some cross-domain categories, such as DBpedia (Faralli et al., 2017; Flati et al., 2014), which can represent more general user interests. In addition, it is also worth noting that the style of user interest representation can be different. It can be a vector, taxonomy or graph which retains the general relationships among categories. Our published work (Zhu et al., 2019) uses a three-level category to represent user interest profiles.

Multi-faceted representation approaches have created more attention. They model multiple aspects for a target user based on different information sources, in order to derive a comprehensive view of that user. The assumption here is that the different aspects of multi-faceted representations can more comprehensively express user interests and improve recommendation performance in social marketing practices. At last, these different aspects of user interests can be either combined to construct a single user interest profile or maintained separately for a target user.

3.3. Construction and enhancement of user interest profiles

This process provides details on how user interest profiles with a specific representation can be constructed. The weighting schemes and temporal dynamics of user interests are two main considerations in the construction of user interest profiles. Weighting schemes are mainly to determine the weights of user interest formats denoting the importance of these interests, such as words or concepts in user interest profiles. A common and simple weighting scheme is using some heuristic approaches, such as the *term frequency-inverse document frequency* (TF·IDF) (Jiang and Sha, 2015). In the context of OSNs, specific approaches have to be devised for constructing user interest profiles by aggregating their followees' normalized weights of interest.

The temporal dynamics of user interests is another important consideration when constructing user interest profiles. There are mainly two types of approaches. *Constraint-based approaches* extract user interest profiles based on some specified constraints (e.g., using a temporal constraint based on a target user's posts in a period of time or using an item constraint based on a certain number of posts). In contrast, *interest decay functions* aim at including all interests of a user but permit the old ones to decay. The intuition is that a higher weight should be given to recent interests than older ones. Interest decay functions mainly include: exponential decay (Orlandi et al., 2012), time-sensitive interest decay (Abel et al., 2011) and some other modified interest decay functions (Ahmed et al., 2011). The exponential interest decay function modified with the timed weights of interest nodes is adopted in our previous work (Zhu et al., 2019). To sum up, constraint-based approaches are more suitable for representing a user's relatively stable preferences, while the interest decay functions are more helpful to capture a user's temporary interest

3.4. Evaluation strategies

Evaluation strategies are one of the necessary parts in user interest modeling mainly include: questionnaires, interviews and extrinsic evaluation. First, questionnaires can be used for collecting users' explicit feedback about their interest profiles (Nishioka and Scherp, 2016). This may be the most direct and accurate way of evaluating an interest model. To this end, this approach generally recruits a number of users to get their authorization for building their interest profiles. They explicitly provide the scores on the interest topics in their profiles in the experiments that we conducted (Zhu et al., 2019).

Interview methods also can be used to gather user opinions such as their satisfaction and the accuracy of the inferred user interest profiles. Compared to questionnaires, the interview method is more applicable to evaluate a user interest model qualitatively. However, the two methods require recruiting many volunteers and impose an extra burden on users. Other extrinsic approaches can indirectly evaluate user interest profiles with respect to the performance of some applications where these models are employed, for example, some recommender systems emphasize the extent to which a user actually liked a recommended item (Bellogn and Said, 2017).

4. Future research opportunities and challenges

Finally, there are several opportunities and challenges for future research with respect to the study of user interest modeling oriented to social marketing:

- Some assumptions, for example, different topics may decay at different speeds, and the interest weights of each user should have different weights in different contexts, need to be considered further. Thus, more dynamicity needs to be flexibly incorporated in user interest profiles.
- There are various aspects and views based on different dimensions in the course of user interest modeling. They include the data sources, temporal dynamics of interests, representation forms, more comprehensive user interest, and modeling strategies, which should be investigated. Therefore, we believe that multi-faceted user interest profiles should be given more attention going forward.

- The lack of common benchmarks and datasets always hinders comparison with other methods. Thus, it is important to evaluate inferred user interest profiles in terms of different tasks or settings to understand the advantages and weaknesses of different models. In addition, some standard accuracy and ranking metrics such as precision and recall have been widely employed, while non-accuracy metrics such as coincidence and novelty should be developed more fully for some specific applications.

CRediT authorship contribution statement

Zhiguo Zhu: Conceptualization, Methodology, Writing - original draft. **Liping Kong:** Writing - review & editing. **Xiaoyi Deng:** Investigation. **Bo Tan:** Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors appreciate the editors of ECRA, who provide us a rare chance to give a 2020 perspective on an earlier-published article. This work was supported and funded by the National Natural Science Foundation of China (No. 71672023, 71772033, 71401058, 71602021); the Humanities and Social Sciences Foundation of the Ministry of Education of China (No. 16YJAZH083); and the Innovative Talent Support Plan for Universities in Liaoning Province (No. WR2019011) and Scientific Research Funds of Education Department of Liaoning Province (No. LN2019J34).

References

Abel, F., Gao, Q., Houben, G.J., Tao, K., 2011. Analyzing temporal dynamics in Twitter profiles for personalized recommendations in the social web. In: Proc. 3rd Intl. Web Sci., Conf. ACM Press8, New York, pp. 1.

Ahmed, A., Low, Y., Aly, M., Josifovski, V., Smola, A.J., 2011. Scalable distributed inference of dynamic user interests for behavioral targeting. In: Proc. 17th Intl. Conf. on Knowl. Disc. and Data Mining. ACM Press, New York, pp. 114–122.

Bellogn, A., Said, A., 2017. Rate CTRCT, gain DCGDC, error MAEMA, precision MAPMA, learning MLM, error RRMS: Recommender systems evaluation. Working paper, Information Retrieval Group, Universidad Autonoma de Madrid, Madrid, Spain.

Besel, C., Schlötterer, J., Granitzer, M., 2016. Inferring semantic interest profiles from Twitter followees: Does Twitter know better than your friends? In: Proc. 31st Ann. ACM Symp. on Appl. Comp. ACM Press, New York, pp. 1152–1157.

Bhattacharya, P., Zafar, M.B., Ganguly, N., Ghosh, S., Gummadi, K.P., 2014. Inferring user interests in the Twitter social network. In: Proc. 8th ACM Conf. on Rec. Sys. ACM Press, New York, pp. 357–360.

Faralli, S., Stilo, G., Velardi, P., 2017. Automatic acquisition of a taxonomy of microblogs users' interests. J. Web Sem. 45, 23–40.

Flati, T., Vannella, D., Pasini, T., 2014. Navigli, R.: Two is bigger (and better) than one: theWikipedia bitaxonomy project. In: 52nd Annual Meeting of the Association for Computational Linguistics, ACL, Baltimore,MD, USA, pp. 945–955.

Gong, W., Lim, E.P., Zhu, F., 2015. Characterizing silent users in social media communities. In Proc. 9th Intl. Conf. on the Web and Soc. Media, Menlo Park, CA: AAAI Press. Available at: https://www.aaai.org/ocs/index.php/ICWSM/ ICWSM15/paper/view/10462.

Han, H., Xu, H., Chen, H., 2018. Social commerce: a systematic review and data synthesis. Electron. Comm. Res. Appl. 30, 38–50.

Jiang, B., Sha, Y., 2015. Modeling temporal dynamics of user interests in online social networks. Proc. Comput. Sci. 51, 503–512.

Nguyen, P.T., Tomeo, P., Di Noia, T., Di Sciascio, E, 2015. Content-based recom mendations via DBpedia and Freebase: a case study in the music domain. In: International SemanticWeb Conference, pp. 605–621.

Nishioka, C., Scherp, A., 2016. Profiling vs. time vs. content: what does matter for top-k publication recommendation based on Twitter profiles? In: Proc. 16th ACM/IEEE-CS Joint Conf. on Dig. Lib. ACM Press, New York, pp. 171–180.

Orlandi, F., Breslin, J., Passant, A., 2012. Aggregated, interoperable and multi-domain user profiles for the social web. In: Proc. 8th Intl. Conf. on Sem. ACM Press, Sys., New York, pp. 41–48.

Perera, S., Mendes, P.N., Alex, A., 2016. Sheth, A.P., Thirunarayan, K.: Implicit entity linking in tweets BT—the semantic web. In: Latest Advances and New Domains: 13th International Conference, ESWC 2016, pp. 118–132.

Trikha, A.K., Zarrinkalam, F., Bagheri, E., 2018.Topic-association mining for user interest detection. In: The 40th European Conference on Information Retrieval.

Zhu, Z., Wang, J., Wang, X., Wan, X., 2016. Exploring factors of user's peer-influence behavior in social media on purchase intention: Evidence from QQ. Comput. Hum. Behav. 63, 980–987.

Zhu, Z., Zhou, Y., Deng, X., Wang, X., 2019. A graph-oriented model for hierarchical user interest in precision social marketing. Electron. Comm. Res. Appl. 35, 100845.