Personalized Mobile Information Recommendation Based on Fine-Grained User Behaviors

YILEI WANG [a], XUEQIN CHEN [a,1,2]

[a] *College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350108,China*

Abstract. At present, the enthusiasm of users to score actively in mobile information recommendation system is generally poor. Moreover, the existing research works rarely start with the analysis of fine-grained reading behaviors of mobile terminal users, but mostly based on the analysis of reading content and the improvement of model. It is difficult to find out the objective, short-term and local behavioral preferences of users. To solve the above problems, we propose six kinds of explicit fine-grained reading behaviors and integrate them into the user reading interest model to form the SVR-ALL model. The effectiveness of these six fine-grained behaviors is verified by ablative experiments. On the basis of SVR-ALL model, four implicit fine-grained reading behaviors are further mined by considering the difference of user reading habits, and then propose the user reading preference model called F-AFC. The updating mechanism for user preference designed in F-AFC can fully reflect the changes of users' reading habits in different periods. Experiments show that the accuracy of the user interest model considering user's reading preference and its update can be improved to some extent.

Keywords. fine-grained reading behavior, user reading interest model, user reading behavior preference model

1. Introduction

At present, the service mode of the information push application is centered on the mobile Internet. And the intelligent mobile terminal device is the main carrier of information in the mobile Internet environment[1]. However, the traditional user personalized information acquisition method in the Internet era cannot be directly transplanted to the mobile terminal[2,3].

Meng Xiangwu et al. [4] proposed that the disadvantages of mobile devices lie in poor input ability and small screen size, and the explicit scoring method adopted by users in the process of reading has a certain impact on user experience. Therefore, the implicit

scoring method should be adopted in the acquisition, that is, there is no need for users to make additional operations for scoring. Xu Hailing et al. [5] proposed that the acquisition of users' reading behavior can be obtained from the frequency of users' reading pages, explicit score and the time stayed on the page. By introducing more detailed object characteristics and more precise user preferences, the accuracy of predicting user ratings would also be higher. Ming Z et al. [6] analyzed the diversified purchasing behaviors of users before online shopping and confirmed the critical shopping behaviors that determine the precursors of purchases by adopting the 5-fold cross verification method. Meng W et al. [7] proposed a novel session-based recommendation(SR) model MKM-SR, which incorporates user micro-behaviors and item knowledge into multi-task learning for SR, thus improving the recommendation accuracy. The above researches fully demonstrate that the discussion of more detailed and precise behaviors of users can effectively improve the accuracy of recommendation. "Fine-grained" is often used to represent further subdivision of an object[8,9,10], we use the term to refer to more granular user behaviors.

Due to the rapid development of machine learning, most scholars tend to focus on the modeling of complex models and the mining of users' implicit behaviors. Zhou C et al. [11] proposed an attention based user behavior modeling framework called ATRank, which is dedicated to modeling heterogeneous user behaviors with Multi-headed attention. Feng Y et al. [12] proposed a new framework named Adaptive Target-Behavior Relational Graph network(ATBRG) to effectively capture structural relations of target user-item pairs over knowledge graph. Minh-Duc Nguyen et al. [13] proposed a mixture model based on variational autoencoder, which is dedicated to mining users' hidden interests and discovering users' invisible behaviors. Although these studies have indeed improved the performance and accuracy of recommendation, they neglected to explore the fine-grained behavior of users.

Claypool M et al. [14] proposed that when acquiring user interest can according to the length of time that user browsing a certain page and the time spent on draging the page with mouse. [15] showed that the user's accurate points of interest can be obtained by getting the user's average reading speed and time. But none of the above documents showed how to quantify users' reading interest based on various types of behavior data. Teng Y et al. [16] caculated users' interest based on the statistical relationship between the user's average active rating and operation behavior. This method has a better effect than others, but its accuracy is still poor. Claypool M et al. [14] took into account the influences of different reading behaviors of users, and proposed a model calculating users' interest based on their behaviors. But it is only applicable to the calculation and analysis of users' web browsing behaviors on the desktop computer.

The operation mode of the mobile terminal is completely different from that of the computer. There are also big differences in the format of the information and how the page is displayed. As a result, the users' reading behaviors on the mobile terminal and on the computer terminal are quite different[17]. Therefore, the existed analysis of the users' reading behaviors cannot be directly used in the mobile-side recommendation system.

We aim to study a personalized mobile information recommendation model based on fine-grained user behaviors. The main contributions of this paper are summarized as follow:

In order to explore the objective short-term and local behavioral preferences of users, we propose six kinds of explicit fine-grained reading behaviors and integrate them into

the user reading interest model to form the SVR-ALL model. The effectiveness of the proposed six explicit fine-grained user reading behaviors is proved by ablation experiments. Based on the users' reading interest model and considering the different reading habits of different users, four implicit fine-grained user reading behaviors are further explored. A user reading behavior preference model named F-AFC is established. We also design a reasonable updating mechanism for the user's reading preference model, and further improves the accuracy of the model through continuous iteration in the updating process.

The remainder of this paper is organized as follows. In Section 2 the related work is presented. Section 3 illustrates the proposed user reading interest model and user reading behavior preference model. The experimental studies are presented in Section 4. Finally, Section 5 concludes the paper.

2. Related Works

2.1. Support Vector Regression Model(SVR)

The basic idea of regression is to determine a function that accurately approximates the target values using the input value[18]. Support Vector Regression(SVR) model is obtained by introducing loss function in Support Vector Machine(SVM) and expanding it[19]. There are two major training strategies for SVR. One is ε-SVR, which employs an ε-insensitive loss function to solve the quadratic optimization problem. Another training strategy is called v-SVR[20], and the parameter v can effectively control the number of support vectors. In this paper, we adopt the first strategy ε-SVR for the subsequent related work.

2.2. Volatility of Reading Behavior

In order to avoid the burden of frequent updating of the model, we propose the Volatility of Reading Behavior(VRB). VRB refers to the change of a certain behavior b_j of user u_i in a certain period of time compared with its historical state. The formula is as follows:

$$VRB_{i,j,T} = \frac{1}{n_{i,j,T}} \sum_{t \in T} \frac{\left| X\left(u_i, b_j, t\right) - \overline{X\left(u_i, b_j\right)} \right|}{X\left(u_i, b_j\right)}, 1 \leq j \leq m \tag{1}$$

Where T is the update time window of model; $VRB_{i,j,T}$ represents the volatility of behavior b_j of user u_i in time period T; $n_{i,j,T}$ represents the actual times that behavior b_j of user u_i happened in time period T; $\overline{X\left(u_i, b_j\right)}$ is the average b_j behavior of user u_i calculated based on historical data, and the statistical cut-off is the last global update; $X\left(u_i, b_j, t\right)$ is the behavior value of b_j uploaded by user u_i at time t.

2.3. Coulomb's Law

In the early days, Marin L et al. [21] innovatively introduced a physical concept - Coulomb's law in user preference parameter learning. Recently, Haixin Z et al. [22] proposed a new method to calculate the fractal dimension of complex networks. Lai Y et

al. [23] proposed the concept of coulomb's law of urban traffic to model the relationship between urban taxis and passengers, and on this basis proposed a path recommendation scheme for taxis and passengers as positive and negative charges. Inspired by the above research, we introduce Coulomb's law into the F-AFC model for weight updating, and the details will be explained in the section 3.2.2.
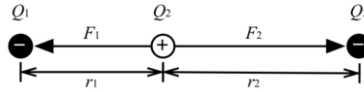

Figure 1. Coulomb's law diagram

Figure 1 shows the distribution of charges at three fixed points on a straight line in a vacuum environment. The charges are represented by Q_1, Q_2, Q_3, and the corresponding electric quantities of them are represented by q_1, q_2, q_3. The following formula can be used when calculating the resultant force of charges:

$$F_e = F_1 - F_2 = k_e \frac{q_1 q_2}{r_1^2} - k_e \frac{q_1 q_2}{r_2^2} \tag{2}$$

k_e in Eq.(2) represents the Coulomb constant[24].

3. The Proposed Method

3.1. User Reading Interest Model

Taking the user reading on the mobile terminal as the research background, the reading behaviors is generated under the guidance of user interest preferences[25]. Therefore, the users' reading behaviors can make the most direct reflection on the their information needs and reading preferences, and also reflect the users' reading interest to a large extent.

3.1.1. Mobile Fine-grained User Reading Behavior Data Acquisition

Compared with the computer, the amount of user behavior generated by mobile devices is much larger. Fine-grained behavior means that more reasonable and scientific behavior objects can be generated by subdividing user behavior objects. For example, the scrolling behavior that superficially displays the users' browsing information can form a series of objects after the scrolling behavior is subdivided. Including the scrolling displacement and the scrolling duration. This behavior can directly express the user's reading preferences.

After analysis, we mainly collected and pre-processed fifteen kinds of fine-grained user behaviors of mobile terminals. Some of them are shown in Table 1.

Table 1. Fine-grained user behaviors collected

Symbol	Meaning
readCmp	Page read completeness
dragCnt	Number of effective page drags
meanV(msg)	User's average reading speed on a page
meanV(u)	User's average reading speed
scrollPt	The point in time when the page first scrolled
loadedPt	The point in time at which the page is loaded
slContentOffset	The displacement of a page generated by one slide

(1) User reading time

User reading time is a key content to measure user interest and it can't be equal to the time user staying on the page. Therefore, we consider the following factors when calculating this item:

1) Starting point for reading time

In addition to the text content, the page of mobile terminals often has objects such as pictures or emoticons. Due to the differences in the hardware and software configuration of each mobile device, there will inevitably be differences in the loading speed of each page. And most users have already started reading the content of the page without loading all the content. So, accumulating immediately after the user enters the page is improper.

Generally, comparing the height of the information content with the height of the reading area of the mobile terminal screen. When the former is higher than the latter (*contentHt*>*readZoneHt*), the user-generated page scrolling behavior indicates that the user has read part of the content and continues reading. So the starting point of reading time set to be the time that first scrolling behavior happened minus the time spent scrolling the page. When *readZoneHt*>*contentHt*, we use the first two formulas in Eq.(3) to compute the starting point of reading time:

$$
startPt = \begin{cases} 0.5, & contentHt \leq readZoneHt \text{ and } u \text{ is a new user} \\ meanPt(u), & contentHt \leq readZoneHt \text{ and } u \text{ is not a new user} \\ scrollPt - Time(scroll), & contentHt > readZoneHt \end{cases} \quad (3)
$$

Where Time(scroll) represents the time required for page scrolling and *meanPt(u)* represents the average value of the starting point of reading time of user u on each page.

2) Reading time pause

The user may has a sudden state during actual reading and the reading page is shelved for a long period of time. When the system detects that the user has not operated the terminal screen for more than a certain period of time, the user is deemed to have suspended reading, and the length of time before the user resumes reading should not be counted as the user's reading time. The timeout threshold is set to the average time the user *u* needs to read text at the height of *readzoneHt*:

$$
threshold(msg) = readzoneHt(msg)/meanV(u) \quad (4)
$$

Where *readzoneHt* represents the height of the reading area of the mobile terminal screen, $meanV(u)$ represents the average reading speed of the user.

(2) Page reading completeness

Whether a user has read a certain page completely can reflect the user's level of interest. This parameter can be obtained simply by determining whether the page is scrolled to the bottom. In the case that the $contentHt(P)$ of the page height is less than the height of the readable area readHt that can be displayed on the screen, if the user's reading duration $t(P)$ reaches the operation timeout threshold $ths(P)$, the page reading completeness set to be "1". However, in the vast majority of mobile information push systems, the tail of most pages contains extended reading and advertisements. In this case, the page may not be scrolled to the end. Therefore, so long as the page is scrolled to reaches a certain page ratio, the user is considered to have read it completely.

(3) Effective number of page drags

Due to the large screen size of traditional desktop computers, users only need to swipe a few times to read the entire content. Therefore, in the process of calculating users' Web page reading time, the impact of sliding/dragging pages is generally ignored. At present, most pages of the mobile terminal are in the form of "scroll", which is limited by the screen size. The user needs to drag the finger on the page multiple times when reading on the mobile device.

Users' behaviors of scrolling the page includes dragging and sliding. Dragging means the pages will stop rolling after the finger leave screen. Sliding means pages will keep rolling for one to two seconds after the finger leave screen. The former action reflects the user's slow reading speed or high concentration, while the latter reflects the user is browsing the page hastily.

We record the drag behavior generated by the page, including the start time of the behavior, the displacement *drContentOffset* generated by the page during the drag, and the end time *drEndTime* of the behavior. The meaningful drag behavior of the page should be close to the average reading speed of the user, so the actual drag time is obtained by using the ratio between the average reading time of the user and the displacement of the content information. So, if the time interval between the two dragging actions before and after the dragging threshold is lower than the dragging threshold, they can be combined into one "effective dragging operation":

$$scrollThs_i = drContentOffset_{2t+1}/meanV(u), \quad i = 1,\ldots,n \qquad (5)$$

Where $scrollThs_i$ means that the i-th group continuous page dragging behavior produces a corresponding drag threshold. Using Eq.(5), the user's drag time series is corrected to obtain the number of effective drag operations during the user's reading process.

(4) Effective number of page slides

Taking the iOS operating system as an example. When acquiring the user's sliding page behavior, determine whether the page is a decelerated sliding based on the parameter *decelerate* generated by the user's finger touching and leaving the screen. We collect and use the following data: sliding page start time *slStarTime*, information page displacement distance *slContentOffset* and sliding page end

time *slEndTime*. The calculation of the number of slides is similar to the drag operation, and the "effective operation" must also be judged.

(5) Number of user reviews

The retrospective behavior means that the user adopts the reverse continuous mode when quickly dragging the page, so that the page can be scrolled to a certain position that has been read in the shortest time. The number of reviews can intuitively indicate the user's preference for information. We find the following rules by analyzing the drag behavior and sliding behavior. During the user's search for the location, multiple drag behavior and sliding behavior will be generated. The above mentioned continuous behavior does not have any meaning and is regarded as an "effective operation". However, the information that the user sees during the review process is the information that has been browsed, so all the content of the information before reaching the target position is not meaningful. So after analyzing a large number of test data, we set 0.45s as the threshold of review time.

(6) Average reading speed

We divide the average reading speed into the average speed of reading page and average reading speed of user.

1) Average speed of reading page

The average speed of reading page *meanV(msg)* represents the average speed of a user reading a page. The same user will have different average speeds on different pages. In Eq.(6), when the height of the reading area is greater than the height of the information, the average speed of the reading page is the ratio of the height of the information to the total length of the time the user use in reading the information; while when the height of the reading area is less than the height of the information, the user must have a drag or slide operation. The average speed of the "effective operation" generated by the user's continuous forward reading is summed, and the average value is further calculated as the average speed of reading the page.

$$meanV(msg) = \begin{cases} \frac{contentHt(msg))}{readTime(msg))}, contentHt \leq readZoneHt \\ \frac{1}{n}\sum_{i=1}^{n}\frac{contentOffset_i}{Time(contentOffset_i))}, contentHt > readZoneHt \end{cases} \tag{6}$$

Among them, *n* represents the number of "effective operations" generated by continuous forward reading, $contentOffset_i$ represents the actual page scrolling displacement distance in the i-th "effective operation", and $Time(contentOffset_i)$ represents the time required for the i-th "effective operation" to scroll the page.

2) Average reading speed of user

This parameter is expressed by *meanV(u)*, which refers to the average value of all *meanV(msg)* sums generated by user *u* in all the content read. It can be calculated by Eq.(7):

$$meanV(u) = \frac{1}{m}\sum_{i=1}^{m}meanV(msg_i) \tag{7}$$

Where *m* represents the amount of all reading information of user *u*.

3.1.2. User Reading Interest Modeling

We conducte a statistical analysis of fine-grained user reading behavior and find that there is a linear relationship between multiple fine-grained reading behavior and the user's reading interest score. Considering the sample distribution characteristics of the regression model, data approximation accuracy and limited sample optimization issues. Therefore, support vector regression model (SVR) is used for subsequent experiments.

After all the sample sets are trained, SVR can calculate the predicted value by Eq.(8):

$$f(x) \equiv \sum_{i=1}^{n} \omega_1 \phi(x, x_i) + b \tag{8}$$

In this paper, the values of all support vector regression models are calculated separately, and the average value is calculated after completion. This can make the generated values more reliable:

$$b = \frac{1}{N_{SVR}} \left\{ \sum_{0 < \alpha_i < C} \left[y_i - \sum_{x_j \in SVR} (\alpha_j - \alpha_j^*) x_j \cdot x_i - \varepsilon \right] \right.$$
$$\left. + \sum_{0 < \alpha_i^* < C} \left[y_i - \sum_{x_j \in SVR} (\alpha_j - \alpha_j^*) x_j \cdot x_i + \varepsilon \right] \right\} \tag{9}$$

N_{SVR} in Eq.(9) is the number of support vector models.

3.2. User Reading Behavior Preference Model

The definition of the user reading behavior preference represents the personalized operation behavior and habits of the user when reading. Analysis of the current user behavior data obtained through normalization processing shows that different users have different reading behavior preferences.
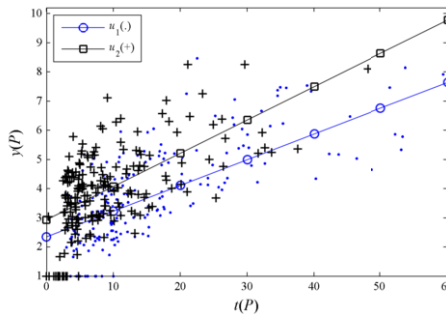

Figure 2. Differences in the linear relationship between reading time and interest score of two users

Figure 2 shows that in a short and same period of time, different users have different reading behavior preferences. For example, under the same $t(p)$, the score corresponding

to user u_1 is significantly lower than the score of user u_2. The larger $t(p)$ is, the more gradually the rating gap between the two users increases. So, using the same regression equation to calculate the reading interest of different users is less effective. Therefore, considering different reading preferences of users and establishing corresponding preference models can further improve the accuracy of calculating user interest.

3.2.1. Establishment of User Reading Behavior Preference Model

(1) Representation of the preference model

We use Eq.(10) to express user reading behavior preference[26,27]:

$$\{(behavior_1, weight_1), (behavior_2, weight_2), ..., (behavior_m, weight_m)\} \quad (10)$$

Where $behavior_j(1 \le j \le m)$ represents the user's reading behavior, m represents the specific number of behavior types, and $weight_j(1 \le j \le m)$ represents the preference weight corresponding to the behavior $behavior_j(1 \le j \le m)$.

Use the multivariate group in Eq.(11) to represent the user reading behavior preference model.

$$M = \{u_i, B, W\} \quad (11)$$

Where u_i represents user, $B = \{behavior_1, behavior_2, \cdots, behavior_m\}$, which is the fine-grained user reading behavior vector, and $W = \{weight_1, weight_2, \cdots weight_m\}$ represents the preference weight vector corresponding to the reading behavior in B.

After calculation, the reading interest degree $\hat{y}(P)$ generated on page P is obtained. The formula is as follows:

$$\hat{y}(P) = X^{(P)} \cdot W \quad (12)$$

Where $X^{(P)} = \left(x_1^{(P)}, x_2^{(P)}, \cdots, x_m^{(P)}\right)$ represents the corresponding vector of each fine-grained user reading behavior value generated on page P.

(2) Determination of implicit behavior

We divides fine-grained user behavior into two categories, explicit and implicit, according to the relationship between the user's reading interest and reading behavior. If the user 's reading interest score and some fine-grained user behaviors show a significant or less obvious potential relationship, they are classified as explicit behaviors, such as the six explicit behaviors mentioned in Section 3.1.1. Otherwise, it is implicit behavior.

Based on experimental verification and data analysis, we propose four implicit fine-grained reading behaviors: delete page operation(del), page repeat reading(rp), page collection(cl), and page content copy(cp) and add to user reading behavior preference model. The fine-grained user reading behavior vector after adding these four implicit behaviors can be expressed as:

$$B = (readTime, readHt, meanV, dragCnt, slideCnt, reviewCnt, del, cl, rp, cp) \quad (13)$$

(3) Establishment of preference model

The establishment of the user reading behavior preference model *M* means that after the behavior vector B is determined, the initial value corresponding to the preference weight vector is selected. In the process of calculating the implicit user reading behavior preference weight, we use the establishment of a linear regression model to determine the residuals of the support vector regression model. Using this regression model, the behavior preference weights of all users are set to a uniform initial value. In the process of continuously updating the user model, each parameter corresponding to the reading preference model of different users will increasingly reflect the user's own real reading habits.

Eq.(14) represents the quaternary linear regression model composed of the residuals and $\{del, cl, rp, cp\}$ in the support vector regression model. The linear regression model uses the residuals corresponding to each set of data in subsequent support vector regression experiments as its dependent variable.

$$r_{R-4}(P) = a_5 \cdot del + a_6 \cdot cl + a_7 \cdot rp + a_6 \cdot cp + \delta \tag{14}$$

3.2.2. User Reading Behavior Preference Model Update

The update of the user reading behavior preference model is to make the model fully reflect the change of the user's reading habits under the change of time, and make the model more accurate in predicting the user's reading interest.
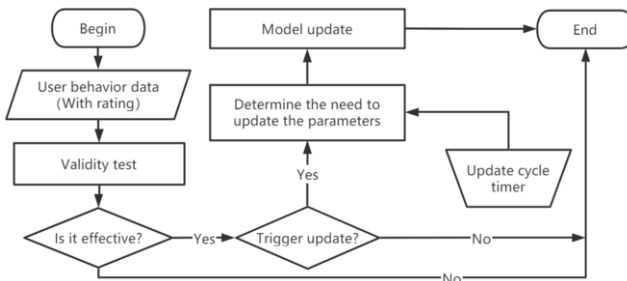

Figure 3. Flowchart of updating user reading behavior preference model

In this paper, the two strategies of system global update and update time window are applied to the update of the preference model at the same time. The update method is shown in Figure 3. When the newly added effective data reaches a certain amount, the user reading interest model based on the support vector regression model is rerun, and the user preference model is reanalyzed and adjusted based on the data generated by the model. Considering that a large number of users in the recommendation system are accustomed not to submit reading scores actively for a long time, the preference model adds a periodic timer in updating process to start mandatory updates. This avoids that the model cannot be updated effectively for a long time.

(1) Trigger condition for update

In this paper, the trigger conditions for model update are set as follows. For all users, when the error value $|\hat{y}(P) - y(P)|$ reaches a certain range compared with

the mean $\bar{r}_{SVR-ALL}$ of the residual value of the historical support vector regression model, the model update can be triggered. The formula is shown below:

$$|\hat{y}(P) - y(P)| \geq \alpha \cdot \bar{r}_{SVR-ALL} \tag{15}$$

Where the α is set to 1.3 on the basis of the experiment; $\bar{r}_{SVR-ALL} = \frac{1}{n}\sum_{i=1}^{n}|r_i|$, r_i represents the residual of each group of data in the SVR model experiment.

(2) Determine update parameters

The change of user reading behavior preference lies in the overall change or the change of a certain behavior habit. Both of the above cases need to evaluate all the behavior parameters one by one and judge whether they need to be adjusted to effectively update the model. For the change of the corresponding user preference model caused by time changes, Shi Yancui et al. [28] proposes a new concept, called behavior volatility, which is used to quantify the change of user behavior.

Changes in user behavior habits will increase the volatility. When it exceeds a certain threshold, the user preference model should be updated. We set the volatility threshold $VRB_{j,ths}$ for all fine-grained user behaviors. If $VRB_{i,j,T} > VRB_{j,ths}$, that is, the behavior b_j of the user u_i has changed greatly, resulting in the volatility corresponding to the behavior exceeding the threshold. At this time, the weight ω_j corresponding to the behavior b_j should be updated. The setting of $VRB_{j,\,ths}$ is related to factors such as system load, operating capacity, and computing power. For example, the system configuration is low and the computing resources are insufficient. In this case, the volatility threshold corresponding to certain behaviors can be slightly increased to reduce the frequency of model updates and ensure the normal operation of the system.

(3) Law of behavioral forgetting

Usually the speed of human forgetting is proportional to time, and after a certain period of time, only a small part of memory remains. Although the characteristics of human behavior habits and human memory content are not exactly the same, they also have the law of forgetting[29,30,31,32]. Ming-Sheng S et al. [33] proposes a human behavior dynamics model.

In order to better achieve the simulation of the forgetting rules of user reading behavior preference, we make appropriate modifications to the model based on the characteristics of the mobile information recommendation system based on the theory of [33]. The formula is shown as Eq.(16):

$$w'_j\left(t'\right) = \frac{\alpha}{\beta + (t' - t)} w_j(t) \tag{16}$$

Where $w'_j\left(t'\right)$ represents the new preference weight updated at time $t', w_i(t)$ represents the value generated by updating the weight at time t, or the corresponding preference weight value may be generated according to the initial settings, and α and β are forgetting parameters, respectively.
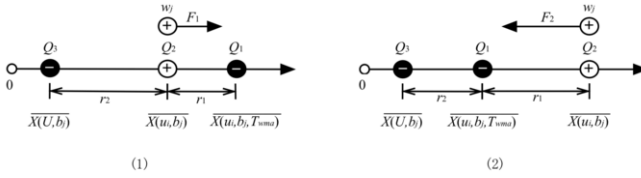
Figure 4. Schematic diagram of the relationship between Coulomb's law and user behavior

(4) Weight update function

As shown in Figure 4, we refer to Coulomb's law and regards the average value $\overline{X(u_i, b_j, T_{wma})}$ of the behavior b_j of the user u_i in the time window T_{wma} of the preference model update as the charge Q_1. The historical average value $\overline{X(u_i, b_j)}$ of the behavior b_j of the user u_i as the charge Q_2. The charge Q_3 is the average $\overline{X(U, b_j)}$ of b_j behavior of all users. The "resultant force" received by "charge" is mapped to the preference weight value ω_j corresponding to the b_j behavior of user u_i, showing the update mode used by ω_j.

A large change in the "resultant force F_e" experienced by the charge Q_2 means that the user has a large deviation in behavior preference within the model update time window T_{wma}, and the credibility of this change is low. In this case, preference weights should not be adjusted significantly. The small drift of preference will result in a small change in the resultant force. In this case, the credibility of the change is higher. Therefore, we divide the user's behavior change into two "directions". "Forward" means getting closer to the average of all users' behavior, and "reverse" means moving away from the average. When users change their behavior habits in the "forward" direction, we believe that the adjustment range of their behavior preferences should be greater than the "reverse" situation.

3.2.3. F-AFC Model Based on the Forgetting Law of User Behavior Preference

Based on the above analysis, we propose a user reading behavior preference update model named F-AFC based on the user behavior preference forgetting law.

$$
\begin{aligned}
\omega'_{i,j} &= f_{AFC}(u_i, b_j, \omega_j, t_{trig}, t_j) \\
&= f(t_{trig}, t_j) \cdot \omega_j - (1 - f(t_{trig}, t_j)) \cdot k(b_j) \cdot \left(\frac{d_1}{r_1^2(b_j)} - \frac{d_2}{r_2^2(b_j)} \right)
\end{aligned} \tag{17}
$$

Where ω_j is the preference weignt corresponding to the behavior b_j of the user u_i, t_j is the previous update time of ω_j, $\omega'_{i,j}$ is the updated value of ω_j, t_{trig} is the time when the update is triggered in the model update window T_{wma}, and $f(t_{trig}, t_j) = \frac{\alpha}{\beta + (t_{trg} - t_j)}$ is the forgetting function. $k(b_j)$ is the weight coefficient corresponding to behavior b_j. d_1 and d_2 are the behavior fluctuation proportion coefficient, which are used to adjust the behavior fluctuation relationship to determine the weight update direction. And set the upper limit $r_{max1}(b_j)$, $r_{max2}(b_j)$ and the lower limit $r_{min1}(b_j)$, $r_{min2}(b_j)$ in $r_1(b_j)$ and $r_2(b_j)$ to avoid small fluctuations in behaviors that result in large-scale updating of behavioral weights, or behavioral weights cannot be effectively updated after substantial changes in behaviors.

We use the following six-tuple to represent the user reading behavior preference update model F-AFC considering the user's fine-grained behavior forgetting laws.

$$F - AFC = \{u_i, B, W, \bar{X}, T, f_{AFC}\} \tag{18}$$

Where B is the behavior vector of user u_i, W is the preference weight corresponding to each behavior in B, $\bar{X} = (\bar{x_j} : \bar{x_j} = X(u_i, b_j) \wedge 1 \leq j \leq m)$ is the historical average value of each behavior in B, $T = (t_1, t_2, \ldots, t_m)$ is the update time corresponding to each behavior preference weight in $F - AFC$ model, and f_{AFC} is the preference weight update function in Eq.(17).

4. Experiments and Analysis

4.1. User Reading Interest Modeling

4.1.1. Experimental Data and Environment

The real fine-grained data of user reading behavior is used as the data in the experiment to ensure that the experimental results are accurate and effective. A total of 1537 pieces of information are selected, and the number of users is 116. The number of fine-grained reading behaviors with explicit reading interest rating of users is 18013. Set up the experimental environment: Intel $Core^{TM}$ i5-4570 CPU 3.20 GHz processor; 8 GB memory. Table 2 lists the data format information of some users' fine-grained behaviors.

Table 2. User behavior data table

Symbol	user 1	user 2	user 3
topicID	1	2	3
readTime(s)	23.65	47.82	12.10
readCmp	1	0.3	1
meanV(row/min)	20.13	28.34	5.49
dragCnt(time)	4	8	5
slideCnt(time)	6	12	7
reviewCnt(time)	1	2	0
rating	5	8	5.5

4.1.2. Experimental Results and Analysis

We propose a variety of explicit fine-grained reading behaviors of different mobile terminals, including user reading time, page reading completeness, effective number of page drags, effective number of page slides, number of user reviews and average reading speed. In this paper, by selecting two, three, four and six independent variables from six different explanatory variables, the corresponding binary model (SVR-2), ternary model (SVR-3), quaternary model (SVR-4) and six-metamodel (SVR-ALL) are established. Based on three indicators: average absolute error (MAE), goodness of fit (R^2) and accuracy (precision), the regression models using partial fine-grained user reading behavior and all fine-grained user reading behavior are compared.

Figure 5 shows the experimental results of four support vector regression models under the evaluation of goodness of fit R^2. After considering more fine-grained reading behaviors, the regression model's goodness of fit R^2 is even better. What's more, it can be seen that the results of the goodness of fit R^2 of the four support vector regression models are all higher than 0.7. In particular, the R^2 of SVR-ALL, which takes into account all fine-grained reading behavior, is closer to 0.8, which fully reflects the advantages of the model in representing the relationship between user's reading interest and fine-grained reading behaviors.

Figure 6 shows the MAE results of four different support vector regression models. Under the four vector regression models, the results are all less than 0.9. The optimal one is the SVR-ALL curve, whose final MAE value is close to 0.6.

The accuracy rate curve shown in Figure 7 shows an upward trend, and the absolute error has the highest accuracy when $\varepsilon = 1$. Keep other conditions invariant, with the continuous introduction of six kinds of explicit fine-grained reading behaviors(from SVR-2 to SVR-ALL), the accuracy of the model is also constantly improving. The experimental results fully demonstrate the effectiveness of the six behaviors. It also shows that the regression analysis can explain the functional relationship between user reading interest and fine-grained reading behaviors.
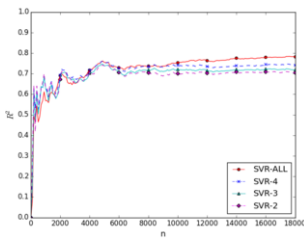

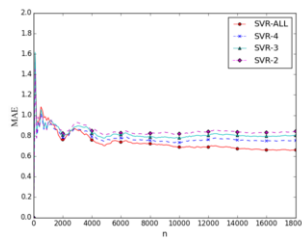**Figure 5.** Goodness of fit of SVR model at Multiple data scales

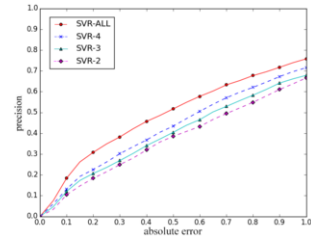**Figure 6.** The average absolute error of the SVR model at various data scales

Figure 7. Accuracy of SVR model under different absolute errors

4.2. User Reading Behavior Preference Model

4.2.1. Experimental Data and Environment

This section uses the same experimental conditions as 4.1 above, and the source of the user reading behavior data is also the same. The F-AFC model is trained according to the training set, and a personalized user reading behavior preference model is generated for all users. During training, the SVR-ALL experimental data is compared and analyzed, and then the test set is used to verify the validity of the F-AFC model.

In the process of training the F-AFC model, first initialize the user reading behavior preference model. Then set the global update cycle for ten days and run the F-AFC model. During the system operation, the user's reading preference model can be automatically updated to make the model more in line with the user's personalized characteristics. On the basis of the model obtained in the previous step, the test set is used to calculate the users' reading interest. The actual scores of users are compared, and the changes of accuracy and average absolute error are analyzed. Meanwhile, the experimental results displayed by SVR-ALL model are compared and analyzed.

4.2.2. Preference Model Parameter Setting

Table 3 lists the initial weight values of the implicit behavior preference model.

Table 3. Initial weight table of implicit fine-grained behavior preferences

	Page Delete	Page Collection	Page Repetition	Page Content Copy
Symbol	$del(P)$	$cl(P)$	$rp(P)$	$cp(P)$
Weight	-1.0936	2.0105	1.7087	1.3263

After a lot of experiments in this paper, the parameters of the F-AFC model are obtained. $\alpha = 25.5, \beta = 28.1, d_1 = 0.142, d_2 = 0.105$, The remaining parameters are shown in Table 4-7.

Table 4. Explicit fine-grained behavior volatility threshold table

b_j	b_1	b_2	b_3	b_4	b_5	b_6
Corresponding Behavior Name	Page Reading Time	Page Reading Completeness	Page Drag Times	Page Swiping Times	Review Times	Average Page Reading Speed
$VRB_{j,ths}$	0.06	0.026	0.12	0.083	0.054	0.023

Table 5. Implicit fine-grained behavior volatility threshold table

	b_1	b_2	b_3	b_4	b_5	b_6
$k(b_j)$	0.15	0.0014	0.0016	0.00093	0.0036	0.0279
$r_{\min 1}(b_j)$	0.18	0.08	0.32	0.28	0.14	0.36
$r_{\max 1}(b_j)$	1.5	0.2	1.1	2.2	0.26	1.03
$r_{\min 2}(b_j)$	0.18	0.05	0.23	0.31	0.15	0.28
$r_{\max 2}(b_j)$	2.0	0.15	1.2	2.2	0.28	1.15

Table 6. Corresponding parameters of explicit behavior in the F-AFC model

	b_1	b_2	b_3	b_4	b_5	b_6
$k(b_j)$	0.15	0.0014	0.0016	0.00093	0.0036	0.0279
$r_{\min 1}(b_j)$	0.18	0.08	0.32	0.28	0.14	0.36
$r_{\max 1}(b_j)$	1.5	0.2	1.1	2.2	0.26	1.03
$r_{\min 2}(b_j)$	0.18	0.05	0.23	0.31	0.15	0.28
$r_{\max 2}(b_j)$	2.0	0.15	1.2	2.2	0.28	1.15

Table 7. Corresponding parameters of implicit behavior in F-AFC model

	b_7	b_8	b_9	b_{10}
$k(b_j)$	0.0091	0.027	0.021	0.016
$r_{\min 1}(b_j)$	0.06	0.08	0.08	0.05
$r_{\max 1}(b_j)$	0.11	0.15	1.3	1.6
$r_{\min 2}(b_j)$	0.06	0.08	0.09	0.06
$r_{\max 2}(b_j)$	0.11	0.15	1.2	1.9

4.2.3. Experimental Results and Analysis

The experiment compares two models: SVR-ALL model and F-AFC model. Figure 8 shows the average absolute error of them. It can be clearly seen from the figure that when the data volume rises to 600 groups, the fitting error is significantly reduced. When the data volume rises to 3600 groups, the value tends to be stable, around 0.6 and the overall performance is more stable. Which indicate that the F-AFC model has a certain effect on improving the accuracy of the user's reading preference model.

Figure 9 compares the accuracy of the above two models in calculating user reading interest. The experimental results show that after running the F-AFC model, the accuracy is significantly increased. When the absolute error δ_{prc} is set to 0.5, the accuracy rate increases from approximately 0.4 to 0.5. When δ_{prc} is set to 1, the accuracy rate is improved from 0.71 to 0.88 .

The model is futher tested by using the test data set. Figure 10 shows the average absolute error value before and after the F-AFC model is run. The results also show that the F-AFC model has certain advantages. Figure 11 shows the accuracy rate curve of the user's reading interest before and after the F-AFC model is run. It can be seen from the figure that after the F-AFC model is run and the accuracy of the predicted value also rises significantly. When the value of δ_{prc} is 0.5, the accuracy rate can reach 0.49. If the limit of δ_{prc} is relaxed to 1, the accuracy rate is about 0.86. The experimental results show that based on the SVR-ALL model, the four implicit fine-grained user reading behaviors are effective in improving the accuracy of user reading interest prediction. Moreover, on the basis of the former, the F-AFC model considering the updating mechanism of user preference has a certain improvement in recommendation performance and accuracy compared with the SVR-ALL model.
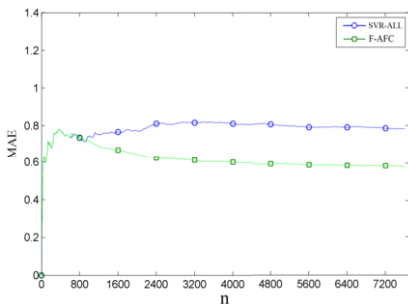

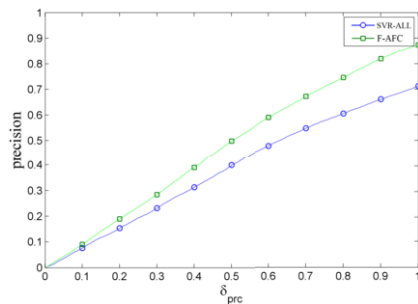**Figure 8.** Comparison of average absolute error under two models

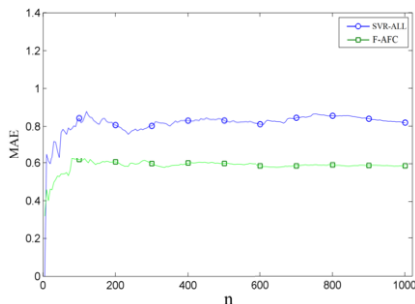
Figure 9. Comparison of accuracy under two models

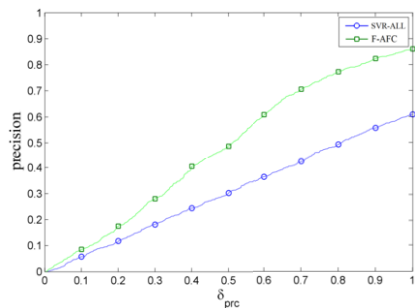**Figure 10.** Comparison of MAE under test data

Figure 11. Comparison of accuracy under test data

5. Conclusions

In this paper, the normalization method of user's fine-grained behavior acquisition and preprocessing is studied for mobile information push system, which can carry out unified normalization and quantification of fine-grained user behavior. A support vector regression model is introduced to model the fine-grained user's reading behavior, and the relationship between fine-grained user reading behavior and user reading interest is successfully established. And the validity of the proposed six explicit fine-grained behaviors is verified by ablation experiments. On this basis, considering the different reading habits of different users, four implicit fine-grained behaviors are introduced and the corresponding updating mechanism is designed to establish and update the user reading behavior preference model. Experiments show that the change of the user's reading interest can be explained by the user's fine-grained reading behavior. The proposed preference modeling and updating method have a certain improvement effect on the accuracy of reading interest calculation of users.

References

[1] Kwok R. Personal technology: Phoning in data. Nature, 2009, 458(7241): 959.

[2] Roitman H, Carmel D, Mass Y, et al. Modeling the uniqueness of the user preferences for recommendation systems. In: Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval. New York: ACM, 2013. 777-780.

[3] Said A, Jain B J, Narr S, et al. Estimating the magic barrier of recommender systems: a user study. In: Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval. New York: ACM, 2012. 1061-1062.

[4] Meng X W , Hu X , Wang L C , et al. Mobile Recommender Systems and Their Applications. Journal of Software, 2014, 24(1):91-108.

[5] Hai-Ling X U , Xiao W U , Xiao-Dong L I , et al. Comparison Study of Internet Recommendation System. Journal of Software, 2009, 20(2):350-362.

[6] Zeng M, Cao H, Chen M, et al. User behaviour modeling, recommendations, and purchase prediction during shopping festivals. Electronic Markets, 2019, 29(2): 263-274.

[7] Meng W, Yang D, Xiao Y. Incorporating user micro-behaviors and item knowledge into multi-task learning for session-based recommendation. In: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2020: 1091-1100.

[8] Ma Q J, Ng C W W, Mašín D, et al. An approach for modelling volume change of fine-grained soil subjected to thermal cycles. Canadian Geotechnical Journal, 2017, 54(6): 896-901.

[9]   He X, Peng Y, Zhao J. Fast fine-grained image classification via weakly supervised discriminative local-
      ization. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 29(5): 1394-1407.
[10] Huang F, Wu W, Yang M, et al. A fine-grained permission control mechanism for external storage
      of android. IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 2016:
      002911-002916.
[11] Zhou C, Bai J, Song J, et al. Atrank: An attention-based user behavior modeling framework for recom-
      mendation. arXiv preprint arXiv:1711.06632, 2017.
[12] Feng Y, Hu B, Lv F, et al. ATBRG: Adaptive Target-Behavior Relational Graph Network for Effective
      Recommendation. In: Proceedings of the 43rd International ACM SIGIR Conference on Research and
      Development in Information Retrieval. 2020: 2231-2240.
[13] Nguyen M D, Cho Y S. A Variational Autoencoder Mixture Model for Online Behavior Recommenda-
      tion. IEEE Access, 2020, 8: 132736-132747.
[14] Claypool M, Le P, Wased M, et al. Implicit interest indicators. In: Proceedings of the 6th international
      conference on Intelligent user interfaces. New York: ACM, 2001. 33-40.
[15] Liang T P, Lai H J. Discovering user interests from web browsing behavior: An application to internet
      news services. In: Proceedings of the 35th annual Hawaii international conference on system sciences.
      IEEE, 2002: 2718-2727.
[16] Teng Y, He L. Rating prediction algorithm and recommendation based on user beahavior in IPTV. In:
      2012 2nd International Conference on Consumer Electronics, Communications and Networks (CEC-
      Net). Yichang: IEEE, 2012. 3373-3378.
[17] Song Y, Ma H, Wang H, et al. Exploring and exploiting user search behavior on mobile and tablet devices
      to improve search relevance. In: Proceedings of the 22nd international conference on World Wide Web.
      2013: 1201-1212.
[18] Rho S, Han B, Hwang E. SVR-based music mood classification and context-based music recommenda-
      tion. In: Proceedings of the 17th ACM international conference on Multimedia. 2009: 713-716.
[19] Vohr, Hans-Werner. Systemic Vascular Resistance (SVR). Springer Berlin Heidelberg, 2005.
[20] Hao P Y. New support vector algorithms with parametric insensitive/margin model. Neural Networks,
      2010, 23(1): 60-73.
[21] Marin L, Isern D, Moreno A. Dynamic adaptation of numerical attributes in a user profile. Applied
      intelligence, 2013, 39(2): 421-437.
[22] Zhang H, Wei D, Hu Y, et al. Modeling the self-similarity in complex networks based on Coulomb's
      law. Communications in Nonlinear Science and Numerical Simulation, 2016, 35: 97-104.
[23] Lai Y, Lv Z, Li K C, et al. Urban traffic coulomb's law: A new approach for taxi route recommendation.
      IEEE Transactions on Intelligent Transportation Systems, 2018, 20(8): 3024-3037.
[24] Nichols D M. Implicit Rating and Filtering. In: Proceedings of the Fifth Delos Workshop on Filtering
      and Collaborative Filtering, 1997:31-36.
[25] Bu J, Tan S, Chen C, et al. Music recommendation by unified hypergraph: combining social media in-
      formation and music content. In: Proceedings of the 18th ACM international conference on Multimedia.
      2010: 391-400.
[26] Salton G, Wong A, Yang C S. A vector space model for automatic indexing. Communications of the
      ACM, 1975, 18(11): 613-620.
[27] Lin Hongfei, Yang Yuansheng. Representation and update mechanism of user interest model. Computer
      Research and Development, 2002, 39 (7): 843-847.
[28] Shi Yancui, Meng Xiangwu, Zhang Yujie, etc. An adaptive learning method for contextual mobile user
      preferences. Journal of Computers, 2012, 23 (10): 2533-2549.
[29] Matuszyk P, Vinagre J, Spiliopoulou M, et al. Forgetting techniques for stream-based matrix factoriza-
      tion in recommender systems. Knowledge and Information Systems, 2018, 55(2): 275-304.
[30] Barabasi A L. The origin of bursts and heavy tails in human dynamics. Nature,2005, 435(7039): 207-
      211.
[31] Fan Chao, Guo Jinli, Han Xiaopu, etc. A review of studies on human behavior dynamics. Complex
      Systems and Complexity Science, 2011, 8 (2): 1-17.
[32] Vazquez A. Impact of memory on human dynamics. Physica A: Statistical Mechanics and its Applica-
      tions, 2007, 373: 747-752.
[33] Ming-Sheng S, Guan-Xiong C, Shuang-Xing D, et al. Interest-driven model for human dynamics. Chi-
      nese Physics Letters, 2010, 27(4): 250-252.