Collaborating Filtering Community Image Recommendation System Based on Scene

Tao HE, Ning ZHU, Gui-Xi XIONG, Zai-Rang ZHAO

Academy of Computer Science and Technology, Beihang university, Beijing, China

15801146625@163.com, zhuning730@qq.com, xionguixi@buaa.edu.cn, zhaozairang@foxmail.com

Abstract—With the advancement of smart city, the development of intelligent mobile terminal and wireless network, the traditional text information service no longer meet the needs of the community residents, community image service appeared as a new media service. "There are pictures of the truth" has become a community residents to understand and master the new dynamic community, image information service has become a new information service. However, there are two major problems in image information service. Firstly, the underlying eigenvalues extracted by current image feature extraction techniques are difficult for users to understand, and there is a semantic gap between the image content itself and the user's understanding; secondly, in community life of the image data increasing quickly, it is difficult to find their own interested image data. Aiming at the two problems, this paper proposes a unified image semantic scene model to express the image content. On this basis, a collaborative filtering recommendation model of fusion scene semantics is proposed. In the recommendation model, a comprehensiveness and accuracy user interest model is proposed to improve the recommendation quality. The results of the present study have achieved good results in the pilot cities of Wenzhou and Yan'an, and it is applied normally.

1 Introduction

Internet information technology to guide the city to the direction of intelligent development, with the wisdom of the city to promote, urban transport, environment, security and other aspects to improve and enhance. Smart City's goal is to rely on information technology infrastructure, the use of Internet of things technology, data technology, mobile Internet technology research and application, so that urban residents to fully enjoy the wisdom of urban life brought about by information technology. As the basic unit of intelligent urban development, community management and service occupy a very important position, and directly relate to the daily life experience of urban residents. How to manage the community and how to provide community residents with information services is an important issue for community development.

During the construction of the intelligent community, the community managers and community residents have produced rich information data reflecting the community status through the camera equipment, the mobile terminal equipment and the community website. The information is mainly divided into three categories: text, image and video. text is simple, image is not likely to quickly access content in a short time; image expression information intuitive,

visual impact, to meet people in a short time to master the information needs; video content expression information is full of vivid, but takes up the storage space big, and often needs to look from the beginning to the end, the time consuming, therefore the image data auxiliary with the text data becomes the community resident most willing to accept the information expression form. However, the content of the image data is difficult to access, and there is a semantic gap, the image association is low, and the low efficiency of image retrieval leads to the inability of the community residents to acquire the community image rapidly, so that they can not grasp or understand the community dynamics. With the popular web2.0, similar to Flickr, nice, in, ink weather such as picture sharing site has been popular, people began to use social labels to image tagging, image tag is between the image content and The granularity feature of image vision can solve the problem of "semantic gap" of image, and organize and manage the image with label. In the application of intelligent community, residents can use the mobile phone side of the community entities and community events anytime, anywhere to take pictures of the upload at the same time to help tag text information to describe the community image data, these community image data in the source, structure and type of decentralized But they often come with rich tag information and associated text information, but also have relevance in a certain time

and space, which provides a new idea for image management and service.

Information overload era, passive information retrieval is often replaced by active recommendation, the wisdom of the community building process is also true. Residents in the access to community information, image data retrieval is often unnecessary, the residents will not search for images to the picture, but the need for residents based on individual preferences or the popularity of the event, the initiative to reflect community image to residents to achieve community image data recommendation become an important business community services.

2 Community image scene semantic model

The semantic data of community image is multi-level. If all the semantic features of community image data are extracted at different levels, there are problems in the implementation of technology and application scenarios. Firstly, the features of community images are diverse, and different image features have different feature extraction algorithms, and some features of community images are high-dimensional, such as SIFT features, and require additional techniques to reduce dimensionality, resulting in community image semantic data access difficulties. Second, for the users' image acquisition needs, all of these features can not be one to one with the needs of users, it is difficult in the community to use the actual scene. Therefore, combined with community image scene elements and image tags, the community image scene-based semantic model is constructed to provide the foundation for community image acquisition service.

2.1 Community Image Acquisition

In the intelligent community application, the community image data comes from two parts: the user upload pictures and Internet crawl pictures. The user perceives the community image through the mobile terminal, and generates the image data with the metadata description. The user uploads the picture data often has the controllable characteristic of the information format information, which can meet the requirements of community management and service. Network crawl technology, access to all types of Internet image data and related text description data, usually on the Internet to crawl the image type and information format is inconsistent, the lack of many other aspects of information, such as the lack of access to effective user information, picture utilization low.

2.2 Community Image Scene Abstract Model

Community events reflect the status of the community, in time, space and subject have relevance, community image records of community events, is a more intuitive reflection of community status. Community image data reflects the state

of the community, more intuitive than the text description credible, the same picture can also reflect the state of the community time information, spatial information, event information and even the nature of the event information. In this respect, community image data and community event scene information have great analogy, in order to make community residents a better understanding and master community dynamics, community image data and scene information associated with the use of community image is critical The meaning. Community image data scene provides a more simple and effective image content description method and community state expression mode, while the value of community image data is utilized, community residents and community managers can understand and master community dynamics.

$$S = \sum_{time=0} \sum_{place=0} I(time, place, theme, event, person, quality) \quad (1)$$

I representing an image data resource; $time$ representing a community event occurrence period label; $place$ representing an area label; $theme$ representing a subject representative of an image, selected by an uploader; $event$ representing an uploader or person involved in the image; $person$ representing an event reflected by the current picture; $quality$ representing the nature of the incident, that is, the severity of the event or the extent of the use of 5-point description.

Community Image Scene Abstract Model as follows, shows the image scene represent the community events.

Figure 1. Community image scene abstract model.

2.3 Community Image Scene Semantic Model

Based on the source of community image data, two kinds of data sources are studied. Under the support of tag semantic library and the model of community image abstraction, interaction between scene user and system and Internet image data based on link relation are used to crawl , To achieve tag recommendation, semantic annotation and scene semantic mapping, the formation of community image scene semantic model. The realization of the community image scene semantic model completes the document representation of the community image, and solves the difficult problem of the image representation. It is very

important for the organization, management, retrieval and recommendation of the image.

Community Image Scene Semantic Model as follows:

Figure 2. Community image scene semantic model.

The semantic model of community image scene includes the image tag recommendation, image semantic annotation and scene semantic mapping, and the image semantics and user information are stored in the database for the scene user uploaded image and the Internet crawling image.

2.3.1 Image Tag Recommendation

In the mass annotation mode, the tag is the link between the image data and the user's link, the label can express the image content, can also represent the user's preference type. In order to accurately describe the image content and express the user's preference, the system is provided with the filter of the temporal and spatial theme scene elements. In order to provide a more accurate description of the image content and the expression of the user's preference, the system provides the content of the markup content to the same community events, different people's concerns, Label Recommendation. After the tagged image data is tagged, the tag vector set of tagged image resource is converted to a synonym word group, and the tags in the tag vector of each image resource are found corresponding to "Class, as a feature, a number of feature items constitute the feature space, statistical feature space set for each feature in the frequency of the emergence of the highest frequency of five features as the recommended candidate feature, and then statistical features of the highest candidate The five candidate label sets corresponding to the five candidate feature items are given.

When the scene user uploads a picture, it acquires the time-space scene element and the theme scene element, sets the scene factor filter condition, sets the time dimension to 30 minutes, the space dimension is 500m, the theme equals the user- The set of label vectors for all the tagged image resources can be represented in the following form.

$$S = \sum_{t=-15}^{t=15} \sum_{d=0}^{d=500} \sum_{i=1}^n I_i \quad (2)$$

S represents a collection of all image resources satisfying the scene filter condition. It is a set of label vectors. Each

image resource in the set is represented by a series of labels, and the representation is as follows.

$$I_i = (tag_{i1}, tag_{i2}, \dots, tag_{ij}, \dots, tag_{im}) \quad (3)$$

A tag-based vector space model (VSM) is proposed for the tagged image resource set.

1. Select the feature

For the labeled resource set S , we need to fetch the feature item corresponding to each label by image resource, take the image resource as an example to select the feature item, and for every label of the image resource. The characteristic items are shown below.

$$\begin{aligned} tag_{i1} &\Rightarrow X_{i1} \\ tag_{i2} &\Rightarrow X_{i2} \\ tag_{ij} &\Rightarrow X_{ij} \\ tag_{im} &\Rightarrow X_{im} \end{aligned} \quad (4)$$

2. Construct feature space

The construction of feature space is to transform the label vector of tagged image resource, to associate the tag with the synonym class information, and to count the number of tags contained in different feature items as the feature weight of each feature item. After the feature set is processed according to the selected feature, the feature set and the corresponding image label will be obtained.

$$\begin{aligned} tag_1, tag_2, \dots, tag_x &\Rightarrow X \\ tag_1, tag_2, \dots, tag_y &\Rightarrow Y \\ tag_1, tag_2, \dots, tag_z &\Rightarrow Z \end{aligned} \quad (5)$$

The resulting feature space is as follows

$$K = \{(X, x), (Y, y), \dots, (Z, z)\} \quad (6)$$

Among them, X, Y, Z is the characteristic item, x, y, z , is the characteristic item weight.

3. Candidate feature items

The feature space model K is sorted according to the weight, and the five feature items with the highest weight are selected as candidate feature items

$$T = \begin{cases} X_1 = x_{11}, x_{21}, \dots, x_{n1} \\ X_2 = x_{21}, x_{22}, \dots, x_{2n} \\ X_3 = x_{31}, x_{32}, \dots, x_{3n} \\ X_4 = x_{41}, x_{42}, \dots, x_{4n} \\ X_5 = x_{51}, x_{52}, \dots, x_{5n} \end{cases} \quad (7)$$

4. Candidate label set

For the candidate feature items, the labels in each feature item are statistically calculated, and a label with the highest number of occurrences in each feature item is selected to constitute the candidate label set, as shown below.

$$RecommendTag = \{tag_1, tag_2, tag_3, tag_4, tag_5\} \quad (8)$$

2.3.2 Image Semantic Annotation

This paper uses the mass annotation method to annotate the community image semantics. In the event of a community event, the user will select a system-recommended image tag and label the community image according to his preference and willingness to form the label semantics of the image data. The label semantics of the image data can be described as follows form.

$$I = (tag_1, tag_2, tag_3, \dots, tag_n) \quad (9)$$

2.3.3 Scene Semantic Mapping

The time field of the image data is usually presented in the form of year, month, day, minute, and second. The location field is presented as a GPS value. These specific time values and GPS values are the original scene semantics. There is no semantic feature. The user specific time for the image data Value and GPS value is not concerned, so the actual application of the project, the need to map the specific time value into the morning, noon, afternoon, evening and other semantic, the specific GPS values into a certain district, a certain road, Grid and so on. However, the semantic mapping rules of different scene elements under different topics are inconsistent, and the different vocabulary words are extended and classified, and the subject vocabulary is added to the different semantic map. The topic time vocabulary that should be added is breakfast, lunch and afternoon tea. The time vocabulary of the subject should be added, such as morning peak, day peak, afternoon peak and evening peak. , Dinner, supper and so on.

3 Community Image Recommendation model

The traditional collaborative filtering recommendation technique is based on the user-item scoring matrix to get the user information, and the other feedback information of the user, such as the user's feedback information, the user's feedback information, and the user's feedback information. In addition, the collaborative filtering recommendation algorithm is the most mature recommendation algorithm applied to complex objects such as images, Not taking full account of the user's interest will change over time and the user's location of the event is relatively fixed and fixed image data to produce these two issues, resulting in low efficiency of the recommendation. Aiming at the problems in the practical application of the intelligent city, this paper proposes a collaborative filtering image recommendation model which integrates the scene semantics, introduces the label semantics and the scene semantics of the image, filters out some image data which does not meet the condition, and combines the time scene semantic space scene Semantics to express the interest of the target user in a certain spatial region more accurately, and then use the collaborative

filtering recommendation algorithm for image recommendation.

Figure 3. Community image recommendation model.

The synergistic filtering model of scene fusion semantics includes scene information filtering mechanism, user interest modeling, similar user acquisition and recommendation result generation. Scene information filtering mechanism mainly refers to the use of the current time and current location to filter away from here and now far image data, the remaining image data to better reflect the user's interest here and now preferences. User interest modeling is different from traditional scoring data modeling. In this paper, user interest preference based on scene information filtering mechanism is more accurate, and the user interest is represented by label and label weight information. Image recommendations The next steps are critical. Similar user acquisition refers to the similarity calculation between the user and the target user in the user space using the improved similarity calculation method, and selects the first k users to constitute the set of neighbors of the target user based on the similarity value sorting. Recommendation result generation refers to the evaluation of the images based on the uploaded records of the neighboring users of the target user, and generates the recommendation values, thereby recommending the ranking and generating the recommendation result.

3.1 Scene filtering mechanism

Scene information includes temporal scene elements, spatial scene elements and so on. Both time scene elements and space scene elements are related to user's interests and preferences. User's interests will change with time. The same geographical elements also reflect the users' Of interest. Everyone's perception of things in a short time does not change too much, has a certain stability, and similar geographical location of the personal preferences of things similar, so you can according to a certain period of time within a certain location of the user's preferences to predict the user's preferences. This is the core idea of collaborative

filtering. It is very important to perform scene-based resource filtering of image data resources before user personalized modeling and generating recommended images, in order to reduce the amount of computation and to express the interest in the most recent spatial range of user's most recent time, to improve the quality of image recommendation.

In the actual scene, the behaviors of similar users at different time and in different places affect the formation of user interest model. Adding scene information filtering mechanism can improve the accuracy of recommendation algorithm.

3.2 User Interest Model

In the real life, the human forgetting curves exist objectively, and the user's preferences are changing constantly. The closer to the present time point, the closer the image is to the distance. Nowadays, the user interest model based on scene filtering is more realistic. Can change on behalf of the user's preferences. A user's interest in a particular area at this point in time is much greater than that of other similar users within the same time frame. User interest model needs to take into account the user's recent behavior changes, to reflect the user's recent hobbies. The user's interest vector is generated by using the tag information when the user uploads the image, not just the user interest model formed by monotonous user rating data.

The tag data at the time of image upload also represents the user's interest preference. The frequency and type of the tag use the user's hobby distribution to a great extent. The user personalized modeling is performed using the tag data generated by the user uploading the image data. First of all, for a target user who needs to be recommended, we need to obtain the time and space scene information, get the tagged data set of the tagged image in the scenario, and count the frequency of each tag. The tag data set represents the user's interest Type, label frequency represents the degree of user preferences. So the user interest model can be expressed as the following formula.

$$U = \{(tag_1, w_1), (tag_2, w_2), \dots, (tag_i, w_i), \dots, (tag_n, w_n)\} \quad (10)$$

The order of the labels is sorted according to the weight, that is, the more you can reflect the interests of the user preferences, on the contrary, the more the ability to reflect the user's interest in the weaker.

Figure 4. User interest model.

Compared with the traditional user interest model, the label information reflects the comprehensive user preference, and the tag weight information quantizes the user preference, so that the user's interest preference degree is expressed accurately.

3.3 Similar User Acquisition

The similarity between the two users is the key step for the users to get. Considering the difference of different users, some users have high scores on the project and some users are generally low on the project. Another improved cosine similarity calculation method is shown below.

$$\text{sim}(u_i, u_j) = \cos(u_i, u_j) = \frac{\sum_{k=1}^n (r_{i,k} - \bar{r}_i) \times (r_{j,k} - \bar{r}_j)}{\sqrt{\sum_{k=1}^n (r_{i,k} - \bar{r}_i)^2} \sqrt{\sum_{k=1}^n (r_{j,k} - \bar{r}_j)^2}} \quad (11)$$

3.4 Sorting produces recommendations

The ranking is based on the likelihood that the target user will like a particular item he did not like, called the predictive score. The average weighting strategy is used to calculate the predicted score. Based on the historical score of the target user and the nearest user, The average weighting strategy is composed of two parts: the average score and the similarity value of the user, and the calculation method is as follows.

$$P_{u,i} = \bar{r}_u + \frac{\sum_{s=1}^k \text{sim}(u, u_s) (r_{s,i} - \bar{r}_s)}{\sum_{s=1}^k \text{sim}(u, u_s)} \quad (12)$$

Results

Because it is the actual project, the end result to the mobile side of the app and the Web version of the recommendation, in the wisdom of Yan'an and wisdom of Wenzhou have achieved effective results. The following are the results of Wenzhou.

Figure 5. User upload picture.

Figure 6. Community image shows in the map and its semantic.

Figure 7. Community images shows the Wenzhou traffic dynamically change with the time and place.

The following are the results of Yan'an.

Figure 8. Yan'an picture wall by topic.

Figure 9. Map - based tourism image recommendation.

Conclusions

This paper mainly solves the following two problems, for the image data content is difficult to express, and the underlying eigenvalue for the user is difficult to understand and so on. Through the analysis of the community image scene, abstract the image elements, according to the specific scene to draw the image scene semantic formation process, constructs the community image scene semantic model. Based on the semantic model of community image scene, aiming at the problem that the interest of users in the traditional collaborative filtering changed with the passage of time, the user interest model of the fusion scene was proposed, and the good recommendation effect was achieved.

Acknowledgment

This research derived from National Natural Science Foundation of China whose name is : Large Data Computing Theory and Software Technology (61421003).

References

- [1] J. Eakins, M. Graham, Content-based image retrieval, Technical Report, University of Northumbria at Newcastle, 1999.
- [2] Bhusari P, Gupta R, Sinahal A. Personalized Image Search from Photo Sharing Websites Using Ranking Based Tensor Factorization Model (RMTF)[J]. International Journal, 2013, 3(8).
- [3] Sharma S, Kaur D. Location based context aware recommender system through user defined rules[C]//Computing, Communication & Automation (ICCCA), 2015 International Conference on. IEEE, 2015: 257-261.
- [4] Champiri Z D, Shahamiri S R, Salim S S B. A systematic review of scholar context-aware recommender systems[J]. Expert Systems with Applications, 2015, 42(3): 1743-1758.
- [5] Zheng N, Li Q. A recommender system based on tag and time information for social tagging systems[J]. Expert Systems with Applications, 2011, 38(4): 4575-4587.
- [6] Kim H N, Alkhalidi A, El Saddik A, et al. Collaborative user modeling with user-generated tags for social recommender systems[J]. Expert Systems with Applications, 2011, 38(7): 8488-8496.
- [7] Tamura H, Yokoya N. Image database systems: A survey[J]. Pattern recognition, 1984, 17(1): 29-43.
- [8] Szummer M, Picard R W. Indoor-outdoor image classification[C]//Content-Based Access of Image and Video Database, 1998. Proceedings., 1998 IEEE International Workshop on. IEEE, 1998: 42-51.
- [9] Mori Y, Takahashi H, Oka R. Image-to-word transformation based on dividing and vector quantizing images with words[C]//First International Workshop on Multimedia Intelligent Storage and Retrieval Management. 1999: 1-9.
- [10] Dye J. Folksonomy: A game of high-tech (and high-stakes) tag[J]. EContent, 2006, 29(3).
- [11] Sun A, Bhowmick S S, Nguyen N, et al. Tag-based social image retrieval: An empirical evaluation[J]. Journal of the American Society for Information Science and Technology, 2011, 62(12): 2364-2381.
- [12] Vadivu P S, Sumathy P, Vadivel A. Image Retrieval from WWW using Attributes in HTML TAGs[J]. Procedia Technology, 2012, 6: 509-516.
- [13] Kantor P B, Rokach L, Ricci F, et al. Recommender systems handbook[M]. Springer, 2011.
- [14] Pazzani M J, Billsus D. Content-based recommendation systems[M]//The adaptive web. Springer Berlin Heidelberg, 2007: 325-341.