iNewsBox: Modeling and Exploiting Implicit Feedback for Building Personalized News Radio

Yanan Xie
Zhejiang University
Hangzhou, China
xyn@zju.edu.cn

Liang Chen
Zhejiang University
Hangzhou, China
cliang@zju.edu.cn

Kunyang Jia
Zhejiang University
Hangzhou, China
jiaky777@zju.edu.cn

Lichuan Ji
Zhejiang University
Hangzhou, China
jilichuan@zju.edu.cn

Jian Wu
Zhejiang University
Hangzhou, China
wujian2000@zju.edu.cn

ABSTRACT
Online news reading has become the major method to know about the world as web provide more information than other media like TV and radio. However, traditional online news reading interface is inconvenient for many types of people, especially for those who are disabled or taking a bus. This paper presents a mobile application iNewsBox enabling users to listen to news collected from the Internet. In order to simplify necessary interactions of getting valuable news, we also propose a framework for using implicit feedback to recommend news in this paper. Experiment shows our algorithms in iNewsBox are effective.

Categories and Subject Descriptors
H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval—clustering, relevance feedback; H.4 [Information Systems Applications]: Miscellaneous; H.5.2 [Information Interfaces and Presentation]: User Interfaces

General Terms
Algorithms, Design, Human Factors

Keywords
News recommendation, implicit feedback, news clustering, mobile application

1. INTRODUCTION
With the increasing number of media that provide their contents through the Internet, more Americans have been getting their news from the Web than from newspapers since

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).

ACM 978-1-4503-2263-8/13/10.
http://dx.doi.org/10.1145/2505515.2508199.
Section 3 introduce the way we cluster news. The detail of the news recommender is showed in Section 4. After an introduction of iNewsBox user interface in Section 5, our experiment is demonstrated in Section 6. As the last part of this paper, our discussion is given in Section 7.

2. RELATED WORK

Recently, various clustering algorithms, topic models and recommender algorithms are used to personalize news reading. Abhinandan S. Das et al. propose to utilize MinHash clustering, PLSI and covisitation counts to build their news recommender in Google News [5]. Jiahui Liu et al. present their personalized news recommender system based on click behavior [4]. Other attempts to advance the personalization of news reading stay at improving screen reading but not radio styled reading like iNewsBox.

3. NEWS CLUSTERING

In this part, we describe the methods we use for news clustering. As discussed before, news has to be clustered into events to avoid recommending similar news on same event. As a matter of fact, news focusing on one same event will have little differences even they are from various source. So we transform the original form of news to a bag-of-words representation to extract the intrinsic key words of it. This is introduced in TF-IDF classifier method. Since news comes in a stream way, clustering algorithm such as K-means will not fit this scenario. We adopt a two-stage clustering method to cluster news into events. A similar method is first introduced in [5].

4. MODELING RECOMMENDER

After the collected articles are clustered into events, the problem of recommending news turns into recommending events for users. Once our recommender decided to recommend an event, the latest article of the event would be played to the user. No more than one article of the same event will be broadcasted unless the event has been updated. In the remainder of this section, we will focus on how we employ implicit feedback to recommend event.

4.1 Implicit feedback

In order to learn user preferences about news, following 4 types of feedback are collected to model our recommender:

- **Complete.** Complete is the commonest implicit feedback in iNewsBox though there is no corresponding buttons in iNewsBox. A Complete response is generated once item is played to the end for one user, which means that the user would like to hear, at least tolerate us to continue recommend this kind of items. We regard Complete as positive feedback to an item.

- **Like.** Users are enabled to tell us they want to hear more items like current one by touch Like button in iNewsBox and send us Like feedback. Like is also regarded as a kind of positive feedback but with higher effect than Complete’s.

- **Dislike.** As the opposite response to Like, a button in iNewsBox interface is connected with Dislike response which can be used to express their strong detestation of some items. Dislike is negative feedback. If a user touches the Dislike button, a Skip response defined as follow is also sent to us.

- **Skip.** Unlike real radio station, iNewsBox allows users to skip current item into next one by touch Skip button. This kind of user behaviors are treated as negative feedback since in most cases Skip simply implies that user is not interested in current item.

4.2 Similarity-based personalization

LDA is a powerful topic model learning topics from some documents.[1] We perform LDA on all events in our database to get the relevance of every event to generated topics. Both the results of LDA and the numbers of co-responses are employed to measure the similarities between each pair of events.

The main reason we employ two different approaches to deal with news clustering and event similarity measurement is that TF-IDF is sensitive for event word feature like specific persons and countries while LDA is more likely to find out the topic similarity rather than event similarity.

We propose to utilize MinHash’s. For the later-collected events, we estimate their vectors by employing TF-IDF similarities between them and the events that have been processed by LDA.

In order to model users’ behaviors, we use four 0-1 vectors to represent 4 responses given in Section 4.1 as follow:

\[
\begin{bmatrix}
R_{\text{Complete}} & R_{\text{Like}} & R_{\text{Dislike}} & R_{\text{Skip}}
\end{bmatrix} = 
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

Based on the assumption that similar items get similar responses for specific user. A prediction of user u’s response to item i is defined as follow:

\[
R'_{\text{u},i} = \frac{\sum_{j \in D} R_{\text{u},j} \times \text{Sim}(d_i,d_j)}{\sum_{j \in D} \text{Sim}(d_i,d_j)}
\]

while \( R_{\text{u},j} \in \{R_{\text{Complete}},R_{\text{Like}},R_{\text{Dislike}},R_{\text{Skip}}\} \) is the response given by user u to item j, D is the items set containing all items that have been recommended to user u before and Sim is the similarity between item i and j defined as above.

In the last of this part, a scoring function considering the user’s implicit feedback as well as items’ content is given as follow:

\[
\text{Score}(u,i) = R'_{\text{u},i}^T \begin{bmatrix}
1 & \alpha & -\beta & -1
\end{bmatrix}^T
\]

while \( \alpha, \beta \geq 0 \) are coefficients determined by experiments demonstrating the relative effective of Like and Dislike responses.

However, our recommender doesn’t always broadcast the news with highest score but also recommend the latest one by a small probability \( \epsilon \) to enhance recommendation recall. It is similar to \( \epsilon \)-greedy algorithm in reinforcement learning. We take \( \epsilon = 0.1 \) in iNewsBox.

5. USER INTERFACE

iNewsBox is available on iOS devices now. Users can download and install iNewsBox of latest version by searching iNewsBox in the App Store. Figure 1 shows two screen captures of iNewsBox while the user is listening to our news radio. From the left screen capture of Figure 2, we can find
that there are two channels, named Private Radio and Real-time News, that can be chosen on the top. Users can listen to the personalized news radio in Private Radio while in Real-time News, users are always getting the latest news. A picture of the current news is showed below so that users can pause or play by touching the picture. The present time of current track and its title are displayed too. On the bottom of this screen capture, 3 main buttons are given, enabling users to label current news as Like, label current news as Dislike and skip respectively.
iNewsBox can also run in the background of iOS device. During the usage of iNewsBox, the device should be locked in most cases. We employ lock screen to make it convenient for users to give their feedback in those circumstances. The right screen capture of Figure 1 shows the lock screen while iNewsBox is working. We can see the title and the picture of the current playing news, followed by Last, Play/Pause and Next buttons. Users can also use the remotes on their headset wires to control iNewsBox.

6. EXPERIMENT

To demonstrate the effectiveness of our algorithm, we ask 20 volunteers to try the two channels of iNewsBox for 20 minutes. The reason we fix the listening time is in order to eliminate the effects of volunteers’ impatience and get more data from such a small number of volunteers. These 20 people are divided into 2 groups without being notified. Volunteers in the first group of 10 get the latest news without any personalization (i.e., Realtime News channel, RN) in the first 10 minutes, then listen to the personalized radio using our proposed recommender (i.e., Private Radio channel, PR) in the last 10 minutes. 10 volunteers in the second group firstly listen to Private Radio for 10 minutes, and then listen to Realtime News for another 10 minutes. No volunteers are informed about the purpose of the experiment.

As shown in Table 1, our proposed recommender lengthen the average listening time per article (#avg. listening seconds) and successfully reduce the the number of touches during the listening time. Even though many of volunteers indicate that they have felt no obvious differences between our two channels during those 20-minute-long tryouts, our algorithm did help them avoid many operations. We believe people who are disabled or having other troubles with standard interface would benefit from the features of our iNewsBox.

7. DISCUSSION

In this paper, we propose a mobile application that allows users to listen to personalized news radio synthesized by TTS. Our system is effective and extensible. A story can be broadcasted to users by our system within 2 minutes after the media published the story through RSS. The recommendations of iNewsBox are effective in our tests. In our future work, we will design recommender algorithms for this novel scenario and do experiments to compare those algorithms in our iNewsBox with massive datasets.

8. ACKNOWLEDGMENTS

This research was partially supported by the National Technology Support Program under grant of 2011BAH15B05, the National Natural Science Foundation of China under grant of 61173176 and 71171077, Science and Technology Program of Zhejiang Province under grant of 2008C03007, National High-Tech Research and Development Plan of China under Grant No.2009AA110302, National Key Science and Technology Research Program of China (2009ZX01043-003-003).

9. REFERENCES