

Recommendation System of Program Based on REST Style

Jin Bao SONG^{1,a}, Jin Hong SONG², Jian Ping CHAI¹ and Ying LIN¹

¹Information Engineering School, Communication University of China, Beijing, China

²Shandong Gold Mining Jiaojia Gold Mine (Laizhou) Co., Ltd

Abstract. With the popularity of digital TV, TV programs have been on the increase no matter in both the number and species, which brought many choice to the users. Although the digital TV has increased largely in the selectivity, it has become a fussy process that users search for programs which they are interested in. So there is need to have an efficient program recommendation system to solve the problem that is "information overload" for users. It can not only help users to get the program which they require, but also bring convenience to people's life. The program recommendation system named MyView is planned and designed, aimed at providing an efficient information platform. The system also involves intelligent recommendation. The information guide will trigger the recommendation engine after users registering information, the engine will accord to the data in the guide information to make the personalized program recommendation. The system was deployed in the Tomcat and Apache integration servers on my localhost, so it also belongs to the Web application based on J2EE platform. AJAX is used that can achieve a good user experience to develop web presentation layer on MyView PC browser with flexible interface performance. The background of business services uses the hierarchical form. MyView System uses the CXF framework and Hibernate to equip controller and data persistence layer in the Spring container. The overall framework of the system uses the REST style, in order to extend the performance and function later. Background service layer with uniform interface, marked by the URI resource. At the same time, HTTP requestion is submitted by the AJAX to obtain services provided by resources. Finally, we can analyze and summary the features of MyView System.

1 Instruction

Along with a variety of video sites increasing in recent years, the way of people watching TV programs converts from the traditional TV to set-top box and then to the use of mobile phones and tablet computers to access the Internet to watch a variety of programs through different video sites without subjecting to the limitations of time and space. In order to gain more traffic and revenue, major video sites will provide more video resources and more personalized services as their competitive advantage. How to effectively integrate the user's own needs and existing video resources together becomes an exceptionally arduous task in the field of today's information technology services. Currently, Youku Tudou, iQIYI and Sohu video lead in the video program system, which basically complete the functional requirements including the user information sharing, personalized recommendation, and multi-display terminals. Especially, personalized recommendation is not only applied on the video site, but also even applied to the major online shopping platforms such as Taobao, Jingdong, etc. In order to allow users to have a better experience, and strive to be accurate and efficient on personalized recommendation [1].

Internet is based on HTTP protocol at the beginning. The HTML document is located through URL. Thus mutual communication is allowed between the various groups within the scope of the world. Web services are often used to solve problems between different interrelated platforms. But it seems to ignore the distributed programming platform features which support the Web success to improve the technology. The RESTful architecture allows Web service back up the essence of Web. With the rapid development of the mobile Internet and cloud computing today, REST style has dominated the field of Web development. More and more Web services begin to use design and implementation of REST style such as this Internet giant like Google, Amazon. The report notes that 95% of businesses are using the REST service from lightweight in Amazon bookstore. At the same time, the service system needs some individual business processes in order to attract users. A variety of complex services based on HTTP protocol survive on the Internet. In order to continue, these popular services not only attract basic groups of users by the unique function, but also promises the ability to open interfaces to other service systems. In order to achieve a personalized treatment pluggable module replacement, we need to decouple service

^a Corresponding author: songjinbao@cuc.edu.cn

processing module and personalized processing module [2].

2 Development environment

2.1 Maven

In order to be able to build a large system to facilitate the release of project information and a shared JAR in multiple projects, the traditional architecture of Ant is not used when to the build cross-platform system, but the software project management tools –Maven is used. Maven is the software project management tools of the project construction management, reporting and documents by a short description information based on the Project Object Model (POM). When you use Maven, you use a well-defined project object model to describe your project and run the plug-in target to generate the site, manage JAR files according to their need.

2.2 Navicat

In order to make database operations easier, a visual database management tool Navicat is used. Navicat is a fast, reliable database management tools to simplify the management of database and reduce cost of management system. It is designed to meet the needs of developers. Navicat is based on intuitive graphical interface. So data resources can be created, organized, accessed and shared in a simple way.

2.3 Apache Http Server 2.0

Apache HTTP Server is open-source web server which can run on most computer operating systems. It can construct Web server with a good cross-platform and security. It also provides Web server management function. The professional Apache HTTP server can be used to handle static resources of our entire system.

2.4 Tomcat 7

Throughout the program recommendation system, a lightweight web application server Tomcat is used to deal with the occasions of small and medium systems and not many concurrent access users to develop and debug JSP program. Apache server is used to respond to HTML page access request. Tomcat is actually part of an extension of the Apache server. So when Tomcat is run solely, Tomcat and Apache actually run separately as a separate process. Dynamic resources are dealt with by Tomcat. Servlet and JSP support to improve the ability of the processing of requests. When Tomcat is configured correctly, Apache serves HTML pages, and Tomcat actually runs JSP pages and Servlet.

2.5 MongoDB

In the whole MyView system, the most popular MySQL is used as the relational database to store the platform

basic information and data. It is worth mentioning that, a non-relational database MongoDB is also used mainly to store user authentication information and guidance information. The guidance information is very important for the start of the behind recommendation function. Mongo DB is attracted by much of the current IT professionals because of its flexible data storage and a good object-oriented thinking achievement. All data persistence operations needn't have to manually write SQL statements. You can directly call the method to save the user session state. The server achieves stateless, which can ensure the entire system uses a unified REST architectural style.

3 System technology selection

3.1. REST architectural style

REST-style Web services are primarily based on three basic Web technologies: XML markup language, HTTP application protocol and URI naming standard, also known as RESTfulweb services. RESTful web services and traditional services are somewhat different in interface design and use, loose coupling and expandability. But REST-style Web services have reached a simpler, more convenient and good performance [3].

REST refers to Representational State Transfer. When a well-designed Web composed by the network pages runs, the user selects the appropriate URL link is referred to state transfer. It refers to the state to advance in network applications and lead to a page, which is presented to users. The representation of REST refers to the next state in the transfer process, including the common HTML pages, XML pages, or certain data and services, these are the representations.

The most obvious REST several design criteria in Web applications are:

- All things on the network are abstracted to resource;
- Each resource has a unique resource identifier, which is the URL;
- Resource identifier will not change because of a variety of operations;
- Resource is operated through a common connector interface, which is the HTTP protocol PUT, DELETE, HEAD, GET, POST and other actions;
- All operations are stateless.

As it can be seen, URI is not only the resource name on the Internet, but also the address of the resources on Web. The statements involved in REST resources refers to resources can transfer between the client side and the server side described in the status of a specific moment [4].

3.2 RESTfulWeb service

RESTfulWeb service means to build Web services following the REST style. We mentioned earlier that the modern Web services are more and more away from the original Web development concept. The appearance of REST is to enable it to return to the resource-oriented

architecture. RESTfulWeb identification and location of resources are by URI. The operation of these resources are performed using HTTP methods (common HTTP methods DELETE, POST, GET, HEAD, and PUT). That is say, in RESTfulWeb, the place for scope information is URI path of the resource. The method information of data operation is placed in the HTTP method. From a technical point of view, RESTfulWeb service is related to basic URI and XML technology in addition to HTTP, known as three basic technologies of REST [5].

3.3 REST operational features

In the traditional Java web architecture, people often prefer to use MVC design pattern with Spring and Hibernate technology together development. The biggest feature MVC is which Model member to handle user requests and which View to display the returned data are determined by the Controller.

For REST, on the client and server assignment, it follows a client - server design specification. The service is the responsibility of offering resources and service methods of operating resources. The client only need to make page rendering in the front end according to resource data and status.

In order to enrich Web service dynamically personalized user experience, due to the characteristics of REST itself bounding the transfer of resources and data resources to the client and separated from the server separation, AjaxWeb with the most advantageous features is used in the front page rendering, which can interact with the server without needing to require a full refresh. So that it resides in the browser to execute application logic and state, to achieve partial refresh and to avoid the screen flicker.

4 System implementation.

As a RESTful Web project, MyView uses Maven build tool. Domain model is developed in stratified mode. Background service layer is abstracted to have a unified interface. Resource is identified by a URI. Service layer is divided into two layers DAO and Service is. DAO mainly located in the lower layer to package some common methods of operating entity Bean which waits for the call from the upper. Service is located in the upper layer, calls the DAO layer methods, and provides specific business logic as external manifestations of REST-style Web Service. Web front-end uses Ajax technology. The presentation layer uses Spring + MVC, using its flexible, expressive rich interface to get a good user experience.

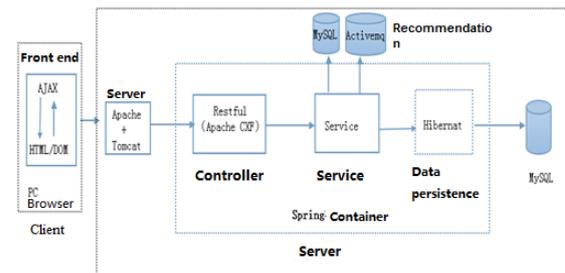


Figure 1. System structure diagram.

5 Recommendation system analysis.

After the birth of SlopeOne, the Internet publishers paid more and more attention to recommendation function as a mainstream business so as to bring more traffic. At present, the more extensive used personalized recommendation technology is primarily Collaborative Filtering Recommendation and Content -based Recommendation. In order to produce better recommendation, some mixture of different recommendation algorithms will also be used. The technology such as big data is used to optimize the algorithms. MyView mainly uses recommendation algorithm based on content to avoid the Matthew effect.

5.1 Classic recommendation algorithm

SlopeOne algorithm is essentially a kind of algorithm based on the project, but in conventional cognition is a kind of classic algorithms based on score predictions. SlopeOne makes predictions on the basis of the principle of linear regression. The predictor is in the form of $f(x) = x + b$. Free parameter b refers to the users' average deviation of two projects score. Based on the available parameters, for example, user B's score prediction of project j is represented as $r_{Bj} = r_{Bi} + (r_{Aj} - r_{Ai})$.

	Item <i>i</i>	Item <i>j</i>	
	1	3	UserA
	2	?	UserB

Figure 2. Basic SlopeOne algorithm.

Slopeone mainly computes mean deviation $\{dev\}_{jk}$ and the target project score $\{P(u)\}_{j}$.

$$dev_{jk} = \sum_{u_i \in I_{jk}} \frac{r_{ij} - r_{ik}}{card(I_{jk})} \tag{1}$$

$$P(u)_j = \frac{\sum_{k \in R_j} (dev_{jk} + u_k)}{card(R_j)} \quad (2)$$

In the above basic formula, the mean deviation [dev]_{jk} is based on the benchmark of target project i_j with other project i_k. Score [P (u)]_j is used to predict the current active users u's score for the target project i_j. I_{jk} is collection for customer evaluation of project i_j, i_k. Card (S) stands for the number of elements in the collection S. u_j is user U's score for project i_j. R_j is project collection that is evaluated with project i_j at the same time. For a simple recommendation system, under the premise of not requiring the accuracy of recommendation, in order to avoid ignoring the number of the user's evaluation, a weighted SlopeOne algorithm is considered to use to improve the accuracy. In weighted algorithm, the weight of deviation is represented by the number of users that evaluate project i_j and project i_k at the same time. c_{jk} stands for the weight of deviation. c_{jk} = card (I_{jk}). The final prediction result is as follows:

$$P^w(u)_j = \frac{\sum_{k \in R_j} (dev_{jk} + u_k) c_{jk}}{\sum_{k \in R_j} c_{jk}} \quad (3)$$

5.2 Similarity calculation formula

In various kinds of similarity calculation methods with different benefits, the current common used methods are based on the vector. The calculation size stands for the similarity. In the MyView system, the size is used to calculate the similarity based on the Cosine Similarity.

Cosine Similarity is widely used in the document vector similarity calculation. Formula is shown as follows:

$$T(x, y) = \frac{xy}{\|x\|^2 \times \|y\|^2} = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2} \sqrt{\sum y_i^2}} \quad (4)$$

T (x, y) function measures similarity by computing two vectors x, y angle Cosine. The similarity relation between vector and the change of internal elements are reflected. The calculated distance between the vectors is more to distinguish the differences from the direction. Cosine Similarity is used to distinguish user interest similarity and difference, as it is not sensitive to the absolute value [6].

5.3 Recommendation algorithm based on the content

In the process of recommendation, each user needs to create a user model. In order to be intuitive to recommend explanation and change the weights of different interest items according to the need, programs and users can be processed for domains. Firstly, the user preference information and programming model can be divided into n different areas and each domain corresponds to an interest item. It can be concluded that the user preference model.

$$\text{User_Profile} = \{(k_1, w_1), (k_2, w_2), \dots, (k_n, w_n)\}. \quad (5)$$

$$\text{Film_Model} = \{(l_1, v_1), (l_2, v_2), \dots, (l_n, v_n)\}. \quad (6)$$

Then score between user and program can be calculated according to the Cosine similarity formula. When programs are recommended, system will sort the programs according to the score size in order to achieve the goal of recommending program. Sim (l, k) is for the similarity between the user and i domain of the program. After the cosine similarity formula, it is available:

$$sim(l_i, k_i) = \frac{\sum_{p=1}^j w_p' v_p'}{\sqrt{\sum_{p=1}^j w_p'^2 \cdot \sum_{p=1}^j v_p'^2}} \quad (7)$$

w_p' refers to the weight of the i domain the p word in user preference model. v_p' refers to the weight of the i domain the p word in program model. For the sake of accuracy, we can consider this kind of circumstance: without a word, then the weight of the word is 0.

For the film show, relative to the user, under the recommendation algorithm based on content, the score is Score_{cb}, then the calculation method of Score_{cb} represented by formula (8).

$$\text{Score}_{cb} = \frac{\sum_{i=1}^n sim(l_i, k_i) w_i v_i}{\sqrt{\sum_{i=1}^n (w_i v_i)^2}} \quad (8)$$

5.4 Lucene index mechanism

Our core algorithm module mainly realize the similarity algorithm based on content, by introducing the heat factor to get the recommended list. But in the core algorithm module, we need the data retrieval and filtering, using Lucene tools library to build indexes and create a query.

In recent years, Lucene has become a popular Java open source information retrieval tools libraries. It is often used to increase the index and search functions. Lucene provides simple and easy API to use index and search. A Lucene database is created in the database. The data unit is converted to inverted index through the database.

MyView establishes Lucene database in Navicat Lite. The data structure is usually stored in the file system as a set of index files. Lucene libraries join the search function to MyView recommendation engines to build indexes and create a query. Those programs that do not conform to the user interests will be screened out, greatly reducing the subsequent data calculation amount.

```
public void index_0(IndexWriter writer) throws ClassNotFoundException, SQLException, IOException
{
    Class.forName("com.mysql.jdbc.Driver");
    Connection connection = DriverManager.getConnection("jdbc:mysql://localhost:3306/lucene", "root", "123456");

    PreparedStatement stmt = connection.prepareStatement("select * from OCCUPATION");
    stmt.execute();
    ResultSet rs = stmt.getResultSet();

    while(rs.next())
    {
        Document document = new Document();
        Field idField = new IntField("ID", rs.getInt("ID"), Store.YES);
        document.add(idField);
        Field occupation_name = new TextField("OCCUPATION_NAME", String.valueOf(rs.getString("OCCUPATION_NAME")), Store.NO);
        Field work_address1 = new TextField("WORK_ADDRESS1", String.valueOf(rs.getString("WORK_ADDRESS")), Store.NO);
        Field work_address2 = new TextField("WORK_ADDRESS2", String.valueOf(rs.getString("WORK_ADDRESS")), Store.NO);
        Field company_field = new TextField("COMPANY_FIELD", String.valueOf(rs.getString("COMPANY_FIELD")), Store.NO);

        occupation_name.setBoost(1F);
        document.add(occupation_name);
        work_address1.setBoost(1F);
        document.add(work_address1);
        work_address2.setBoost(1F);
        document.add(work_address2);
        company_field.setBoost(0.8F);
        document.add(company_field);
    }
}
```

Figure 3. The main program for program data index.

6 Summary

With the popularity of digital TV, TV programs have been on the increase no matter in both the number and species, which brought many choice to the users. Although the digital TV has increased largely in the selectivity, it has become a fussy process that users search for programs which they are interested in. people gradually accept and rely on using the system such as video website to playback TV shows on demand, because the way of watching TV on television is accepted by the development of all kinds of the video website. So there is need to have an efficient program recommendation system to solve the problem that is "information overload" for users. It can not only help users to get the program which they require, but also bring convenience to people's life.

The program recommendation system named MyView is planned and designed, aimed to providing an efficient information platform. It mainly serves for audience and the users who publish video on Internet. Our users are called the Viewer in this system. Users can create your own personal resume list, and program list, searching programs, publishing program and reviews, managing programs so on. The system also involves intelligent recommendation. The information guide will trigger the recommendation engine after users registering information, the engine will accord to the data in the guide information to make the personalized program recommendation.

References

1. J.B. Song, The characteristics and application of recommendation system, Journal of China digital TV, 46,2009, 10.
2. C. Chen, K. Y. Qi, J. B. Chen. Podcast show recommendation based on user clustering, Journal of computer applications and software, 2009 (01):7-10.
3. H.Tian, Employment system design and implementation based on REST style, Communication university of China, 2014.
4. J. Zhang, The research and implementation of Web services based on REST architecture, Beijing university of posts and telecommunications, 2011.

5. X. F. Cheng, The research and design of Web Services based on the REST architecture, Wuhan university of technology, 2010.
6. J. T. Ru, Christine. the position recommendation optimization research based on information retrieval, Communication university of China, 2014.