

Link and Content Based Similarity Integration in Visa Clearance System

H.A.Shaikh¹, M.B.Waghale², V.V.Darade³, Prof.R.A.Mahajan⁴

Department of Information Technology, JSPM'S BSIOTR(w),Wagholi, Pune, India

For Correspondences: Shaikhheena543@gmail.com

ABSTRACT-

In this paper we introduce the concept of heterogeneous information network and problem of how to perform information retrieval and recommendation in such networks. We propose an algorithm integrated weighted similarity learning (IWSL) to account for both link based and content based similarities by considering the network structure and mutually reinforcing link similarity and feature weight learning. Experimental results on VISA Clearance System show that our approach is significantly better than traditional methods in terms of relevance and speed.

Keywords— Information retrieval, data mining, information network, link & content based similarity

1. INTRODUCTION

VISA is a simply document issued by country's government allowing the holder to Enter or Leave the country. We introduce the concept of heterogeneous information network. In this project we introduce the concept of heterogeneous information network and the problem of how to perform information retrieval and recommendation in such networks. We proposed a fast algorithm heterogeneous minimum order Integrated Weighted Similarity Learning to account for both link based and content based similarities by considering the network structure. In data Content based retrieval,

most methods such as Google's Visual Rank [4] and systems [5], [6], [7], [8], [9] compute data similarity based on data content features.

We propose an algorithm Integrated Weighted Similarity Learning (IWSL) to account for both link based and content based similarities by considering the network structure and mutually reinforcing link similarity and feature weight learning. Link based Similarity:-For Evaluating Similarity between node's in information network. Two nodes are similar if they link by similar node in the network.Content based similarity:-Information similarity can be estimated from Data content features. Content features like color, shape etc

2. EXISTING SYSTEM

Current VISA system consist of following issues,

- 1) Same Name Issue:-Rejected if just Name of User and criminal matches.
- 2) More Processing Time for approval.
- 3) No central database is available to track all types of clearances.

2. PROPOSED SYSTEM

To solve this issue, solution is VISA CLEARANCE SYSTEM by giving Groups, Tags, and Categories to the information of that user. After that there is a Criminal Clearance, which will be clear by Groups, Tags, and Categories.

2.1 INFORMATION NETWORK

Experimental result on VISA Clearance System is done by generating information network.

Information network categories into 3 steps:-

- 1) CATEGORIES
- 2) GROUPS
- 3) TAGS

Information network can be generated by following table

CATEGORIES	GROUPS	TAGS
New user	Citizen	Permanent citizen
Existing user	Student	Temporary citizen
Fraud user	Criminal	Long duration
Duplicate user	Employee	Short duration
Eligible user	Tourist	Higher Studies
Ineligible user	Religious	
	Businessman	

3. PROPOSED WORK

Flow of project it consist of following steps:-

- 1) Input:-In this project we gives input Visa Application form & all required document
- 2) Validation:-After giving input there will be a primary validation done. Primary Validation is based on Valid Details.
- 3) Identifying groups, tags and categories: Based on the all information of user ,the details of user can be categorized into Groups,Tags,Categories.It means user belongs to which groups,category,tags.
- 4) Check for Duplicate Application:-Duplicate Application is check, if documents or

duplicate entry occur then VISA get rejected if not then continue the procedure.

- 5) Enriching information into system:- Information of user can be enriching into Criminal Clearance System.
- 6) Searching:- Search is based on Groups, tags, name and category on the basis of Criminal clearance and Bank Clearance.
- 7) Check all other legal parameter:-Check all other Documents and legal parameter.
- 8) Approve VISA:- If all Documents and Clearance are done and result is fine then approval of VISA take place.

4. ALGORITHM USED

We propose an algorithm Integrated Weighted Similarity Learning (IWSL) to account for both link based and content based similarities.

Input:-G, is information network.

Step 1. Find top K similar candidates of each object;

Step 2. Initialization;

Step 3. Iterate {

Step 4. Compute link similarity for all category pairs;

Step 5. Compute link similarity for all group pairs;

Step 6. Compute link similarity for all tag pairs;

Step 7. } until converge or stop criteria satisfied;

Step 8. Perform feature learning to update increment by 1.

Step 9. Update similarities;

Output: - S, pair-wise node similarity scores.

5. APPLICATION

1) Use both content and link based similarities

To combine content and link information to achieve more robust performance IWSL is used.

2) Better performance than traditional algorithms

Algorithm IWSL further improves the performance by introducing a new way of integrating content and link information via mutual reinforcement with feature learning. IWSL obtains the best results in terms of the relevance for both semantic and visual appearances. We can use such tag similarity to help find more relevant information for a keyword query.

3) Better in terms of relevance and speed

We have implemented search and recommendation system to find both visually similar and semantically relevant products based on our algorithm. IWSL algorithm we plan to study the problem of how to get an optimal combination of both local and global learning to achieve a balance on time and quality performance.

4) Data Retrieval & recommendation

We introduce the concept of heterogeneous information networks and the problem of how to perform information retrieval and recommendation in such network.

- a) Used in weighted heterogeneous information networks.
- b) Used in Criminal Record Clearance.
- c) Used in Bank Clearance.
- d) Better than traditional approaches.
- e) By using Groups, Tags and Category Data mining can done relevantly.
- f) Required less processing time.

CONCLUSIONS

We propose the algorithm IWSL for computing link and content similarity integration in information network. IWSL algorithm improve the performance than traditional approaches. We propose IWSL to efficiently compute weighted link based and content based similarity in weighted heterogeneous information networks. In this project we introduce the concept of heterogeneous information network and the

problem of how to perform information retrieval and recommendation in such networks.

REFERENCES

- [1] X.Jin, J.Luo, J.Yu, G.Wang, D.Joshi, and J.Han, "iRIN: Image Retrieval in Image-Rich Information Networks," Proc.19th Int'l Conf.World Wide Web (WWW '10),pp. 1261-1264, 2010.
- [2] R.L.Cilibrasi and P.M.B.Vitanyi, "The Google Similarity Distance,"IEEE Trans. Knowledge and Data Eng.,vol.19, no. 3,pp.370-383, Mar .2007.
- [3] L.Wu, X.-S.Hua, N.Yu, W.-Y.Ma, and S.Li, "Flickr Distance,"Proc.16th ACM Int'L conf.Multimedia, pp.31-40, 2008.
- [4] Y.Jing and S.Baluja, "VisualRank: Applying Pagerank to Large-Scale Image Search," IEEE Trans. Pattern Analysis and Machine Intelligence, vol.30, no.11, pp.1877-1890, Nov. 2008.
- [5] R.C. Veltkamp and M. Tanase, "Content-Based Image Retrieval Systems: A Survey," technical report,Dept. of Computing Science,Utrecht Univ.,2002.
- [6] H. Tamura and N. Yokoya, "Image Database Systems: A Survey,"Pattern Recognition, vol. 17, no.1, pp. 29-43,1984.
- [7] W.I. Grosky, "Multimedia Information Systems," IEEE Multi-Media, vol.1, no.1, pp.12-24, Spring, 1994.
- [8] V.N. Gudivada and V.V. Raghavan, "Content-Based Image Retrieval Systems," Computer, vol. 28, no.9, pp.18-22, Sept. 1995.
- [9] Y. Rui, T.S. Huang, and S.-F. Chang, "Image Retrieval: Current Techniques,Promising Directions,and Open Issues," J. Visual Comm. and Image Representation, vol. 10, no. 1, pp. 39-62, 1999.
- [10] R. Datta, D. Joshi, J. Li, and J.Z. Wang, "Image Retrieval: Ideas,Influences, and Trends of the New

Age,” ACM Computing Surveys, vol. 40, no. 2, pp. 1-60, Apr. 2008.

[11] T. Deselaers and H. Mller, “Combining Textual- and Content-Based Image Retrieval, Tutorial,” Proc. 19th Int’l Conf. Pattern Recognition (ICPR’08), http://thomas.deselaers.de/teaching/files/tutorial_icpr08/04-combination.pdf, 2008.

[12] S. Sclaroff, M. La Cascia, and S. Sethi, “Unifying Textual and Visual Cues for Content-Based Image Retrieval on the World Wide Web,” Computing Vision Image Understanding, vol. 75, pp. 86-98, July 1999.