Improving Video Recommendation Systems from Implicit Feedback in the E-marketing Environment

Zhang J. H., Chong W., Liu O., and Man K.L.

***Abstract*— Because of the overload of information, it is necessary for online video websites to develop effective recommendation systems to help video users find out the videos of interest efficiently. Furthermore, due to the lack of explicit feedback, implicit feedback will play an important role in the development of video recommendation systems. Based on past research, this paper tries to discover the implicit interest indicators that can indicate video users' interest based on gender. These implicit interest indicators are broadly comprised of cursor movements, scrolling activities and mouse speed.**

***Index Terms*—Implicit feedback, User interest, Recommendation systems**

I. Introduction

As online videos become more popular, many video content online video websites try to provide as many videos as possible to attract video users, so there are often massive number of videos in large online video websites and it is often difficult for the users to find a video of interest, which makes them impatient and degrade their overall experience [3, 12, 16]. In order to solve this problem, the concept of video recommendation systems has been introduced that can help extract user preferences from various interaction information with recommendation algorithms at the core [23]. Then, recommendation lists including videos that are likely to spark video users' interest are generated and sent to video users in order to help them discover videos of interest with greater efficiency, thus potentially increasing user stickiness [12].

Currently, most of recommendation algorithms use explicit feedback, like user ratings, to infer user interest [10, 16]. However, some inherent problems about explicit feedback degrade the performance of recommendation systems. One of the key problems is that the provision of explicit feedback requires online users to alter their normal ways of reading or viewing because of the interruptions [10], so it is likely that they are not willing to offer a rating after watching a video when they do not fully understand the benefit of the submission of the feedback [4, 10]. Besides, it may be cognitively burdensome for online users to provide such feedback because they are just knowledge-building users who may not know relevant requirements [4]. Thus, enough explicit ratings for a video may not be collected and this influences the performance of recommendation algorithms such as collaborative filtering (CF) recommendation algorithm [10].

In order to increase the accuracy of personalized recommendations, implicit feedback has been considered a promising alternative to explicit feedback [1]. Implicit feedback is comprised of various implicit interest indicators, like scrolling activities and cursor movements, and it is also obtained from user interaction with websites [10]. In contrast with explicit feedback, this feedback has some advantages. A huge amount of implicit feedback can be generated naturally when browsing, so user experience cannot be affected and relevant costs are removed accordingly such as the cost of rating items [10]. Secondly,

Manuscript received January 08, 2017; revised February 16, 2017.

J.H. Zhang is with the International Business School Suzhou, Xi'an Jiaotong-Liverpool University, Suzhou, China. (e-mail: Junhui.Zhang15@student.xjtlu.edu.cn).

W. Chong is with the International Business School Suzhou, Xi'an Jiaotong-Liverpool University, Suzhou, China. (phone: +86(0)512 8816 1717; fax: +86 (0)512 8816 1730 e-mail: Woonkian.chong@xjtlu.edu.cn).

K.L. Man is with the Department of Computer Science & Software Engineering, Xi'an Jiaotong-Liverpool University, China. (e-mail: ka.man@xjtlu.edu.cn).

O. Liu is with the International Business School Suzhou, Xi'an Jiaotong-Liverpool University, Suzhou, China. (e-mail: Owen.Liu@xjtlu.edu.cn).

ISBN: 978-988-14047-7-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

the dynamic information needs of online users can help offer more accurate personalized recommendations and the needs can be captured by gaining implicit feedback in time [14, 19, 24]. Finally, implicit feedback can be as reliable as explicit feedback, and a more accurate prediction can be obtained by combining the two types of feedback [10]. Thus, an accurate prediction of user interest can be benefited from the use of implicit feedback, especially with the lack of explicit feedback.

In addition, [11, 20] show that the impacts of demographics on users' online behavior should receive much attention like gender, as different types of online users may react differently to marketing stimuli, so targeting strategies that can affect the user effectively can be different [18, 27]. Besides, browsing tendencies of online users can be predicted by demographic characteristics [20, 21], so the characteristics can contribute to the heterogeneity in online behavior that can be reflected by implicit feedback.

However, although there are many studies [15, 20] exploring the effectiveness of implicit feedback in terms of predicting user interest, most of these studies focus on other industries such as news, search engines and online shopping, rather than online video, not to mention the studies exploring the effectiveness in online video according to gender. Furthermore, some of implicit interest indicators have still remained controversy in terms of the inference of user interest.

In order to address the two gaps, this study builds on prior research on implicit feedback and on gender differences to make the exploration of the significance of implicit feedback for video recommendation systems. This study aims to find out the implicit interest indicators that can reflect video users’ interest based on gender.

The remainder of this paper is organized in the following way. Section 2 introduces relevant past literature about implicit feedback in different industries and about gender differences followed by detailed description of the research methods in Section 3. Finally, Section 4 shows the academic and practical implications and indicates directions for future work.

II. LITERATURE REVIEW

Table I: Implicit Interest Indicators

Cursor Movements	Maxy: maximal y coordinate
	Miny: minimal y coordinate
	Curcnt: amount of cursor movements
Mouse Speed	Curspeed: speed of cursor movements
Scrolling Activities	Scrofre: frequency of scrolling

A. *Eye Movements and Cursor Movements*

Cursor movements include implicit interest indicators related to the cursor, like total distance the cursor travels [15], while eye movements are comprised of the fixations and the movements themselves [22]. [7] shows that the two types of movements correlate strongly with each other, so if cursor movements can indicate user interest, eye movements can also infer the interest, and vice versa. [5] shows that eye movements can help identify the parts of a document which are read, skimmed or skipped. More specifically, online users are likely to be interested in the read or skimmed parts of a document, while they seem to dislike the skipped parts [5]. However, by studying emotional states, [17] drew a different conclusion that it is possible that eye movements do not correlate with user interest because viewing preferences may be spontaneous rather than selective. Thus, we expect that video users’ interest may be inferred by the eye movements or cursor movements.

B. *Mouse Speed*

Mouse speed records how fast online users move their cursors and some scholars studied its correlation with user interest. [14] explores searcher interest by capturing mouse acceleration and speed and this study contributes to the provision of an accurate result list. In [15], it is found that mouse speed correlates negatively with searcher interest. Thus, we reason that this indicator correlates with user interest in online video.

C. *Scrolling Activities*

Scrolling activities are associated to wheel operations, such as frequency of scrolling behavior. In an experiment, [15] observed that there is a significantly negative correlation between scrolling frequency and user interest. However, the same conclusion cannot be drawn by [13] which could not find a significant correlation between them, so we expect that a significant correlation between scrolling activities and video users' interest may not exist.

D. *Consumer Heterogeneity - Gender*

Prior studies have shown that demographics should be taken into account when analyzing users' online behavior, as they are predictive of the browsing tendencies [20, 21]. For example, [21] found that higher-income male users with children are more likely to use home pages when browsing. Besides, male users may spend less time browsing per page while female users are inclined to spend more time browsing per page [8]. Furthermore, during their browsing, the two groups tend to display different navigation patterns that consist of different types of implicit interest indicators [8]. Thus, we expect that male and female video users' interest can be indicated by different implicit interest indicators, so the two groups will be explored separately.

According to above analysis, in relation to the online video industry, two gaps are identified in the current academia.

(1)As most of the relevant research focus on other industries like search engine and online shopping, rather than video websites, the correlation between implicit feedback and video users' interest needs to be further studied.

(2)Except mouse speed, there is still much disagreement about the effectiveness of other indicators in predicting user interest, so the exploration of the effectiveness of other indicators in online video is needed.

Consequently, this paper aims to fill the two gaps by identifying the implicit interest indicators that can indicate online video users' interest according to gender.

III. METHODOLODY

This research aims to discover the implicit interest indicators which infer user interest in online video according to gender and the indicators are the embodiment of online users' browsing behavior, so an experiment is designed to extract and collect the implicit data including cursor movements, scrolling activities and mouse speed. Then, as there is large amount of implicit data that needed to be gathered and analyzed, the quantitative method will be adopted.

44 college students (26 males and 18 females) are asked to browse an online video website, Youku, to search for and watch videos for some time until they can determine their attitudes towards the videos and the attitudes are then delivered verbally to researchers, just like or dislike it. Throughout their viewings, about 200 sets of implicit data can be generated and collected by free software, IOGraph, MouseMonitor and Quick Macro. In addition, the students provide their explicit ratings on a two-point scale, because [6] found that it is likely that people are able to offer more reliable and consistent judgments on a two-point scale rather than a five-point scale.

According to a survey, online video users aged from 20 to 29 make up the highest percentage of all the video users, at about 37% [9], so experimental subjects surveyed in this experiment are among this age group and they are all from Shaoguan University. As for the experimental platform, Youku, the largest online video website in China [25], has been developing rapidly and steadily in recent years. In 2011, its monthly users increased to above 300 million mark and it was granted in 2014 'the most influential online video service' [26]. In summary, the sample is typical and representative of the total population. Also, it is probable that the experimental subjects are already Youku's users and they are familiar with this website, so the reliability and quality of the implicit data collected can be high.

However, there are still limitations associated with data collection and these can potentially affect the analysis results. In the experiment, only free software is used to collect the implicit data and all the data is collected only from the client side, thus reducing the accuracy of the data to some extent. Additionally, as the data is collected from Chinese online users in a Chinese video website, it is likely that the findings may not be applicable for video users from other countries.

IV. CONCLUSION AND FUTURE RESEARCH

This research concentrates on the exploration of the effectiveness of implicit feedback in predicting video users' interest according to gender. Thus, the implicit interest indicators that can indicate male and female video users' interest can be identified respectively. During the literature

ISBN: 978-988-14047-7-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

review, some past research associated with implicit feedback and gender differences have been reviewed and some hypotheses based on the research and limitations have been presented. Not only does this study can provide a better understanding of the correlation between video users' interest in online video and implicit feedback in the academic word, but also it opens a new way of how to improve their video recommendation systems to increase user stickiness for video content providers.

In addition to identifying the correlation between implicit feedback and user interest in online video, how these findings have an impact on practical applications a worthy topic to explore. For example, based on the essential features of video recommendations, how to improve video recommendation systems in order to provide more accurate recommendation lists by using the significant implicit interest indicators is worthy of consideration. Additionally, as the implicit interest indicators identified in this study may not be applicable for other types of recommendation systems, future studies on other different types of recommendation systems involved with online video users from other countries besides China are needed.

REFERENCES

[1]. Adomavicius, G & Tuzhilin, A. (2005) 'Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions', *IEEE Transactions on Knowledge and Data Engineering*, 17(6), pp.734-749.

[2]. Arcand, M., Nantel, J. & Sénécal, S. (2011) 'The effects of sex and Internet usage on search efficiency and effectiveness', *International Journal of Computer Information Systems and Industrial Management Applications*, 3, pp.663-670.

[3]. Banister, E. N. (2003) 'Using the Internet for Customer Behaviour Research: A Guide to the Challenges and Opportunities of the Internet as a Survey Medium', *Journal of Customer Behaviour*, 2(1), pp.105-123.

[4]. Beaulieu, M. & Jones, S. (1998) 'Interactive searching and interface issues in the okapi best match retrieval system', *Interacting with Computers*, 10(3), pp.237–248.

[5]. Buscher, G., Dengel, A. & Van, E. L. (2008) 'Query expansion using gaze-based feedback on the subdocument level', *International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp.387-394.

[6]. Carterette, B., Bennett, P. N., Chickering, D. M. & Dumais, S. T. (2008) 'Here or there: Preference judgments for relevance', *Proceedings of European Conference on Information Retrieval (ECIR)*, pp.16-27.

[7]. Chen, M. C., Anderson, J. R. & Sohn, M. H. (2001) 'What can a mouse cursor tell us more?: correlation of eye/mouse movements on web browsing', *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, ACM, pp.281-282.

[8]. Chen, S. Y. & Macredie, R. (2010) 'Web-based interaction: a review of three important human factors', *International Journal of Information Management*, 30(5), pp.379-387.

[9]. China Netcasting Services Association. (2015) *Chinese internet audio-visual research report of 2015*, Beijing: China Netcasting Services Association.

[10]. Claypool, M. et al. (2001) 'Implicit Interest Indicators', *Proceedings of the 6th international conference on Intelligent user interfaces,* pp.33-40.

[11]. Danaher, P. J., Mullarkey, G. W. & Essegaier, S. (2006) 'Factors Affecting Web Site Visit Duration: A Cross-Domain Analysis', *Journal of Marketing Research*, 43(2), pp.182-194.

[12]. De Bruyn, A., Liechty, J., Huizingh, E. & Lilien, G. (2008) 'Offering Online Recommendations with Minimum Customer Input Through Conjoint-Based Decision Aids', *Marketing Science*, 27(3), pp.443-460.

[13]. Goecks, J. & Shavlik, J. (2000) 'Learning users' interests by unobtrusively observing their normal behavior', *Proceedings of the 2000 International Conference on Intelligent User Interfaces*, New Orleans, LA., pp.129-132.

[14]. Guo, Q. & Agichtein, E. (2010) 'Ready to Buy or Just Browsing? Detecting Web Searcher Goals from Interaction Data', *Proceeding of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, Geneva, Switzerland, pp.130-137.

[15]. Guo, Q. & Agichtein, E. (2012) 'Beyond dwell time: estimating document relevance from cursor movements and other post-click searcher behavior', *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, Atlanta, Georgia, USA, pp.3601-3606.

[16]. Jawaheer, G., Weller, P. & Kostkova, P. (2014) 'Modeling User Preferences in Recommender Systems: A Classification Framework for Explicit and Implicit User Feedback', *Acm Transactions on Interactive Intelligent Systems*, 4(2), pp.1-26.

[17]. Kahneman, D. (1973) *Attention and Effort*. Prentice-Hall Inc, Englewood Cliffs, New Jersey.

[18]. Li, H. & Kannan, P. K. (2014) 'Attributing Conversions in a Multichannel Online Marketing Environment: An Empirical Model and a Field Experiment', *Journal of Marketing Research*, 51(1), pp.40-56.

[19]. Lu, M. et al. (2014) ‘Scalable news recommendation using multi-dimensional similarity and Jaccard-Kmeans clustering’, *Journal of Systems & Software*, 95(9), pp.242-251.

[20]. Mallapragada, G., Chandukala, S. & Qing, L. (2016) 'Exploring the Effects of "What" (Product) and "Where" (Website) Characteristics on Online Shopping Behavior', *Journal of Marketing*, 80(2), pp.21-38.

[21]. Montgomery, A. L. et al. (2004) ‘Modeling Online Browsing and Path Analysis Using Clickstream Data’, *Marketing Science*, 23(4), pp.579-595.

[22]. Rayner, K. (2009) ‘Eye movements and attention in reading, scene perception, and visual search’, *Quarterly Journal of Experimental Psychology*, 62(8), pp.1457-1506.

[23]. Sun, P. et al. (2013) ‘An Overview of Collaborative Filtering Recommendation Algorithm’, *Advanced Materials Research*, pp.756-759.

[24]. White, R. W., Jose, J. M. & Ruthven, I. (2006) ‘An implicit feedback approach for interactive information retrieval’, *Information Processing & Management*, 42(1), pp.166-190.

[25]. Yamada, K. I. (2014) ‘Booming Internet Video Sites in China : Current Situation and Challenges’, *Nhk Monthly Report on Broadcast Research*, 64, pp.80-92.

[26]. Youku. *Youku millstone* [Online]. Available from: http://c.youku.com/aboutcn/milestone?spm=a2hc9.20008155.m_69776.5~1~3!2~A (Accessed: 29 December 2016)

[27]. Young, Martin. R., DeSarbo, W. S. & Morwitz, V. G. (1998) ‘The Stochastic Modeling of Purchase Intentions and Behavior’, *Management Science*, 44(2), pp.188-202.

ISBN: 978-988-14047-7-0
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)