International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 5 Issue 1 ǁ January. 2017 ǁ PP. 07-15

Based on the Influence Factors in the Heterogeneous Network t-path Similarity Algorithms

HU Zhaoqi[1], ZHENG Liming[1]
[1](*College of Information Science and Technology, Jinan University, Guangzhou 510632, China*)

ABSTRACT: In view of the existing heterogeneous network largely depends on the meta-path similarity calculation method of direct link, ignores the influence factors of different entities and the time difference of different paths. The hetero-geneous network is proposed in this paper, based on the factors affecting the time path of similarity algorithm Hete-DS. In HeteRecom algorithm on the basis of the algorithm, make up for the path of only consider different types of weights, different object relations by in the path to the weight computation of time factor, strengthen the effectiveness of the path to weight. At the same time, through the influence factors of different entities, object relational matrix was established considering the difference between different nodes under the same path type links. In multiple data sets on the experimental results show that, compared with HeteRecom algorithm, Hete-DS algorithm has higher accuracy.
Keywords -heterogeneous networks; similarity; meta-path; time score; personalized weights; path weights

I. INTRODUCTION

The coming of the era of big data, in providing the user with rich information resources, it's also increasing the difficulty of a large number of information processing. Information network is a huge network structurewhich is through the relationships of objects. Through the information network of similarity measure, we can get a lot of important information.

At present most of the information network is heterogeneous, however, The based on heterogeneous network path similarity measure methodswhich have been proposed have shortcomings, such as ObjectRank [1], PopRank [2]. Although they noticed the heterogeneous relation may affect similarity, but did not take into account the composed of different types of object semantic difference of the path. PCRW [3] algorithm is a random walk model based on path, measure the similarity of nodes in the directed graph, but computing complexity is larger. At present HeteSim [4] algorithm has high esteemed, which is based on the two-way path of random walk model, and overcome the PathSim [5] algorithms which is only consider symmetry paths. HeteRecom [6] algorithm is on the basis of the original HeteSim algorithm, joined the path weight. But it only considered the weights of different types of paths, and neglected the path difference between the nodes in the same type path.

But the similarity search algorithm without considering the correlation between entities will change over time, and ignored the authority of different entities on the influence of the heterogeneous network. Such as the literature [7] [8] is pointed out that the authority of the object is different, which will lead to impact on the network, and proposes the algorithm which combines node weight and path weight. Literature [9] mentioned the time factors influencing recommend prediction. Although in this paper, it considered the time property separately, it expounded the relationships between objects have a certain degree of timeliness, that is to say, time will affect the relation network. Considering the above two factors, for the Digital Bibliography & Library Project (DBLP), the path through the node which has the higher authority of the paper or the higher author profile may represent the more important meaning; the closer time the relationship between nodes established by, the higher the similarity between nodes at the moment, and more likely to have the same activity, which has a higher similarity. Aiming at the shortcomings of the above algorithm, this paper is based on HeteRecom algorithm, proposing the based on the influence factors in the heterogeneous network time-path similarity algorithms(Hete-DS).It uses to dynamic analysis of the relationship between objects in heterogeneous information network, the improvements include: 1) introducing time weight to improve HeteRecom algorithm path weight; 2) calculating the authority of different entities to obtain different nodes of the node weights.

II. BASED ON THE META-PATH HETEROGENEOUS NETWORK RELATED CONCEPTS

In the information network, objects are regarded as nodes, the relationships between the object are regarded as the links between the nodes. When the information network of object types|A>1| or link relations

|R>1|, this information network is called a heterogeneous information network, otherwise, it is homogeneous information network. Unlike homogeneous network, heterogeneous networks can through different types of path, the paths represent the different meaning, as shown in figure 1 of DBLP.

Definition 1: In the information network, there are a link relation R from the entity type A to the entity type B, which is expressed as $A \xrightarrow{R} B$. Relationship between path is the network mode of a path S=(A,R), expressed as $A_1 \xrightarrow{R_1} A_2 \xrightarrow{R_2} ... \xrightarrow{R_l} A_{l+1}$, the length of the path P is the number of links on the relationship P.

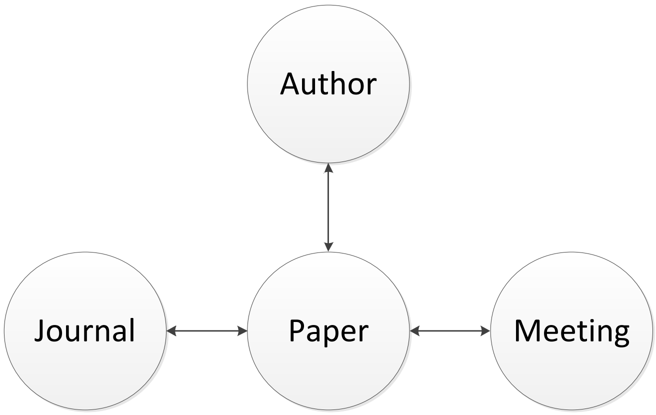

Fig.1 DBLP network topology data sets

Definition 2 (HeteRecom [6]) : meta-path of heterogeneous information network contains different semantic backgrounds, lead to the path with differences between nodes. Even though there may be thousands of path link waybetween two entities, because of the longer path, the less important [5], so all the relevant path length is less than the threshold. HeteRecom algorithm expression such as formula (1) - (2) :

$$Sim(A,B) = \sum_{i=1}^{N} w_i * HeteSim(A,B|P_i) \quad (1)$$

$$HeteSim(s,t|R_1 \circ R_2 \circ ... \circ R_l|) = \frac{1}{|O(s|R_1)||I(t|R_l)|} \quad (2)$$
$$\sum_{i=1}^{|O(s|R_1)|} \sum_{j=1}^{|I(t|R_l)|} HeteSim(O_i(s|R_1), I_j(t|R_l)|R_2 \circ ... \circ R_l)$$

$|O(s|R_1)|$refers to the number of the out-neighbor relationship links R_1 , $|I(t|R_l)|$ refers to the number of neighbors link R_l.

Importance of path (I) are related to the intensity of the path (S), and negatively related to the path length (L). The definition of important degree (I) expression such as formula (3) - (6) :

$$I(P) = f(S,l) = e^{S-l} \quad (3)$$

$$w_i = \frac{I_i}{\sum_{i=1}^{N} I_i} \quad (4)$$

$$S(P) = \prod_{i=1}^{l} S(R_i) \quad (5)$$

$$S(R_i) = (O(A|R_i) I(B|R_i))^{-\alpha} (\alpha \in [0,1]) \quad (6)$$

$O(A|R_i)$ refers to the average out-degree of node type A, $I(B|R_i)$refers to the average in-degree of node type B.

2.1 Time factors impact on meta-path

In macro point of view, time factor impact on the meta-path [10]:1) the link relationships which are etween the entity objects and their same intermediate node have a time lag. In theory, the smaller the time interval between them, the higher correlation they are; 2) There are the time intervalbetween the path building time and now, the establishment of the time interval is smaller, that the path is timeliness. As shown in figure 2, the node a and b share common neighbor c, here are the latest time of having a relationship between a and c and the latest build relationship between b and c. These two factors are just like in the literature information network, a field or topic of research background is changing over time. If the time interval is smaller, the stronger the correlation between the literature is. At the same time, if the literature before presentis smaller, which is said that literature content is novel and likely to have greater influence.

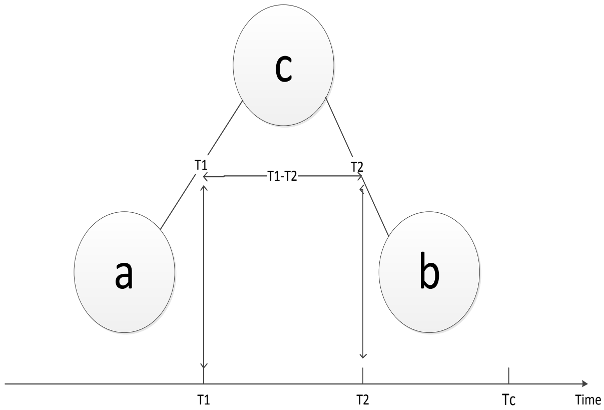

Fig.2 The diagram of node path set up time

2.2 personalized weights impact on meta-path

For the heterogeneous information networks, due to the status of each node in the network may not be the same. The higher authority the nodes have, the greater weight the nodes have, which shows the status of the node in the network also becomes more important. For example, Literature [14]showed that the identity degree of node object is referred to as the network influence, and expoundedthat the nodes had higher authority in the network,which were often more likely to affect the other node objects. Literature [15] thoughtthat without the influence of external factors, the influence of the nodes in a network is formed spontaneously through the interaction between nodes. There is a similar PageRank method to calculate the influence factors of the heterogeneous information networks in the objects.

Node influence factor is not an absolute value, is used to measure the influence relatively high and low between objects in the same type subject. In DBLP, for example, it contains three node type such as the author (A), paper (P) and meeting (C). Due to different authors have published different number of papers, at the same time, the influence factors of published papers are different. So the influence factors of the authors' authority are the number of published papers and their papers' influence factors. With similar, the different meetings also have different influences. The more number of excellent papers the meeting has, correspondingly, thehigher the importance of the meeting is. Finally, there are many factors impact the paper's influence, such as: the number of citing, the influence of author, the influence of the journal or conference which is published in and so on.

III. BASED ON THE INFLUENCE FACTORS IN THE HETEROGENEOUS NETWORK TIME-PATH SIMILARITY ALGORITHMS

Because of the existing algorithms did not take into account the above two aspects, in this paper, it proposed the algorithm which is based on HeteRecom algorithm, and took into account the time factor and weight of the node to specify the semantic information of the role and influence. Therefore, the following is divided into two parts in detail in this paper.

3.1 Time factors impact on meta-path

In this paper, the path weight of time factor is calculated by the relation path building time between two nodes. But the path intensity in the literature [10] is the same as HeteRecom path weights. By improving its computation about time factor, letthe time factors to combine with HeteRecom algorithm. The original path weight measuring factors (strength (S) of the path, the path length (L)) put in time factor (T), the time factor is regarded as one of the factor weights in the path.

1) For the whole path importance, if the last time in path between nodes keeps away from now, then the path will become thin. The formula is k^{-a}. As the literature network, the published time is closer to now, then thesis research content is novel and the path is more valuable.

2) If the establishing time interval between the two nodes and their common neighbor is smaller, then the similarity between the two nodes is higher. The formula is $(|t_1-t_2|+1)^b$. As literature network, if the published time interval between two papers is closer, then they are more likely to research under the same field.

Combined with the above consideration, it has established the time weight (TS) on links. For two time factors exist in the path, at the same time, because the link between node a and node b may be more than one, for example the authors of papers were published in the same meeting at different time, so its average time weight is defined as formula(7) :

$$TS(a,b) = \frac{1}{n}\sum_{n}\frac{k_n^{-\alpha}}{(|t_{1_n} - t_{2_n}| + 1)^{\beta}} \quad (7)$$

α 、 β is an attenuation factor $(0 < \beta < 1、0 < \alpha < 1)$, k refers to the time interval between the real time t_c and the most close to real time(between t_1 and t_2), it is defined as formula (8) :

$$k = t_c - \max(t_1 - t_2) \quad (8)$$

About the part $k_n^{-\alpha}$, its value increases with the decrease of k_n , $|t_{1_n} - t_{2_n}| + 1$ is used to prevent the situation. When two nodes at the same time interval, time weight (TS) become infinite.

For the path has only a single time factor, namely $|t_{1_n} - t_{2_n}|$ is 0. For k part, it refers to the time interval between the real time t_c and node a time t_a :

$$k = t_c - t_a \quad (9)$$

$$TS(a,b) = \frac{1}{n}\sum_{n} k_n^{-\alpha} \quad (10)$$

Let time weight (TS) combined with path weight to get a new dynamic path weight. The relational expression is defined as formula (11) :

$$W_t = TS * w_i \quad (11)$$

3.2 personalized weights

Via the interaction between nodes, this paper puts forward the calculation to result authority. For the traditional algorithm, the link relation between two nodes is determined by whether there is a connection or not. Only if it has, then it equals to 1, otherwise 0. However, for the node weights which joined link relations, is no longer a simple 1 or 0, but taking into account the different types of nodes' influence.

Through using iterative calculation and the basic idea of PageRank or HITS algorithm, influence factors are formed by the interaction between nodes. According to the DBLP data set network structure in this paper, the DBLP data set has three types of nodes respectively, the author, papers and conference. In figure 1, the author's authority is calculated by the number of published papers and average authorityof papers, and the authority of paper is affected by the average authority of authors who wrote it and its published conference'sauthority. In the face, this influence factors of three node types are expressed respectively :

$$influence_A = \frac{n}{M} + \frac{1-\beta}{n} * \sum_{n} influence_P \quad (12)$$

$$influence_P = \frac{\beta}{m}\sum_{m} influence_A + (1-\beta) * influence_C \quad (13)$$

$$influence_C = \frac{p}{M} \quad (14)$$

$influence_A$ 、$influence_P$ 、$influence_C$ are regarded as the author influence, the paper influence and the meeting influence respectively, β is damping factor, n is a number of the published papers, m represents a number of the authors who wrote this paper, p represents a number of papers in a meeting, Mrepresents the total number of papers on data set.

It is assumed that there are links between node A and B. Because three node types authority are between [0, 1], for $e^{influence}$, it can guarantee the node weights between [1, e], the link relations as :

$$U'_{AB} = e^{influence_A * influence_B} \quad (15)$$

3.3 Hete-DS algorithm

This paper uses the method of transition probability matrix between the node types, and can quickly calculate Hete-DS algorithm, and reduce the computational complexity.For the transition probability matrix $U'_{A_1A_2}$ of the node type A_1A_2, if there is no link between nodes, the transition probability is 0, if existing relationships between nodes, transition probability is equal to $e^{influence_a * influence_b}$, it combines with HeteRecom algorithm model and its formula is :

$$\begin{aligned} &\text{Hete-DS}(A_1, A_{l+1} | P) = \text{Hete-DS}(A_1, A_{l+1} | P_L\ P_R) \\ &= U'_{A_1A_2} \ldots U'_{A_{mid}M} V'_{MA_{mid+1}} \ldots V'_{A_lA_{l+1}} = PM_{P_L} PM'_{P_R^{-1}} \end{aligned} \quad (16)$$

Due to the different node types of dimension is not the same, so they need to proceed the normalized processing, the formula (17) is :

$$Hete\text{-}DS(a,b|P) = \frac{PM_{P_L}(a,:)PM'_{P_R^{-1}}(b,:)}{\sqrt{\|PM_{P_L}(a,:)\| \|PM'_{P_R^{-1}}(b,:)\|}} \quad (17)$$

For the whole Hete-DS algorithm is :

$$Sim(a,b) = W_t * Hete\text{-}DS(a,b|P) \quad (18)$$

IV. EXPERIMENT AND RESULT ANALYSIS

This section used the digital bibliography and library project (DBLP) data sets to make some related experiments, this article's algorithm were comparedby HeteRecom algorithm. The experiment simulation was by Python, and it verified the effectiveness of the proposed algorithm.

Experimental environment and the data set :

Experimental environment: processor for @ 2.00 GHz Intel Core i7-2630, 4 GB RAM, Windows 7 operating system.

Data set: this paper used the heterogeneous information network data set to test algorithm. In the experiments, it selected the DBLP network subset, included 23 major field of study in the meeting, the author 11411, 5343 papers, each paper published time and their time attributes in links.

4.1 The effectiveness of the time factor in Hete-DS algorithm

Assuming that a given symmetrical meta-path P, it analyzed on HeteRecom and Hete-DS algorithm respectively. For example: the meta-path APCPA (i.e., author – paper - meeting - paper - author) found the authors who had released papers in a same meeting. For meta-path APA (i.e., author - paper - author), it foundthe authors who had written the same paper. According to meta-path,it could calculate relationship matrix and could be applied to the HeteRecom or Hete-DS algorithm. In the experiment, it set the parameters of the time factor $\alpha = \beta = 0.1$ in the algorithm, the node weight parameter is $\beta = 0.5$.

First, for the meta-path APA, it calculated the co-author similarity which isthe published papers of Oguz Ergin, and Hete - DS was comparedwith HeteRecom and HeteSim, as shown in table 1. Due to the DBLP data set, time variable refers to the time of paper which has published by the authors. That is to say, the establishment time of A - P or P - A. For the meta-path APA, it need to calculate the time interval which is between the published papers and now.

Table 1 The author similarity of Oguz Ergin under the APA path

Author	HeteSim	HeteRecom	Hete-DS
Oguz Ergin	1.0	0.905866	0.757267
TimothyM. Jones	1.0	0.905866	0.757267
MichaelF.P. O'Boyle	1.0	0.905866	0.757267
Jaume Abella	1.0	0.905866	0.757267
Antonio González	0.5	0.452933	0.094658

In table 1, the author Timothy M. Jones, Michael F. P. O'Boyle, Jaume Abella have the same similarity with the source node, because out-degrees and in-degrees between them are the same. That is to say, their published papers are mutual cooperation. In addition, it can be seen that similarity of Hete-DS algorithm compares with HeteRecom, which has the same rate of decrease. It can be seen that the longer distance from the reality, the lower similarity may be.

Then for the mate-path APCPA, in addition to consider the distance from now, it also need to analyze the path building time intervals both A - P and P – A. Then it finds out the authorswho are similar to Oguz Ergin. Finally, itcompared HeteRecom and Hete-DS comparison, as shown in table 2.

Table 2 The top-10 similar authors to Oguz Ergin under the APCPA path

Author	HeteSim	HeteRecom	Hete-DS
Oguz Ergin	1.0	0.094134	0.073422
Timothy M. Jones	1.0	0.094134	0.073422
Michael F. P. O'Boyle	1.0	0.094134	0.073422
Jaume Abella	1.0	0.094134	0.073422
Antonio González	0.816497	0.076860	0.059949
Manuel E. Acacio	1.0	0.094134	0.061378
José González	0.894427	0.084196	0.060284
José M. García	0.707107	0.066563	0.043401
José Duato	0.196116	0.018461	0.012037
BasilioB.Fraguela	0.707107	0.066563	0.043401

In table 2, HeteRecom compared to HeteSim, which is no longer just consider whether there is a path between nodes link or not to determine the similarity between them. Moreover, it introduced the weight path between the nodes and considered the importance of the different paths, so the accuracy has raised. On the original path weights,this article join the parameters of time factor, which make the path weight not only considering the difference between different path, but also the path building time. Because even the same type paths may also have difference. If the building interval between two paths is smaller, their activities may be happened at the closer time, that is to say, they existcontact more possible at this moment.

Finally considering the DBLP information network, it analyzed all paths between the nodes in the network and their path weight. Through the Hete-DS algorithm calculation, it concluded the top-10 authors who were the most similar to Oguz Ergin, compared with HeteRecom algorithm, as shown in table 3.

Table 3 The top-10 similar authors to Oguz Ergin

Author	HeteSim	HeteRecom	Hete-DS
Oguz Ergin	2.0	1.0	0.830689
Timothy M. Jones	2.0	1.0	0.830689
Michael F. P. O'Boyle	2.0	1.0	0.830689
Jaume Abella	2.0	1.0	0.830689
Antonio González	1.316497	0.529793	0.154607
Ilya Ganusov	1.0	0.094134	0.073422
Martin Burtscher	1.0	0.094134	0.073422
Gabriel H. Loh	1.0	0.094134	0.073422
Lisa R. Hsu	1.0	0.094134	0.073422
Steven K. Reinhardt	1.0	0.094134	0.073422

In table 3, due to the papers in data sets had a long time from reality, so the similarity by Hete-DS algorithm was relatively decreased. Path weight included the influence of time factor, which made the result more accurate.

4.2 The effectiveness of personalized weights in Hete-DS algorithm

Under the given meta-path P, for DBLP, different papers, or different authors, even different meetings, have different authority influence, that is to say, the status of each node in the network is not the same. For the path in the network, it considers no longerwhether there is a connection or not, but the higher influence of node is, the more important the meta-path P is in the network. In other words, the transfer matrix of relationship between nodes is no longer just 0 or 1, but is controlled by the influence of the corresponding node.

First, for the influence of papers, conference and author in DBLP data set, they Interact with each other, which can be seen by the influence of formula (12) - (14). So they need to iterate through the three operations. Through the formula $\frac{|The\ current\ iteration - The\ last\ iteration|}{The\ last\ iteration}$, it observed three influence convergence in the process of iteration, as shown in figure 3, the horizontal axis represents the number of iteration, the vertical axis represents the number of authors, papers.

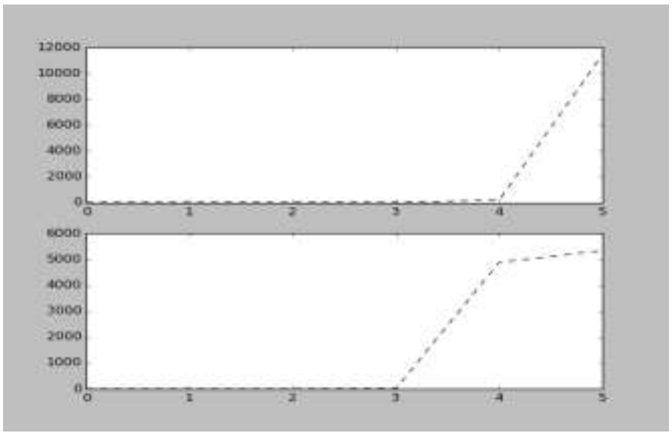

Fig.3 The author influence iterative process (up) and paper iterative process (down)

In figure 3, it can be seen that absolute difference values of three nodes type are into decline after each iteration.With the increase of the number of iterations,drop amplitude will decrease. After the sixth iteration, the change is below10^{-3}. Due to the meeting node type of authority is unchanged, its calculation formula of authority is fixed and is not change in the iteration. Therefore, in this experiment, we adopt the way of 6 iterations to calculate each node weights.

Under meta-path APCPA, the experimental analysis just join node weights in Hete-DS algorithm, and compared with HeteRecom algorithm, which judged the effectiveness of the algorithm, as shown in table 4.

Table 4 The author similarity of Antonio González under the APCPA path

Author	HeteRecom	Hete-DS
Antonio González	0.094134	0.094134
John B. Carter	0.088750	0.088811
José González	0.085932	0.084368
Julita Corbalán	0.081522	0.079809
José M. García	0.081522	0.079769
Chen Ding	0.081522	0.079721
Gurindar S. Sohi	0.081522	0.079701
Franck Cappello	0.081522	0.079693
Santosh Pande	0.081522	0.079690
Michael L. Scott	0.081522	0.079688

In table 4, only considering personalized weight, Hete-DS algorithm of node similarity compared with HeteRecom algorithm. Since HeteRecom algorithms only consider the influence of different paths on the nodes similarity, but it is unable to distinguish between nodes under the same path similarity. Hete-DS algorithm considered the weights of different nodes, the nodes weight will affect the similarity. As shown in table 5, according to the formula (12) - (16),the formula $e^{influence_A + \frac{\sum influence_P}{number\ \ of\ \ the\ \ paper} + influence_C}$ can estimate the situation of different authors' average node weighted in the same path.

5 The node weights are compared between the author Antonio Gonzalez and John b. Carter, Jose Gonzalez

Author	In the meeting path PACT	In the meeting path IPDPS	In the meeting path ICPP
Antonio González	1.030504	1.137623	1.133705
John B. Carter	1.034168	1.143472	1.139123
José González	1.029929	1.134830	——

In table 5, relative to the author Antonio Gonzalez,the average node weight of John b. Carter is improved under the three paths of meeting. The author Jose Gonzalez and Antonio Gonzalez only has contact in pathswhich pass the meeting IPDPS or PACT, so it only considers the average personalized weights under these two paths. It can be seen that the average node weight of Jose Gonzalezis lower than Antonio Gonzalez.In table 4 and table 5, they can obtain that the higher the node weight is, the more similar the node is.On the contrary, the lower the node weight is, the more similar the node is.

Finally, the whole Hete-DS algorithm is compared to HeteRecom algorithm, as shown in table 6. Because Hete-DS joined the time factor, so the whole similarityof Hete-DS is reduced.

Table 6 The author similarity of Oguz Ergin

Author	HeteRecom	Hete-DS
Antonio González	1.0	0.833713
Qiong Cai	0.529793	0.445097
Ryan Rakvic	0.529793	0.445097
Pedro Chaparro	0.529793	0.445097
Timothy M. Jones	0.529793	0.444899
Michael F. P. O'Boyle	0.529793	0.444899
Jaume Abella	0.529793	0.444899
Oguz Ergin	0.529793	0.444899
Fernando Latorre	0.491363	0.427290
Teresa Monreal	0.491363	0.391109

V. CONCLUSION

In this paper, for the problem of the similarity between entities in heterogeneous network problems, it is presented that the algorithm which is based on the influence factors in the heterogeneous network time-path similarity algorithms (Hete-DS). On the basis of the original HeteRecom algorithm, increased the time factor weights in different paths between. At the same time, fordifferent influence factors of different nodes, it joinsthe influence factors of each node to the original transfer matrix, and implements comprehensive algorithm. Not only has dynamic time similarity search been realized in the heterogeneous information network, but also improve the accuracy of the similarity search. The experimental results also prove the validity of Hete-DS. In the next stage,wewill consider to realize the fast calculation of parallel structure in order to improve the operation speed of the algorithm and apply the algorithm in the multiple data sets, which include commodity information network, social networking, etc.It can also be used as a node distance calculation and combines with popular mining algorithm to implementthe recommendation system function.

REFERENCE

[1]. Balmin A, Hristidis V, Papakonstantinou Y. Objectrank: authority-based keyword search in databases[C]// Thirtieth International Conference on Very Large Data Bases. VLDB Endowment, 2004:564-575.

[2]. Nie Z, Zhang Y, Wen J R, et al. Object-level ranking: bringing order to Web objects[C]// International Conference on World Wide Web, WWW 2005, Chiba, Japan, May. 2005:567-574.

[3]. Lao N, Cohen W W. Fast query execution for retrieval models based on path-constrained random walks[C]// ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, Dc, Usa, July. 2010:881-888.

[4]. Shi C, Kong X, Huang Y, et al. HeteSim: A General Framework for Relevance Measure in Heterogeneous Networks[J]. IEEE Transactions on Knowledge & Data Engineering, 2014, 26(10):2479-2492.

[5]. Sun Y, Han J, Yan X, et al. PathSim: Meta Path-Based Top-K Similarity Search in Heterogeneous Information Networks.[J]. Proceedings of the Vldb Endowment, 2011, 4(11):992-1003.

[6]. Shi C, Zhou C, Kong X, et al. HeteRecom: a semantic-based recommendation system in heterogeneous networks[C]// ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2012:1552-1555.

[7]. CHEN Yunzhi. Prestige based information ranking model in heterogeneous information networks[J]. Computer Engineering and Applications, 2014, 50(1):121-126.

[8]. Shi C, Zhang Z, Luo P, et al. Semantic Path based Personalized Recommendation on Weighted Heterogeneous Information Networks[C]// The, ACM International. 2015:453-462.

[9]. Xiang L, Yuan Q, Zhao S, et al. Temporal recommendation on graphs via long- and short-term preference fusion[C]// ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2010:723-732.

[10]. Munasinghe L, Ichise R. Time score: A new feature for link prediction in social networks[J]. IEICE TRANSACTIONS on Information and Systems, 2012, 95(3): 821-828.

[11]. ZHU Lei. Person Correction Based on Temporal Sematic Path Similarity in Heterogeneous Networks[D]. Central South University, 2014.

[12]. Mmele N, Ichise R, Werthner H. Exploring Supervised Methods for Temporal Link Prediction in Heterogeneous Social Networks[C]// International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, 2015:421-5.

[13]. CHEN Xiangtao, DING Pingjian, WANG Jing. Meta path-based dynamic similarity search in heterogeneous network[J]. Computer Application, 2014, 34(9):2604-2607.

[14]. ZHANG Weizhe, WANG Bailing, HE Hui, TAN Zhuopeng. Public Opinion Leader Community Mining based on the Heterogeneous Network[J]. Acta Electronica Sinica, 2012, 40(10): 1927-1932.

[15]. GE Jun. Similarity Search in Heterogeneous Information Network[D]. Xidian University, 2014.