SDIA: A dynamic situation driven information fusion algorithm for cloud environment

Guo Shuhang *, Wang Tong, Wang Jian*

* Central University of Finance and Economics & School of Information, China

Abstract. Information fusion is an important issue in information integration domain. In order to form an extensive information fusion technology under the complex and diverse situations, a new information fusion algorithm is proposed. Firstly, a fuzzy evaluation model of tag utility was proposed that can be used to count the tag entropy. Secondly, a ubiquitous situation tag tree model is proposed to define multidimensional structure of information situation. Thirdly, the similarity matching between the situation models is classified into three types: the tree inclusion, the tree embedding, and the tree compatibility. Next, in order to reduce the time complexity of the tree compatible matching algorithm, a fast and ordered tree matching algorithm is proposed based on the node entropy, which is used to support the information fusion by ubiquitous situation. Since the algorithm revolve from the graph theory of disordered tree matching algorithm, it can improve the information fusion present recall rate and precision rate in the situation. The information fusion algorithm is compared with the star and the random tree matching algorithm, and the difference between the three algorithms is analyzed in the view of isomorphism, which proves the innovation and applicability of the algorithm.

1 Introduction
The traditional research of information fusion is that building the user association rules library based on the analysis results of the user behaviours to and classify users according to the associated rules, and then recommending users the information which users might have interests on according to the association rules library.

Aiming at the situation feature of the information, we propose an information integration presentation method which applies ubiquitous situation. After the conversion the situational features of each dimension from sibling relationship further becomes directed sequence, this can greatly improve the matching efficiency. The proposed method of situations presented can be used as an effective method of information visualization.

2 Related work
The present perceptual structure which is centred with the users’ behaviour preference, often does not apply to the user’s specific situation, such as position, user identity, entity attributes and condition [1]. One of the important research directions of context-awareness is how to effectively utilize the accessed situation information and transfer it into rational representation [2]. A proper context model is the basis of situation information representation and storage management.

According to its time characteristics, the situation information can be divided into static situation and dynamic situation [3]. The situation information modelling can also solve the problem of information sharing, eliminate the inconsistency of information and process the imperfection of information [4]. In terms of information presentation, exploring the present technology with high
universality and high accuracy has become the core theory issue [5]. The traditional retrieval techniques such as basing on keyword matching and full text search, failed to fully embody that situation features play an important role in information fusion present. The current research mainly focuses on star alignment algorithm, which is a fast algorithm for solving the multiple sequence alignment problem [6]. Due to the time complexity of star alignment algorithm is square level, and cannot meet the actual need, it must solve to begin from the perspective of its multi-dimensional structured to improve and put forward the multidimensional star situation sequence alignment algorithm based on tree.

3 Situation tag and fuzzy evaluation

Actual application analysis shows that the current tags belong to plane and static classification system. Its main function is to achieve multi-dimensional browsing and searching of tag. This method has not yet reflected the sequence of tag utility, and then brings the tag system of casualness, chaotic attribution and other issues. In addition, tag utility is time-efficient and dynamic. This paper makes use of fuzzy comprehensive evaluation method to establish music tag utility model. Considering massive factors associated with the music tag utility, we adopt two-level fuzzy synthesis evaluation to measure tag utility. The first layer of indicators consist of the capacity, click and duration three elements. The second layer is a kind of refinement and concretization of the first layer, and they are all indicators can be computed based on the access record. For $TAG\rho$, we build the indicator system as follows.

For example, in the field of Internet music information resources, the data is below in Table 1.
Table 1. Underlying data of tag utility measurement

| Tag Name   | A1 | A2 | A3 | B1   | B2   | B3   | C1   | C2   | C3   |
|------------|----|----|----|------|------|------|------|------|------|
| World cup  | 9.7| 113| 4  | 34842| 8478 | 1995 | 8250.8| 44092| 750.7|
| Oldie      | 58.4| 1000| 96 | 71173| 4150 | 1692 | 24244.4| 8799 | 201.3|
| Travel     | 54.4| 848| 32 | 68959| 2769 | 967  | 15068.9| 22990| 548.7|
| Japan      | 37.5| 1000| 75 | 54853| 3260 | 1293 | 18750.7| 16694| 607.3|
| Europe     | 65.9| 1000| 66 | 96377| 6545 | 748  | 43114.9| 41003| 547.9|
| Light Music| 64.1| 1000| 71 | 40635| 5260 | 1403 | 33722.9| 24414| 405.4|
| Chinese style | 32.4| 519| 28 | 62674| 4165 | 894  | 13474.1| 13183| 474.5|
| Network Song| 57.8| 841| 39 | 30143| 6267 | 1481 | 36245.8| 16871| 293.1|
| Teana Music | 15.4| 344| 10 | 77614| 5948 | 1403| 33722.9| 13183| 474.5|
| Love Song  | 67.5| 1000| 83 | 32958| 9255 | 1622 | 62441.9| 45980| 216.2|
| Madden     | 27.4| 376| 18 | 55536| 3572 | 1176 | 9773.0| 11123| 575.6|
| Nursery Song | 64.6| 1000| 70 | 72917| 7815 | 625  | 50474.0| 44119| 212.9|
| The 1990's | 13.6| 145| 4  | 90182| 7784 | 843  | 10583.4| 31087| 345.8|
| Classical  | 60.6| 1000| 7  | 96337| 4424 | 1870 | 26796.4| 29744| 531.7|
| Popular    | 68.2| 1000| 96 | 59249| 6110 | 728  | 41647.0| 11523| 642.1|
| Rock       | 67.3| 1000| 44 | 23436| 1682 | 1308 | 11323.7| 32993| 639.3|
| Military Song | 53.9| 821| 40 | 24635| 1485 | 743  | 7998.2| 8418  | 138.8|
| Hip-hop    | 47.0| 1000| 80 | 33176| 5777 | 617  | 21764.9| 20793| 430.5|
| Country Song | 49.9| 1000| 77 | 22476| 8013 | 1164 | 40001.6| 43387| 394.9|
| Famous     | 34.9| 513| 19 | 33176| 9407 | 1739 | 32836.2| 28096| 402.5|
| After 70s  | 41.7| 1000| 24 | 85665| 2169 | 506  | 9040.4| 23081| 300.4|
| After 80s  | 48.5| 1000| 49 | 32885| 6119 | 1663 | 29667.2| 15796| 446.1|
| After 90s  | 67.3| 1000| 76 | 83286| 5224 | 1943 | 35180.6| 13276| 448.6|
| Graduate   | 16.4| 200| 5  | 20709| 1107 | 660  | 1811.9| 10549| 669.2|
| Romantic   | 67.8| 1000| 45 | 31917| 4034 | 1726 | 27352.7| 15798| 653.6|

According to the above tag utility fuzzy evaluation model, getting tag utility values and their rankings are as follows.

![Figure 2. Tag utility values and their rankings](image)

Calculation results are basically consistent with the actual situation. This fully shows the fuzzy comprehensive evaluation system is a scientific and practical model.

4 Information Fusion Algorithm Based on Ubiquitous Situation

In order to make situation features play a better role of information fusion based on information fusion, we present a situation-based tree similarity matching algorithm of information fusion.
4.1. Model design of ubiquitous situations

The information contained in each situation dimension is diverse, and if using the same approach to all the situational elements, so in the process of situation recognition access and situation tree similarity evaluation. Situation hierarchical organization method makes information situation description has good scalability and adaptability. First line of text or heading.

![Ubiquitous Situation Tag Tree Model Example](image)

Situation dimensions is a classification tags collection integrated by atom situation tags from different dimension. In this article, situation dimensions include network situation, commercial situation, user situation, time situation, physical information, social situation, traffic situation. Each classified dimension contains unique properties of atoms, different classifications are complementary, and that is, atomic situation property between two situation dimensions is not overlapping and crossing, so as not to cause confusion of tag data.

4.2. Situation tree’s comparison and the ordering design

The similarity measure problem in the situation translates into similarity calculation problem between the trees from graph theory’s view. For a given ubiquitous situation tree Q, the tree model in assessment information of situation database with its similarity in the story, which called the Q is main situations tree and comparative object T is customer situations tree. The existing disordered tree matching model can be divided into three classes; tree inclusion, tree embedded and tree compatible. Tree embedded refers to Q is a sub-tree of T; Tree inclusion refers to T is a sub-tree of Q; Tree compatibility refers to the Q and T have the same sub-tree. In the following figure 3, from up to down said tree embedded, tree inclusion and tree compatible.
The above three trees matching model is divided into three grades in matching situation to the subject situation tree and the object situation tree. Tree compatible improved above the definition basis of the tree embedding and the tree contain have higher relaxation ability, which makes the matching model based on tree compatible have higher precision and recall, at the same time, making the price of the algorithm in calculating tree matching’s time complexity limited in polynomial level and ensure the situation tree similarity matching efficiency.

The purpose of this study is to measure the tag utility and its drifting from the underlying data of music tag. In order to achieve this goal, firstly building a hierarchical music tag system. Taking Baidu music (http://music.baidu.com) for case analysis and research, to build a multi-perspective and hierarchical structure. Building the tag utility model from capacity, click and visit three aspects. We chose fuzzy comprehensive evaluation model to measure the tag utility and use the weights constructed by triangular fuzzy numbers, rather than the traditional AHP, to calculate the weight of each layer index weight, which makes the results more in line with people's awareness.

4.3. Fast matching method for situation tree models

The following is given in this article the cluster partition solution of the information situation. For each situation information ordered tree clustering center $CC_i (1 \leq i \leq d)$, using calculation method and order of isomorphism degree $STIC_{Embed} \geq STIC_{Include} \geq STIC_{Compat}$ that we proposed, among that
(1 ≤ i ≤ d, gtreeToSituation$^{i}$ represents the object of the matching query of the ubiquitous situation feature in the ordered tree), the homogeneous degree of the ubiquitous feature ordered tree gtreeToSituation and cluster centers $CC_{i}$ are obtained.

5 Algorithm validation and experimental analysis
Comparing with three other ways include star comparison algorithm, disordered tree algorithm and the random information fusion, it can be found that the method presented in this paper can better conduct orderly integration of information presented according to ubiquitous Situation Tag similarity. As shown in Figure 3.

![Figure 7. Ordered tree isomorphism level result](image)

![Figure 8. Star Contrast isomorphism level result](image)

![Figure 9. Disorder tree graph isomorphism level result](image)
From the above, in order to better verify the innovation of algorithm which we proposed in this paper, comparing and verifying with the stars comparison algorithm, disordered tree graph comparison algorithm, we select the degree of isomorphism to judge. From the perspective of efficiency of the algorithm, compared to the whole similarity matching selection, its complexity of directed graph's choice analysis is much smaller. So the proposed method in this paper can quickly find the most network information that applies the ubiquitous situation. It will be seen from this that the proposed method not only has better applicability but also can increase the effectiveness and efficiency of the choice of the directed graph, improve the quality of the information integration presentation system.

**6 Conclusions and expectation**

Aiming at the situation feature of the information, we propose an information integration presentation method which applies ubiquitous situation. This method combines described hierarchical characteristics of situational features. And make full use of the information entropy of existing situational feature of each dimension, as an important parameters of situational feature and use this parameter to achieve a conversion of disordered tree to ordered tree. After the conversion the situational features of each dimension from sibling relationship further becomes directed sequence, this can greatly improve the matching efficiency. The proposed method of situations presented can be used as an effective method of situations presented. In the practical aspects, it has higher retrieval effectiveness and quality (precision and recall), in practice this method is applied to network information integration and fusion presentation, the fast presentational construction of ubiquitous multi-network in information fusion is realized.

**References**

[1] Popovici D, Desertot M, Lecomte S, Delot T. A framework for mobile and context-aware applications applied to vehicular social networks[J]. Soc Netw Anal Min 2012.3:1–12

[2] Li eiPing, Wang WuSheng, Mo Tong, ZhangZhiChao, Zhu WeiJie, Wu ZhongHai. Summary of the research on context computing[J]. Computer research and development, 2015, 52(2): 542-552.

[3] Urzaiz, Gabriel. Context-Awareness and Context-Driven Interaction[C]. Proceedings of the 7th International Conference on Ubiquitous Computing and Ambient Intelligence, UCAmI 2013, held in Guanacaste, Costa Rica, in December 2013.

[4] Zeng ZiMing, Li Xin. Personalized information recommendation based on context awareness in mobile environment[J]. Journal of information.2012.31(8):166-170.

[5] Rendle S. Context-Aware Ranking with Factorization Models[J]. Berlin: Springer-Verlag, 2010. 1-180.

[6] Adomavicius G, Tuzhilin A. Context-Aware recommender systems[J]. In: Recommender Systems Handbook. Berlin: Springer-Verlag, 2011. 217-253.