International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 05 | May 2020 www.irjet.net p-ISSN: 2395-0072

Hierarchical Multi-Clue Designing for POI Popularity Prediction with Heterogeneous Tourist Instruction

Aneesha M.S

Student, Dept. of Dual Degree Computer Applications, Sree Narayana Guru Institute of Science and Technology, Kerala, India

---

Abstract - Predicting the recognition of Point of Interest (POI) has become increasingly crucial for location-based services, like POI recommendation. Most of the prevailing techniques can seldom attain quality overall performance thanks to the shortage of POI's information, which tendentiously confines the recommendation to famous scene spots, and ignores the unpopular sights with probably treasured values. In this paper, we propose a completely unique approach, termed Hierarchical Multi-Clue Fusion (HMCF), for predicting the recognition of POIs. Specifically, so on have an effect on the count number of know-how sparsity, we endorse to comprehensively describe POI the usage of a number of varieties of person generated content (UGC) (e.g., textual content and image) from a couple of sources. Then, we devise an environment friendly POI modelling technique all through a hierarchical manner, which concurrently injects semantic expertise additionally as multi-clue consultant electricity into POIs. For evaluation, we assemble a multi-source POI dataset through amassing all the textual and visible content material of numerous specific provinces in China from 4 main-stream tourism structures all through 2006 to 2017. Extensive experimental results show that the proposed method can significantly improve the performance of predicting the attractions' popularity as compared to many baseline methods. Weather forecasting is an other provider supplied by way of this.

***Key Words*: POI Popularity Prediction, Hierarchical Structure, Multiple Sources, Multi-view Learning.**

1. INTRODUCTION

TOURISM is an important industry to national economy and world economy, and its comprehensive contribution accounts for 11% of GDP in China.1. However, through the analysis of data from several travel websites, we notice that the distribution of travelers is uneven among the scenic spots. This characteristic of scenic spots' popularity exhibits the long tail effect. In other words, nearly 90% scenic spots are untraversed, while the rest attract almost all of the tourists. As a matter off act, most of these unpopular spots actually receive high evaluation from the tourists, which implies the precious values of these unpopular spots may be substantially underestimated .

POI popularity prediction aims at analyzing the popularity condition in the next period of time based on the features of scene spots. This research is valuable and important, for it not only improves the accuracy of attractions recommendation and route planning for visitors, but also provide reference information to mining commercial value hidden in the unpopular spots. In addition, there exist many problems in traditional popular tourist attractions, such as congestion, disappearance of characteristics, rising of entrance fee, excessive commercialization and so on. Therefore, more and more tourists prefer to visit some unpopular scenic spots with exceptional features.

In recent years, several research endeavors have been dedicated to popularity prediction and its related applications. Most researches pay little attention to the field of popularity prediction of tourist sites. Related In application fields involving forecasting popularity of scenic spots, most approaches mainly focus on establishing data-based prediction model by mining popular travel POIs through social media, descriptions from a large number of travelers and highly ranked results from search engines. Although these studies have raised some innovative approaches of applying popularity prediction to attractions, they ignore the significant influence of emerging scenic spots and undiscovered attractions on tourism tendency, i.e., attractions with freshness and uniqueness are likely to be discovered by travelers and become popular in the near future.

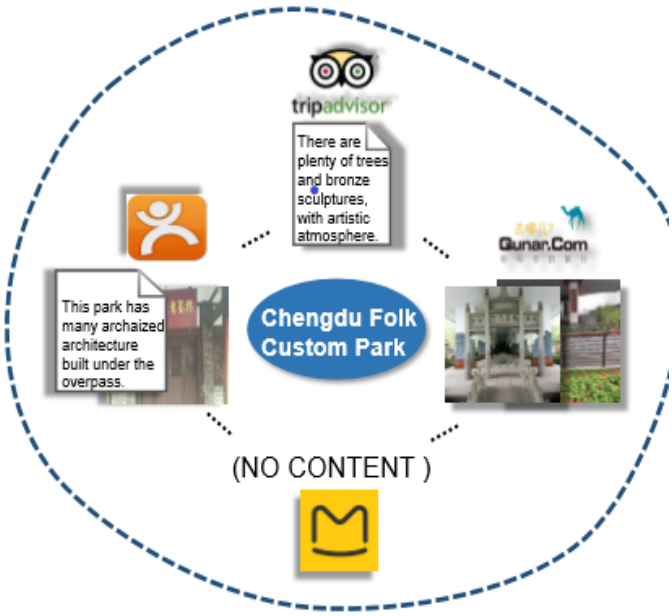
1.1 Challenges for POI

1) the description information of POI in social network is very sparse. Even in famous and widely-used

websites, a large proportion of POIs only have few photos and/or associated text.

2) Visual ambiguity is common in different types of scene spots. For instance, it is difficult to distinguish picking garden from natural scenery only through their images. Therefore, it is partial to predict popularity by utilizing only low-level features.

3) Seldom researches have been made to effectively fuse multi-modal features from multiple sources to model POI data. It is non-trivial to effectively integrate multiple social clues for modelling POIs.



2. EXISTING SYSTEM

Most of this strategies can seldom attain first-class overall performance thanks to the shortage of POI's information, which tendentiously confines the suggestion to famous scene spots, and ignores the unpopular sights with probably treasured values.In this paper, we promote a completely unique approach, termed Hierarchical Multi-Clue Fusion (HMCF), for predicting the popularity of POIs. Specifically, so as to deal with the matter of knowledge sparsity, we propose to comprehensively describe POI using various sorts of user generated content (UGC) (e.g., text and image) from multiple sources. Then, we devise an fantastic POI modelling approach during a hierarchical manner, which concurrently injects semantic information as properly as multi-clue consultant energy into POIs. For evaluation, we assemble a multi-source POI dataset with the help of gathering all the textual and visual content material of variety of unique provinces in China from 4 main-stream tourism systems at some stage in 2006 to 2017.Extensive experimental effects exhibit that the proposed technique can considerably enhance the general performance of predicting the attractions' reputation as in contrast to numerous baseline methods.

Specifically, so as to deal with the effort of records sparsity, we advise to comprehensively describe POI the utilization of variety of sorts of consumer generated content material (UGC) (e.g., textual content and image) from a few of sources. Then, we devise an tremendous POI modelling technique during a hierarchical manner, which concurrently injects semantic expertise as nicely as multi-clue consultant energy into POIs. For evaluation, we assemble a multi-source POI dataset via amassing all the textual and visual content material of varied unique provinces in China from 4 main-stream tourism structures at some point of 2006 to 2017.Extensive experimental outcomes exhibit that the proposed technique can considerably enhance the general performance of predicting the attractions' reputation as in contrast to varied baseline methods.

3. METHODOLOGY

In this paper solving the sparse problem of real-world POI data. To enhance the comprehensiveness of POI description, we establish our real-world POI dataset by collecting several types of user generated content from multiple sources. Our improved dataset contains richer information and more balanced distribution of POI data.

We propose a POI modelling method based on a hierarchical structure with four specific layers, where we take POIs' semantic information and multi-modal representation into consideration simultaneously. Thus, we are able to complete popularity prediction through double classification. The proposed hierarchical method in this version improves the quality of topics based on new dataset, which is detailed in subsequent discussions. We achieve different kinds of multi-clue feature fusion methods mainly based on multi-view learning to achieve multi-clue feature representation. SVM (Support Vector Machine) algorithm is used in this paper. It is a supervised machine learning algorithm which can be used for both classification or regression challenges. However, it is mostly used in classification problems

4. FUTURE WORK

In future work, we will examine the applicability of proposed method by utilizing different types of dataset. We plan to extend this work to optimization of existing POI recommendation system, thus we are possible to recommend scenic spots with definite potentiality to visitors. Besides, more types of data in multi-media, such as video, is likely to be integrated into POI popularity prediction and POI recommendation in our following work.In future, we can add Weather forecasting details, which will assist travelers a lot for their explore.

5. CONCLUSION

In this work, we tried to solve the problem of POI popularity prediction, which facilitates existing POI recommendation by mining unpopular attractions with potential values. We proposed a novel hierarchical POI modeling strategy which simultaneously utilizes semantic content as well as multi-clue representation of POIs. Specially, in multi-clue fusion, we fully exploited various types of UGC content from multiple sources and utilized different types of feature fusion strategies based on early fusion as well as multi-view learning. Besides, we optimized the HMCF method in dataset and hierarchical strategy, where the chosen topics, the distribution of POIs and feature extraction are adjusted grounded on realities of situation. All experimental studies were conducted on our real-world dataset collected from four main-stream tourism platforms. We especially set several comparison schemes to discuss the relative advantages and differences as compared to the previous paper and demonstrate the superior performance of the proposed approach as compared to baseline methods.

This work provided some new ideas and interesting directions for future POI researches. It differentiated from existing studies of popularity prediction and POI recommendation due to the method of multi-clue fusion based on hierarchical POI modelling. Considering that popularity is a key factor in POI recommendation, it is valuable to explore the combination between popularity prediction and recommendation for POIs, which provides a wide application prospect for visitors, enterprises and tourism industries.

ACKNOWLEDGEMENT

In the name of almighty, i would like to extend my heartfelt thanks to our HoD, Mrs. Kavitha C.R, Department Of Dual Degree Master of Computer Applications for the helps extended to me throughout my course of study. I am deeply grateful to my guide Mrs Tintu P.B, Assistant Professor, Department of Dual Degree Master Of Computer Applications for the valuable guidance.

REFERENCES

[1] H. T. Shen, X. Hua, J. Luo, and V. Oria, "Guest editorial: content, concept and context mining in social media," WWW, vol. 15, no. 2, pp. 115–116, 2012

[2] Q. Zhao, M. A. Erdogdu, H. Y. He, A. Rajaraman, and J. Leskovec, "SEISMIC: A self-exciting point process model for predicting tweet popularity," in SIGKDD, 2015, pp. 1513–1522.

[3] Y. Duan, X. Wang, Y. Yang, Z. Huang, N. Xie, and H. T. Shen, "POI popularity prediction via hierarchical fusion of multiple social clues," in SIGIR, 2017, pp. 1001–1004.

[4] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: user movement in location-based social networks," in SIGKDD, 2011, pp. 1082–1090.

[5]] K. H. Lim, "Personalized recommendation of travel itineraries based on tourist interests and preferences," in UMAP, 2016.

[6] M. Ye, P. Yin, W. Lee, and D. L. Lee, "Exploiting geographical influence for collaborative point-of-interest recommendation," in SIGIR, 2011.

[7] C. Xu, D. Tao, and C. Xu, "A survey on multi-view learning," CoRR, 2013.

[8] C. Xu, D. Tao, and C. Xu, "A survey on multi-view learning," CoRR, 2013.

[9] X. Zhu, Z. Ming, X. Zhu, and T. Chua, "Topic hierarchy construction for the organization of multi-source user generated contents," in SIGIR, 2013, pp. 233–242.

[10] A. Ahmed, L. Hong, and A. J. Smola, "Hierarchical geographical modeling of user locations from social media posts," in WWW, 2013, pp. 25–36.