To Design and Implement a Recommender System based on Brainwave: Applying Empirical Model Decomposition (EMD) and Neural Networks

Chia-Yu Lai  
National Pingtung University of Science and Technology  
chiayulai@mail.npust.edu.tw

Zhe-Lun Jhang  
National Pingtung University of Science and Technology  
M10456016@mail.npust.edu.tw

Deng-Neng Chen  
National Pingtung University of Science and Technology  
dnchen@mail.npust.edu.tw

Abstract

Recommender systems collect and analyze users’ preferences to help users overcome information overload and make their decisions. In this research, we develop an online book recommender system based on users’ brainwave information. We collect users’ brainwave data by utilizing electroencephalography (EEG) device and apply empirical mode decomposition (EMD) to decompose the brainwave signals into intrinsic mode functions (IMFs). We propose a back-propagation neural networks (BPNN) model to portrait the user’s brainwave preference correlations based on IMFs of brainwave signals, thereby designing and developing the book recommender system. The experimental results show that the recommender system combined with the brainwave analysis can improve accuracy significantly. This research has highlighted a future direction for research and development on human-computer interaction (HCI) design and recommender system.

1. Introduction

In contemporary society, the rapid development of new technology and the Internet facilitate the acquisition of information and render the process more simple and convenient. The rapid-growth amount of online streaming information force people to face big challenges to filter out the worthless data or identify the valuable knowledge. Thus, to solve those problems, recommender systems have been widely used in many business fields or academic areas to help users overcome information overload and make their decisions.

A recommender system is defined as an information system that can recommend related information or products for users by identifying their preferences on purchasing behaviors, entered keywords, or browsing history on the Internet. Recommender systems can be applied to a variety of areas, i.e., movies, news, music, and books [1]. Drawing upon the widely adopted information filtering technologies, recommender systems can be divided into content-oriented filtering, collaborative filtering, demographic filtering [2], and hybrid filtering [3]. A content-oriented filtering method constructs a model of a user’s preference by integrating their historical records such as shared contents and reviews, comparing the similarities between the products and the user’s preferences. The product with the highest similarity will be recommended to the user. Collaborative filtering is to identify the preferences of a group of similar users and makes recommendations to a single user based on group preferences. Hybrid filtering merges content-oriented filtering and collaborative filtering with indicators of weight, exchange, and mixture [4].

With regards to recommender systems, using the traditional collaborative filtering methods limits profile information for the new beginner. In recommender systems, the cold start problem is most prevalent with limited or no information about the user’s profile. To address this issue, some recommenders try to lessen a beginner’s time to learn enough to recommend the new user an item in her/his interest [5, 6]. For instance, Lee, Cho and Kim [7] have put forward to a suitable method called implicit rating using collaborative filtering for the mobile music market. Their findings show that this new method has better accuracy than traditional methods. Some famous e-commerce sites or video websites, such as Amazon, eBay, and YouTube, all have introduced recommendation technologies for users’ preference [8-10].

Another work done by Castro-Schez, Miguel, Vallejo and López-López [11] have introduced the concept of fuzzy logic in B2C e-commerce, which allows the system to provide the search results of
products that are potentially related, even when the keywords are not accurate. Similarly, Barragáns-Martínez, Costa-Montenegro, Burguillo, Rey-López, Mikic-Fonte and Peleteiro [12] propose recommender systems for Web 2.0 TV programs, allowing users to add and tag content, evaluations, review items. Their work provides TV recommendations combined with content-oriented filtering and collaborative filtering, as well as the typical advantages of social networking. In addition, Crespo, Martínez, Loveless, García-Bustelo, Gayo and De Pablos [13] attempt to build a recommender system based on the known preferences of other users or other users with similar characteristics.

Regardless of which method is adopted, users’ preferences must be identified before developing the preference model to provide recommendations. Users’ preferences can be distinguished in many ways such as entered keywords, browsing history, purchase records, questionnaires, and product reviews. The mechanism for collecting users’ preferences is the major challenge in designing a recommender system.

Recently, with the proliferation of sensor-based technologies, many researchers use sensing devices to measure people’s emotion and feelings for understanding human behaviors or detecting their preferences. Hence, we can easily measure users’ physical conditions via such equipment. Sensing devices display data, reflecting the user’s physiological and psychological status to obtain the overall picture of the situation, which is helpful for the establishment of the user’s preference model; for instance, Lin, Wang, Wu, Jeng and Chen [14] utilize different types of music to produce four kinds of emotions, i.e., happiness, anger, sadness, and joy. The research analyzes brainwaves signals to indicate the users’ psychological reaction at that time.

Furthermore, due to the advancements in sensing technology for brainwave signals, neuroscience, and information systems, a new discipline has been developed. Combines with the above fields into a neuro-information-system (NeuroIS) [15, 16], it has become an emerging research field and attracts researchers’ attention. Hence, many scholars of information systems and neuroscience attempt to operate a variety of non-invasive devices to measure brain activity. Currently applied devices include magnetoencephalography (MEG), electroencephalography (EEG), functional magnetic resonance imaging (fMRI), positron emission tomography [17], and eye fixation related potentials (EFRP) [16].

At present, a single channel head-mounted EEG has been launched on the market, which is widely available at a reasonable price, comfortable to wear, and user-friendly to operate. In particular, as its weight is low, it allows users to be free from perceiving the experiment during the experiment [18].

However, as brainwave signals are nonlinear, subsequent research on the treatment of brainwave signals is necessary. The empirical mode decomposition (EMD) can specifically process non-linear and unstable brainwave signals [19]. Therefore, this study adopts this EMD approach for solving the non-linear issue to analyze brainwave signals.

In previous studies of brainwave signals, many scholars apply machine learning algorithms, such as artificial neural networks, to classifying brainwave signals to detect whether a driver is likely to be struck by drowsiness or epilepsy, and the experimental results show the model is successfully predicting the classification result and fairly good for implementing [20-22]. Hence, this study adopts artificial neural networks to classify brainwave signals.

Previously, recommender systems are designed to collect the users’ external behaviors or other users with similar preferences; however, neither of these systems consider the users’ psychological status. Consequently, we attempt to construct a recommender system cooperating with the brainwave-preference correlational model, which is developed by transforming the users’ brainwave signals as their preference indicators. The brainwave signals are analyzed by EMD first. By using analyzed EMD data, we construct the brainwave-preference correlational model by implementing the neural network classifiers to provide better understand in predicting users’ preference. In this study, books are selected as the commodities for recommendations. We propose a book recommender system to identify the users’ preferences by analyzing their brainwave signals and provide recommendations to the users.

2. Literature review

2.1. Electroencephalography (EEG)

Electroencephalography (EEG) is a non-invasive neuroimaging technique, using electrodes placed on the scalp to record the electrical activity of the brain [23]. The amplitude of the EEG signals measured by the scalp surface ranges from 10 to 200 μV with a frequency in the range 0.5–40 Hz. The recorded EEG signals, also known as brainwaves, are composed of
multiple oscillations at different frequencies, which have been classified in five major frequency bands: delta (1-3 Hz), theta (4-7 Hz), alpha (8-13 Hz), beta (14-30 Hz) and gamma (31-50 Hz) \[24\]. Generally, the slowest brain rhythms, delta, are dominant during an inactive state and the fastest, gamma, are observed during active information processing. In addition, EEG, as a biometric trait, represents the one-dimensional signal which can be recorded efficiently and processed in a short computational time, thereby achieving promising recognition rates \[25\].

Recent EEG measurements are widely used for various fields. Merzagora, Bunce, Izzetoglu and Onaral \[26\] use EEG to yield more reliable results for detecting deception in clinics and legal areas. Azali and Nadiah \[27\] propose to classify EEG signals for monitoring wheelchair for severe impairment users. Furthermore, several works exist that are related to emotion recognition by using EEG device \[28-30\]. Experiments show that emotions can be recognized by using music videos or audiovisual pictures to induce emotions. Therefore, several studies in personalized music recommendation or customer preference are conducted by EEG device \[31-33\]. Follow this research line, we adopt the EEG signals to identify users’ preferences.

Specially, EEG measures voltage fluctuations externally direct from the scalp surface with no risk and limitation. In this research, only EEG measured from the head surface is considered and an EEG device developed by NeuroSky company is used in our experiment \[34\] \[35\].

2.2. Empirical Mode Decomposition (EMD)

Empirical mode decomposition (EMD), proposed by Huang \[19\], can specifically deal with non-linear and unstable data series. Brainwaves are often characterized by their prominent frequency and energy content, representing non-stationary and non-linear signals. Thus, to analyze EEG signals, many previous studies have confirmed that EMD can be applied effectively to address the nonlinear issue of brainwave signals \[17, 36, 37\]. For instance, Pachori \[36\] use the mean frequency (MF) measure of intrinsic mode functions (IMFs) as a feature to identify the differences in EEG signals.

The EMD method is an adaptive algorithm without any condition about the stationarity and linearity of the signal. The main purpose of the EMD method is to decompose the nonlinear and nonstationary signal \(x(t)\) into a sum of intrinsic mode functions (IMFs) through the iterative process, as shown in Figure 1. The iterative process can be described as followed:

- First, detect the extrema (maxima and minima) of the brainwave dataset \(x(t)\)
- Then, calculate the upper and lower envelopes respectively by connecting the maxima and minima separately with cubic spline interpolation.
- Determine the local mean \(Mn(t)\) and decide the detail \(h(t)\) is an IMF by checking the two basic conditions including the number of maxima and the number of zero crossings must be the same or differ at most by one and the mean value of the local maxima envelope and the local minima envelope is zero, where \(m\) is the number of IMFs.
- Finally, to determine the rest of the IMFs, generate the residue \(r(t)\). The residue can be treated as the new signal, repeating the above iterative process until the final residue is a constant or a function from which no more IMFs can be derived.

\[\text{Figure 1. The Flowchart of EMD}\]
3. System architecture

To develop the recommender system, two-stage experiments are conducted for training recommendation model and providing recommendation sequentially. The first-stage experiment involves recording the users’ book review scores while collecting their current brainwave signals. After collecting signals, we transform the EEG signals by employing EMD to extract useful features. In this stage, we build the personal brainwave preference correlational model for emotions classification from EEG signals using an artificial neural network classifier.

The second stage is designed to develop a personal book recommender system cooperating with the brainwave preference correlational model built in the first stage. The aim of the recommender system is to predict the users’ preferences and recommend books to them based on their brainwave analysis. Figure 2 shows our proposed system architecture and all the phases of this study.

The main purpose of the first stage is to establish the relationship between the brainwave signals and the users’ preferences; therefore, an online bookstore website is developed for this experiment. While wearing the EEG device on their heads, users are invited to browse the website. Next, users are asked to grade the books. The system records the book abstracts read by each user and their preference scores. Meanwhile, the system collects users’ brainwave signals. The brainwave signals are processed by EMD for the acquisition of IMF, where four indicators in each section of IMF are maximum, minimum, average, and standard deviation. Those indicators represent input parameters and the preference scores represent the output parameters of the artificial neural network for establishing the brainwave preference correlational model.

The first stage is designed to establish the brainwave preference correlational model, which is the core model of the recommender system in the second stage. We develop an online bookstore website for users to browse during the experiment. The aim of the website is to store the book types and browsing time of each user. In addition, this study establishes a module of book preferences, where two items, “like” and “dislike”, are set to indicate the scale of the users’ preferences for the books. After reading book abstracts, users are required to grade their preferences.

We also collect subjects’ brainwave signals with MindWave, as developed by NeuroSky. This EEG device is a portable electroencephalograph. EEG signals are transmitted to the computer via Bluetooth. The signals captured during the experiment involve the original brainwave signals, the degrees of concentration and meditation. Additionally, eSense indicators are developed by NeuroSky, namely alpha, beta, Delta, theta, and gamma. This study only uses original brainwave signals as the experimental data, not eSense indicators.

Due to the nonlinearity of brainwave signals. We adopt EMD to process nonlinear and unstable wave signals. EMD can convert any unstable and nonlinear signals into IMFs[21]. Four IMFs statistical values: minimum, maximum, average, and standard deviation represent features as eigenvalues [21]. The artificial neural networks classifier is trained with eigenvalues to establish the brainwave preference correlational model.

The artificial neural networks classifier applied to this study is back-propagation neural networks (BPNN) [37]. The main purpose is to input the users’ preferences and four statistical values from IMFs into the neural network to develop the brainwave preference correlational model. In the second stage, the trained brainwave preference correlational model is implemented to predict the users’ preferences and make book recommendations.
The brainwave signals preference correlational model, as established in the first stage, can effectively display the degree of correlation between brainwave signals and the users’ preferences. Through observing the changes in brainwave signals, this study acquires the users’ preferences scores, as well as the brainwave signals, which undergo EMD conversion to the brainwave signals preference correlational model. This system can effectively and timely predict users’ preferences, search the similar books in the database with the preferred book types, and make recommendations to the user.

4. Experiment design

This research is divided into two stages. The first stage is to collect four statistical values in IMFs after the EMD conversion of brainwave signals and the users’ preferences to construct the brainwave signals preference correlational model using the artificial neural networks. The second stage is to establish a personal book recommender system cooperating with brainwave signals preference correlational model. Both stages require to collect the users’ brainwave signals. The proposed processes of the two stages are illustrated as follows.

4.1 The flow of first stage experiment

The purpose of this stage is to establish the brainwave signals preference correlational model. The methods of data collection are designed in the specialized website and experimentation. First, the experiment is carried out in a quiet environment, where the interference of external noises on the brainwaves could be reduced. The users are asked to wear the equipment. After the devices are successfully connected, the users are asked to browse the online bookstore website. The system randomly recommends a book for the users to read and keeps track of the two scale items of “like” and “dislike”, and the brainwave signals are recorded. Each user should read about 1000 words (in Chinese) for each book’s abstract. The abstracts are written by the e-book store staff, not by the authors of the books; therefore, the abstracts have the similar writing style. Five rounds are conducted in this stage of the experiment.

When the experiments collect sufficient samples, this study conduct EMD conversion process. Next, the converted signals and users’ book preferences are put into the artificial neural networks classifier for training. As the artificial neural networks converged to the threshold set by this study, the correct rate of classification would be greater than 90%. Thus, this study can establish the artificial neural networks classifier as the brainwave signals preference correlational model for the second stage of the experiment. The flow chart of the first-stage research design is shown in Figure 3.
4.2 The flow of second stage experiment

At this stage, we collect data and design our experimentation mainly on the specialized website. The brainwave signals preference correlational model, as created in the previous stage, is used to predict new users’ book preferences. At first, similar to the previous stage, the experiment is carried out in a quiet environment for reduced external noise, the users are invited to put on the equipment, and the connection is checked to ensure success. In the case of connection failure, the users are assisted to adjust the equipment until successfully connected.

After successfully connecting with the equipment, the subjects are required to browse the online bookstore. The system also randomly provides a book for the users to read and score, while their brainwave signals are recording. When the users finish reading, the system can conduct EMD conversion of the brainwave signals and convert them to the brainwave signals preference correlational model to predict users’ preferences and make recommendations. Later, system performance and accuracy would be evaluated by the users’ preference scores. The flow chart of the second-stage research design is shown in Figure 4.

5. Data collection and analysis

5.1 Data collection

First, in this study, we design a specified website for personal book recommendations. All the contents of the website are from an online bookstore and are classified into 8 categories of suspense/reasoning, science fiction/fantasy, horror/thriller, warmth/healing, love, homosexual love, romance, and other literature. Based on the popularity ranking by the e-book store in descending order, we select 10 books for each type, amounting to the total of 80 books, as shown in Table 1. The book’s title, authors, introduction, and summarized contents are collected in the database.

<table>
<thead>
<tr>
<th>Types of books</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspense/Detective</td>
<td>10</td>
</tr>
<tr>
<td>Science fiction/Fantasy</td>
<td>10</td>
</tr>
<tr>
<td>Horror/Thriller</td>
<td>10</td>
</tr>
<tr>
<td>Warm/Healing</td>
<td>10</td>
</tr>
<tr>
<td>Love</td>
<td>10</td>
</tr>
<tr>
<td>Homosexual love</td>
<td>10</td>
</tr>
<tr>
<td>Romance</td>
<td>10</td>
</tr>
<tr>
<td>Other literature</td>
<td>10</td>
</tr>
</tbody>
</table>

There were 20 users invited in the first-stage experiment, and each user went through 5 to 10 rounds of experiments. In each round, after the users read a book’s information, they were asked to fill out the books preference scores. In the second stage of the

---

1 http://www.books.com.tw
experiment, 50 new users were invited, and each underwent 3 rounds of experiments, amounting to the total of 150 samples. In each round, each user was required to read a book and complete the book preference scores. The male ratio is 66%, and the female is 34%; the age range of the users is 21-26 years old; 64% have college degrees, while 36% have postgraduate degrees. Each experiment is about 20-30 minutes. The reason that we drop some samples is due to the difficulty in setting up the subject for the signal acquisition process. Although recent portable EEG devices have solved this problem to some extent, the quality of EEG recorded from portable devices is still poor. For instance, transmitting signals or the connecting with devices might be interrupted. Thus, some recorded signals are too short for segmentation and further analysis. After excluding the invalid samples, a total of 122 effective samples were collected.

5.2 Model training

In the first-stage experiment, the original brainwave signals are processed. The EMD conversion of brainwave signals produces 8 IMFs and a complementary wave. This study uses 4 statistical values of 8 IMFs: maximum, minimum, average, and standard deviation as the eigenvalues to be input in the neural networks [21]; while the users' preferences for books are the artificial neural network's output values. The values are put into the artificial neural networks for training, and to establish the brainwave signals preference correlational model. As multiple attempts may be required to train the artificial neural networks to display satisfying convergence, this study uses various parameters, such as the learning rate and numbers of neurons in the hidden layer, for the adjustment of the training artificial neural networks, as shown in Table 2.

Table 2. The Parameters of Artificial Neural Networks

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nodes in input layer</td>
<td>32</td>
</tr>
<tr>
<td>Number of hidden layer</td>
<td>1</td>
</tr>
<tr>
<td>Nodes in hidden layer</td>
<td>20</td>
</tr>
<tr>
<td>Nodes in output layer</td>
<td>2</td>
</tr>
<tr>
<td>Learning rate</td>
<td>0.1</td>
</tr>
</tbody>
</table>

5.3 Data analysis

The experimental results of second-stage experiment are shown in Table 3.

Table 3. The experimental results

<table>
<thead>
<tr>
<th></th>
<th>Categorized as &quot;Like&quot; by the system</th>
<th>Categorized as &quot;Dislike&quot; by the system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Categorized as &quot;Like&quot; by</td>
<td>32 (a)</td>
<td>21 (b)</td>
</tr>
<tr>
<td>the subject</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Categorized as &quot;Dislike&quot;</td>
<td>32 (c)</td>
<td>37 (d)</td>
</tr>
<tr>
<td>by the subject</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The accuracy, precision, and recall rates are used to measure the effectiveness of the recommender systems. Accuracy denotes the number of samples that the system correctly predicted against the original sample of this study; precision denotes the sample number of correct predictions of "Like" against the total number of "Like" predicted by the system; recall denotes the number correctly predicted in the category of "Like" against the number of "Like" in the original sample of this study. The equations of each index are shown in followings.

\[
\text{Accuracy} = \frac{a + d}{a + b + c + d} \times 100\%
\]
\[
\text{Precision} = \frac{a}{a + c} \times 100\%
\]
\[
\text{Recall} = \frac{a}{a + b} \times 100\%
\]

According to Table 3, the accuracy rate is 56.6%, the precision rate is 50%, and the recall rate is 60.4%, where outperform other similar brainwave recommendation models. Given our experimental design, eight categories of books are selected to predict users’ preferences, while the recommendation results are made to users based on eight categories. Without any recommendation mechanism, the book is recommended to the user randomly, yielding only 12.5% probability rate hitting user’s preference. The 56.6% accuracy rate is significantly higher than the 12.5%, which also proves that the recommender system based on brainwave signals is feasible.

6. Conclusions

In this research, we propose a recommender system cooperating with brainwave analysis in this research. A major challenge is the limited availability of users’ preferences for developing recommender systems. We address this limitation by utilizing brainwave information. The brainwave signals are applied to indicate the personal preferences, processing by EMD first and analyzed by BPNN to
construct the brainwave preference correlational model. The model is used as the core of the recommender system. The experimental results show that the system has an acceptable effectiveness of recommendation.

This is an ongoing research and hence, has substantial room for improvement. Our study comes with some limitations. Most notably, due to the instability of brainwave signals, we also found that it is difficult to get stable and consistent results in our preliminary experiment. Secondly, the writing style of the book abstract might confound the outcome of recommendation. Third, the demographic factors would affect the brainwave signals when the users are under different stimulated conditions [38]. Therefore, to design a better brainwave-based recommender, demographic factors, such as age and gender, could be concerned in the future [39].

In summary, our results provide a promising step towards inferring the impacts of brainwave analysis on recommendations. EMD is an effective method to preprocess the brainwave signals. Our brainwave signals analysis could be applied to many other fields, such as snippets of audio or short YouTube clips. For future research direction, we have highlighted a novel HCI research direction about brainwave-based recommender system.

7. References


[38] Corsi-Cabrera, M., J. Ramos, M. A. Guevara, C. Arce, and S. Gutierrez, "Gender Differences in the EEG During Cognitive Activity", International Journal of Neuroscience, 72, 3-4, 1993, pp. 257-264.