AN ADAPTIVE GENETIC ALGORITHM FOR GENERATING PERSONALIZED TRAVEL ITINERARY RECOMMENDATION

DR.B.RATNAKANTH[1]

ASSOSCIATE PROFESSOR, DEPARTMENT OF CSE, BHOJ REDDY ENGINEERING COLLEGE FOR WOMEN, VINAY NAGAR, HYDERABAD-59

GIRI SRILEKHA[2]

UG SCHOLAR, DEPARTMENT OF CSE, BHOJ REDDY ENGINEERING COLLEGE FOR WOMEN, VINAY NAGAR, HYDERABAD-59

PASHIKANTI UDAYSREYA[3]

UG SCHOLAR, DEPARTMENT OF CSE, BHOJ REDDY ENGINEERING COLLEGE FOR WOMEN, VINAY NAGAR, HYDERABAD-59

AVULA VENKATA NEHA[4]

UG SCHOLAR, DEPARTMENT OF CSE, BHOJ REDDY ENGINEERING COLLEGE FOR WOMEN, VINAY NAGAR, HYDERABAD-59

ABSTRACT

The massive growth of internet users nowadays can be a big opportunity for the businesses to promote their services. This opportunity is not only for e-commerce, but also for other e-services, such as e-tourism. In this paper, we propose an approach of personalized recommender system with pairwise preference elicitation for the e-tourism domain area. We used a combination of Genetic Agorithm with pairwise user preference elicitation approach. The advantages of pairwise preference elicitation method, as opposed to the pointwise method, have been shown in many studies, including reducing incosistency and confusion of a rating number. We also performed a user evaluation study by inviting 24 participants to examine the proposed system and publish the POIs dataset which contains 201 attractions used in this study

Keywords: Recommender system, Pairwise preferences, Genetic algorithm, Preference learning

INTRODUCTION

The rapid growth of internet users nowadays has impacted many countries in the world. This growth also occurred in Southeast Asia. In a most recent published annual report by Google, Temasek and Bain & Company on Southeast Asia's (SEA) digital economy, titled e-Conomy SEA 2020 [1], it has been revealed that there is a massive increasing amount of internet users with 40 millions of new users in the year 2020 alone. The total of active internet users in SEA has reached 400 million in 2020. The outbreak of COVID19 was considered to be the major reason for the digital market acceleration. The new users tried to use digital services to fullfil their need during the pandemic and the countries' lockdown policy. This new phenomenon can have a good impact on the country's economy as this new trend seems to be sticky. Of all those new digital internet users, it has been reported that the 94% of them intend to continue with the current service even after the pandemic situation is over. This is a good opportunity for businesses to keep improving their digital services. A smart way to improve digital businesses has been proven by big web services companies like Youtube, Amazon, Netflix, Instagram and Facebook through the implementation of the recommender systems. They are a technology to suggest the most suitable contents to their users and match with the user preferences. During the last couple of decades, it was no doubt that the implementation of this technology was able to make them stand out significantly from their competitors. In addition to that, recommender systems are not only beneficial for the businesses, but also for the users [2]. For the users, they can minimize the cost of

choosing the most suitable product in the context of internet shopping [3]. The example of costs that can be minimized is the cost of buying the wrong and unnecessary services from the wrong providers and costs of browsing. Another benefit for the users, as mentioned by Pathak et al. [4], is that the recommender systems can improve the decision making process and quality. Currently, there are many various techniques used to build the recommender systems such as collaborative filtering, content-based filering and hybrid filtering [5]. The Collaborative Filtering (CF) technique is the most well-known and most commonly implemented in industries. It works by recommending the items based on the other users with similar taste. The second most commonly used is Content-Based (CB) filtering. This technique works by recommending similar items based on the particular user's information without taking into consideration the other users. There is also a hybrid filtering that combines more than one technique of recommender system algorithms. Most of the common approaches use single point feedback as the input to the system such as rating score with a linear scale of 1 to 5 (1 for the worst and 5 for the best). Based on these ratings, the system can calculate the recommendation score for each other item. although has been very well implemented, this single point feedback approach still has some drawbacks as mentioned in [6], such as the ratings are quite personal. Even though two users have similar preferences, it is very likely they will give different ratings to the same items. The other drawback is inconsistency. Users easily forget what ratings did they give for the items with similar properties in the past, so whenever they are asked to give a rating to the similar items, they can give a different score. A user may also feel difficult to give a slightly lower preference to some items because the rating scale does not normally have a half score. Therefore, there is another approach, called pairwise elicitation, which has been introduced by the researchers in this domain which shows pair choices to the users. The problem of eliciting preference by using the pairwise method itself still become an interest for the researchers, such as in [6–13]. By using this approach, a user will be shown with a series of pair options and a preference can be expressed by selecting only one of the most preferred item between the two items (a pair) at a time. This technique can reduce the confusion of a rating number. In this paper, we want to focus on the use of the pairwise elicitation method to learn about user preferences. Some studies have also shown the combination of the pairwise elicitation with the use of CF as well as the CB filtering method in a recommender system, such as the one introduced by Liu et al. [7]. Our study will propose a different approach, i.e. using a genetic algorithm to optimize the searching strategy combined with pairwise preferences as the elicitation method. Recommender systems are very popular nowadays due to the benefit being offered as explained earlier. In an e-commerce context, they can reduce the cost of searching and finding suitable items and direct the customers to buy products they mostly liked. Similar to that, in the tourism domain, the travellers often face the problem of spending too much time browsing the possible destinations before visiting a new place which can waste their time and energy. The searching action can be more complicated in the case of the limited budget and the trip duration which can be spent by the travellers. They need to choose the destinations wisely so that they still can spend the money and time in an effective way. One possible solution is by using Google Maps and choose the destination manually. We can see there is a gap between the existing solution and the research to address this problem. Often with the same difficulties as faced by the e-commerce users, the travel application users also feel overwhelmed with many choices available. The pairwise choices can simplify the options and generate a list of recommendation even before they provide ratings to the items. This is one of the advantages of our proposed approach. In 2020, we proposed a system design of an e-tourism mobile application as a solution to promote tourism in Indonesia which has been worsened due to the COVID-19 pandemic situation, called SONIA (pariwiSata ONline Indonesi A or Indonesia Online Tourism) [14]. In this paper, we will discuss in more detail the recommender system module of the system. In summary, the contributions of this study are as follows: (1) publish a dataset of tourist attractions in Jakarta, Indonesia to be used for further study, (2) propose an approach of a trip recommender system based on a genetic algorithm with pairwise options, and (3) conduct a user evaluation study on the implemented pairwise trip recommender system with real-life data. Related work Genetic Algorithm (GA) has been around for decades, but the use of this optimization algorithm in recommender system is still limited. GA has been used in many domains with optimization and search problems. In general, a recommendation problem can also be considered as a searching problem for the best items. All the recommender system algorithms aim to give as least error as possible. GA is believed to be an effective algorithm that can provide a near-optimal solution in a reasonable time. Henceforth, there is no reason why GA is not suitable for solving a recommendation problem. In the Recommender System (RS) domain, GA has been utilized for clustering, such as the study performed by Kim and Ahn in 2008 [15], Zhang and Chan in 2006 [16], and Mohammadpour et al. in 2019 [17]. The researchers in this area also used

GA to increase the accuracy of recommendation proposal generated by classic RS algorithms, such as collaborative filtering and content-based filtering. GA works in improving the population of recommendation solutions in each iteration. A study by Kilani et al. [18] proposed the GA-based matrix factorization hybrid approach of RS. They use the approach to predict items for the active user. This is an improvement work of a study by Navgaran et al. [19]. They show that their approach can perform faster than the previous work with better recall and precision values in some datasets. Recently, Alhijawi and Kilani in 2020 [20] proposed a novel GA-based collaborative filtering that aims to select the best items which meet the active user's preferences based on multi-filtering criteria. Another recent study by Gasmi et al. in 2021 [21] also proposed a user-based collaborative filtering combined with the GA based meta-heuristic. A study by Xiao [22] proposed a combination of item-based collaborative filtering with GA which is called itemCFGA. This study also proposed a novel similarity function that uses the average rating of each user. For a hybrid RS model, GA is used in [23] and [24]. Another related work in this area has been introduced in [25] and [26]

EXISITNG SYSTEM

Many methods such as Dynamic Programming, Branch and Bound , Heuristic method have been developed for solving Travelling Salesman Problem. It is solved very easily when there is less number of cities, but as the number of cities increases it is very hard to solve, as large amount of computation time is required.

DISADVANTAGES:

- It is time taking process.
- The result obtained may not be optimum.
- Can not be used to process large datasets.

PROPOSED SYSTEM

The genetic algorithm is an instant based learning that mimics the process of natural evolution. Genetic Algorithm uses various problem solutions and after the termination of genetic algorithm, an optimal solution is obtained.

ADVANTAGES:

- Genetic algorithm is used for personalized itinerary planning for travelers to plan their itineraries better.
- The robustness of genetic algorithm is due to its capacity to locate the optimum in a multi-model landscape.
- They are readily amenable to parallel implementation.
- Using this method repeatedly population will hopefully evolve good solutions

GENETIC ALGORITHM

Genetic Algorithms(GAs) are adaptive heuristic search algorithms that belong to the larger part of evolutionary algorithms. Genetic algorithms are based on the ideas of natural selection and genetics. These are intelligent exploitation of random search provided with historical data to direct the search into the region of better performance in solution space. **They are commonly used to generate high-quality solutions for optimization problems and search problems.**

Genetic algorithms simulate the process of natural selection which means those species who can adapt to changes in their environment are able to survive and reproduce and go to next generation. In simple words, they simulate "survival of the fittest" among individual of consecutive generation for solving a problem. **Each generation consist of a population of individuals** and each individual represents a point in search space and possible solution. Each individual is represented as a string of character/integer/float/bits. This string is analogous to the Chromosome.

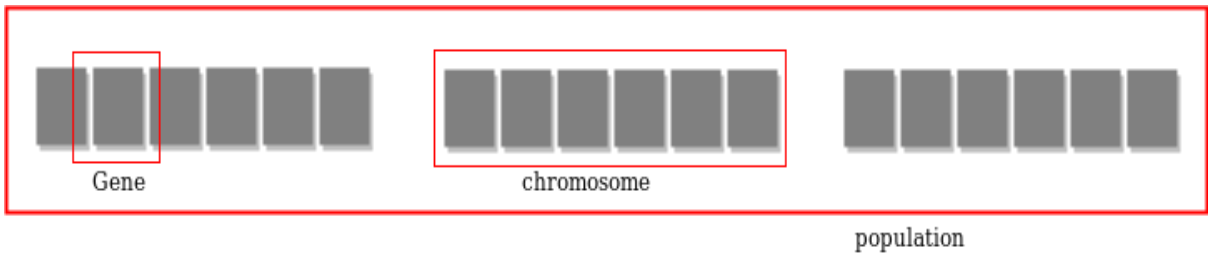
Foundation of Genetic Algorithms

Genetic algorithms are based on an analogy with genetic structure and behavior of chromosome of the population. Following is the foundation of GAs based on this analogy –

1. Individual in population compete for resources and mate
2. Those individuals who are successful (fittest) then mate to create more offspring than others
3. Genes from "fittest" parent propagate throughout the generation, that is sometimes parents create offspring which is better than either parent.
4. Thus each successive generation is more suited for their environment.

Search space

The population of individuals are maintained within search space. Each individual represent a solution in search space for given problem. Each individual is coded as a finite length vector (analogous to chromosome) of components. These variable components are analogous to Genes. Thus a chromosome (individual) is composed of several genes (variable components).



Fitness Score

A Fitness Score is given to each individual which **shows the ability of an individual to "compete"**. The individual having optimal fitness score (or near optimal) are sought.

The GAs maintains the population of n individuals (chromosome/solutions) along with their fitness scores.The individuals having better fitness scores are given more chance to reproduce than others. The individuals with better fitness scores are selected who mate and produce **better offspring** by combining chromosomes of parents. The population size is static so the room has to be created for new arrivals. So, some individuals die and get replaced by new arrivals eventually creating new generation when all the mating opportunity of the old population is exhausted. It is hoped that over successive generations better solutions will arrive while least fit die.

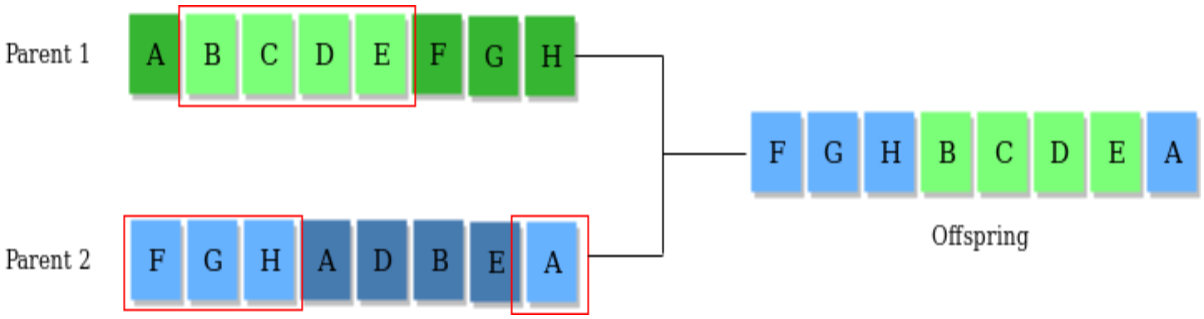
Each new generation has on average more "better genes" than the individual (solution) of previous generations. Thus each new generations have better **"partial solutions"** than previous generations. Once the offsprings produced having no significant difference than offspring produced by previous populations, the population is converged. The algorithm is said to be converged to a set of solutions for the problem.

Operators of Genetic Algorithms

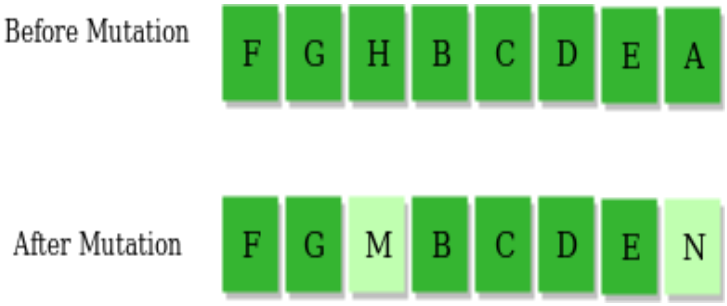
Once the initial generation is created, the algorithm evolve the generation using following operators –

1) Selection Operator: The idea is to give preference to the individuals with good fitness scores and allow them to pass their genes to the successive generations.

2) Crossover Operator: This represents mating between individuals. Two individuals are selected using selection operator and crossover sites are chosen randomly. Then the genes at these crossover sites are exchanged thus creating a completely new individual (offspring). For example –



3) Mutation Operator: The key idea is to insert random genes in offspring to maintain the diversity in population to avoid the premature convergence. For example –



The whole algorithm can be summarized as –

1) Randomly initialize populations p

2) Determine fitness of population

3) Untill convergence repeat:

a) Select parents from population

b) Crossover and generate new population

c) Perform mutation on new population

d) Calculate fitness for new population

CONCLUSION

In this paper, we proposed an adaptive genetic algorithm for personalized itinerary planning for travelers. Firstly, desired starting POIs and destination POIs were considered in our approach. Secondly, we also took some general factors into account that travelers would consider in their preferences of an itinerary, which are mandatory POIs, the total number of POIs, the overall POI popularity, the overall cost, and the overall rating. Thirdly, we viewed this kind of recommendation task as a Multi-Objective Optimization problem, and we proposed the AGAM for solving this problem, which is based on an adaptive genetic algorithm with the crossover and mutation probabilities to better find the best global solution. Fourthly, we allocated different weights to every factor for generating the personalized itinerary planning to better meet many kinds of preferences of tourists. Finally, we compared our approach against baselines on real-world datasets which include six touristic cities, and the experimental results showed that the AGAM outperforms better than baseline methods in terms of the mandatory POIs, total POI visits, overall POI popularity, total travel time (including travel time and visit duration), overall cost, and overall rating. We simply allocated a larger integer to each factor to represent the users' preference, a better weight allocation rule could be employed in the

future. Also, we will consider the various modes of transport and sentiments such as opinions and reviews for user preferences and route planning.

REFERENCES

[1] P. Yochum, L. Chang, T. Gu, and M. Zhu, ‘‘Linked open data in locationbased recommendation system on tourism domain: A survey,’’ IEEE Access, vol. 8, pp. 16409–16439, 2020.
[2] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera, ‘‘Personalized trip recommendation for tourists based on user interests, points of interest visit durations and visit recency,’’ Knowl. Inf. Syst., vol. 54, no. 2, pp. 375–406, Feb. 2018.
[3] Y. Yu, Y. Zhao, G. Yu, and G. Wang, ‘‘Mining coterie patterns from Instagram photo trajectories for recommending popular travel routes,’’ Frontiers Comput. Sci., vol. 11, no. 6, pp. 1007–1022, Dec. 2017.
[4] R. Gaonkar, M. Tavakol, and U. Brefeld, ‘‘MDP-based itinerary recommendation using Geo-tagged social media,’’ in Proc. IDA, 2018, pp. 111–123.
[5] K. Taylor, K. H. Lim, and J. Chan, ‘‘Travel itinerary recommendations with must-see points-of-interest,’’ in Proc. Companion Web Conf. Web Conf. (WWW), 2018, pp. 1198–1205.
[6] A. Fogli and G. Sansonetti, ‘‘Exploiting semantics for context-aware itinerary recommendation,’’ Pers. Ubiquitous Comput., vol. 23, no. 2, pp. 215–231, Apr. 2019.
[7] K. H. Lim, J. Chan, S. Karunasekera, and C. Leckie, ‘‘Personalized itinerary recommendation with queuing time awareness,’’ in Proc. 40th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr., Aug. 2017, pp. 325–334.
[8] X. Peng and Z. Huang, ‘‘A novel popular tourist attraction discovering approach based on geo-tagged social media big data,’’ ISPRS Int. J. GeoInf., vol. 6, no. 7, p. 216, 2017.
[9] P. Bolzoni, F. Persia, and S. Helmer, ‘‘Itinerary planning with category constraints using a probabilistic approach,’’ in Proc. DEXA, vol. 2, 2017, pp. 363–377
[10] P. Padia, B. Singhal, and K. H. Lim, ‘‘User-relative personalized tour recommendation,’’ in Proc. IUI Workshops, 2019, pp. 1–6.
[11] X. Wang, C. Leckie, J. Chan, K. H. Lim, and T. Vaithianathan, ‘‘Improving personalized trip recommendation by avoiding crowds,’’ in Proc. 25th ACM Int. Conf. Inf. Knowl. Manage. (CIKM), 2016, pp. 25–34.
[12] D. Jiaoman, L. Lei, and L. Xiang, ‘‘Travel planning problem considering site selection and itinerary making,’’ in Proc. Conf. Res. Adapt. Convergent Syst. (RACS), 2018, pp. 29–36.
[13] Y. Zhang and J. Tang, ‘‘Itinerary planning with time budget for risk-averse travelers,’’ Eur. J. Oper. Res., vol. 267, no. 1, pp. 288–303, May 2018.
[14] D. Nurbakova, L. Laporte, S. Calabretto, and J. Gensel, ‘‘Itinerary recommendation for cruises: User study,’’ in Proc. RecTour RecSys, 2017, pp. 31–34.
[15] H. Liu, C. Jin, and A. Zhou, ‘‘Popular route planning with travel cost estimation from trajectories,’’ Frontiers Comput. Sci., vol. 14, no. 1, pp. 191–207, Feb. 2020.

