

Association for Information Systems AIS Electronic Library (AISeL)

PACIS 2018 Proceedings

Pacific Asia Conference on Information Systems
(PACIS)

6-26-2018

Intelligent Point-of-Interest Recommendation for Tourism Planning via Density-based Clustering and Genetic Algorithm

Chih-Kun Ke

National Taichung University of Science and Technology, ckk@nutc.edu.tw

Mei-Yu Wu

Chung Hua University, mywu@chu.edu.tw

Wang-Chi Ho

National Taichung University of Science and Technology, s1310634003@nutc.edu.tw

Suz-Cheng Lai

National Taichung University of Science and Technology, s1810531013@nutc.edu.tw

Li-Te Huang

Service Systems Technology Center, Industrial Technology Research Institute, huanglite@itri.org.tw

Follow this and additional works at: <https://aisel.aisnet.org/pacis2018>

Recommended Citation

Ke, Chih-Kun; Wu, Mei-Yu; Ho, Wang-Chi; Lai, Suz-Cheng; and Huang, Li-Te, "Intelligent Point-of-Interest Recommendation for Tourism Planning via Density-based Clustering and Genetic Algorithm" (2018). *PACIS 2018 Proceedings*. 140.
<https://aisel.aisnet.org/pacis2018/140>

This material is brought to you by the Pacific Asia Conference on Information Systems (PACIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in PACIS 2018 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Intelligent Point-of-Interest Recommendation for Tourism Planning via Density-based Clustering and Genetic Algorithm

Research-in-Progress

Chih-Kun Ke

Department of Information Management,
National Taichung University of Science
and Technology
No.129, Sec.3, Sanmin Rd, North Dist.,
Taichung, Taiwan 40401, R.O.C.
ckk@nutc.edu.tw

Mei-Yu Wu

Department of Information Management,
Chung Hua University
707, Sec.2, WuFu Rd., Hsinchu, Taiwan
30012, R.O.C.
mywu@chu.edu.tw

Wang-Chi Ho

Department of Information Management,
National Taichung University of Science
and Technology
No.129, Sec.3, Sanmin Rd, North Dist.,
Taichung, Taiwan 40401, R.O.C.
s1310634003@nutc.edu.tw

Szu-Cheng Lai

Department of Information Management,
National Taichung University of Science
and Technology
No.129, Sec.3, Sanmin Rd, North Dist.,
Taichung, Taiwan 40401, R.O.C.
s1810531013@nutc.edu.tw

Li-Te Huang

Service Systems Technology Center, Industrial Technology Research Institute
195, Sec. 4, Chung Hsing Rd., Chutung, Hsinchu, Taiwan 31057, R.O.C.
huanglite@itri.org.tw

Abstract

In recent years, geographic information service and relevant social media become more popular, some geographic point may interest people, e.g. scenic spot or famous store, naming as a point-of-interest (POI). However, the number of POI contributing by social media grows exponentially which causing a searching problem. How to recommend a POI to a user/tourist becomes a challenge. This study proposes an intelligent system using density-based clustering and genetic algorithm to recommend a POIs solution for tourism planning. Density-based clustering identifies candidate POIs. Skyline method decides a superior POI from candidate POIs by dominant of multiple attributes. Genetic algorithm optimizes the recommendation solution. The contribution is to get a tourism POI solution from a huge amount of candidate POIs based on user/tourist preferences. An experimental system implementation is in progress. In future, we will use open data from Google map and Foursquare to proof the proposed system mechanism effectiveness.

Keywords: Tourism planning, point-of-interest, density-based clustering, genetic algorithm, intelligent recommendation

Introduction

Location-based social network (LBSN) covers various points with geographic information including altitude, latitude and longitude. Besides, if the point is a scenic spot or famous store, the social

information may be presented by name, address, phone, opening hours, score, price, commitment, and website attribute, etc. People may pay attention to some geographic point with high interesting, naming as a point-of-interest (POI) (Shenglin et al., 2016). In recent years, advanced information and communication technology assists people to get and share a POI location information in a geographic information service and relevant social media, e.g., Twitter, Google Map, Foursquare, Instagram, and Youtube (Bao et al., 2015; Wang et al., 2015). However, the number of POIs contributing by various social media grows exponentially which causing a searching problem. Therefore, a user/tourist cannot easy to get the desired POI from a huge amount of candidate POIs. How to discover the desired POI for a user/tourist becomes a critical challenge.

In relevant literature, Takeshi et al. (2010) combined Markov and Topic model to calculate the probability of user visiting a scenic spot according to the upload order of scenic spot photos' position in a social platform. They recommended Top-k tourism solutions to assist user making decision. Based on tracking tourist's location, it discovered candidate POIs using to plan various tourism solution. Yu et al. (2016) used LBSN to customize a heuristic tourism recommendation mechanism. They simulated user requirements and position according to collected data from an LBSN. Wei et al. (2012) proposed a concept that a tourism solution plan can be inferred by user's personal tourism experience and POI. Zheng et al. (2009) consider that a tourism solution is not only a user experience but also POI connection. Therefore, they used data mining techniques to discover POIs in order to compose a tourism solution based on historical GPS tracking data.

Traditional recommendation system searching by keyword to get a result is insufficient. For example, a user may use location or price to search a POI, the result may be a huge number or messy scene spots information. The information may be useless to a user making decision to select a POI. Advanced recommendation system uses a novel mechanism combining information system with artificial intelligence to help a user making a decision, e.g., genetic algorithm (GA), particle swarm optimization algorithms (PSO), and ant colony optimal algorithm, etc. The advanced recommendation system becomes more powerful and intelligent. Therefore, according to the LBSN's information, the advanced recommendation system can handle and response user specific and personal requirements. Besides, based on user requirements, the personal recommendation is well to provide reasonable solutions (Mulvenna et al., 2010; Riecken, 2010; Schafer et al., 2001).

In 1960, John Von Neumann proposed a self-reproducing theory to construct a basis of genetic algorithm (GA). John Holland (1962) solved a dynamic adaptive adjustment problem between nature and artificial system in cellular automata. He proposed GA with his students at the University of Michigan in 1975. GA is a famous evolutionary algorithm (EA) and an optimal space searching method. It is famous and widely used in various domains, including problem optimization, data search, artificial intelligence, and machine learning (Goldberg, 1989; Chang and Lilly, 2004; de Oliveira et al., 2018). In the US, Wal-Mart and Amazon used GA to optimize a warehouse management and web distribution. Zhang et al. (2013) proposed a memetic algorithm based on GA to plan vehicle routing.

This study proposes a novel intelligent system framework to enforce point-of-interest recommendation (POIR). A user/tourist shall finish some configurations, including user profile, tourism requirements, and conditions. Data preprocess process collects various POIs information to construct a POI knowledge map. Based on the predefined radius and density parameters in a density-based clustering technique, the proposed system discovers a candidate POIs group. Skyline method filters out low priority candidate POIs by domination and keeps high priority candidate POIs for a recommendation. If some candidate POIs have equal priority, GA is used to determine which candidate POI will be used solving the decision problem. A candidate POI solution will be constructed to recommend a user/tourist for tourism planning. This work plans to collect the open data as data sources. The proposed system will use the web crawler technology and official application programming interface (API) to get experimental data from Google Map and Foursquare platform. The system framework mechanism will be implemented in the experimental platform. We hope the experiment results will recommend a novel tourism solution to help a tourism planning or make a decision to select a POI.

The remainder of this paper is organized as follows. Section 2 introduces presents a novel system framework for a point-of-interest recommendation over a tourism-based social network. The

mechanism of the novel system framework is illustrated in Section 3, and the experiment design and plan are contained in Section 4. Finally, Section 5 presents our conclusions.

A novel system framework for a point-of-interest recommendation (POIR)

This section describes a novel system framework for a point-of-interest recommendation (POIR) (Ricci, 2002; Zhang et al., 2013; Zhiwen et al., 2016), including a user/tourist configuration and representation, data preprocess, point-of-interest (POI) knowledge map, recommendation, and knowledge repository. The procedure of the proposed system framework is shown in Figure 1. At the start of a POIR, a user/tourist uses the proposed system via a configuration and representation module to setup user profile, tourism requirements, and tourism relevant conditions. The input data is sent to the data preprocess module to compose with tourism knowledge getting from the knowledge repository and all candidate POIs getting from the POI knowledge map module. The collected tourist comprehensive data builds a basis which provided to the recommendation module. The recommendation module carries out a GA-based POIR mechanism to produce an optimal tourism solution. The optimal solution is recommended to the user/tourist in the configuration and representation module. The user/tourist sends feedback to configuration and representation module and stored in knowledge repository for improving system mechanism in future.

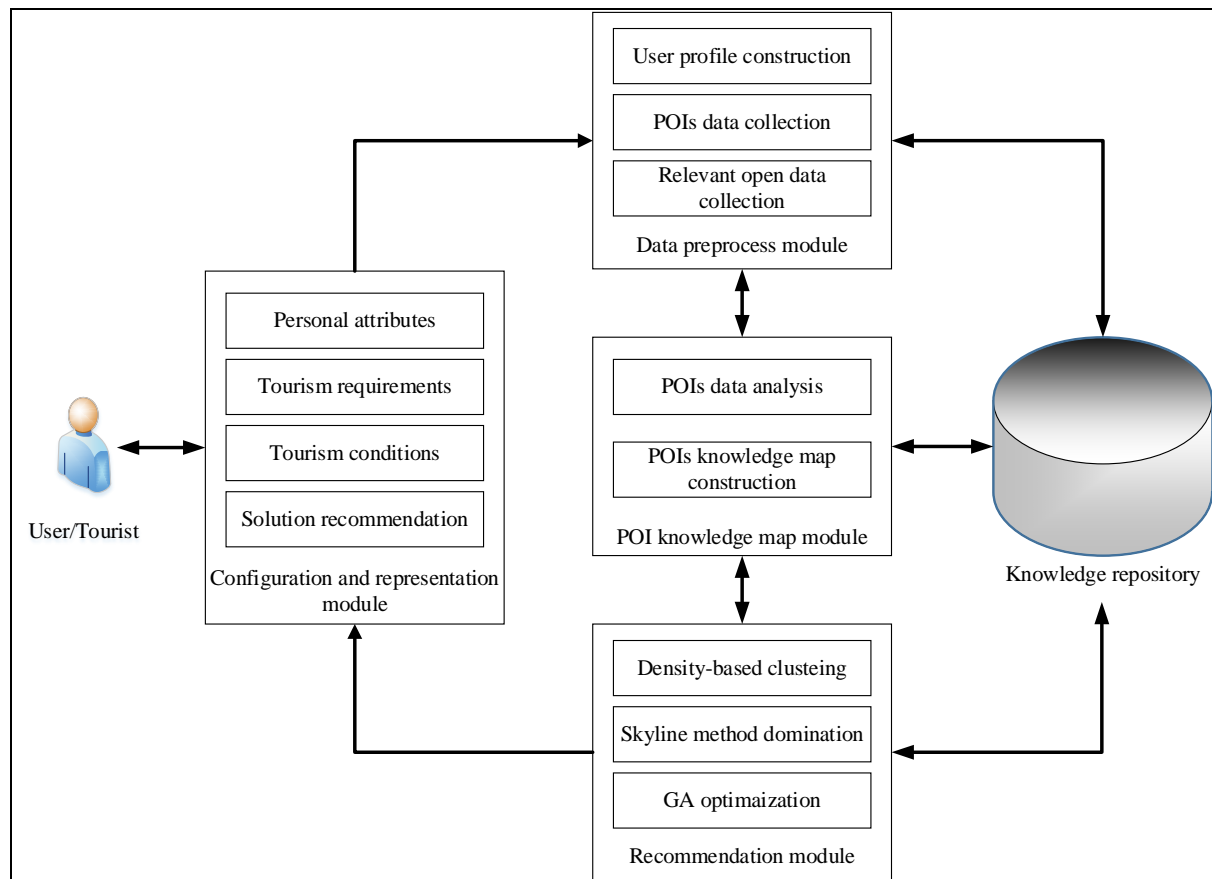


Figure 1. The proposed novel POIR system framework overview

The module functionality is introduced as follows.

- **Configuration and representation module:** This module provides a user interface of the system for a user/tourist setup personal information, tourism requirements, and relevant conditions. The input data will be automatic send to data preprocess module to process for further using. Besides, the user/tourist gets a recommendation solution or sends feedback from this module.
- **Data preprocess module:** This module uses advanced information technology, e.g., web crawler and data mining techniques, to process and compose data, including data collect, select, clean, normalize, integrate and merge with other datasets, e.g., open data or historical data, etc. The data output from

this module will be stored in the knowledge repository and provided to POI knowledge map and recommendation module using.

- **POI knowledge map module:** This module constructs a POI knowledge map offline based on POI information collected from various service or social media platform, e.g., Google Map and Foursquare, etc. The POI knowledge map includes not only geographic information but also various social network information, e.g., POI's altitude, latitude, longitude, name, address, phone, opening hours, score, price, commitment, and website attribute, etc. The comprehensive POI information is considered as tourism knowledge and is useful for tourism solution construction in the recommendation module.
- **Recommendation module:** Based on the comprehensive data collected from the POI knowledge map module and knowledge repository, this module enforces user requirement analysis, POI knowledge map integrated with density-based clustering (Ester et al., 1996; Ankerst et al., 1999), skyline method (Chao et al., 2017) domination, and GA optimization to get a recommendation solution. The recommendation solution will send to the configuration and representation module helping a user/tourist to make a decision in a tourism or a tourism planning.
- **Knowledge repository:** The knowledge repository stores relevant information produced from data preprocess, POI knowledge map, and recommendation module. It also plays a role as a data warehouse for further huge POI data analysis enhancing the quality of POI knowledge map and recommendation solution.

The GA-based POIR mechanism

This section describes the GA-based POIR mechanism of the proposed system framework. Figure 2 shows an overview of the GA-based POIR mechanism working process. Based on a user/tourist profile, tourism requirements, tourism conditions, and POI knowledge map, the GA-based POIR mechanism carries out the procedures are illustrated in follows.

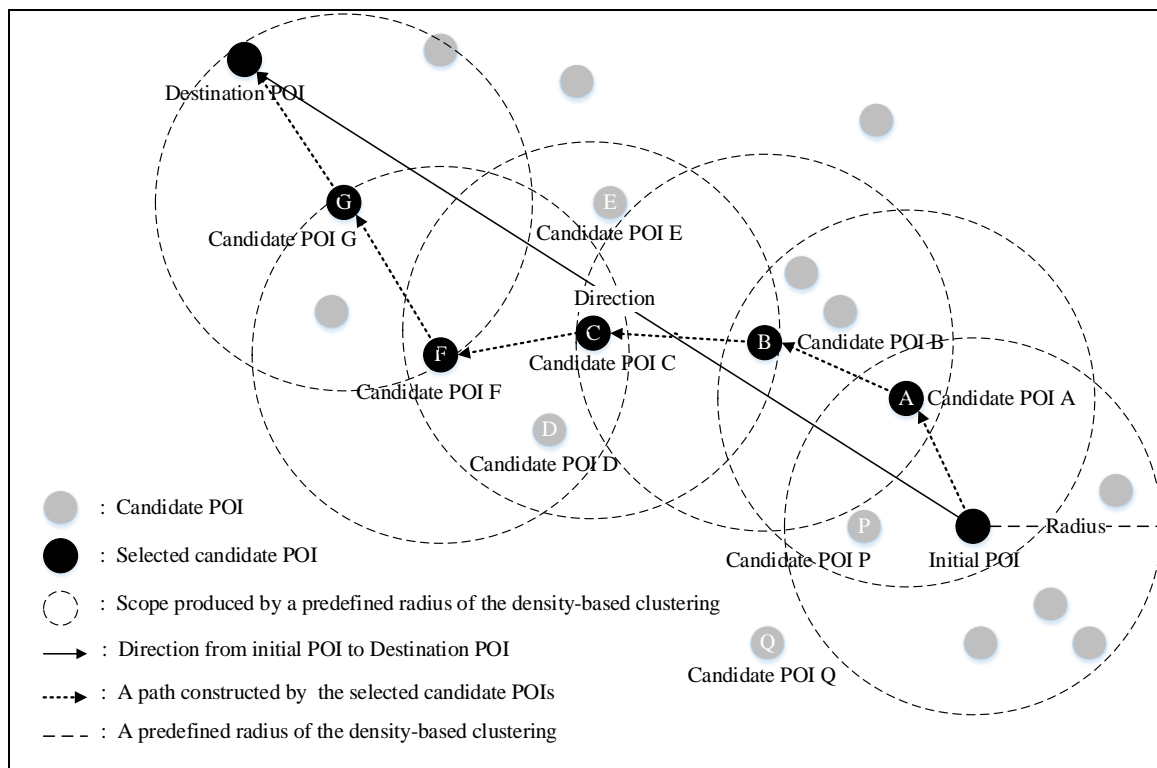


Figure 2. The GA-based POIR mechanism working process

- **Density-based clustering:** The user profile indicates user location as initial POI, and the tourism condition indicates a destination POI. The initial POI and destination POI will be marked in POI knowledge map. From initial POI to destination POI can get a direction which guides a user/tourist

to select a candidate POI. A density-based clustering technique is used to construct several candidate POIs groups. Two parameters of the density-based clustering: radius and density are used to construct a specific candidate POI group. The radius parameter produces a scanning scope, the Euclidean distance between a candidate POI and center candidate POI calculates by the x-coordinate and y-coordinate value, if the Euclidean distance is equal or smaller than the radius value, the candidate POI belongs to the center candidate POI's group, otherwise not. For example, in Figure 2, candidate POI P belongs to the initial POI's group and candidate POI Q does not belong to the initial POI's group. The density parameter determines a limited number of candidate POIs which stay in the center candidate POI's group. The candidate POIs filtering is enforced by the skyline method domination.

- Skyline method domination: if the number of candidate POIs in a candidate POIs group is smaller or equal the density value, the system will recommend all POIs in the candidate POI group to the user for selecting next tourism POI. If the number of candidate POIs in a candidate POIs group is larger than the density value, it violates the parameter definition of density-based clustering. We must filter out some low priority POIs in order to recommend high priority POIs to a user/tourist. The domination of skyline method solves this problem. In general, after skyline method domination executed, the system will get the number of POIs which satisfy the density value to recommend a user the candidate POIs. In a special situation, skyline method cannot process multiple candidate POIs with the same priority. To solve the special situation, the proposed mechanism enforces the GA to optimize the candidate POIs selection from multiple candidate POIs with the same priority.
- Optimize candidate POIs selection by GA: GA simulates the process of a biological competition in nature. The steps of GA include initialization, selection, crossover, mutation, and termination. Based on the GA working process, the study designs a GA-based POIR algorithm. The algorithm's pseudocode is presented in Figure 3. The steps of a GA-based POIR mechanism are shown in follows.
 - Srep 1. Initialize the candidate POIs (*PointCandidate*).
 - Srep 2. Determine the number of generation (*NumberOfGeneration*) to control the generation frequency.
 - Srep 3. Get the number of crossover (*NumberOfCrossover*) from the number of candidate POIs (*PointCandidate*) multiplied by the number of crossover (*RateOfCrossover*).
 - Srep 4. Use the number of crossover (*NumberOfCrossover*) to control the crossover frequency.
 - Srep 5. Calculate the fitness value (*Fitness*) of each candidate POI.
 - Srep 6. Random select a candidate POI with higher fitness value (*Fitness*) from a population(*Population*). If a candidate POI with higher fitness value (*Fitness*), it has a higher probability to be selected, otherwise not.
 - Srep 7. Generate new candidate POIs (*Individuals*) via crossover.
 - Srep 8. If the random probability is smaller than the probability rate of mutation (*RateOfMutation*), the proposed mechanism carries out the mutation of the candidate POIs (*PointCandidate*).
 - Srep 9. Calculate new candidate POIs' (*Individuals*) fitness value (*Fitness*).
 - Srep 10. Add new candidate POIs (*Individuals*) into a population (*Population*), and replace the old candidate POIs (*Individuals*) with lowest fitness value (*Fitness*) in a population (*Population*).
 - Srep 11. If the optimal candidate POI (*BestPoint*) equals to the candidate POIs (*PointCandidate*) which has the best fitness value(*Fitness*) in a population (*Population*), the proposed mechanism returns the optimal candidate POI (*BestPoint*).

For example, in Figure 2, if candidate POI C, D, E have the same priority, the candidate POIs solution includes {Initial POI, A, B, C, F, G, Destination POI}, {Initial POI, A, B, D, F, G, Destination POI}, and {Initial POI, A, B, E, F, G, Destination POI}, the GA-based POIR mechanism will be enforced to optimize the three candidate POIs solutions to get that {Initial POI, A, B, C, F, G, Destination POI} is the best recommendation solution.

Parameter definition:

<i>PointCandidate</i>	The candidate POIs;
<i>NumberOfGeneration</i>	The number of generation in GA;
<i>Population</i>	The population of candidate POI;
<i>Fitness</i>	The value of fitness in GA;
<i>BestPoint</i>	The optimal candidate POI from GA;
<i>NumberOfCrossover</i>	The number of crossover in GA;
<i>Individuals</i>	The number of candidate POIs;
<i>RateOfCrossover</i>	The probability rate of crossover;
<i>RateOfMutation</i>	The probability rate of mutation.

Pseudo Code of GA-based POIR algorithmInput: *PointCandidate*Output: *BestPoint*POIR(*PointCandidate*)

```

{
    Generate initial population of size PointCandidate;
    for(  $i = 1; i \in \text{NumberOfGeneration}; i++$ )
    {
        NumberOfCrossover equal  $\text{PointCandidate} * \text{RateOfCrossover}$ 
        for(  $j = 1; j \in \text{NumberOfCrossover}; j++$ )
        {
            Evaluate the Fitness of the Individuals;
            Random selection Individual that have higher Fitness;
            Generate new Individuals via crossover
            If(  $\text{RandOnRange}(0.0, 1.0) < \text{RateOfMutation}$ )
                Mutate Individuals;
            End If
            Evaluate the Fitness of new Individuals;
            Add the new Individuals into Population;
            Replace the worst Individuals of the Population with new Individuals.
        }
    }
    Return BestPoint which equals the fittest individual from the population
}

```

Figure 3. Pseudocode of GA-based POIR algorithm

- Recommendation solution: After carry out the GA-based POIR mechanism, the system will get an optimal candidate POI solving the selection from multiple POIs with the same priority problem

occurred in the skyline method domination. Then the proposed system recommends the reasonable number of candidate POIs for user selecting which follows the predefined density value of the density-based clustering. The important contribution of this work is to filter out huge amount number of candidate POIs, and provide some candidate POIs with higher priority to help user making a decision. An example of a user/tourist uses the recommended solution is like {Initial POI, A, B, C, F, G, Destination POI}, is shown in Figure 2.

- Feedback and evaluation: The GA-based POIR mechanism carries out a novel algorithm. The evaluation criteria are not easy to define. Therefore, user feedback is very important to assist improving the proposed system. An evaluation process will build to verify the candidate POI recommendation effectiveness.

The experimental system implementation

This section introduces an experimental system implementation for the proposed system framework and core mechanism of GA-based POI recommendation. The study follows software development life cycle (SDLC) to build up the experimental system. We have finished the user requirement analysis and system requirement analysis. The experimental system construction is in progress.

- User requirement analysis: The study explores user requirements from some well-famous searching platform and social media, e.g., e.g., Twitter, Google Map, Foursquare, Instagram, and Youtube, to identify user tourism preferences, POI's attributes, and relevant information. The user requirement analysis helps us to define a format of a user profile, collect the key attributes and metadata of POI knowledge map, define a data schema of the knowledge repository, evaluate the software/hardware and relevant resources to construct the proposed system and carry out the experiments in the proposed system.
- System analysis: This study uses Uniform Modeling Language (UML) to enforce object-oriented system analysis. Each module's functionality will be analyzed by use case diagram, activity diagram, sequence diagram, and class diagram, etc. The database of the proposed system must satisfy the normalization forms.
- System construction: The work is in progress.

Conclusions

This study proposes an intelligent system framework using density-based clustering and genetic algorithm to recommend a point-of-interest (POI) in a geographic information service. Density-based clustering, skyline method, and genetic algorithm are composed to build a GA-based point-of-interest recommendation (POIR) mechanism. The GA-based POIR mechanism discovers and optimizes a tourism solution. The proposed system framework automatic recommends the tourism solution to a user/tourist decision making. The contribution of this work is to propose a novel system framework with the GA-based POIR mechanism which helps a user/ tourist getting a tourism POI solution from a huge amount of candidate POIs. The tourism POI solution discovered based on the user/tourist preferences. We plan to implement an experimental system using open data collected from Google map, Instagram, and Foursquare. The system construction is in progress. In future work, we will enforce collected data from several famous social media and searching platform to evaluate the proposed system framework effectiveness. Besides, this study proposes a novel intelligent POIR system which provides a user/tourist new operating model. The unified theory of acceptance & use of technology (UTAUT) will use to test the novel system's acceptance.

Acknowledgements

This research was supported in part by the Industrial Technology Research Institute and the Ministry of Science and Technology, R.O.C. with a MOST grant 106-2221-E-025-011-.

Reference

- Ankerst, M., Breunig, M. M., Kriegel, H. P., Sander, P. Sander, J. 1999, "OPTICS: ordering points to identify the clustering structure," *ACM Sigmod record*, (28:2), pp. 49-60.
- Bao, J., Zheng, Y., Wilkie, D., & Mokbel, M. 2015. "Recommendations in location-based social networks: a survey," *GeoInformatica*, (19:3), pp.525-565.
- Chang, X., & Lilly, J. H. 2004. "Evolutionary design of a fuzzy classifier from data," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, (34:4), pp. 1894-1906.
- Chao, C., Xia, C., Leye, W., Xiaojuan M., Zhu W., Kai, L., Bin, G., & Zhen Z., 2017. "MA-SSR: A memetic algorithm for skyline scenic routes planning leveraging heterogeneous user-generated digital footprints," *IEEE Transaction on Vehicular Technology*, (66:7), pp. 5723-5736.
- de Oliveira, L. L., Freitas, A. A., & Tinós, R. 2018, "Multi-objective genetic algorithms in the study of the genetic code's adaptability," *Information Sciences*, (425), pp. 48-61.
- Ester, M., Kriegel, H. P., Sander, J., & Xu, x. 1996. "A density-based algorithm for discovering in large spatial database with noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, Portland, Oregon, pp. 226-231.
- Goldberg, D. E. 1989, *Genetic algorithms in search, optimization and machine learning*, Reading, MA: Addison Wesley.
- Holland, J. H. 1962. "Outline for a logical theory of adaptive systems," *Journal of the ACM (JACM)*, (9:3), pp. 297-314.
- Holland, J. H. 1975, *Adaptation in natural and artificial systems*, Ann Arbor, MI: The University of Michigan Press.
- Mulvenna, M. D., Anand, S. S., & Büchner, A. G. 2000. "Personalization on the Net using Web mining: introduction," *Communications of the ACM*, (43:8), pp.122-125.
- Ricci, F., 2002. "Travel recommender systems," *IEEE intelligent systems*, (17:6), pp.55-57.
- Riecken, D. 2000. "Personalized views of personalization," *Communications of the ACM*, (43:8), pp.26.
- Schafer, J. B., Konstan, J. A., & Riedl, J. 2001. "E-commerce recommendation applications," *Data mining and knowledge discovery*, (5:1-2), pp. 115-153.
- Shenglin Z., Irwin K., &Michael R. L., 2016. "A survey of Point-of-interest Recommendation in Location-based social network,"*arXiv: 1607.00647v1[cs.IR]*.
- Takeshi, K., Tomoharu, I., Go, I., & Ko, F., 2010. "Travel route recommendation using geotags in photo sharing sites," in *Proceedings of the 19th ACM international conference on Information and knowledge management*, ACM, Toronto, ON, Canada, pp. 579-588.
- Yu, Z., Xu, H., Yang, Z., & Guo, B. 2016. "Personalized travel package with multi-point-of-interest recommendation based on crowdsourced user footprints," *IEEE Transactions on Human-Machine Systems*, (46:1), pp. 151-158.
- Wang, Z. S., Juang, J. F., & Teng, W. G. 2015. "Predicting poi visits with a heterogeneous information network," in *Proceedings of the 2015 Conference on Technologies and Applications of Artificial Intelligence (TAAI)*, IEEE, Tainan, Taiwan, pp.388-395.
- Wei, L. Y., Zheng, Y., & Peng, W. C. 2012. "Constructing popular routes from uncertain trajectories," in *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge discovery and data mining*, ACM, Beijing, China, pp. 195-203.
- Zhang, Z., Che, O., Cheang, B., Lim, A., & Qin, H. 2013. "A memetic algorithm for the multiperiod vehicle routing problem with profit," *European Journal of Operational Research*, (229:3), pp. 573-584.
- Zheng, Y., Zhang, L., Xie, X., & Ma, W. Y. 2009. "Mining interesting locations and travel sequences from GPS trajectories," in *Proceedings of the 18th International Conference on World Wide Web*, ACM, pp. 791-800.
- Zhiwen, Y., Huang, X., Zhe, Y. & Bin G. 2016. "Personalized travel package with multi-point-of-interest recommendation based on crowdsourced user footprints," *IEEE Transaction on Human-Machine Systems*, (46:1), pp. 151-158.