Passenger Travel behavior Model in Railway Network Simulation

Kulkarni Gauri Ramakant
M. B. E. Society College of Engineering, Ambejogai
Maharashtra, India

B. M. Patil
M. B. E. Society College of Engineering, Ambejogai
Maharashtra, India

ABSTRACT
People usually travel to the same destination and same purpose together with the other people in groups.Inferring the travel purpose of passenger groups help us to better understand passengers and bring meaningful changes for personalized travel service.First we construct cotravel network by extracting social relations between passengers from their historical travel records that are available in passenger information system.We generate series of sophisticated features for each passenger group and use the overlapping relation between passenger groups to capture relations.At last we collectively infer the labels of all the groups in iterative way.

Keywords
Collective inference, cotravel networks, iterative classification.

1. INTRODUCTION
People usually travel to the same destination together with the other people in groups.Group based travelling for a large portion of people travels[1],[2] is very different from individual travel in terms of its nature[3].

All the members in a travel group usually book transportation tickets together, leave for same destination and having same purpose. Different purpose of groups are like business, sightseeing or visiting relatives.Travel purpose of group should be single and clear.Inferring the travel purposes of passenger groups may help the carriers or airports provide precise and personalized services or recommendations for passengers; for example, carriers can provide meeting places for business groups and recommend travel routes for tourist groups.

If two passengers once appeared together in one or more tourist groups in history, then the probability that they appeared together later in another tourist group becomes higher. Traditional classifiers neglect this fact, thus leading to the inaccuracy of results. Consequently, we need to simultaneously decide on the labels (i.e., the travel purposes) of all the passenger groups together, rather than classify each group separately, which is so called collective inference [18], [19].

2. LITERATURE SURVEY
For performing the collective inference and generating more sophisticated features for passenger groups, we need to build relations between passengers and construct passenger social network.

As a matter fact, social networks provide a perspective for enterprise to better understand their customers and have attracted substantial attention in industry. Many previous works of social network analysis have been presented of all kinds of social networks such as online social networks[7], email social networks[8], mobile social network, academic social netwok[10] and terrorist social network[10].From the cotravel network, the

collective inference method is proposed. We find out relationship between passenger groups for finding the overlaps. After that counting of overlapping edges occurring in different groups.

By appling the method on data set,it demonstrates that iterative classification method can efficiently infer the travel purpose of passenger groups in cotravel network.

3. PROBLEM STATEMENT
Some necessary definitions are given as follows:

Definition 1-Passenger Group:

A passenger group $g = (P_g , ori_g , dst_g , dis_g , d\_date_g , r\_date_g)$ is a collection of pas-sengers who book tickets together for a certain travel, where $P_g = \{p_i\}$ is a set of passengers, ori_g and dst_g are the common origin and destination of the travel, dis_g is the mileage of the travel, and d_date_g and r_date_g are the departure date and return date of the travel, respectively. Obviously, we have $|P_g| \geq 1$, where $|P_g| = 1$ means a passenger travels alone.

In practice, passenger carriers record the data of all their passenger groups in their information systems.

Given a large set of passenger groups

$S = \{g_1, g_2, \ldots, g_n\}$ in a period of time, our goal is to infer the travel purpose of each passenger group in the set, from all kinds of information mentioned in Definition 1.

Problem 1—Inferring the Travel Purposes of Passenger Groups: Given a travel record set $S = \{g_1, g_2, \ldots, g_n\}$, the objective is to learn a function such as following equation

$$f : S \rightarrow L \quad \ldots\ldots\ldots (1)$$

where L is the label space of the problem, i.e., all the possible types (travel purposes) of passenger groups. Our goal in this paper is to classify each passenger group into one of the categories: tourist group, business group,educational group.

Firstly we need to construct a cotravel network from a given set of travel records, by extracting the relations between passengers in each group.

Definition of cotravel networks as follows.

Definition 2—Cotravel Networks:

A cotravel network is a graph $G = (V, E)$, where V is a node set, and each node $p_i \in V$ represents a passenger; E is an edge set, and each edge $e_{ij} \in E$ indicates that passengers p_i and p_j have traveled together at least once.

Algorithm for Constructing Cotravel Network is given as

follows in the Table I in Algorithm1[9].

TABLE I
Algorithm I

Algorithm 1.Constructing Cotravel Network

Input: $S = \{g_1, g_2, \ldots, g_n\}$
Output: $G = (V, E)$
for each passenger group $g \in S$
for each passenger pair(p_i, p_j)($p_i, p_j \in p_g, p_i \neq p_j$)
if p_i V then
$V \leftarrow V \cup \{p_i\}$;
end if;
if p_i V then
$V \leftarrow V \cup \{p_j\}$;
end if;
if $e_{ij} \notin$ E then
$w_{ij} = 1$;
$E \leftarrow E \cup \{e_{ij}\}$;
else
$w_{ij=} w_{ij} + 1$;
end if;
end for;
end for;

4. PROPOSED WORK

For the problem of inferring travel purposes in detail. First, we construct cotravel networks from historical travel records. Then, various features of passenger groups, including some basic features and network-

based features, are constructed for classification. Finally, we propose a novel iterative classification method.

A. Constructing Cotravel Networks

The cotravel network is constructed where edges represent social relationship between passengers.

Given a set of travel records $S = \{g_1, g_2, \ldots, g_n\}$, we simply construct a cotravel network G = (V, E) by extracting relations from each passenger group, as outlined in Algorithm 1 in Table I.

B. Generating Features

After constructing cotravel network we generate series of features for the classification of passenger groups.

Features are divided into two categories:

1.Basic Feature.

2.Network-Based Feature.

1.Basic Feature:The basic features of passenger groups are some simple charecteristics that can be directly generated from the groups themselves without cotravel networks including the holistic attributes of the current travel, the demographic characteristics of group members, and the historical travel statistics of group members.

Holistic group attributes: Different types of passenger groups may have different scales of group sizes, travel distances, and travel durations. In addition, whether the travel is on holidays and whether it has a return trip are also very important attributes.

Furthermore, the seat proximity, which can reflect the closeness between the group members, may also help us identify the type of a passenger group.

Demographic characteristics of group members:Demographic charecteristics consists of average age,age variance,ratio of minors,ratio of seniors,ratio of adults,ratio of male.these compositions are useful for inferring the travelling purpose of passengers.

Historical travel statistics of group members:It includes total mileage,average mileage,total travel times,average travel times.

2.Network-Based Feature.

All the given basic features are directly generated from a passenger group itself. Network based feature includes edge density, average weight of edge, ratio of components, average clustering coefficient, average embeddedness

5. IMPLEMENTATION

We propose an iterative classification algorithm(ICA)[9] in Table 2 in Algorithm 2

Table 2

Algorithm 2

Algorithm 2.Iterative Classification Algorithm

Input: $S=\{g_1, g_2, \ldots, g_n\}$;
$G = (V, E)$;
F;//F ie set of features
Output: $P(Y)=\{P(_{y1}), P(_{y2}),\ldots\ldots\ldots P(_{yn})\}$;
for each edge $e_{ij} \in G$
for each possible label value $l \in L$
$p^{(0)}(y_{ij}=l)=1/|L|$;
end for;
end for;
//Initiation
repeat
for each passenger group $g \in S$
for each possible label value $l \in L$

$$p^{(t)}(y_g=l)= \frac{\sum_{e_{ij} \in E_g} p^{(t-1)}(y_{ij} = l)}{| E_{ij} |};$$

$P^{(t)}(y_g=l)=f(F,p^{(t)}(y_g=l))$;
end for;
end for;
for each edge $e_{ij} \in G$
for each possible label value $l \in L$

$$p^{(t)}(y_{ij}=l)= \frac{\sum_{G_g : E_g \ni e_{ij}} p^{(t-1)}(y_g = l)}{w_{ij}};$$

end for;
end for;
until $| p^{(t)}(y_g = l) - P^{(t-1)}(y_g = l) | < \xi;$
// ξ is threshold

6. EXPERIMENTS

A.Data set

We collected passenger group data set. The data set consists the information of passengers(e.g. passenger ID,passenger name,address,city,gender,Mobile-no,E-mail ID,category).

B. Experimental Results

To evaluate the contributions of different types of features and the performance of our proposed ICA, we perform three sets of experiments:

1) ICA with the basic features.
2) ICA with the network features.
3) ICA with the basic and network-based features.

1) ICA with the basic features:

The graphical representation of ICA basic feature with different groups is given as follows:

The Demographic feature analysis with business group is given in following Fig 1.

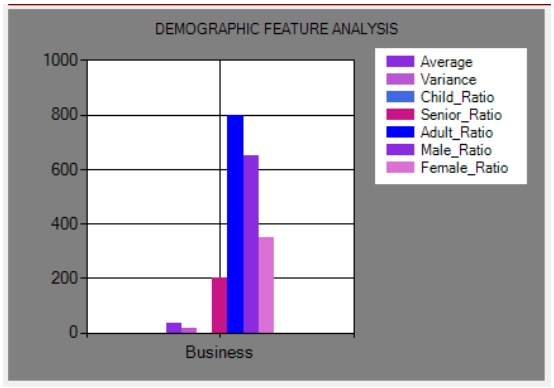

Fig 1:Demographic Feature analysis with Business Group.

The Demographic feature analysis with tourist group is given in following Fig 2.

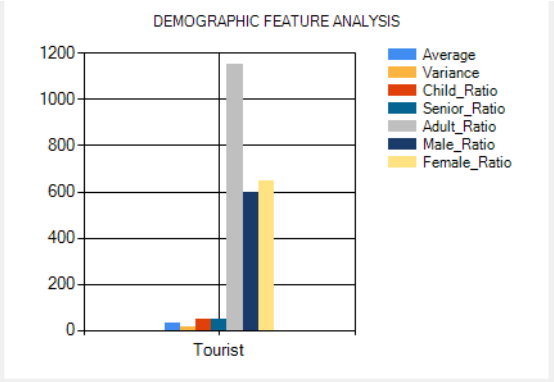

Fig 2:Demographic Feature analysis with Tourist Group.

The Demographic feature analysis with tourist group is given in following Fig 3

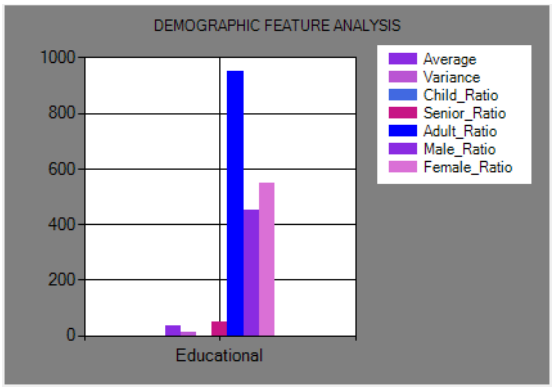

Fig 3:Demographic Feature analysis with Educational Group.

2) ICA with the network features:

The graphical representation of ICA network feature with different groups is given as follows .

The Network Feature analysis with Business group is given in Fig 4..

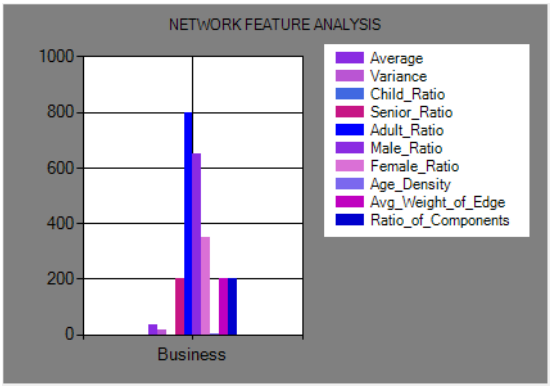

Fig 4:Network Feature analysis with Business Group.

The Network Feature analysis with Tourist group is given in Fig 5.

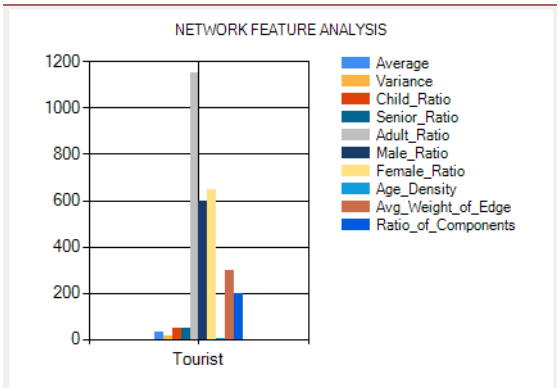

Fig 5:Network Feature analysis with Tourist Group.

The Network Feature analysis with Educational group is given in Fig 5.

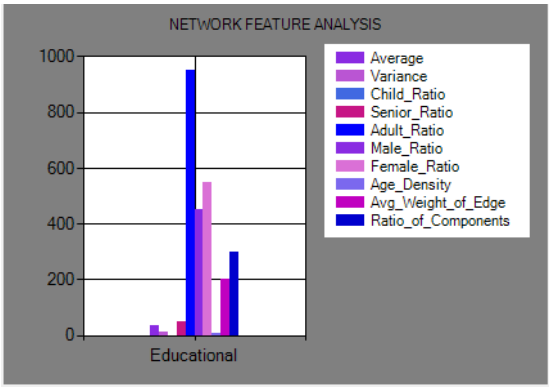

Fig 5:Network Feature analysis with Educational Group

7. CONCLUSION

In this paper, we have studied the problem of inferring the travel purposes of passenger groups and proposed a novel iterative classification approach based on the idea of collective inference. Firstly constructed cotravel networks by extracting social relations between passengers from their historical travel records. Then, we generated a series of network-based features and used the overlapping relations between passenger groups to represent the dependence relations between their labels. Finally, we collectively inferred all the labels by using an ICA.

The experimental results showed that our proposed iterative classification approach per-formed well in terms of the precision of inferring the travel purposes of passenger groups. The ICA approach significantly improves the accuracy than logistic regression.

8. REFERENCES

[1] R. Nicole, J. Carlson, and P. J. Rosenberger, III, "Factors affecting group-oriented travel intention to major events," *J. Travel Tourism Marketing*, vol. 29, no. 2, pp. 185–204, 2012.

[2] L. Wu, J. Zhang, and A. Fujiwara, "Representing tourists heterogeneous choices of destination and travel party with an integrated latent class and nested logit model," *Tourism Manage.*, vol. 32, no. 6, pp. 1407–1413, Dec. 2011.

[3] R. March and A. G. Woodside, "Testing theory of planned versus real-ized tourism behavior," *Ann. Tourism Res.*, vol. 32, no. 4, pp. 905–924, Oct. 2005.

[4] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee, "Measurement and analysis of online social networks," in *Proc. 7th IMC*, 2007, pp. 29–42.

[5] S. Wuchty and B. Uzzi, "Human communication dynamics in digital footsteps: A study of the agreement between self-reported ties and email networks," *PLoS One*, vol. 6, no. 11, p. e26972, Nov. 2011.

[6] D. Jensen, J. Neville, and B. Gallagher, "Why collective infer-ence improves relational classification," in *Proc. 10th SIGKDD*, 2004.

[7] R. Xiang and J. Neville, "Collective inference for network data with copula latent Markov networks," in *Proc. 6th WSDM*, 2013, pp. 647–656.

[8] D. Jensen and J. Neville, "Linkage and autocorrelation cause fea-ture selection bias in relational learning," in *Proc. 19th ICML*, 2002.

[9] Youfang Lin, Huaiyu Wan, Rui Jiang, Zhihao Wu,Xuguang Jia."Inferring the travelling purpose of passenger groups for better understanding of passengers".

[10] S. Pike and C. Ryan, "Destination positioning analysis through a compar-ison of cognitive, affective, conative perceptions," *J. Travel Res.*, vol. 42, no. 4, pp. 333–342, May 2004.

[11] . Tang *et al.*, "ArnetMiner: Extraction and mining of academic social networks," in *Proc. 14th SIGKDD*, 2008, pp. 990–998.