JCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Volume-3 Issue- 4 E-ISSN: 2347-2693

Review paper on Automatic Itinerary Planning for Traveling Services

Bhosale Pallavi[1*], Holkar Sarika [2], Takale Shailaja [3], Divekar Shweta[4] and Ambole Rajaram [5]

[1234] *Department of Computer Engineering, SP Pune University, India*

www.ijcseonline.org

Received: Mar/30/2015 Revised: Apr/09//2015 Accepted: Apr/22/2015 Published: Apr/30/ 2015

Abstract— The trip planning is very difficult task for peoples, for the places which are not known. So creating an efficient and economic trip plan is the most essential. Although some travel agency can provide some predefined planning of days, which is not suitable for a specific or individual customer**.**

A route search is an enhancement of an ordinary geographic search. Instead of merely returning a set of entities, the result is a route that goes via entities that are relevant to the search [2].For Constructing Travel Itineraries from Tagged Geo Temporal Breadcrumbs develops an end-to-end approach for constructing intra-city travel itineraries automatically by tapping a latent source reacting geo-temporal breadcrumbs left by millions of tourists[3]. A nature-pushed figuring ,imitating the without any preparation creation system for music players, has been beginning late made and named Harmony Search (HS)[4].The extraction of Event and Place Semantics from Flickr Tags that contains, Tags usually manifest in the form of a freely-chosen, short list of keyword associated by a user with a resource such as a photo, web page, or blog entry[5].

Keywords— *Harmony Search, Generlized Orienting Problem, Tag Semantics*

I. INTRODUCTION

It is fun to travel but painful to arrange the trip. Travel itinerary planning is often it difficult and time consuming task of for traveler visiting a location for first time when we start planning, flights accommodation and attraction are spread across multiple website on the internet. We spend time in searching each of them for the best deals and get the attraction reviews from established the performers in the market wouldn't it be good if we just specify the destination and some platform magically did everything for us and gave out itinerary to us? So a travel agent needs more planning as consider all the points or details of customer. Such as time factor , money, hotel selection, list of point of interest (POI) etc. Therefore there is need for aromatic planning for traveling services to attract more customers. Customer gives the list of interested POI's and specifies the time, money, and budget**.**

Yaron Kanza ,Roy Levin ,Eliyahu Safra described method of Interactive Route Search in the Presence of Order Constraints[2] that contains, A route search is an enhancement of an ordinary geographic search. Instead of merely returning a set of entities, the result is a route that goes via entities that are relevant to the search. The main task is to find a route that visits at least one satisfying entity of each type. Interactive route search in the presence of order constraints that specify some types of entities should be visited before others.

Munmun De Choudhury, Moran Feldman, Sihem Amer-Yahia described method Constructing Travel Itineraries from Tagged Geo-Temporal Breadcrumbs[3] that contains, Vacation planning is a frequent laborious task which requires skilled interaction with a multitude of resources. For Constructing Travel Itineraries from Tagged Geo-Temporal Breadcrumbs develops an end-to-end approach for constructing intra-city travel itineraries automatically by tapping a latent source reacting geo-temporal breadcrumbs left by millions of tourists. In particular, the popular rich media sharing site, Flickr, allows photos to be stamped by the date and time of when they were taken.

Zong Woo Geem, Chung-Li Tseng, and Yongjin Park gives the Detail about Harmony search algorithm Generalized Orienteering Problem by using harmony search algorithm and Generalized Orienteering Problem[4]. To defeat the disadvantages of numerical change strategies, soft computing estimations have been energetically introduced in the midst of the earlier decade. Regardless, there are still a couple of possible results of imagining new figuring in perspective of analogies with regular phenomena. A nature-pushed figuring ,imitating the without any preparation creation system for music players, has been beginning late made and named Harmony Search (HS).

Tye Rattenbury, Nathaniel Good and Mor Naaman described method of Automatic Extraction of Event and Place Semantics from Flickr Tags that contains[5], Tags usually manifest in the form of a freely-chosen, short list of keyword associated by a user with a resource such as a photo, web page, or blog entry. Unlike category- or ontology-based systems, tags result in unstructured knowledge they have no a-priori semantics. However, it is precisely the unstructured nature of tags that enables their utility. For example, tags are probably easier to enter than picking categories from an ontology; tags allow for greater flexibility and variation; and tags may naturally evolve to reflect emergent properties of the data.

II. Problem Defination

In general trip planning it is not possible to generate trip planning according to single user preference and within the specified time. Although the travel agencies provide efficient and convenient services, for experienced travelers, the itineraries provided by the travel agents lack customization and cannot satisfy individual requirements.

.

III. Methods

A. Interactive Route Search in the Presence of Order Constraints:

A route search is a development of an usual geographic search.[2] The input to the problem consists of a number of search queries, and each query defines a characterizes a kind of geological entities. This paper investigates interactive route search in the existence of order constraints that specify that some types of entities should be visited before others. The geographic entities are considered as the user needs and are specified by search queries.

Interactive route search in the presence of order constraints, for two cases. In one case, the constraints define a complete order over the types of entities that should be visited, and in the other they define a partial order. For each case, there are three algorithms, having two goals: computing an effective route (i.e. a route that is as short as possible) and doing it efficiently (i.e. finding the next object on the route as quickly as possible). The Greedy algorithm is the most efficient, yet the route it computes is the least effective.The MED algorithm, that provides the most effective route; however, its efficiency is the lowest. The Optimistic algorithm that provides a route with effectiveness and efficiency that are between those of MED and Greedy. The differences between the running times of the three algorithms are just in the preprocessing phase. The time needed to find the next object is about the same in all of them (less than 1 millisecond). If efficiency is important, then the best may be a hybrid approach that determines the first object using the Greedy algorithm, and then switches to the MED (or Optimistic) algorithm in order to find subsequent objects. The time it takes the user to get to the first object is more than enough for completing the preprocessing. Thus, the hybrid approach is both efficient and effective.

B. Constructing Travel Itineraries from Tagged Geo-Temporal Breadcrumbs

This paper[3] builds up an end-to-end approach for building Inter City travel schedules consequently by drawing from a dormant source re-eating geo-transient breadcrumbs left by a large number of tourists.

In this paper, we build up a way to naturally develop travel agendas at a huge scale from photographs transferred by users. All the more particularly, by examining floods of photographs taken by users, one can deduce the cities visited by a person, which POIs that individual took photographs at, to what extent that individual spent at every POI, and what the travel time was between POIs went by in progression.

The primary step of our methodology is to change over the raw user photographs into single person timed ways for a given city. Instinctively, these ways, which join different POIs, are built from individual photograph streams and depict the developments of individual tourists.

This paper addressed the problem of automatic generation of travel itineraries for popular touristic cities from large-scale user contributed rich media repositories. We plan to explore many directions such as applying different filtering and aggregation techniques to accommodate different types of travelers, and constructing "off the beaten track" itineraries that cater to niche audiences rather than mainstream crowds.

C. Harmony Search For Generlized Orienting Problem:

1) Harmony Search Algorithm:

Harmony Search (HS) calculation was as of late grown in a similarity with music act of spontaneity methodology where music players Improvise the pitches of their instruments o acquire better harmony [7].The HS algorithm has been effectively connected to different benchmarking and real world issues including travelling salesman problem [7], parameter optimization of river flood model [8], c design of pipeline network [9], and design of truss structures [10]. Thus, the HS calculation gives a probability of achievement in a TSP like NP-hard issue. According to the above algorithm concept, the steps of HS for the generalized orienteering problem are as follows:

Step1. Initialize the Parameters for Problem and Algorithm.
Step2. Initialize the Harmony Memory (HM).
Step3. Improvise a New Harmony.
Step4. Update the Harmony Memory.
Step5. Check the stopping criterion.

2) Generalized Orienteering Problem

On this research, HS is usually applied to generalized orienteering problem (GOP). The target connected with GOP is usually to search for the optimum trip beneath the concern connected with entire distance limit although rewarding several targets.

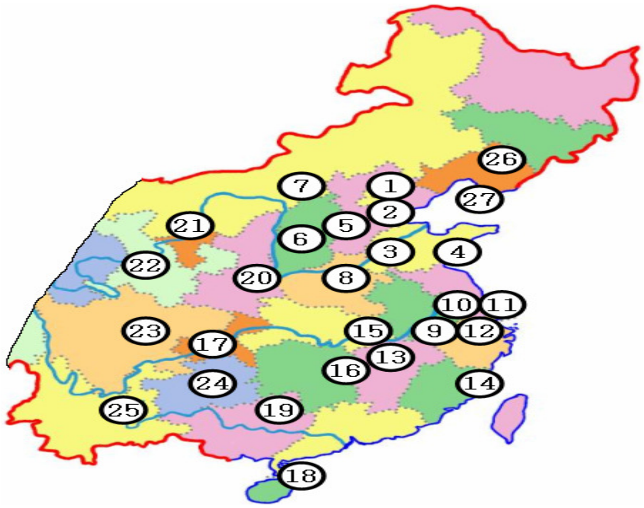

Fig. 1. Map of 27 cities in eastern part of China[4]

In the event that an traveler visits eastern piece of China, as indicated in Figure 1, and he/she needs to go however many urban areas as could reasonably be expected with the motivation behind best satisfying numerous components such as 1) natural beauty, 2) historical interest, 3) cultural event, and 4) business opportunities under the restricted aggregate moving separation, his/her travel can Generalized orienteering issue where every city has particular measured scores for all components and the estimation of a visit is performed in view of the summation of those scores in the visit. The GOP is a generalization of the orienteering issue (OP) and the principle distinction between the two is that every city in GOP has numerous scores while every city in Operation has one and only scores [14-16].

Let V be the set of N focuses and E the set of edges between focuses in V. G = {V, E} is a complete diagram. Every edge in E has a symmetric, non-negative expense d(i, j) which turns into the separation or travel time between point i and j. Expect the beginning stage is point 1 and the end point will be point N. Every point i in V has a non-negative score vector T S(i) = (S1(i),S2 (i),K, Sm (i)) , where m is the quantity of individual objectives, and Sg (i) is the score of point i concerning objective g. A differentiable objective function that defines total score of a path P, which starts at point 1 and ends at point N can be formulated as follows[4]:

$$Z=\sum_{g=1}^{m} W_g\left[\left\{\sum_{i\in P}\left[S_g(i)\right]^k\right\}^{1/k}\right].$$

Where *Wg* is the weight of goal g, and the exponent *k* is set to 5 in this problem.

Table 1 presents city data such as city number, longitude, latitude, and score vector.$S1$, $S2$, $S3$, and $S4$ are the scores approximately scaled from 1 to 10 in the aspects of natural beauty, historical interest, cultural event, and business opportunities, proposed by Wang et al. [6].

Table 1. Physical location and score vector for each city

No.	Name	Longitude	Latitude	S_1	S_2	S_3	S_4
1	Beijing	116.40	39.91	8	10	10	7
2	Tianjin	117.18	39.16	6	5	8	8
3	Jinan	117.00	36.67	7	7	5	6
4	Qingdao	120.33	36.06	7	4	5	7
5	Shijiazhuang	114.50	38.05	5	4	5	5
6	Taiyuan	112.58	37.87	5	6	5	5
7	Huhehaote	111.70	40.87	6	6	5	5
8	Zhengzhou	113.60	34.75	5	6	5	5
9	Huangshan	118.29	29.73	9	3	2	2
10	Nanjing	118.75	32.04	7	8	8	6
11	Shanghai	121.45	31.22	5	4	9	9
12	Hangzhou	120.15	30.25	9	8	7	6
13	Nanchang	115.88	28.35	7	6	5	5
14	Fuzhou	119.30	26.10	6	5	5	7
15	Wuhan	114.30	30.55	6	6	8	6
16	Changsha	113.00	28.20	6	6	6	5
17	Guangzhou	113.15	23.15	6	6	5	10
18	Haikou	110.35	20.02	7	3	4	8
19	Guilin	110.29	25.28	10	4	4	4
20	Xi'an	108.92	34.28	5	9	8	6
21	Yinchuan	106.27	38.48	5	7	5	5
22	Lanzhou	103.80	36.03	7	6	5	6
23	Chengdo	104.07	30.66	6	7	6	5
24	Guiyang	106.00	26.59	8	5	4	5
25	Kunming	102.80	25.05	9	7	7	6
26	Shenyang	123.40	41.80	5	8	5	6
27	Dalian	121.60	38.92	7	5	6	7

D. *Event/Place Identification*

The goal of this method is to determine, for each tag in the dataset, whether the tag represents an event, and whether the tag represents a place. The intuition behind the various methods we present is that an event (or place) refers to a specific segment of time (or region in space).This method first present adaptation of two well-known techniques to the problem . Then present a new method for event and place identification: Scale-structure Identification.

1) *Borrowed Methods*

The steps for the modified, burst-detection methods are the following:

1. Scale Specification – Choose an ordered set of scale values, R = {rk | k = 1 . . .K, rk1 > rk2 () k1 > k2}. Generally choose an exponentially increasing set of scales (e.g., rk = _k for some _ > 1.0).

2. Segment Specification – For each scale rk define a finite set of time segments to search over, say Yrk . We use a regularly spaced grid where the grid size is based on the scale; but overlapping or arbitrary segments are possible.

3. Partial Computation – For each scale rk and each time segment in Yrk, compute a statistic on Tx that captures some aspect of the tag's usage pattern in time (likely, although not necessarily, based on some relationship between the usage occurrences within a time segment versus outside of the segment).
4. Significance Test – Aggregate the partial computation statistics for each time segment at each scale to determine Whether tag x is an event.
5. Identify Significant Segments – Provided a significant pattern for x is found, determine which scales and time segments correspond to the event.

1 Naive Scan Methods:
The Naive Scan methods are an application of a standard burst detection method used in signal processing [11]. The method computes the frequency of usage for each time segment at each scale. The method identifies a "burst" when the frequency of data in a single time segment is larger than the average frequency of the data over all segments plus two times the standard deviation of segment frequencies.

2 Spatial Scan Method:
The Spatial Scan methods are a standard application of the Spatial Scan statistic [12], a burst detection method used in epidemiology. These methods assume an underlying probability model of observing some phenomenon over some domain. The methods then test whether the number of occurrences of a phenomenon in a segment of the domain (e.g. segment of time) is abnormal relative to the underlying probability model. This abnormality test is performed for each segment.

2)Scale-structure Identification:
The Scale-structure Identification method performs a significance test (Step 4 above) that depends on multiple scales simultaneously and does not rely on a-priori defined time segments. Accordingly, the Scale-structure Identification method performs all the steps listed above except the Segment Specification step (Step 2). The key intuition behind Scale-structure Identification is the following: if tag x is an event then the points in Tx, the time usage distribution, should appear as a single cluster at many scales. The clustering mechanism used in Scale structure Identification is similar to the clustering mechanism in the scale-space method developed by Witkin [13].
However, whereas Witkin was interested in any structure that exhibited robustness over a range of scales, we are interested in the robustness of a single type of structure – a single cluster.
Consider the graph over Tx where edges between points exist if and only if the points are closer together than r(recall that r is the scale variable). Let Yr be the set of connected subcomponents of this graph. The Partial Computation step (Step 3 above) computes the entropy of Yr for each scale r. Specifically, the partial computation statistic is defined as

$$E_r \triangleq \sum_{Y\in\mathcal{Y}_r} (|Y|/|\mathcal{T}_x|)\log_2(|\mathcal{T}_x|/|Y|).$$

We use the entropy value as a measurement of how similar the data is to a single cluster since entropy increases as data becomes more distributed. We are interested in low entropy structures, Yr (note that Er = 0 when the usage distribution is a single cluster, i.e. |Yr| = 1). For place identification we simply replace Tx with Lx in the calculation of Er (we compute the distance between points in Lx as the L2 distance between the points as they lie on a sphere).
A caution to the partial computation statistic concerns periodic events. Periodic events have strong clusters, at multiple scales, that are evenly spaced apart in time. Practically, because tags occur in bursts, we also require that a periodic tag exhibit at least three strong clusters (to rule out tags that just happened to occur in two strong temporal clusters but are not truly periodic). Of course, this assumption could result in some false negatives (e.g. recurring events that only appear twice in our dataset), but it is necessary due to the sparse nature of our data (to mitigate these false negatives we could check whether the two strong clusters were spaced apart at some culturally meaningful distance like one month, one year, etc.).
We check for periodic events by:
(1) identifying "strong"clusters (i.e. clusters that contain at least 2% of the data),
(2) measuring how far apart the strong clusters are.
(3)making sure the cluster variances are not too big relative to the distances between clusters (i.e. the standard deviations
of the usage distributions for each cluster should be, on average, smaller than 10% of the average inter-cluster distance), and
(4) making sure the distances between clusters are "even" (i.e. the standard deviation of inter-cluster distances is smaller than 10% of the average inter-cluster distance). If a tag's temporal distribution passes all of these tests2, we re-compute the scale structure for this tag by treating time as modulo μ, the average inter-cluster distance. Specifically, we re-compute Yr from

$$\mathcal{T}_x' \triangleq \{t \text{ modulo } \mu \mid t \in \mathcal{T}_x\}$$

using the distance metric :

$$\|| t_1 - t_2 \|| \triangleq \min(|t_1 - t_2|, |\mu + t_1 - t_2|, |\mu + t_2 - t_1|).$$

Intuitively, this modulo adjustment to the time dimension aligns the "strong" clusters so that they will be treated as a single cluster. For example, if a tag's temporal distribution has 3 strong clusters that are on average 365 days apart, the modulo adjustment to time corresponds to the cyclical calendar year. Finally, the significance test calculation (Step

4) aggregates the partial computation statistics simply by summing them over the set of scales:

$$\sum_{k=1}^{K} E_{r_k}.$$

This summed value is tested against a threshold to determine if the tag is an event. By recording the scale structures at each scale, we can determine which time segments strongly characterize an event tag (Step 5). In fact, we can then characterize the tag, or rather the event it refers to, at multiple scales.

IV. CONCLUSIONS

We studied the problem of interactive route search in the presence of order constraints, for two cases. In one case, the constraints define a complete order over the types of entities that should be visited, and in the other they define a partial order[2].The question of automatic generation of travel itineraries for popular touristic cities from large-scale user contributed rich media repositories[3].The algorithm, HS, mimics three major behaviors of music players: 1)memory consideration; 2) pitch adjustment; and 3) random choice. These behaviors can be successfully translated in GOP: 'memory consideration' becomes that HS chooses any one city from the cities stored in HM; 'pitch adjustment' is that HS chooses the nearest city as next city; and 'random choice' is that HS chooses any one city from all the possible cities. After applied to GOP, HS could find equal or better solutions when compared with those of others[4]. We have taken a first step in showing that semantics can be assigned to free-form tags using the usage distribution of each tag. The ability to extract semantics can improve current tagging systems, for instance, by allowing more powerful search and disambiguation mechanisms. Additionally, the knowledge that these methods extract can help with tasks that outside the scope of the specific system [5].

V. REFERENCES

[1] Automatic Itinerary Planning for Traveling Services Gang Chen, Sai Wu, Jingbo Zhou, and Anthony K.H. Tung.

[2] R. Levin, Y. Kanza, E. Safra, and Y. Sagiv, "Interactive Route Search in the Presence of Order Constraints," Proc. VLDB.

[3]C.-H. Tai, D.-N. Yang, L.-T. Lin, and M.-S. Chen, "Recommending Personalized Scenic Itinerary with Geo-Tagged Photos," Proc.IEEE Int'l Conf. Multimedia and Expo (ICME), pp. 1209-1212, 2008.

[4] Zong Woo Geem, Chung-Li Tseng, and Yongjin Park: Harmony Search for Generalized Orienteering Problem: Best Touring in China.

[5] T. Rattenbury, N. Good, and M. Naaman, "Toward Automatic Extraction of Event and Place Semantics from Flickr Tags," Proc.30th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR '07), pp. 103-110, 2007.

[6]. Wang, Q., Sun, C., and Golden, B. L.: Using Artificial Neural Networks to Solve Generalized Orienteering Problems. Proceedings of Artificial Neural Networks in Engineering Conference (ANNIE '96). (1996).

[7]. Geem, Z. W., Kim, J. H., and Loganathan, G. V.: A New Heuristic Optimization Algorithm: Harmony Search. Simulation. 76(2) (2001) 60-68.

[8]. Kim, J. H., Geem, Z. W., and Kim, E. S.: Parameter Estimation of the Nonlinear Muskingum Model using Harmony Search. Journal of the American Water Resources Association. 37(5) (2001) 1131-1138.

[9]. Geem, Z. W., Kim, J. H., and Loganathan. G. V.: Harmony Search Optimization: Application to Pipe Network Design. International Journal of Modelling and Simulation. 22(2) (2002) 125-133.

[10]. Kang, S. L., and Geem, Z. W.: A New Structural Optimization Method Based on the Harmony Search Algorithm. Computers and Structures. 82(9-10) (2004) 781-798.

[11] M. Vlachos, C. Meek, Z. Vagena, and D. Gunopulos. Identifying similarities, periodicities and bursts for online search queries. In Proc. SIGMOD, p 131–142. ACM, 2004.

[12] M. Kulldorff. Spatial scan statistics: models, calculations, and applications. In Scan Statistics and Applications, p 303–322,1999.

[13] A. Witkin. Scale space filtering. In Proc. Int'l Joint Conf. Artificial Intelligence, p 1019–1022, 1983.

