Universidad del Norte
Department of Electrical and Electronics Engineering

Intelligent Tourist Recommender System Focused on Collective Profiles

By
Michel Antony Barros Barrios

MASTER THESIS

Advisor
Dr. Christian G. Quintero M.

Barranquilla, Atlántico, Colombia
June 2017

Universidad del Norte
Department of Electrical and Electronics Engineering

Intelligent Tourist Recommender System Focused on Collective Profiles

A dissertation presented to the Universidad del Norte in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE

By

Michel Antony Barros Barrios

Advisor:

Dr. Christian G. Quintero M.

Barranquilla, Atlántico, Colombia
June 2017

ABSTRACT

Intelligent Tourist Recommender System Focused on Collective Profiles

By Michel Antony Barros Barrios

Advisor: Dr. Christian G. Quintero M.

Group recommendation is complex due to the selection procedure, structure and group conduction could conditioning negatively its effectiveness. Aspects like expectations of its components, the group size, time, communication standards, the previous experience or condition of members could have a negative influence. World Tourism Organization (UNWTO) defines tourism as a social, cultural and economic phenomenon which entails the movement of people to countries or places outside their usual environment for personal or business purposes. These people are called visitors (which may be either tourist or excursionists; resident or non-residents) and tourism has to do with their activities, some of which involve tourism expenditure. International tourism now represents 7% of the world's exports of goods and services, up from 6% in 2014, as tourism has grown faster than world trade over the past four years. Holidays, recreation and other forms of leisure have been just over half of all international tourist arrivals in 2015 (53% or 632 million). Business and professional purposes accounted for some 14% of all international tourists, another 27% travelled for other reasons such as visiting friends and relatives (VFR), religious reasons and pilgrimages, health treatment. The purpose of visit for the remaining 6% of arrivals was not specified. Nowadays, the greater part

of tourists around the world plan their vacation, make reservations or buy services, moreover, they share their experiences through the Internet.

In this research is implemented an intelligent system for managing and recommending tourist places to collective profiles, which is able to identify and satisfy preferences of group members.

Acknowledgments

To my God for allowing reach this great opportunity and this great achievement in my career.
To my parents for having done of me the person, I am nowadays. Likewise, for their great love, dedication, determination, patience, trust, support and all the great sacrifices. To be always there, you are the most important in my life.

To my girlfriend, father, family and friends for tolerating my absence at times and provide me all their support to long this professional and personal challenge.

Special thanks to Christian G. Quintero M. thanks a lot to provide me the opportunity to be better and show me each difficulty as a challenge, to make available his knowledge as professor, professional, and friend.

Likewise, I would like to thank all members of the Electric and Electronic Laboratories and GIRSI research group in the Universidad del Norte, for their friendship, support, and encouragement during this time.

To my God, mother, father, family and friends

General Contents

PART I: INTRODUCTION AND RELATED WORK

Motivation, objectives, main contributions and an overview of general concepts used in this thesis dissertation.

A review of relevant related work used as reference and inspiration to develop the proposed approach.

PART II: PROPOSED APPROACH

General considerations and implementation of the proposed tourist recommendation system approach in recommending touristy places to collective profiles, according to the common group preferences.

PART III: EXPERIMENTAL RESULTS AND CONCLUSIONS

Analysis and discussion of the experimental results, final conclusions and future research related to tourist recommender system focused on collective profiles.

Detailed Contents

CHAPTER 1 15

INTRODUCTION 15

1.1. MOTIVATION 15

1.2. OBJECTIVES 17

1.2.1. Thesis Question 17
1.2.2. Approach 17

1.3. CONTRIBUTIONS 18

1.4. READER'S GUIDE TO THE THESIS 19

CHAPTER 2 20

BACKGROUND INFORMATION 20

2.1. TOURIST ACTIVITIES SYSTEM 20

2.1.1. Different Definitions about Tourism 21
2.1.2. Tourist Topologies 22
2.1.3. Evaluation and Characterization Instruments 23

2.2. GROUP MODELLING STRATEGIES 25

2.2.1. Group Model Based (GMB) 25
2.2.2. Individual Recommendation Merging (IRM) 26

2.3. RECOMMENDATION STRATEGIES 27

2.3.1. Content Based 27
2.3.2. Collaborative Filtering 28
2.3.3. Demographic 28
2.3.4. Knowledge Based 28
2.3.5. Community Based 29
2.3.6. Hybrid recommender systems 29

2.4. SIMILARITY METRICS IN RECOMMENDATION SYSTEMS 29

2.5. HYBRIDIZATION OF FILTERING METHODS 33

2.5.1. Weighted 33
2.5.2. Switching 33
2.5.3. Mixed 33
2.5.4. Features Combination 33
2.5.5. Cascade 33
2.5.6. Features Augmentation 34
2.5.7. Meta Level 34

CHAPTER 3 35

RELATED WORK 35
3.1. TOURIST RECOMMENDER SYSTEMS 35
3.2. GROUP RECOMMENDER SYSTEMS 38
3.3. KNOWLEDGE-BASED RECOMMENDER SYSTEMS 40
3.4. COLLABORATIVE FILTERING RECOMMENDER SYSTEMS 40
3.5. CONTENT-BASED RECOMMENDER SYSTEMS 41
3.6. HYBRID RECOMMENDER SYSTEMS 42
3.7. FINAL REMARKS 45
CHAPTER 4 48
HYBRID RECOMMENDATION SYSTEM APPROACH 48
4.1. PROBLEM STATEMENT 48
4.2. CHARACTERIZATION PROFILES 49
4.2.1. Cultural Subprofile 49
4.2.2. Bioecologic Subprofile 50
4.2.3. Adventure Subprofile 51
4.2.4. Urban Subprofile 51
4.2.5. Sport Subprofile 52
4.3. PROPOSED CHARACTERIZATION INSTRUMENT 53
4.3.1. Survey Model 54
4.3.2. Evaluation Model 57
4.4. GROUP MODELLING STRATEGY 61
4.5. HYBRID RECOMMENDATION SYSTEM 62
4.5.1. Tag Characterization 63
4.5.2. Intelligent system 65
4.5.3. Data Preprocessing 66
4.6. GENERAL CONSIDERATIONS 69
CHAPTER 5 71
HYBRID RECOMMENDATION SYSTEM IMPLEMENTATION 71
5.1. DEVELOPED SYSTEM 71
5.2. CHARACTERIZATION PROFILE AND GROUP MODEL 72

5.3. HYBRID RECOMMENDATION ALGORITHM 78
CHAPTER 6 82
ANALYSIS OF THE EXPERIMENTAL RESULTS 82
6.1. EXPERIMENTAL DESIGN 82
6.2. ANALYSIS OF RESULTS 86
CHAPTER 7 94
CONCLUSIONS AND FUTURE WORKS 94
7.1. CONCLUSIONS 94
7.2 MAIN CONTRIBUTIONS 95
7.3 FUTURE RESEARCH AND DIRECTIONS 95
REFERENCES 96
ANNEX 104
RELATIONSHIP BETWEEN TAGS AND SUBPROFILES 104
A.1. TAGS RELATED TO THE ADVENTURE SUBPROFILE 104
A.2. TAGS RELATED TO THE BIOECOLOGIC SUBPROFILE 105
A.3. TAGS RELATED TO THE CULTURAL SUBPROFILE 106
A.4. TAGS RELATED TO THE SPORT SUBPROFILE 107
A.5. TAGS RELATED TO THE URBAN SUBPROFILE 108

List of Figures

Figure 2.2-1 Group Model Based (GMB) 25
Figure 2.2-2 Individual Recommendation Merging (IRM) 27
Figure 4.3-1 Survey Model Proposed 60
Figure 4.3-2 Result of Survey 60
Figure 4.5-1 Hybrid Recommendation System 62
Figure 4.5-2 Algorithm for selecting tags with presence of places 66
Figure 4.6-1 Representation of recommendation process 70
Figure 5.1-1 Index OdinTrip 72
Figure 5.1-2 Create Account OdinTrip 72
Figure 5.1-3 Login OdinTrip 72
Figure 5.2-1 Survey OdinTrip 73
Figure 5.2-2 Individual Profile 74
Figure 5.2-3 Create Group Interface 74
Figure 5.2-4 Collective Profile 75
Figure 5.2-5 List of Groups and Recommendation Section 76
Figure 5.2-6 Individual Section of the Designed Algorithm 77
Figure 5.2-7 Group Process 78
Figure 5.3-1 Recommendation Process 79
Figure 5.3-2 Visualization Process 79
Figure 5.3-3 Recommendation 80
Figure 6.1-1 GTravel System 82
Figure 6.1-2 Example of a group with two (2) members 83
Figure 6.1-3 Example of a group with six (6) members 83
Figure 6.2-1 Result of the Experiment 87
Figure 6.2-2 Manhattan Distance Behavior 88
Figure 6.2-3 Euclidean Distance Behavior 88
Figure 6.2-4 Example of a group with 6 members 89
Figure 6.2-5 Selection of the Destination City and the Number of Places 89
Figure 6.2-6 Recommendation for the Case Study 90
Figure 6.2-7 Case Study with a Different Destination City 90

Figure 6.2-8 Recommendation for the Case Study with other Destination City 91
Figure 6.2-9 Case of Study with a Different Collective Profile......................................92
Figure 6.2-10 Number of Places and Destination City for the Case of Study 92
Figure 6.2-11 Recommendation for the Last Case of Study...93

List of Tables

Table 4.2-1 Description of subprofiles and their characteristic activities.......................53
Table 4.3-1 Relation among Variables to Evaluate the Tourist behavior, the Answer Options of Proposed Survey and the Proposed Individual Profile..................................55
Table 4.3-2 Answer Options Related to Cultural Subprofile..57
Table 4.3-3 Answer Options Related to Bioecologic Subprofile....................................58
Table 4.3-4 Answer Options Related to Adventure Subprofile......................................58
Table 4.3-5 Answer Options Related to Urban Subprofile...59
Table 4.3-6 Answer Options Related to Sport Subprofile 59
Table 6.1-1 Confusion Matrix..84
Table A.1-1 Adventure Tags...104
Table A.2-1 Bioecologic Tags 104
Table A.3-1 Cultural Tags...104
Table A.4-1 Sport Tags...104
Table A.5-1 Urban Tags..104

PART I

INTRODUCTION AND RELATED WORK

Chapter 1

Introduction

This chapter provides an introduction to the work presented in this thesis. Specifically, the motivation in the research area, the pursued aims, and the main contributions are briefly described. Finally, the chapter concludes with an overview of the structure and the content of the thesis.

1.1. Motivation

Many situations in daily life are social activities that can be carried out in groups. These groups could be integrated by friends, family members, coworkers, classmates, members of the same age, and people with different cultures, among others. For these situations, it is better to recommend to a group of users, rather than to an individual. Group recommendation is complex due to the selection procedure, structure and group conduction could conditioning negatively its effectiveness (Torres Bernier, Secall, & Fuentes García, 2006)(Ricci, Rokach, Shapira, Kantor, & Ricci, 2011). Aspects like expectations of its components, the group size, time, communication standards, the previous experience or condition of members could have a negative influence (Ricci et al., 2011). There are several applications in different fields like TV content (Kim & Lee, 2014), films (Fernández, López, Olivera, Rienzi, & Rodríguez-Bocca, 2014), music content (Yoon, Lee, & Kim, 2012) or theme park services (Tsai & Chung, 2012) which are examples of group recommendation systems.

Tourism is defined as a social phenomenon where individuals or groups voluntarily travel looking for entertainment, relaxation, culture or health (Guerrero González & Ramos Mendoza, 2014). World Tourism Organization (UNWTO) defines tourism as a social, cultural and economic phenomenon which entails the movement of people to countries or places outside their usual environment for personal or business purposes. These people are called visitors (which may be either tourist or excursionists; resident or non-residents) and tourism has to do with their activities, some of which involve tourism expenditure. Also, World Tourism Organization defines travel group as a group made up of individuals travelling together: examples are people travelling on the same package tour or youngsters attending a summer camp. International tourism now represents 7% of the world's exports of goods and services, up from 6% in 2014, as tourism has grown faster than world trade over the past four years (Curice, J., Phillips, M. & E., 2016).

Holidays, recreation and other forms of leisure have been just over half of all international tourist arrivals in 2015 (53% or 632 million). Business and professional purposes accounted for some 14% of all international tourists, another 27% travelled for other reasons such as visiting friends and relatives (VFR), religious reasons and pilgrimages, health treatment. The purpose of visit for the remaining 6% of arrivals was not specified (Curice, J., Phillips, M. & E., 2016).

In the tourism industry, the Internet has become the main tool where tourists look for information of their interest. Nowadays, the greater part of tourists around the world plan their vacation, make reservations or buy services, moreover, they share their experiences through the Internet (Xiang, Magnini, & Fesenmaier, 2015)(Li, Yuan, & Jin, 2012). But, online information continues to grow at an exponential rate, this problem is called information overload (Kembellec, Chartron, & Saleh, 2014). People have a non-effective ability to access to a big data set and users are often frustrated by how difficult is to locate the right information quickly and easily (Ma & Uchyigit, 2008b).

In this research, these problems are solved implementing an intelligent system for managing and recommending tourist places to collective profiles, which is able to identify and satisfy preferences of group members. It is proposed the design of characterization instrument and a hybrid recommendation system based on the combination of the Knowledge-based and non-personalized recommendation systems. The proposal was developed based on a qualitative characterization model of individual and collective profiles and a hybrid recommendation system providing places suggested to a group. All of the above for solving problems of group modelling and group recommendation system.

1.2. Objectives

This dissertation is focused on development an intelligent tourist recommendation system capable of recommending tourist places to collective profiles and overcome classic limitations of recommendation systems.

- **Problem**: Make an intelligent recommendation of tourist places to collective profile according to group preferences.
- **General Objective**: Design and implement an intelligent tourist recommendation system focused on collective profiles.
- **Goals:**
  - ✓ To identify and characterize useful information about tourist places, as well involved group profiles.
  - ✓ To develop an intelligent system by using the characterized information, and provide a fitter tourist recommendation for the collective profile.
  - ✓ To test the system operation in different case studies.

1.2.1. Thesis Question

The principal question addressed in this dissertation is:

¿Could a computational intelligence system to recommend tourist places for a group according to their collective profile?

1.2.2. Approach

In this research is proposed the design of an intelligent tourist recommendation system focused on collective profiles. This proposal looks for an intelligent selection according to the group preferences. This approach incorporates a characterization model of tourist collective profiles and an intelligent hybrid recommendation system.

Group modelling is a previous stage before applying recommendation strategy. It is required to build a collective profile based on individual profiles. Group modelling looks for aggregating individual profiles to identify general group preferences and deciding what would be suitable for a collective profile. Individual model is built by means of a qualitative characterization instrument. This instrument evaluates different variables of tourist behavior of group members with a survey, which is based on psychological aspects of members like tourist personality, motivation, favorite tourist activities and tourist experience.

The hybrid recommendation system approach based on the combination of the Knowledge-based and non-personalized recommendation systems is required to take the resulting collective profile and calculate suitable category of places or tags for the active collective profile with a threshold criterion. These tags had been characterized previously evaluating different variables related to the destination. Places associated with these tags are ranked by their reputation. Places reputation is required to calculate the best places in destination according to the group preferences.

When a recommendation system carries out user characterization by group modelling and tag characterization, group recommendation system problems, like group modelling, cold start and ramp up are solved. In this sense, the proposed hybrid recommendation system is based on a suitable group model and tag characterization, these characterizations are compared and the proper tags are selected for the active collective profile.

1.3. Contributions

This thesis makes the following contributions in group recommendation systems problems:

- *An intelligent approach for recommending tourist places to collective profiles using a group modeling strategy and hybrid recommendation system.*
- *A recommendation strategy based on the combination of the Knowledge-based and non-personalized recommendation systems.*

1.4. Reader's Guide to the Thesis

Following is a general description of the contents of this dissertation. This master thesis is organized in three main parts distributed by chapters.

Part I: Introduction and Related Works

Chapter 1 presents a motivational introduction to the main topics, objectives, and contributions regarding this dissertation.

Chapter 2 gives a general overview of background information regarding tourist activities system, group model strategies, recommendation strategies, similarity metrics and hybridization techniques which are required to develop the proposed approach described in chapter 4 and 5.

Chapter 3 provides a general survey of the most relevant work related to the research addressed in this thesis.

Part II: Proposed Approach

Chapter 4 describes the formal aspects of the group model strategy and the hybrid recommendation system presented in this thesis.

Chapter 5 presents the implementation of the proposed approach in chapter 4. The chapter also contributes to complete the description of such proposal.

Part III: Results and Conclusions

Chapter 6 provides experimental results of the implemented approach. An experiment design is presented to evaluate the performance of the proposed hybrid recommendation system in simulation runs.

Chapter 7 discusses and analyzes the results, summarizes the conclusions and contributions of the thesis and outlines the most promising directions for future work.

Chapter 2

Background Information

This chapter introduces and reviews general concepts of tourist activities system, group model strategies, recommendation strategies, similarity metrics and hybridization of filtering methods on intelligent group recommendation systems required for developing the proposed approach.

2.1. Tourist Activities System

The nature of the tourist activity is the result of the interrelation among several elements and factors, which are considered in a systematic form and evolved into a dynamic form (Guerrero González & Ramos Mendoza, 2014).

Under a temporal perspective tourist activities system has two principal features (Torres Bernier et al., 2006):

- **Open system:** this feature incorporates new elements inside doing itself more complex and diverse.
- **Dynamic system:** in this case the system transforms itself, improving the interaction among its elements.

Tourism is a great system of activities in continual growth. This system is divided into three principal subsystems supplementary and interrelated (Jiménez Bulla, 2008).

- **Activities of tourist nature:** these activities are linked directly to essential characteristics of the tourists. Tourist attractions belonging to this subsystem are natural or cultural heritage inside in a particular area (Jiménez Bulla, 2008).
- **Activities of touristic application:** activities of tourist application are linked directly to tourist's motivations. These are non-tourist activities

related to leisure, nature, health, sport, which are applied to tourism by means of an adaptation process (Torres Bernier et al., 2006). These activities do not act as touristic nature activities by themselves, but they contribute added value for tourist (Jiménez Bulla, 2008). All classes of specific tourism are in this field.

- **Support activities for the tourism:** this subsystem is formed by public and private services demanded by tourists (Torres Bernier et al., 2006). These services are shared with the population.

Tourism market is a place where supply and demand interact (Guerrero González & Ramos Mendoza, 2014). Supply is a combination of products, services, and goods proposed to the tourists, which result attractive for them (Jiménez Bulla & Jiménez Barbosa, 2013). Demand is the amount of tourists or companies which are motivated to consume products, services, and goods. Tourist demand is characterized by its high level of segmentation. Demand segmentation is directly proportional to tourism market segmentation (Guerrero González & Ramos Mendoza, 2014). Demarcation of tourism market is determined by tourist's behavior and motivation (Torres Bernier et al., 2006).

2.1.1. Different Definitions about Tourism

Tourism definition has not been unanimous among experts and institutions related to this industry (Jiménez Bulla & Jiménez Barbosa, 2013).

World Tourism Organization in its tourism glossary of February 2014 defines tourism as a social, cultural and economic phenomenon which entails the movement of people to countries or places outside their usual environment for not more than one consecutive year for leisure, business or professional purposes.

Tourism is a human activity that entailing the wish to satisfy diverse motivations of the tourists, which are classified and personal (Guerrero González & Ramos Mendoza, 2014).

Tourism is a complex activity which has become changing according to the general world systems. It has different characteristics in capitalism and others in socialism (Dachary & Arnaiz Burne, n.d.).

These definitions coincide with the concept of tourism as a socio-economic activity which entails travelling to different destinations to the usual environment of the tourist for a short time.

2.1.2. Tourist Topologies

Tourism products topology is a specific characterization of each tourism product subdivisions in its realization process and variables considered for the design of topology (Jiménez Bulla & Jiménez Barbosa, 2013). Demand criterion determines demarcation of tourist topologies (Torres Bernier et al., 2006).

- **Generic Tourism**

Generic tourism has a fuzzy, broad and assorted motivational setting. Concepts like rest, enjoy free time and be bone idle prevail in this topology (Torres Bernier et al., 2006). This tourism class is identified with mass tourism and family demand. Frequently, this tourism is related to the territorial area, the sun, beach, rural tourism and city tourism.

- **Specific Tourism**

Specific tourism catches the attention and interest of a significant people number and it motivates them to move from their habitual residence to the tourist destination (Torres Bernier et al., 2006). Usually, the different classes of specific tourism are classified by areas (Guerrero González & Ramos Mendoza, 2014)(Jiménez Bulla & Jiménez Barbosa, 2013)(Torres Bernier et al., 2006):

- Bioecologic tourism
- Cultural tourism
- Sport tourism
- Adventure tourism
- Urban tourism
- Health tourism
- Professional and scientific tourism
- Resident tourism
- Conditioned tourism by demand
- Touristic micro products

2.1.3. Evaluation and Characterization Instruments

Evaluation and characterization instruments diagnose preferences of tourists inside tourism market (Guerrero González & Ramos Mendoza, 2014). Tourist's data are obtained from these instruments inside tourism market context. Data collection could carry out through qualitative or quantitative techniques (López Bonilla & López Bonilla, 2012). Qualitative techniques go in depth about psychological aspects comprehension of tourists such as beliefs, attitudes, and motivations. Quantitative techniques focus on collecting a large number of data carrying out the statistical analysis and providing information about study phenomenon.

Quantitative techniques have a better structure than qualitative techniques, they use surveys previously designed. Surveys are key elements in quantitative techniques application (López Bonilla & López Bonilla, 2012).

- **Survey**

The survey is defined as a primary source of information over objective, coherent and jointed question combination (Guerrero González & Ramos Mendoza, 2014). This instrument allows for analyzing data by quantitative methods. There are different survey types according to the contact method used between pollster and survey respondent. Oral communication determines contact methods classification. In this way, there are personal surveys and phone surveys if oral communication exists. If there is not oral communication, the pollster could use postal, E-mail or online surveys (López Bonilla & López Bonilla, 2012). Online surveys are getting increasingly used because they are a cheap, fast and easy method.

Surveys could be classified according to the flexibility of their question and answer types. Therefore, there are organized, semi-organized and non-organized surveys (López Bonilla & López Bonilla, 2012). Organized surveys consist of formalized and standardized questions and answers. These surveys are presented to survey respondent in the usual order and they are limited to the answer options predetermined. Semi-organized surveys have more flexibility than organized surveys. In this way, some questions could are drawn up to get open answers or questions order could be changed. Non-organized surveys have the greatest flexibility because many questions could be drawn up during the interview.

Question types of the survey could be gathered in different categories such as flexibility, the degree of freedom of answers, the number of answer options, the objective of questions, the way to carry out questions and the wished information (López Bonilla & López Bonilla, 2012).

According to the degree of freedom of the answers, there are open and closed questions. Open questions have a space to write the answer. Closed questions are circumscribed to specified answer options. Mixed or semi-closed questions are another question type, which has specific answer options and an open answer alternative option.

Depending on the number of answer options, questions could be dichotomous or polytomous. Dichotomous questions only have two answer options, but polytomous questions have more than two answer options.

There are four different question types such as introductory questions, filter questions, tandem questions and control questions. Introductory questions facilitate the developing survey and generate interest. Filter questions guide survey towards a specific direction, i.e., inside survey, there are subordinate questions to a question filter. Tandem questions are a series of questions which allow for getting a variety of interrelated data. Control questions confirm the truthfulness of answers given by survey respondent, i.e., these questions allow for checking the coherence of survey.

- **Multi-Item Scale**

The scale represents the instrument to measure quantitatively characteristics or proprieties of objects, phenomenon or individuals (Torres Bernier et al., 2006). In the design of scales, items and attribute to measure must be interrelated (López Bonilla & López Bonilla, 2012). Items must represent all concepts included in the definition of the attribute. Items must be appropriate to the study population. Items in designing of a scale could be originals or adapted of previous scales. These items could be constructed from experts' knowledge, they could be produced from people's ideas of the study population or from ideas collected from diverse sources of information.

2.2. Group Modelling Strategies

Daily life has many social situations in which is better recommend to a group of users rather than to an individual. Recommendation systems need a target user's model for recommendations (Kim & Lee, 2014). Group modeling strategies are used for recommending to a group of users by aggregating from individual user models (Ricci et al., 2011). There are two major methods to suggest recommendation for group users, which are group model based (GMB) and individual recommendation merging (IRM).

2.2.1. Group Model Based (GMB)

Group model based consists in making a group model from profiles of each member, i.e., this strategy makes aggregation before generating recommendations (Bernier et al., 2010)(see Figure 2.2-1). Group model based imitates group's agreement process (Kim & Lee, 2014).

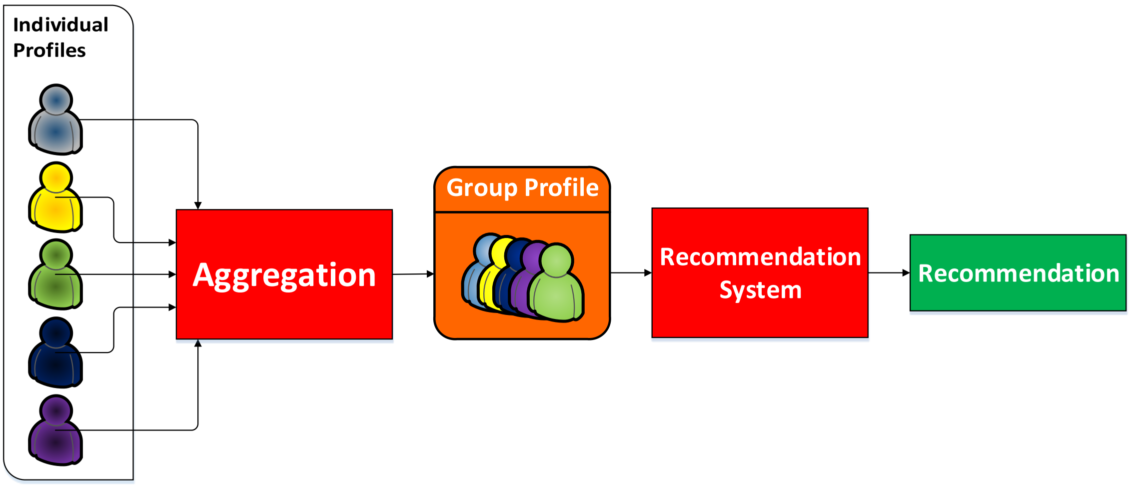

Figure 2.2-1 Group Model Based (GMB)

This strategy could be performed building a group profile using profiles of each member and generating recommendations based on resulting group profile. Also, group model based could be performed making an intersection of individual preferences to know group interests. In this case, the recommendation system computes intersection in order to discover common points and display items which satisfy all group members. However, this strategy is not appropriate for large groups, because the probability to obtain an empty intersection is directly proportional to the size of the group (Bernier et al., 2010)(Kim & Lee, 2014).

- **Consensus Functions**

These functions are called aggregation functions (Bernier et al., 2010)(Kim & Lee, 2014). Aggregation of preferences, criteria or similarities could happen at various stages in recommender systems (Ricci et al., 2011). As aggregation functions are used the arithmetic mean, maximum or minimum functions.

- Minimum Function (Last Misery Strategy)

The minimum function makes a group model using minimum ratings on each characteristic of individual profiles (Kim & Lee, 2014). This function avoids dissatisfaction on each member but could ignore other users' satisfaction (Ricci et al., 2011).

- Maximum Function (Most Pleasure Strategy)

This function calculates group model using maximum ratings on each characteristic of individual profiles (Ricci et al., 2011). It takes an item by a satisfied user but it could ignore other users' dissatisfaction (Kim & Lee, 2014).

- Arithmetic Mean (Average Strategy)

It makes a group model using the arithmetic mean on each characteristic of individual profiles (Kim & Lee, 2014). The arithmetic mean function is appropriate in homogeneous groups because members have similar preferences (Bernier et al., 2010).

2.2.2. Individual Recommendation Merging (IRM)

Recommendation systems treat group members individually with this strategy. Individual recommendation merging consists in making individual recommendations for every group member and then aggregating these results (Bernier et al., 2010) (see Figure 2.2-2). If recommendation system cannot collect or know group user's preference, this strategy cannot be used (Kim & Lee, 2014).

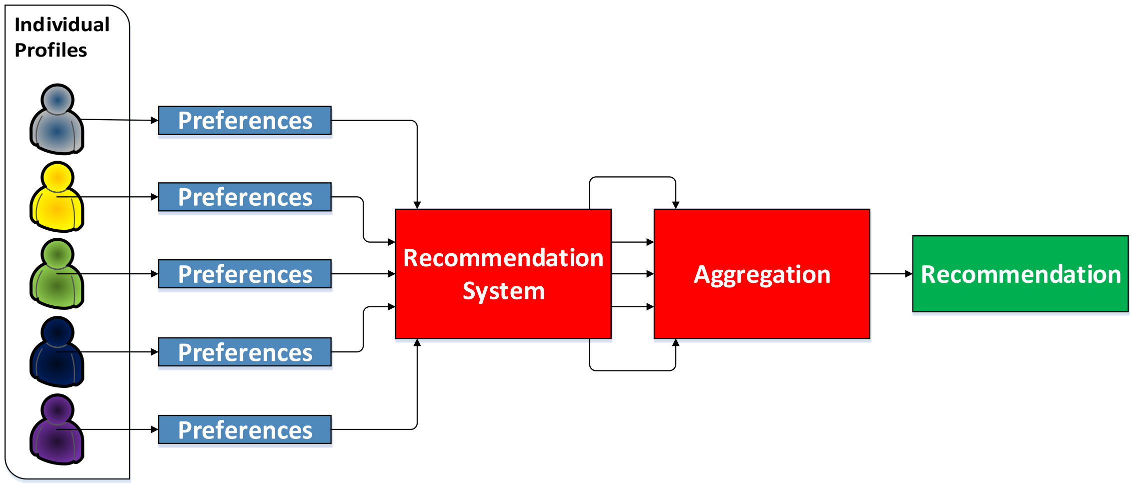

Figure 2.2-2 Individual Recommendation Merging (IRM)

2.3. Recommendation Strategies

Recommender Systems (RSs) provide suggestions for items to be of use to a user (Ricci et al., 2011). "Item" is the general term used to denote what the system recommends to users (Kembellec et al., 2014). These systems appear due to the exponential growth of online information and the poor human ability to access this information effectively, for this reason, users are often frustrated by how difficult it is to locate the right information quickly and easily (Ma & Uchyigit, 2008a). The development of automated recommender systems (RecSys) is, therefore, a foreseeable phenomenon for contributing toward resolving the problem of information overload, valuing content and focusing attention on the user in such a context of overabundance (Kembellec et al., 2014).

In order to implement its core function, identifying the useful items for the user, a recommendation system must predict that an item is worth recommending. In general, authors distinguish between six different classes of recommendation approaches, which are described next.

2.3.1. Content Based

A Content-based algorithm uses user preferences and item information with the purpose of predicting items which could be suitable for users (Quispe & Luna, 2016). The features associated with the compared items determine the similarity of items. (Ricci et al., 2011). For example, if a user has positively rated a song that belongs to the classical genre, then the system can learn to recommend other songs from this genre.

2.3.2. Collaborative Filtering

A collaborative filtering algorithm searches a set of nearest neighbor to a target users. According to the item score of the nearest neighbors, it is proposed an item set to recommend to the target user (Mai, Fan, & Shen, 2009). The similarity in the rating history of the users determines the similarity in the taste of two users (Ricci et al., 2011). Collaborative filtering is called "people-to-people correlation". Collaborative filtering is the most popular and widely implemented technique in recommendation systems.

2.3.3. Demographic

A demographic algorithm recommends items based on the demographic features of the user (Dai, Ye, & Gong, 2009). This algorithm generates different recommendations for different demographic niches (Ricci et al., 2011). Many Web sites use this algorithm because it is a simple and effective personalization way. For example, items could be selected according to the age of the user. These approaches have been popular in marketing, but there is little research into demographic systems.

2.3.4. Knowledge Based

A knowledge-based algorithm recommends items using explicit knowledge. It calculates recommendations based on item features and their usefulness for the user (Zhang, 2012). The knowledge-based recommendation systems are case based. These systems calculate the similarity based on how much the user needs to match the recommendations. The similarity score is interpreted as the utility of the recommendation for the user (Ricci et al., 2011).

The constraint-based algorithm is a variation of knowledge-based recommendation systems (Pawar, Ghorpade, & Shedge, 2016). Constraint-based and knowledge-based systems make use of knowledge in similar form. These systems collect user preferences. These systems repair the inconsistent data and recommendation results are given to users. Constraint-based and knowledge-based systems differ in the form to calculate solutions.

2.3.5. Community Based

Community-based algorithm recommends items based on the preferences of the user's friends (Fatemi & Tokarchuk, 2013). These systems follow the epigram "Tell me who your friends are, and I will tell you who you are". People prefer following recommendations from their friends than recommendations from anonymous individuals. The last concept and the growing popularity of social networks are generating a growth interest in community-based recommendation systems. These systems are called social recommender systems also. These systems model and obtains information about the users and the preferences of the user's friends(J. Dong, Li, & Fang, 2016). Recommendations are calculated based on ratings that user's friends provide previously. The research in this area is in its early phase and results in the performance of these systems are mixed (Ricci et al., 2011).

2.3.6. Hybrid recommender systems

These recommendation systems are based on the combination of two or more of the above-mentioned techniques. When a recommendation system combines techniques X and Y try to use the advantages of X to fix the disadvantages of Y (Kavinkumar et al., 2015). For example, collaborative filtering systems suffer from new-item problems, i.e., they cannot recommend items without ratings. This problem does not limit content-based systems since the prediction for new items is based on their description (Ricci et al., 2011). Given two or more basic techniques, several ways could be proposed for combining them to create a new hybrid system (F. Dong, Luo, Zhu, Wang, & Shen, 2013).

2.4. Similarity Metrics in Recommendation Systems

The similarity is defined as the measure of closeness between objects which could be items to recommend, user profiles or recommendations (Kumar, Gupta, Singh, & Shukla, 2015). Similarity metrics are used in classification methods and clustering techniques (Ricci et al., 2011). Some of the well-known similarity measures are:

- Euclidean Distance

This is the distance between two points or objects, x and y, with i dimensions (Ricci et al., 2011). Euclidean distance is always greater than or equal to zero. The measurement would be zero for identical points and high for points that show little similarity (Kumar et al., 2015). This is the simplest and most common example of a distance measure.

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

Where:

x, y: Data objects.

x_i, y_i: i^{th} attributes or components of data objects.

n: Number of dimensions or attributes.

d: Distance between data objects.

- Manhattan/City Block Distance

Manhattan distance is the addition of absolute differences among i dimensions of data objects, which is similar to movement inside a city where people or cars have to move around buildings instead of going straight through. This distance is always greater than or equal to zero. The measurement would be zero for identical points and high for points that show little similarity (Kumar et al., 2015).

$$d(x,y) = \sum_{i=1}^{n}|x_i - y_i|$$

Where:

x, y: Data objects.

x_i, y_i: i^{th} attributes or components of data objects.

n: Number of dimensions or attributes.

d: Distance between data objects.

- Minkowski Distance

Minkowski distance is a generalization of Euclidean distance (Ricci et al., 2011). Likewise Euclidean and Manhattan distance, Minkowski distance is always greater than or equal to zero. The measurement would be zero for identical points and high for points that show little similarity.

$$d(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i|^r\right)^{\frac{1}{r}}$$

Where:

x, y: Data objects.

x_i, y_i: i^{th} attributes or components of data objects.

n: Number of dimensions or attributes.

r: Degree of distance.

d: Distance between data objects.

- Pearson´s Correlation

The similarity between data objects can also be given by their correlation which measures the linear relationship between them (Ricci et al., 2011). If Pearson's correlation tends to one data objects have a high similarity, if Pearson's correlation tends to zero data objects have a little similarity but if Pearson's correlation tends to minus one data objects have an inverse similarity.

$$c(x, y) = \frac{\sum_i^n (x_i - \bar{x}_i)\,(y_i - \bar{y}_i)}{\sqrt{\sum_i^n (x_i - \bar{x}_i)^2\,(y_i - \bar{y}_i)^2}}$$

Where:

x, y: Data objects.

x_i, y_i: i^{th} attributes or components of data objects.

n: Number of dimensions or attributes.

c: Degree of correlation.

- Cosine Similarity

Also known as vector-based similarity, this formulation considers data objects as document vectors of an n-dimensional space and it defines their similarity as the cosine of the angle that they form (Kumar et al., 2015)(Ricci et al., 2011). This metric is calculated as the quotient between vector dot product and norms product of data objects. The measurement would be zero for identical points and high for points that show little similarity.

$$cos(x,y) = \frac{\overrightarrow{x_i} \cdot \overrightarrow{y_i}}{\|\overrightarrow{x_i}\|_2 * \|\overrightarrow{y_i}\|_2}$$

Where:

x, y: Data objects.

x_i, y_i: i^{th} attributes or components of data objects.

cos: Cosine of the angle that form data objects.

- Tanimoto Coefficient

This similarity metric considers data objects as data sets and calculates the quotient between the intersection and union of data sets (Ricci et al., 2011). If Tanimoto coefficient tends to one data objects have a high similarity but if Tanimoto coefficient tends to zero data objects have a little similarity.

$$tc(x,y) = \frac{\sum_i^n x_i * y_i}{\sum_i^n {x_i}^2 + \sum_i^n {y_i}^2 - \sum_i^n x_i * y_i}$$

Where:

x, y: Data objects.

x_i, y_i: i^{th} attributes or components of data objects.

n: Number of dimensions or attributes.

tc: Degree of similarity.

2.5. Hybridization of Filtering Methods

Hybridization of recommendation systems is the result of a combination of two or more filtering methods (Shambour & Lu, 2010). Hybridization techniques are classified into seven categories, which are weighted, mixed, switching, feature combination, cascade, feature augmentation and meta-level (Ricci et al., 2011).

2.5.1. Weighted

A weighted hybridization strategy combines the result of all available recommendation systems to get the weighted sums of their scores (Lin, 2014). This value is weighted by a confirmation of the users to gain better performance (Kembellec et al., 2014).

2.5.2. Switching

This method employs different algorithms, data-based methods or social filtering depending on the search context of the user (Kembellec et al., 2014)(Lin, 2014).

2.5.3. Mixed

In this strategy, the recommendation system simply takes the union of each recommendation as the output (Lin, 2014). This method facilitates making recommendations from traditional methods with the aim of limiting the drawbacks of each method (Kembellec et al., 2014).

2.5.4. Features Combination

Features combination enriches data which has been integrated a priori into the system with the ratings provided by the users (Kembellec et al., 2014). The calculation of the recommendation is carried out over all of the data set.

2.5.5. Cascade

This method consists of a double analysis of user profiles (Kembellec et al., 2014). The first analysis calculates potential candidates and the second one refines the selection of items (Ricci et al., 2011).

2.5.6. Features Augmentation

This method is similar to cascade method for the first pass-through (Ricci et al., 2011). If the result of a first analysis has a great number of candidates, then a second analysis will carry out a secondary discrimination by integrating the data of recommended items (Kembellec et al., 2014).

2.5.7. Meta Level

Likewise, cascade and features augmentation, Meta level applies filtering users twice in order to determine similarities. But in this method, the first pass-through generates a model or profile type of the user (Kembellec et al., 2014).

Chapter 3

Related Work

This chapter presents an overview of the main works focused on the topics addressed in this dissertation.

3.1. Tourist Recommender Systems

(Ichimura & Tachibana, 2014) proposed a tourist information system which can estimate the user's feelings generated by tourist places. This system is based on an estimation method to calculate the agent's emotion from the contents of expressed emotions which are aroused in the computer agent by using the facial expression. Emotion Generating Calculations (EGC) method is based on the Emotion Eliciting Condition Theory, which can decide whether an event arouses positive feelings or not and quantify the degree of the feelings generated by the event. EGC is made up of 2 or 3 terms such as subject, object, and predicate, which have Favorite Value (FV). The agent can change emotions calculated by ECG according to its own emotions. The authors developed an Android EGC application which the agent works in order to evaluate the feelings in the conversation. Also, this system can guide some places like local food shops and local gifts collected in Hiroshima Tourist Map Android application. This system calculates recommendation lists by an estimate function, which takes the number of hits for a term retrieved by Google search, the importance degree of a term registered in the Hiroshima sightseeing website and the strength of emotion calculated by EGC. Recommendation lists are calculated by the estimate function which consists of information retrieval by Google search, TF-IDF for Hiroshima tourism website and the EGC results.

(Tsai & Chung, 2012) present a system where tourist behaviors are collected through a Radio Frequency Identification (RFID) system and stored in a route database. The route database is segmented into subgroups based on the similarity among tourists' visiting sequences and time lengths. This system generates route suggestions for the visitors based on retrieved visiting behavior data and current facility queuing situation identified by the RFID system. When tourists go to a theme park, they provide information with a wristband embedded with a RFID tag with a unique electronic product code (EPC). Whenever the tourists go into the navigating area of a RFID reader, the reader record the EPC and time, later it transfers this information to the Ride Information Server and the Route Database Server. The route recommendation system is integrated with three modules. The first one, the tourist clustering module, segments tourists' visiting sequences in the route database into sub-groups based on the dissimilarity among the tourists' visiting sequences and time lengths. The subgroup retrieval module (the second module) finds the most similar sub-group to a visitor's input preference. The route generation module (the last module) takes the visiting behaviors data identified in the second module and the queuing information from each ride identified by the RFID system to generate an appropriate visiting recommendation for the visitor.

(Kenteris, Gavalas, & Mpitziopoulos, 2010) developed recommender systems based on collaborative filtering techniques and Wireless Sensor Network (WSN) system installed around tourist places for providing tourist information and ratings about tourist places through mobile devices. MyMytilene is a web-to-mobile tourist framework which allows tourists to use the web in order to build customized mobile standalone guides that run on any mobile device offering tourist information. In the original users chose content items (information about POIs) after browsing all the available tourist content. The recommender system takes the explicit and implicit user data as input and classifies users in separate groups (clusters); thereafter, based on the weighted rating of each POI (depending on the platform used to upload the respective content items), each user is recommended POIs to select. All users' interactions are recorded by the recommendation system. The user is not obligated to explicitly create a personal account; however, the user is encouraged to do so in order to enable personalized recommendations by the system. The content recommendation is provided to users that have declared an explicit profile but also to users that have just used the system long enough to record an implicit profile. If no profile is available then content popular among all users is recommended.

(Baraglia, Frattari, et. al., 2012) proposed a recommender system using item-based collaborative filtering techniques, which has two main components: offline and online. The first one aims to create the knowledge model, which is the basis for computing suggestions. Its execution takes place when new data is available for updating the knowledge model and is based on the trajectory dataset and the points of interest. The online component uses the current user information and the knowledge model in order to produce a list of suggestions. The data processed by the offline component consists of a dataset of trajectories representing the movements of users in a certain period of time, as detected by their GPS devices, and a set of points of interest including their coordinates. User locations are collected by GPS systems and sent to the offline component whenever a new position is detected. The authors evaluate the efficiency of the proposed system by using two trajectory sets: synthetic and real, and a set of predefined points of interest.

(Alghamdi, Zhu, Saddik, & Systems, 2016) developed a system consists of three main parts: a recommender system, the proposed algorithm named the balanced orienteering problem, and a scheduler. To rate each point of interest, the system needs a ranking algorithm, which is done by extracting picture metadata from a social network that uses geolocation tags. By knowing how many users have visited a point of interest, the system will calculate how popular it is. In addition, to get an estimate of the time spent in an individual point of interest, the system calculates the time difference between the next place visited and the place previously visited, in the same trip. The system implements a balanced orienteering problem (BOP), which is a variation of orienteering problem (OP). BOP takes into consideration both the time spent at each location and the time traveling. Moreover, BOP can work on mobile devices, where resources are limited, while the original OP is an NP-hard problem. The scheduler will check if the user is following the trip plan closely and will adjust the plan as needed, to ensure that the user is at his destination in time.

Alternatively, the scheduler provides feedback for the user's preferences. Lastly, because of the limited resources, the scheduler will have to delete out-dated information and obtain missing information.

3.2. Group Recommender Systems

Group recommendation systems represent a new research area due to the recent development of social activities and the special interest in satisfy group preferences. In recent years many authors had developed different strategies for solving this problem. These strategies aim aggregating individual preferences and form a group profile or form a group recommendation based on individual recommendations. Next different dissertations are described relating to this area.

The paper presented in (Fernandez, Lopez, et. al., 2014) describes a system that recommends movies from cinema listing to ephemeral groups in real-time combining a group model based strategy with a collaborative filtering algorithm for individuals. In order to implement the group recommendation strategy, the system uses the multiplicative utilitarian strategy, which considers the preference of each individual in the group by calculating the product of the individual predictions as the happiness value for the group. In order to compute individual predictions is used the slope one algorithm as a collaborative filtering method. Therefore, the individual prediction algorithm strongly determines the final total time for the recommendation.

(Kim & Lee, 2014) designed seven experiments for analysis of group recommendation systems' performance on real dataset. These recommendation systems are focused on group home user in TV domain. They are methods for making the model of the group user. The first experiment is the group recommendation system based on pseudo-user approach. The second, third and fourth experiments are the group recommendation system based on consensus function approaches (Min., Max. and Avg.). The fifth, sixth, seventh experiments are the group recommendation system based on consensus function approaches (Min. with G, Max. with G and Avg. with G). On all of the experiments, the collaborative filtering method is used for the recommendation and neighborhood size is set to 50.

In order to measure the performance of the group recommendation systems, 80% of viewing history were used to the training set and the other 20% of viewing history were used to the testing set. The number of TV programs in the testing set is 133,383. The performance of the group recommendation system is measured with precision metric. In the recommendation system based on consensus function approach, Max. and Avg. were more suitable than Min. for group home user in TV domain. In addition, the pseudo-user approach was more suitable than consensus function approach on the special situation that group users had 300 histories and over.

The purpose of the approach proposed in (Qi, Mamoulis, et. al., 2016) is to study the problem of recommending one or more packages of items to a group of users. The authors proposed two probabilistic models (Group Rating and User Package), both of which incorporate individual ratings by users to items, user impacts, and package viability. In the group rating (GR) model, the probability that the group U will select an item i is defined. Then the probabilities of individual items are combined, to derive the likelihood of a package. The user package (UP) model reverses the above generative process. In UP, the group first chooses a representative user u with probability Pr(u|U, C). The representative user will decide for the whole package. Algorithms were proposed to efficiently implement the two models. In addition, they introduced fairness which is a unique but important feature of the package to a group of users (P2G) problem.

(Pujahari & Padmanabhan, 2015) proposed an approach to group recommender system combining the features of item-item collaborative filtering as well as user-user collaborative filtering to make efficient group recommendation by making homogeneous groups and predicting items that are common for most of the users in the group to generate group recommendation. In order to generate homogeneous group, the system finds the degree of similarity between members. After generating homogeneous group recommendations are generated for the group of users. To achieve this end, the preferences of each user are combined as well as the similarity between the items. A second algorithm generates a top-k recommendation for a group of users, where k is the number of items or things generated by the group recommender system. The value of k is given by the group of users. This algorithm finds the prediction value for each user and then find out the most common instances between the group of users. After finding a common instance it then removes this from the total instances so that same instance is being generated multiple times.

3.3. Knowledge-Based Recommender Systems

The study proposed by (Pawar et al., 2016) aims to suggest recipes using constraint satisfaction problem (CSP) and forward checking algorithm (FCA). So that user will eat their meal and breakfast with a proper nutritional intake which best suits to their disease. It also recommends recipes for all users without diseases. Constraint satisfaction problem depends on following variables and constraints: 1) User properties, 2) Recipe properties, 3) Constraints, 4) Filter conditions and 5) Products. Forward Checking is an improved version of simple backtracking algorithm. To detect inconsistency between current instantiating and future variables limited amount of look-ahead search is done in forward checking algorithm. In FCA, the variable is substantiated to certain value from its domain. Then repeatedly at each step, next variable is instantiated to a value that is consistent with the previous assignments. When the current variable is instantiated, a forward check is performed that detects all values inconsistent with current instantiating from the future domains

3.4. Collaborative Filtering Recommender Systems

The paper presented in (Wang, Yuan, & Sun, 2010) proposes a combination filtering method which firstly constructs a user model offline by combining filtering technologies based on content and demographic information, then makes recommendation online on the basis of the model by using collaborative filtering. The combination is introduced at three different layers: feature layer, model layer, and collaborative filtering algorithm layer. Firstly, item combination features are extracted from item’s detailed description and ratings. Secondly, a hybrid user model is constructed on the basis of combination feature set and demographic information summary. Finally, make a combination of memory-based and model-based filtering methods at the level of collaborative filtering. This study aims to reduce system complexity, shorten the computing time and improve scalability by constructing a new user model and adopting the method of combination filtering. The genetic algorithm learns the weight features in user model and significantly enhances the recommendation accuracy due to the accurate description of user preference.

The study proposed by (Song, Lu, & Zhao, 2011) combines the application of TAG in the Web2.0 era considers the motivation of user's behavior based on item-based collaborative filtering algorithm which uses similarity of user behavior to generate a recommendation, that is, use TAG technology to distinguish user's interest point, and then combines the contents similarity and user's behavior similarity to make recommendation. TAG technology provides a large number of user's feedback which can helps recommendation system to compute items' similarity with considering user different interest point.

3.5. Content-Based Recommender Systems

The study presented in (Bogdanov et al., 2011) proposes a system for music recommendation and user preference visualization. Structurally, its workflow can be divided into data gathering, audio analysis, music recommendation, and preference visualization. This system operates on content-based information extracted from audio, and, more concretely, it exploits high-level semantic descriptions of music tracks automatically inferred by a number of classifiers. The system employs the Last.fm and Soundcloud APIs to generate semantic user models for the members of these music services. To this end, the system extracts audio fragments of the tracks preferred by the users and computes semantic descriptions of these tracks via the Canoris API. Finally, the system generates musical recommendations, relying on a semantic similarity measure between music tracks.

The study proposed by (Miyazawa, Yamamoto, & Kawabe, 2013) presents a context-aware recommendation system that gives optimal information for users based on 1) a content-based image retrieval (CBIR) mechanism to search the similar images aiming to extract the detailed information to the text-inexpressible images

2) the contextual information of such similar images searched the Web, and 3) user's dynamic context or situation considering time-variant factors as well as space factors. It is expected to increase the precision or optimality of recommendation by matching and fusing the context of similar images obtained by CBIR with textual and signal information about user's situation or dynamic context. The system retrieves relevant image information using a similar image search employing content-based information retrieval (CBIR) techniques. Once the image information is retrieved, information is extracted and answered in consideration of user's situation or context.

3.6. Hybrid Recommender Systems

Hybrid recommendation systems approach has its origin in solving problems of different recommendation methods. Hybridization of recommendation methods aims to complement advantages and deleting disadvantages of recommendation systems. The paper presented in (Chen, 2013) describes an item-based collaborative filtering recommendation system to predict the interests of an active tourist by collecting preferences or taste information from a number of other tourists. The authors proposed a mechanism to predict a set recommended tourism places of elicited rating places (e.g., ratings of 1 through 5 stars) for the active tourist pre-traveling places. Furthermore, giving restriction of traveling factors, such as budge and time, the recommendation system will refine the exact set of tourism places by applying genetic algorithm mechanism. The system is based on the minimum cost to schedule traveling path from a set of selected places by the using genetic algorithm approach and hybrid recommendation algorithm focuses on the refining efficiency and provides multifunctional tourism information. The hybrid recommendation system proposed employs three steps: 1) Item-based collaborative filtering 2) Genetic algorithm for restriction time and budget and 3) Genetic algorithm to schedule travel path. The above steps provide a recommendation of places, each step refines the active tourist's preferences and satisfy requirements of place-to-place selection.

The study proposed by (Kavinkumar et al., 2015) presents a multi-criteria recommendation system using Collaborative Filtering techniques with external as well as internally mentioned feedback analysis. This approach can be classified into following steps: 1) Data collection, 2) Feature extraction, 3) User-based collaborative filtering, 4) Item-based collaborative filtering and 5) Feedback analysis. In order to process the reviews, the first step is to convert all words to its root form and then remove stop words like 'the', 'is', 'and' etc. Then all opinion words are extracted and weighted depending on how positive or negative they are. This step also involves identification of negation words and parts of speech tagging which are used to identify opinion. The opinion is then used to classify the comment as very negative, negative, neutral, positive or very positive. This is fed back to the system which alters the weight of recommendations based on the opinion rating.

(Tatli & Birtürk, 2011) present a hybrid approach for music recommendation. The first part of this system performs 6 main tasks, which are web crawling, creating an ontology of musical genres, classifying tags according to the ontology, track profiling, user profiling and enacting the recommendation process. In the web crawling phase of the system, a data set is generated. In the ontology creation phase, the authors created a small ontology-a hierarchical structure -with the help of the data crawled from the Dbpedia. In the tag classification phase of the system, the authors parsed instances existing in the ontology into single words. In the track profiling phase, the size of a tracking vector is the size of mainstream genres (22 in their case). In the user profiling phase, user profiles are represented in 3 different ways: 1) using the users' own tags (personal tags) that they entered, 2) using the users' friends' tags (friends' tags) that their friends entered and 3) using all the tags of the tracks (social tags) that they listened to. In the recommendation phase, the authors use the common cosine similarity method. Jaccard index, correlation, Manhattan distance and Hamming distance are some of the other methods available for finding the similarities, but they find out that cosine similarity gave the best results in their study. In addition to the music domain, the authors also use movie domain acquired from users' Facebook profiles. This is the second part of this recommendation process.

(Vekariya & Kulkarni, 2012) proposed a hybrid filtering to recommend restaurants, which transparently creates and maintains user preferences. It assists users by providing both collaborative filtering and content-based filtering, which are updated in real-time whenever the user changes his/her current page using any navigation technique. The WebBot uses the URLs provided in the restaurant dataset to download restaurant database content from the database. WebBot keeps track of each individual user and provides that a user online assistance. The assistance includes two lists of recommendations based on two different filtering paradigms: collaborative filtering and content-based filtering. WebBot updates the list each time the user changes his/her Current page. Content-based filtering is based on the correlation between the content of the pages and the user preferences. The collaborative filtering is based on a comparison between the user path of navigation and the access patterns of past users. The web crawler uses the URLs provided in the restaurant dataset to download hotel content from the database. After appropriate preprocessing, the downloaded content is stored in the hotel Content Database. The hotel dataset also provides the user-ratings matrix, which is a matrix of users versus items, where each cell is the rating given by a user to an item. The system refers to each row of this matrix as a user rating vector. The user-ratings matrix is very sparse since most items have not been rated by most users. The content-based predictor is trained on each user-ratings vector and a pseudo user-ratings vector is created. A pseudo user-ratings vector contains the user's actual ratings and content-based predictions for the unrated items. All pseudo user-ratings vectors put together to form the pseudo-ratings matrix, which is a full matrix. Now given an active user's ratings, predictions are made for a new item using collaborative filtering on the full pseudo rating matrix.

(Esfahani & Alhan, 2013) developed a hybrid recommendation system focused on E-Commerce, which combines the content-based, collaborative and knowledge-based methods. The key idea of content-based filtering approach is based on similarity of item's features. In collaborative filtering approach, a user will prefer those items that like-minded people prefer. The knowledge-based recommendation systems key idea is to use rules to recommend different items to different users. The system uses C-Means Fuzzy clustering method to cluster the items and user profiles. The system can overcome the cold start problem to recommend items that no one in the community has yet buy or rated. After clustering, the system employs a rule generator engine to make rules from the clusters and take it to a knowledge-based part. After all, the system has an interactive user interface to see what user want and customize what the users want to see and make the most successful suggestions.

3.7. Final Remarks

Tourist recommender systems represent a vast research area. Many authors have developed different approaches, which could be focused on generic or specific tourism, hybrid or conventional recommendation strategies. The vast majority of the approaches mentioned previously are focused on individual profiles, keeping in mind that tourism is defined as a social activity to a large magnitude. These recommender systems need a previous characterization of users, which could be carried out by means of evaluation instruments. Evaluation instruments measure different variables with the purpose of quantifying users' tourist behavior. The user profile is calculated based on the characterization of users. Such profile is used to generate recommendations according to the destination selected. Usually, in the related works, the user profile is based on user history, but this approach needs a considerable interaction to carry out successfully recommendations.

Group recommender systems belong to an emerging research area, which has developed in recent years due to the social Web, social activities and the special interest of different markets around the world in satisfying group preferences. The main problem of this systems is developing a suitable group model in order to satisfy the preferences of all group members. Different strategies of group modeling are used in the related work, these strategies could be applied to individual profiles with the purpose of calculating a group profile or individual recommendations to calculate a group recommendation.

The review of the related work presented above has allowed the identification of the contribution this thesis. This work aims to contribute to the design of a hybrid recommendation system, which calculates a set of suitable places in a selected destination according to the preferences of the collective profile. A consensus function is proposed for group modeling strategy. A group hybrid recommendation system based on knowledge-based and non-personalized recommendation system is proposed. In addition, an intelligent approach is presented for tag characterization, these tags represent places or activities inside tourism market. This approach aims to solve problems of group modeling strategies and group recommendation systems, like cold start, new-user, ramp-up and group consensus.

PART II

PROPOSED APPROACH

Chapter 4

Hybrid Recommendation System Approach

This chapter presents the group modelling and hybrid recommendation system approaches proposed in this dissertation applied to group tourist recommendation systems. The main definitions, general considerations and the algorithms for the characterization instrument, group modelling and hybrid recommendation system in this work are introduced in this chapter.

4.1. Problem Statement

Recommendation systems is a vast research area and they were developed to confront the exponential growing of the available information on Internet, called the information overload problem. However, these systems imply a series of additional problems like cold start, new user, new item, among others. These problems are related to the available information of users as well as items. If the system has not enough information it could not recommend in a successful way. Cold start problem consists in an information shortage about the user, cold start user, or an information shortage of items, cold start item. Cold start problem makes difficult to the system extract inferences for users.

Recently, group recommendation systems have become in a new research area, due to group recommendation systems have the biggest level of difficulty than individual recommendation systems. In addition to individual recommendation systems problems, group recommendation systems confront new problems like user aggregation and group modelling. These problems entail a series of solutions about this theme, but these problems still have not a solution because they imply calculate group profiles or group recommendations with the purpose of satisfying group preferences. Recommend items to groups is a difficult task because it implies aggregate preferences of all group members.

This thesis considers the problem of recommending places to collective profiles in order to satisfy the touristic preferences of all group members. As the system must recommend places to collective profiles a group model is proposed, the goal is to identify group preferences and carry out recommendations according to the collective profile through a proposed hybrid recommendation system. The individual profile is calculated with a characterization instrument, survey, which is used to evaluate psychosocial aspects of each user like favorite tourist activities, tourist personality, motivation and tourist experience. The result of this characterization instrument is the individual profile, which represents the touristic preferences of the user. The group profile is calculated based on individual profiles of group members through a proposed group model. Therefore, a group profile is calculated in order to provide information about the group to the recommendation system and avoid the cold start, new user, and group modelling problems. Recommendations are calculated by the proposed hybrid recommendation system, which is based on both knowledge-based and non-personalized recommendation systems.

4.2. Characterization Profiles

The characterization of individual profiles is based on the tourist topologies and their segmentation (see chapter 2). The proposed profile characterization is based on the most representative tourist segments, according to their demand inside tourism market. These segments are called subprofiles inside the proposed individual profile.

4.2.1. Cultural Subprofile

As a segment of the tourism market, cultural tourism has a large tradition inside history of tourist practice (Secall, Bernier, García, & Rojo, 2006). This segment is one of the most ancient and it continues being a fundamental pillar of the tourism industry (Richards & Munsters, 2010). Inside this tourism segment, the main sectors are art tourism, historical tourism, folk tourism, tourism of cultural events (music festival, cinema, drama, among others), gastronomic tourism, and wine tourism, among others.

A Cultural tourist looks into traditional life, language and local habits. Inside this tourism segment, the tourist observes and participates in regional gastronomy, festivals, folklore and other typical activities of the community (Jiménez Bulla & Jiménez Barbosa, 2013).

Therefore, seven categories of cultural attractions are defined (Duhme, 2013)(Shishmanova, 2015):

1. Archeological places, architecture, museums, and monuments.
2. Art, sculptures, handicrafts, galleries, festivals, cultural events, and theme parks.
3. Music, theatre, and dance.
4. Study of literature and language.
5. Religious ceremonies and pilgrimage.
6. Human settlements and ethnic group.

4.2.2. Bioecologic Subprofile

The increasing importance of activities related to having contact with nature and its diversity of ways in recent decades is related to the income generation to the preservation and protection of green areas (Cobbinah, 2015). This segment of tourism market is focused on activities carried out to having contact with nature under the concept of sustainability (Jiménez Bulla & Jiménez Barbosa, 2013). Inside this tourism segment, the main sectors are ecotourism (Mohamad Danial Md Sabri, Suratman, et. al., 2011), agrotourism (Flanigan, Blackstock, et. al., 2014), biological tourism or tourism of nature inspection and hiking (Torres Bernier et al., 2006).

Bioecologic tourist visits green areas in a responsible way with the purpose of enjoying, appreciating and studying natural attractions like landscapes, flora, and fauna (Jiménez Bulla, 2008).

Therefore, six categories of bioecologic attractions are defined (Weaver, 2001)(Das & Chatterjee, 2015):

1. Natural parks.
2. Hiking.
3. Inspection of fauna and flora.
4. Camping
5. Natural attractions
6. Farms

4.2.3. Adventure Subprofile

Adventure tourism has a high level of homogeneity in all over the world according to different studies (Buckley, McDonald, Duan, Sun, & Chen, 2014). This segment of tourism market consists of tourist activities physically demanding, which includes real or perceived risk elements, due to the tourists test themselves (Jiménez Bulla & Jiménez Barbosa, 2013). Sports competitions are excluded of this segment (Guerrero González & Ramos Mendoza, 2014). Adventure tourism does not require luxurious facilities but it requires the necessary equipment, guide service, meet the security requirements and apply preventive environment measures during the stay (Buckley, 2007).

Adventure tourist looks for new and different sensations continuously (Cater, 2006). Adventure tourist crosses limits looking for enjoyment, freedom and new experiences (Carnicelli-Filho, Schwartz, & Tahara, 2010).

Therefore, three categories of adventure tourism are defined (Guerrero González & Ramos Mendoza, 2014)(Carnicelli-Filho et al., 2010)(Buckley, 2012):

1. Mountain climbing, stroll, parade, bicycle touring, mountain biking, hunting, climbing, rappel and speleism (descent in caves).
2. Diving, diving speleism (combination of speleism and diving), rafting and kayak.
3. Ride in a hot-air balloon, hang-glider flight, paragliding, and skydiving.

4.2.4. Urban Subprofile

Little towns and cities have been the center of tourist activities through history, where accommodation, entertainment, and other services are offered to tourists

(Sharpley, 2006). Inside this segment of the tourism market, there is specific tourism like shopping tourism, leisure places, nightlife, coastal touristic environments, business, expos, conferences, and conventions (Rabbiosi, 2015).

Urban tourist is attracted by services of the city, urban way of life (Cibinskiene & Snieskiene, 2015), leisure and relaxation places, business and academic opportunities (Ashworth & Page, 2011).

Therefore, three categories of urban tourism are defined (Carlisle, Johansen, et. al., 2016)(Edwards, Griffin, et. al., 2008)(Ashworth & Page, 2011):

1. Shopping, restaurant, bar, and casino.
2. Business, conferences, conventions and expos.
3. Relaxation, spa, club, and beach.

4.2.5. Sport Subprofile

Sport tourism has had a significant growth in recent years (Kennelly & Toohey, 2014). This segment has transformed in an investigation area and a popular touristic product (Ritchie, Hall, & Cooper, 2004). These practices are carried out in appropriate places to the sport. Inside this segment of the tourism market, there are specific tourism like amateur sport tourism and show sport tourism. Amateur sport tourism is defined as active sports vacations, this class of tourism is played by sport fans. Show sport tourism is defined as passive sports vacations, this class of tourism is made up by spectators of events or sports places.

Sport tourist looks for active or passive participation in tourist activities, in chance or organized way by commercial or business reasons.

Therefore, six categories of sport tourism are defined (Torres Bernier et al., 2006)(Jiménez Bulla & Jiménez Barbosa, 2013)(Ritchie et al., 2004):

1. Sport activities participation.
2. Sport touristic attractions.
3. Sport tourist areas.
4. Sport touristic cruise.
5. Sport touristic excursions.

6. Sport events.

In Table 4.2-1 are shown the description of subprofiles and their characteristic activities inside the tourism market.

Subprofile	Description	Activities
Cultural	A Cultural tourist looks into traditional life, language and local habits. Inside this tourism segment, the tourist observes and participates in regional gastronomy, festivals, folklore and other typical activities of the community.	1. Archeological places, architecture, museums, and monuments. 2. Art, sculptures, handicrafts, galleries, festivals, cultural events, and theme parks. 3. Music, theatre, and dance. 4. Study of literature and language. 5. Religious ceremonies and pilgrimage. 6. Human settlements and ethnic group.
Bioecologic	Bioecologic tourist visits green areas in a responsible way with the purpose of enjoying, appreciating and studying natural attractions like landscapes, flora, and fauna.	1. Natural parks. 2. Hiking. 3. Inspection of fauna and flora. 4. Camping 5. Natural attractions 6. Farms
Adventure	Adventure tourist looks for new and different sensations continuously. Adventure tourist crosses limits looking for enjoyment, freedom and new experiences.	1. Mountain climbing, stroll, parade, bicycle touring, mountain biking, hunting, climbing, rappel and speleism (descent in caves). 2. Diving, diving speleism (combination of speleism and diving), rafting and kayak. 3. Ride in a hot-air balloon, hang-glider flight, paragliding, and skydiving.
Urban	Urban tourist is attracted by services of the city, urban way of life, leisure and relaxation places, business and academic opportunities.	1. Shopping, restaurant, bar, and casino. 2. Business, conferences, conventions, and expos. 3. Relaxation, spa, club, and beach.
Sport	Sport tourist looks for active or passive participation in tourist activities, in chance or organized way by commercial or business reasons.	1. Sport activities participation. 2. Sport touristic attractions. 3. Sport tourist areas. 4. Sport touristic cruise. 5. Sport touristic excursions. 6. Sport events.

Table 4.2-1 Description of subprofiles and their characteristic activities

4.3.Proposed Characterization Instrument

The characterization instrument proposed in this dissertation is based on a hybrid technique to diagnose preferences of tourists inside tourism market context (see Chapter 2). This hybrid technique is composed of a qualitative technique, which obtains psychological aspects of tourists. Also, this hybrid technique is composed of quantitative technique, because it uses a survey as a key element and this survey was designed previously.

The psychological aspects extracted with the proposed characterization instrument are the most useful variables to evaluate the tourist behavior. These variables are tourist personality (Hsu & Huang, 2010), motivation (López-Guzmán, Vieira-Rodríguez, & Rodríguez-García, 2014), favorite tourist activities (Thiengburanathum, Cang, & Yu, 2015) and tourist experience (Valeri, Baez, & Casati, 2013).

- **Tourist personality:** This variable includes the psychological, cognitive and socioeconomic status of tourist (psychosocial variable).
- **Motivation:** The motivation describes the reason to visit a destination (psychosocial variable).
- **Favorite tourist activities:** This variable describes the favorite activities of tourists and travel characteristics (psychosocial variable).
- **Tourist experience:** Tourist experience includes the activities carried out in their past trips (tourist practice).

4.3.1. Survey Model

In chapter 2 the survey was described as a primary source of information over objective, coherent and jointed question combination. Therefore, in this dissertation is proposed a survey to characterize the tourist preferences of users. This survey is implemented in the process of individual profile construction. The purpose of this characterization instrument is to extract the tourist preferences of users, these preferences are quantified in four variables, which are tourist personality, motivation, favorite tourist activities and tourist experience. These variables represent tourist preferences of users and they are directly related to each characteristic of the individual profile.

The proposed survey is classified as an online survey due to its characteristics, implementation and the absence of oral communication between pollster and survey respondent.

This survey could be classified under different criteria (see chapter 2). In this sense, the proposed survey is an organized survey, because it consists of formalized and standardized questions and answers.

The question type of proposed survey could be also classified under different criteria. Therefore, the questions of the survey are closed, polytomous and tandem questions.

The relation among variables to evaluate the tourist behavior, the answer options of the proposed survey and the proposed individual profile is shown in Table 4.3-1.

<table>
<tr><th rowspan="2">Variables</th><th colspan="5">Subprofiles</th></tr>
<tr><th colspan="2">Cultural</th><th colspan="3">Bioecologic</th></tr>
<tr><td>Favorite tourist activities</td><td colspan="2">Get to know the cultural heritage, regional art, architecture, drama, dance and go to cultural events</td><td colspan="3">Get to know the fauna and flora of the zone, camping, and natural attractions</td></tr>
<tr><td>Tourist personality</td><td>Museums, monuments, and archeological places</td><td>Theme parks, galleries, and festivals</td><td>Farms</td><td colspan="2">Natural parks</td></tr>
<tr><td>Motivation</td><td>Taste regional gastronomy and go to regional events</td><td>Get to know the regional ethnic groups and native settlement</td><td colspan="3">Have contact with nature, open field activities and hiking</td></tr>
<tr><td>Tourist experience</td><td colspan="2">Regional culture, monuments, architecture, archeological places</td><td colspan="3">Inspection of fauna and flora, camping, natural attractions</td></tr>
<tr><td colspan="6"></td></tr>
<tr><th rowspan="2">Variables</th><th colspan="5">Subprofiles</th></tr>
<tr><th>Adventure</th><th>Sport</th><th colspan="3">Urban</th></tr>
<tr><td>Favorite tourist activities</td><td>Extreme sports, new and risky places</td><td>Play recreational sports or go to sport events</td><td colspan="3">Shopping and relax</td></tr>
<tr><td>Tourist personality</td><td>Different and risky places</td><td>Sports arena and local tourney</td><td>Restaurants, bar, club, and spa</td><td>Expos, conference, and convention</td><td>Beach, sun and relax</td></tr>
<tr><td>Motivation</td><td>Play extreme sports</td><td>Recreational physical activity and sport events</td><td colspan="2">Sun, beach and relax</td><td>Get to know the city and go to restaurants</td></tr>
<tr><td>Tourist experience</td><td>Activities of extreme physical demands, new experiences, and emotions</td><td>Sports arenas, regional or international sport events, leisure sport</td><td colspan="3">Get to know the city, shopping, relax, sun and beach</td></tr>
</table>

Table 4.3-1 Relation among Variables to Evaluate the Tourist behavior, the Answer Options of Proposed Survey and the Proposed Individual Profile

Tourist personality is evaluated with next question and answer options:

Which places do you prefer during your stay?

- Museums, monuments, and archeological places
- Theme parks, galleries, and festivals
- Farms
- Natural parks
- Different and risky places
- Restaurants, bar, club, and spa

- Expos, conference, and convention
- Beach, sun and relax
- Sports arena and local tourney

Motivation is calculated with next question and answer options:

Which activities do you prefer in your free time?

- Taste regional gastronomy and go to regional events
- Get to know the regional ethnic groups and native settlement
- Have contact with nature, open field activities and hiking
- Play extreme sports
- Get to know the city and go to restaurants
- Sun, beach and relax
- Recreational physical activity and sport events

Favorite tourist activities are evaluated with next question and answer options:

Which activity do you prefer during a trip?

- Get to know the cultural heritage, regional art, architecture, drama, dance and go to cultural events
- Get to know the fauna and flora of the zone, camping, and natural attractions
- Extreme sports, new and risky places
- Shopping and relax
- Play recreational sports or go to sport events

Tourist experience is evaluated with next question and answer options:

Which of the following has been the reason for your previous trips?

- Regional culture, monuments, architecture, archeological places
- Inspection of fauna and flora, camping, natural attractions
- Activities of extreme physical demands, new experiences, and emotions
- Get to know the city, shopping, relax, sun and beach
- Sports arenas, regional or international sport events, leisure sport

4.3.2. Evaluation Model

Measurement scale of attitudes is implemented to define the survey evaluation model. Nowadays, the attitudes study is frequently used in markets investigation. The vast majority of the survey questions are focused on markets investigation. These questions are designed to measure attitudes. The attitudes study are carried out to foresee the tourist's behavior and have the biggest knowledge about them. The attitudes study are considered as precedents to explain the tourist's behavior.

Attitudes could be measured in comparative or non-comparative form. These are different in the possibility of comparing the study objects.

- **Multi-Item Scale**

The scales used to measure aspects or complex phenomenon require a detailed elaboration process, which contains set of rules to guarantee their precision. The measurement of subprofiles is carried out dividing the number of answer options selected by the user between the total of answer options.

- Cultural Subprofile

$$C: Result\ of\ cultural\ subprofile$$
$$k = Total\ of\ answer\ options$$
$$C_i = Number\ of\ answer\ options\ selected\ by\ user$$
$$C = \frac{\sum_1^k C_i}{k} \times 100\% \qquad (4.3.2-1)$$

Subprofile	Variable	Answer option
Cultural	C_1	Get to know the cultural heritage, regional art, architecture, drama, dance and go to cultural events
	C_2	Museums, monuments, and archeological places
	C_3	Theme parks, galleries, and festivals
	C_4	Taste regional gastronomy and go to regional events
	C_5	Get to know the regional ethnic groups and native settlement
	C_6	Regional culture, monuments, architecture, archeological places

Table 4.3-2 Answer Options Related to Cultural Subprofile

- Bioecologic Subprofile

$$B: Result\ of\ bioecologic\ subprofile$$
$$B_i = Number\ of\ answer\ options\ selected\ by\ user$$
$$B = \frac{\sum_1^k B_i}{k} \times 100\% \qquad (4.3.2-2)$$

Subprofile	Variable	Answer option
Bioecologic	B_1	Get to know the fauna and flora of the zone, camping, and natural attractions
	B_2	Farms
	B_3	Natural parks
	B_4	Have contact with nature, open field activities and hiking
	B_5	Inspection of fauna and flora, camping, natural attractions

Table 4.3-3 Answer Options Related to Bioecologic Subprofile

- Adventure Subprofile

$$A: Result\ of\ adventure\ subprofile$$
$$A_i = Number\ of\ answer\ options\ selected\ by\ user$$
$$A = \frac{\sum_1^k A_i}{k} \times 100\% \qquad (4.3.2-3)$$

Subprofile	Variable	Answer option
Adventure	A_1	Extreme sports, new and risky places
	A_2	Different and risky places
	A_3	Play extreme sports
	A_4	Activities of extreme physical demands, new experiences, and emotions

Table 4.3-4 Answer Options Related to Adventure Subprofile

- Urban Subprofile

$$U: Result\ of\ urban\ subprofile$$
$$U_i = Number\ of\ answer\ options\ selected\ by\ user$$
$$U = \frac{\sum_1^k U_i}{k} \times 100\% \qquad (4.3.2-4)$$

Subprofile	Variable	Answer option
Urban	U_1	Shopping and relax
	U_2	Restaurants, bar, club, and spa
	U_3	Expos, conference, and convention
	U_4	Beach, sun and relax
	U_5	Get to know the city and go to restaurants
	U_6	Sun, beach and relax
	U_7	Get to know the city, shopping, relax, sun and beach

Table 4.3-5 Answer Options Related to Urban Subprofile

- Sport Subprofile

$$S: Result\ of\ sport\ subprofile$$
$$S_i = Number\ of\ answer\ options\ selected\ by\ user$$
$$S = \frac{\sum_1^k S_i}{k} \times 100\% \qquad (4.3.2-5)$$

Subprofile	Variable	Answer option
Sport	S_1	Play recreational sports or go to sport events
	S_2	Sports arena and local tourney
	S_3	Recreational physical activity and sport events
	S_4	Sports arenas, regional or international sport events, leisure sport

Table 4.3-6 Answer Options Related to Sport Subprofile

An example of the survey proposed is shown in figure 4.3-1.

Figure 4.3-1 Survey Model Proposed

The result of the survey shown in Figure 4.3-1 is shown in Figure 4.3-2.

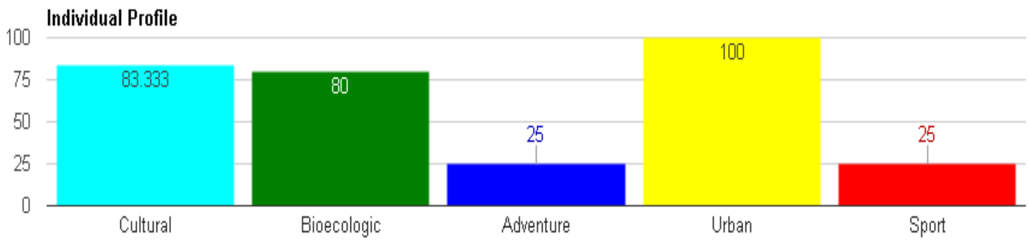

Figure 4.3-2 Result of Survey

Individual profile is represented by (4.3.2-6).

$$P = [C \quad B \quad A \quad U \quad S] \qquad (4.3.2-6)$$

Where:

P: Individual profile.

C: Cultural subprofile calculated with (4.3.2-1).

B: Bioecologic subprofile calculated with (4.3.2-2).

A: Adventure subprofile calculated with (4.3.2-3).

U: Urban subprofile calculated with (4.3.2-4).

S: Sport subprofile calculated with (4.3.2-5).

4.4. Group Modelling Strategy

The group modeling strategy selected in this dissertation is the group model based (GMB) because it offers more flexibility to make recommendations. In group model based (GMB) is calculated a unique group profile to make recommendations. This approach allows to make a recommendation to the group and not make different recommendations for the same group. This process optimizes the efficiency of the system and makes it independent of the number of member of the group. This model builds a group profile based on each profile of group members through a consensus function.

The consensus function proposed in this dissertation is the median. This approach could be described as the median function applied to the individual profiles set. This consensus function calculates the group profile based on individual profiles set. The median function is used due to it is not influenced by extreme values. It can calculate for any type of quantitative data set. This is the most representative measure of central tendency in the case of variables in ordinal scale (William Navidi, 2006). This function is described in (4.4-1).

$$m = \begin{cases} D_{\left(\frac{i+1}{2}\right)}, & for\ i\ odd \\ \left(D_{\left(\frac{i}{2}\right)} + D_{\left(\frac{i}{2}+1\right)}\right)/2, & for\ \ i\ \ even \end{cases} \qquad (4.4-1)$$

Where:

m: Result of the median function.

D: Data set.

i: Number of elements in Data set.

4.5. Hybrid Recommendation System

In this research is proposed the design and implementation of a hybrid recommendation system based on the collective profile construction, tag characterization, and data preprocessing. The collective profile construction consists of building a collective profile in order to represent general preferences of group members, this task is made by means of the consensus function (see chapter 2). The fundamental of this approach is described in section 2.2 and the proposed approach is defined in section 4.4. Tag characterization consists in calculating a tag profile according to different variables to evaluate, this tag profile has the same characteristics or subprofiles defined to the user and collective profile. Each tag represents a type of places or activities available in the destination, these tags are taken of the Foursquare API and these are selected based on the activities described in the characterization profiles (see Section 4.2). The relationship between tags and subprofiles is described in Annex A. The belonging level of tags to each subprofile is defined in following sections. The goal of data preprocessing is setting up similarities between collective profile and tag profiles with the purpose of selecting these tags and choosing the most similar tags. Inside the data preprocessing is carried out a subprocess, this consists in selecting the items of each subprofile by their rating. This is the last step of the proposed recommendation system. In Figure 4.5-1 is shown the process described previously.

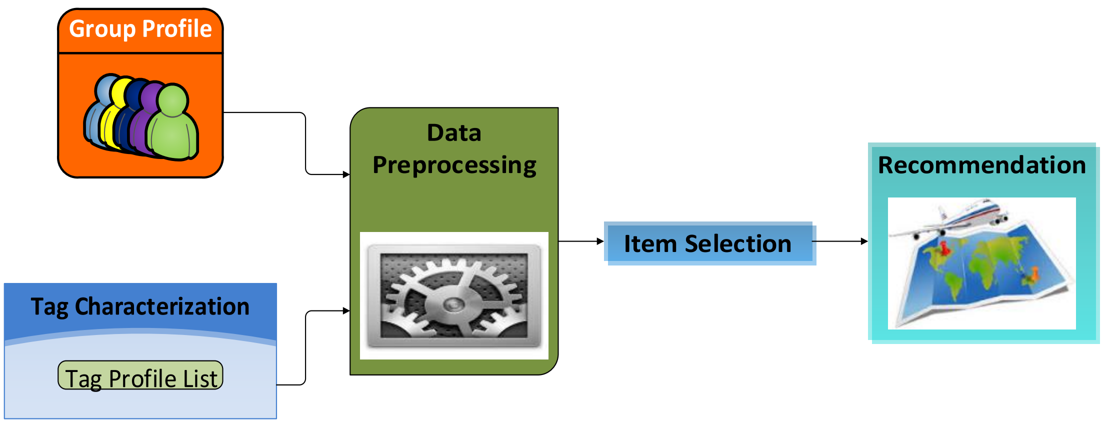

Figure 4.5-1 Hybrid Recommendation System

The proposed hybrid recommendation system consists in a combination of a personalized and non-personalized recommendation system. Personalized recommendation system is based on knowledge-based strategy, due to the proposed approach consists in generating knowledge about the different tags related to tourist activities according to the selected destination. The non-personalized recommendation system is the last step in this system, which filters items by their rating. Non-personalized recommendation system selects the number of items according to the belonging level of the collective profile in the different subprofiles, i.e., the number of items to recommend of each subprofile is directly proportional to the belonging level of the collective profile to each subprofile.

4.5.1. Tag Characterization

The proposed tag characterization consists in a hybrid method of characterization. The proposed method consists in combining an intelligent system and fixed characteristics described in the tourism literature. In this characterization, tags have a completely belonging in a specific subprofile (see section 4.2). But, these tags have a level of belonging in another subprofiles. The level of belonging in another subprofiles is based on different characteristics according to each subprofile.

- Bioecologic

Bioecologic subprofile has twenty-one (21) tags with direct relation (see section 4.2.2). The level of belonging of the remaining tags is determined by the geographical position of the places related to each tag. If all places of a tag are outside of the urban area of the destination, the level of belonging of this tag is one hundred (100) percent. An intelligent system determines if a place is on the outside of an urban area. This intelligent system is trained with a database of urban areas of all over the world. In this sense, the quantity of places outside of the urban area determines the percentage of belonging of the tag to the bioecologic subprofile.

- Adventure

In adventure subprofile, there are twenty (20) tags with direct relation. These tags are activities or places with a complete belonging to this subprofile (see section 4.2.3). The remaining tags have a less level of belonging. The level of belonging of these remaining tags is determined by their own level of risk. The level of risk of these tags is determined in a generic form (Mark Piekarz, Ian Jenkins, 2016)(Hsu, 2006). In this sense, the tags with direct relation have a high level of risk and other tags have a medium or low level of risk. Tags with a high level of risk have a one hundred (100) percent level of belonging, tags with a medium level of risk have thirty (30) percent level of belonging and tags with a low level of risk have ten (10) percent level of belonging. This classification is carried out in by means of a management system of risk (Olcina Cantos, 2012).

- Urban

Urban subprofile has sixty-two (62) tags with direct relation (see section 4.2.4). The level of belonging of the remaining tags is determined by the geographical position of the places related to each tag. If all places of a tag are inside the urban area of the destination, the level of belonging of this tag is one hundred (100) percent. An intelligent system determines if a place is inside of an urban area. This intelligent system is trained with a database of urban areas of all over the world. In this sense, the number of places inside the urban area determines the percentage of belonging of the tag to the urban subprofile.

- Sport

There are nineteen (19) tags with direct relation to sport subprofile. These tags have a complete belonging to the sport subprofile (see section 4.2.5). The level of belonging of the remaining tags is determined by their own level of potential physical activity. The level of potential physical activity is determined by the lack of risk and the presence of physical activity of tags. If the tag implies a physical activity this variable takes a value of one (1), on the contrary, the variable takes a value of zero (0). In this dissertation, it is proposed to calculate this belonging of the remaining tags with (4.5.1-1).

$$S_i = PA * (100 - A_i) \qquad (4.5.1 - 1)$$

Where:

S_i: Level of belonging to the sport subprofile.

PA: Physical activity related to the tag.

A_i: Level of belonging to the adventure subprofile.

i: Position of tags in the list of tags related to the tourism.

- Cultural

In cultural subprofile, there are thirty-nine (39) tags with direct relation (see section 4.2.1). The level of belonging of the remaining tags is determined by the remaining subprofiles of each tag. The level of belonging of the remaining tags is calculated with (4.5.1-2). This calculation is proposed due to the cultural characteristic is determined by the influence of the place or activity in question.

$$C_i = mean(A_i, B_i, U_i, S_i) \qquad (4.5.1 - 1)$$

Where:

C_i: Level of belonging to the cultural subprofile.

A_i: Level of belonging to the adventure subprofile.

B_i: Level of belonging to the bioecologic subprofile.

U_i: Level of belonging to the urban subprofile.

S_i: Level of belonging to the sport subprofile.

i: Position of tags in the list of tags related to the tourism.

4.5.2. Intelligent system

The intelligent system proposed is trained with the urban areas database. This database contains the urban areas in all over the world. The database is taken from Natural Earth, this is a public domain map dataset with maps in different scales. The intelligent system classifies the geolocation of each tourist place and it determines if the places are inside or outside the urban area of the destination. This classification determines the level of belonging of tags to the urban and bioecologic subprofiles, with the exception of tags with a complete level of belonging to these subprofiles. The number of places inside the urban area of the destination determines the percentage of belonging to the urban subprofile. The number of places outside the urban area of the destination determines the percentage of belonging to the bioecologic subprofile.

4.5.3. Data Preprocessing

In many situations of real life, data needs to be preprocessed with the purpose of using this information in analysis techniques. In this dissertation is used an intelligent system to carry out data processing. In this approach are used similarity metrics in data preprocessing in order to filter and select the most similar tags to the collective profile.

○ **Presence of places**

The proposed concept of the presence of places is based on the relation between tags and the places in the destination. If any tag has not relation with places in the destination, this tag has not the presence of places. This concept is implemented in this thesis in order to select tags with related places. If any tag has not the presence of places, it is excluded from the recommendation process. This selection is carried out after the download of places information. In the Figure 4.5-2 is shown the algorithm for selecting tags with the presence of places.

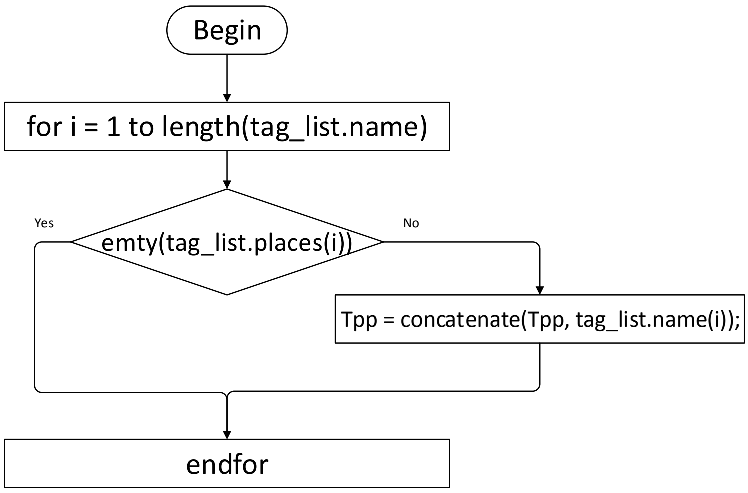

Figure 4.5-2 Algorithm for selecting tags with presence of places

Where:

tag_list: List with information downloaded of places database.

Tpp: List with information of tags with presence of places.

○ **Similarity Metrics**

Similarity metrics used in recommendation systems were described in chapter 2. These metrics are used to carry out comparisons and select items to recommend in different recommendation strategies. Recommendation strategies implement different methods, which are user-user, user-item or item- item comparison. In this research is proposed to carry out a collective profile-tag comparison in order to calculate the most similar tags. The result of these metrics is a scalar, which represents the similarity level in order to determine the most similar tags and discarding the less similar tags. After that, items related to the selected tags are filtering by their rating, this approach is described in following subsections.

In order to select the most similar tags to the collective profile is used a threshold filtering method. This method selects the tags with the highest similarity weight (Ricci et al., 2011)(Mehta & Javia, 2015). The threshold filtering method is more flexible than other methods of neighbor selection, like top-N filtering or negative filtering (Ricci et al., 2011). This technique is unusual in recommendation systems because may be difficult to determine a suitable value of the threshold. The proposed method to determine a threshold of tags number to doing recommendations is calculating the square root of the tags with the presence of places been close to overapproximation (4.5.2-1).

$$Trec \approx \sqrt{Tpp} \qquad (4.5.2-1)$$

Where:

Tpp: Tags with presence of places.

$Trec$: Tags to doing recommendation.

- **Item Selection**

Item selection is the last step of the proposed system. This consists of selecting items according to their rating. Items to recommend are selected in proportion to the belonging level of the collective profile in each subprofile. This proportion is calculated with the function described in (4.5.2-1) for cultural subprofile.

$$N_C = \frac{C}{\sum_{i=1}^{n} CP_i} \qquad (4.5.2-1)$$

Where:

N_C: Number of recommended cultural items.

C: Belonging level of collective profile in cultural subprofile.

CP: Collective profile.

i: Number of subprofiles.

The number of items to recommend related to bioecologic subprofile is calculated by (4.5.2-2).

$$N_B = \frac{B}{\sum_{i=1}^{n} CP_i} \qquad (4.5.2-2)$$

Where:

N_B: Number of recommended bioecologic items.

B: Belonging level of collective profile in bioecologic subprofile.

Adventure items to recommend are given by (4.5.2-3).

$$N_A = \frac{A}{\sum_{i=1}^{n} CP_i} \qquad (4.5.2-3)$$

Where:

N_A: Number of recommended adventure items.

A: Belonging level of collective profile in adventure subprofile.

The number of urban items to recommend is calculated with (4.5.2-4).

$$N_U = \frac{U}{\sum_{i=1}^{n} CP_i} \qquad (4.5.2-4)$$

Where:

N_U: Number of recommended urban items.

U: Belonging level of collective profile in urban subprofile.

The number of items to recommend related to sport subprofile is calculated by (4.5.2-5).

$$N_S = \frac{S}{\sum_{i=1}^{n} CP_i} \qquad (4.5.2-5)$$

Where:

N_S: Number of recommended urban items.

S: Belonging level of collective profile in urban subprofile.

4.6.General Considerations

Let us suppose that a group of people organize a trip. Each group member has different preferences inside the tourism supply. These preferences are represented by an individual profile. The proposed individual profile (P) represents the level of belonging to the different general characteristics of the tourism market, these are called subprofiles. In order to build a suitable collective profile (CP) to represent the preferences of all group members, it is used a group model based strategy based on a proposed consensus function. The consensus function calculates the collective profile based on the individual profiles. The group must select the destination city to begin the download of the places information and later starting the recommendation process. In this sense, the system takes the information of collective profile and the places information to do the tag characterization and data preprocessing. In tag characterization is applied the presence of places concept, i.e., in this step tags without related places are excluded. Inside the tag characterization is built a new table (Tpp) with the places information.

After the places information table is completed, the tag characterization is carried out. The tag characterization consists in calculating the level of belonging to the different subprofiles for each tag. In this sense, the system calculates the similarity between collective profile and all tags with the presence of places, after that the tags are organized according to the result of similarity metric applied inside each subprofile.

With tags organized by their level of similarity, the threshold filtering is applied to select the tags with the highest level of similarity. This selection is carried out by threshold filtering method based on the function proposed in section 4.5.3.

After the threshold filtering method is applied, the number of places to recommend is selected according to the level of belonging to the collective profile in each subprofile and the number of places determined by the group.

With places selected, the recommendation process is concluded and the places to recommend are shown to the group members. The complete recommendation process is represented in Figure 4.6.1.

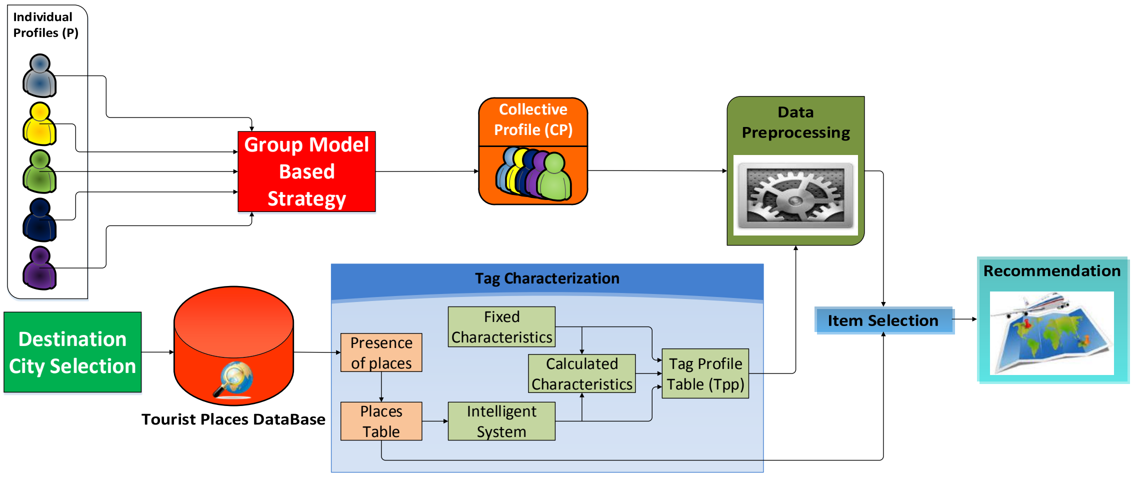

Figure 4.6-1 Representation of recommendation process

Chapter 5

Hybrid Recommendation System Implementation

This chapter presents the application of the proposed approach in a hybrid recommendation system environment to recommend tourist places to collective profiles. It describes some considerations about the environment and simulation system used to develop the experiments, a new proposal in a group modelling strategy to calculate the collective profile according to preferences of group members and the proposal of a hybrid recommendation system. Likewise, hybrid recommendation systems proposed in the literature and in this proposal and the consensus function for the group model based strategy are analyzed.

5.1. Developed System

In this dissertation was developed a Web application (OdinTrip). OdinTrip is a web site designed to manage the user information, group information, and the recommendation process. This web site is developed in PHP, JavaScript, and CSS. Moreover, OdinTrip uses a database based on SQL. OdinTrip uses the API of Google Places, Foursquare and Google Maps inside its recommendation process.

OdinTrip welcomes to users in the index interface (Figure 5.1-1). This web site takes the user information directly. It has an interface which users can create their own account (Figure 5.1-2) or logging in (Figure 5.1-3). The users can provide their tourist information, creating groups and managing their groups to create trips and getting recommendations.

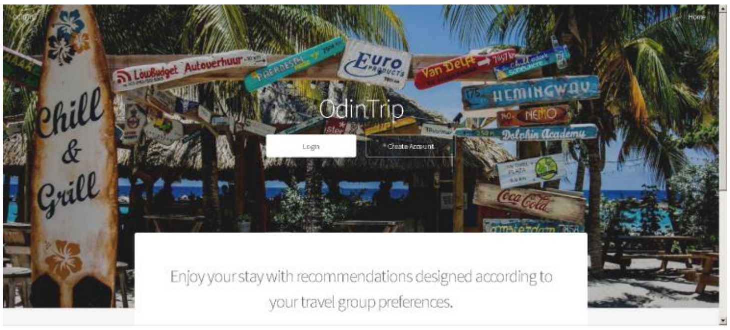

Figure 5.1-1 OdinTrip

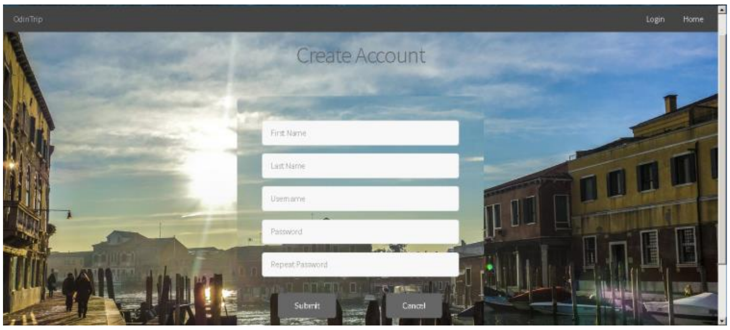

Figure 5.1-2 Create Account

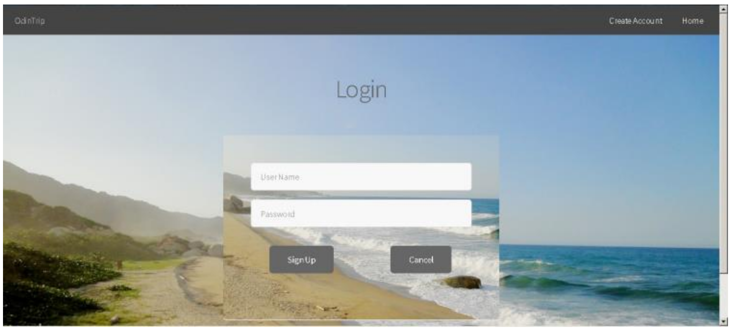

Figure 5.1-3 Login

5.2. Characterization Profile and Group Model

OdinTrip takes the tourist information of users by means of a survey. This survey evaluates different variables related to the tourist behavior (see Chapter 4). It has four questions with multiple-choice, Figure 5.2-1. This survey is part of the characterization profile.

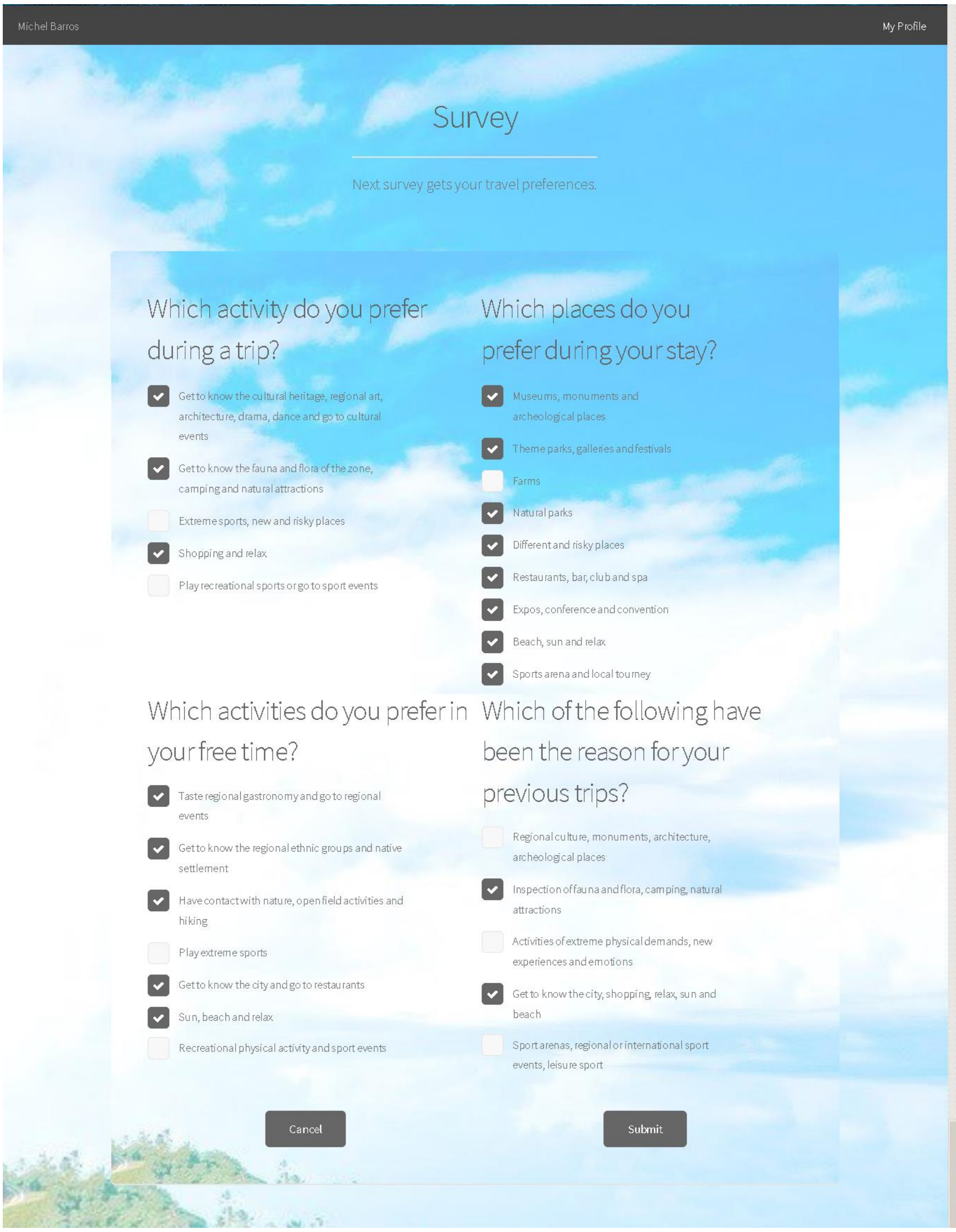

Figure 5.2-1 Survey

OdinTrip takes the user information to calculate the individual profile, Figure 5.2-2. In the individual profile, the users can create groups and recommendations with the created groups previously. The groups could be created with the option "Create Group" in the upper toolbar. The user also can edit the survey according to the recent tourist experience or the change in the preferences. This option also is situated in the upper toolbar.

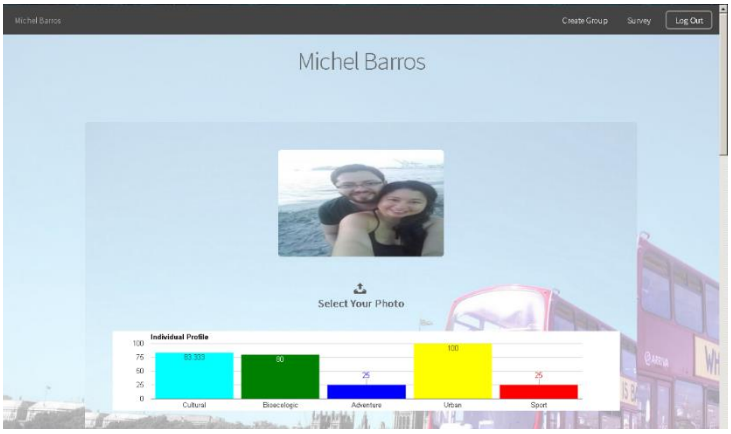

Figure 5.2-2 Individual Profile

OdinTrip has an interface where users can create groups. In this interface, the user must select a group name, a group photo and the remaining group members, Figure 5.2-3. The information of the created group is saved in the database.

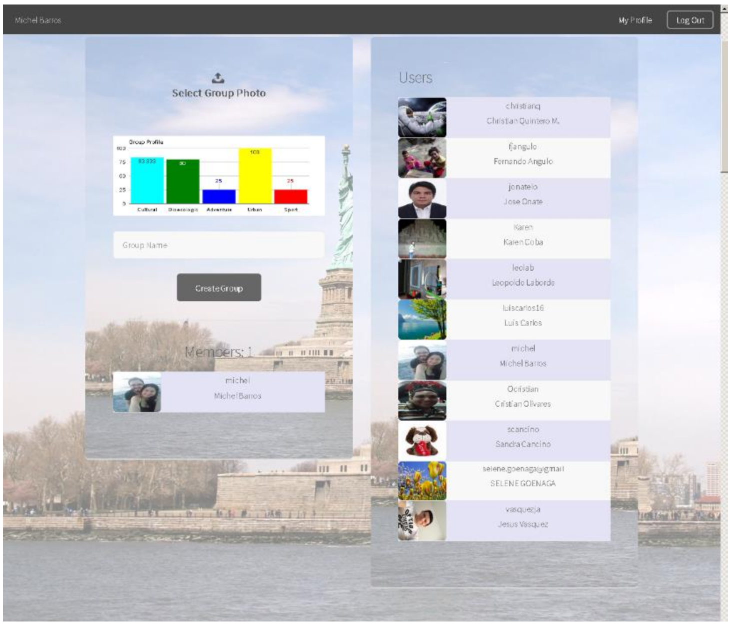

Figure 5.2-3 Create Group Interface

The created group can be edited in the collective profile interface, Figure 5.2-4. In this interface the group members can include new users, deleting members or changing the group photo. This created groups also can be deleted.

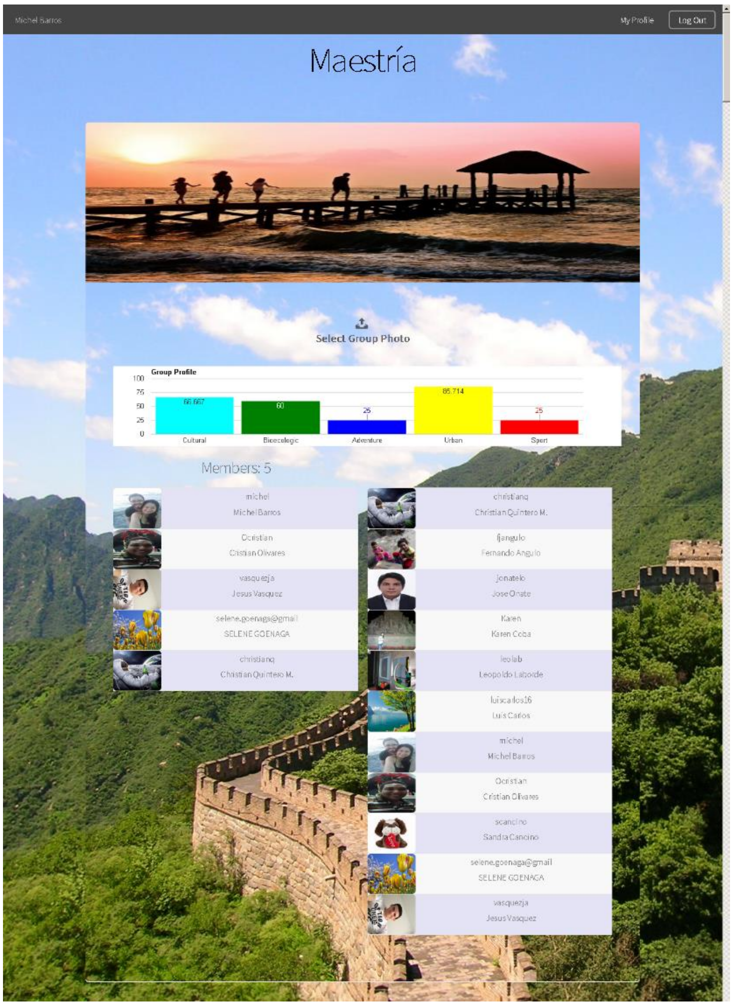

Figure 5.2-4 Collective Profile

To start the recommendation process the user must indicate the destination city, the group and the number of places to the web site. The recommendations can be managed in the individual profile, Figure 5.2-5. In this part are shown the list of groups and the recommendation section with the purpose of managing the recommendation process in an easy way.

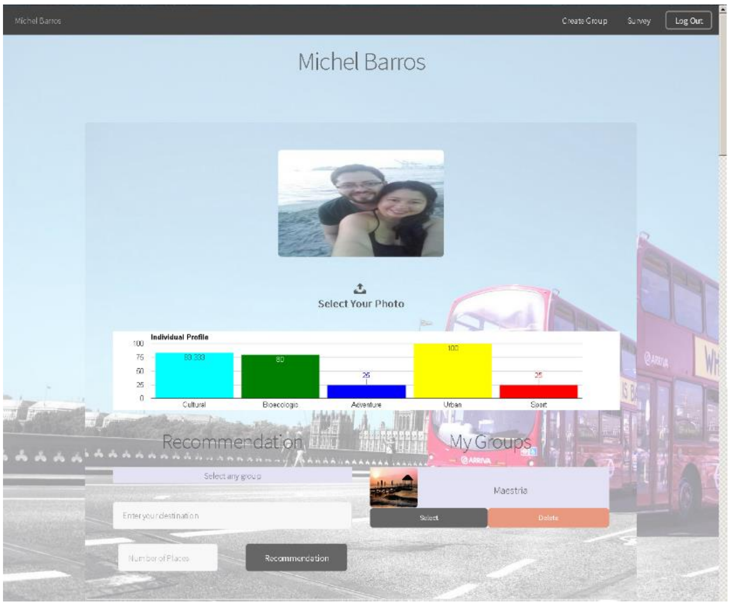

Figure 5.2-5 List of Groups and Recommendation Section

In Figure 5.2-6 is described the process carried out by the user. This is the individual section of the designed algorithm for the web site. The algorithm describes the information flow and functions of the system. This section describes the beginning of the individual process. The individual process implies the account creation or login, the supplying of the individual information and the possibility of creating groups or recommendations. This section is connected with the remaining sections of the algorithm. The other sections manage the creation of groups and the recommendation process. The number four connector carries the user to the create group interface. The number five connector carries the user to the edit group interface. The number seven connector carries the user to the recommendation interface after the selection of the group, the destination city and the number of places.

Group process is another section of the algorithm, Figure 5.2-7. The group process implies the creation and the edition of groups. In the group creation, the users must select the group name, the group photo and the remaining members of the group. In the group edition, the members could include new members or they could delete other members. With these actions, the collective profile is recalculated.

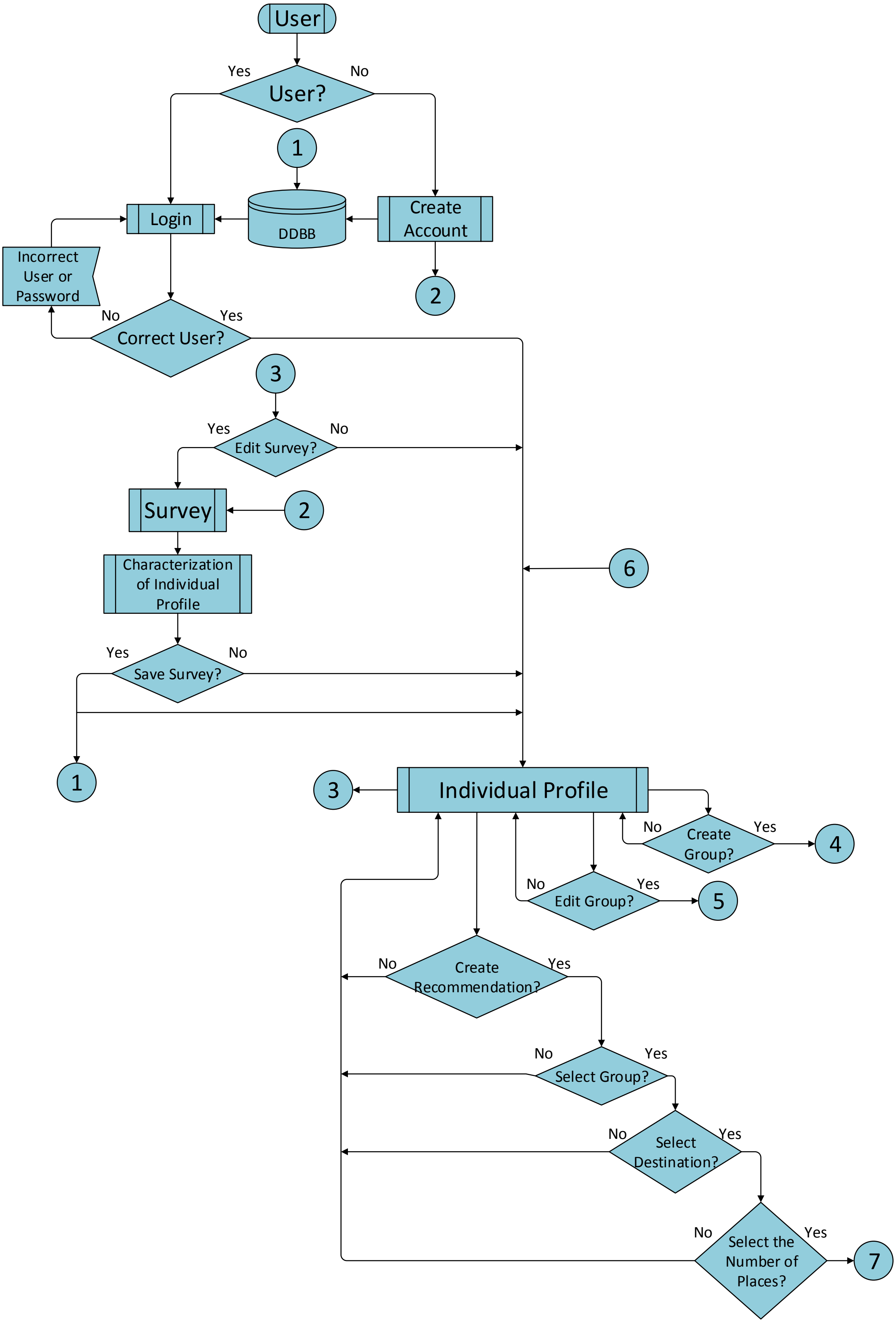

Figure 5.2-6 Individual Section of the Designed Algorithm

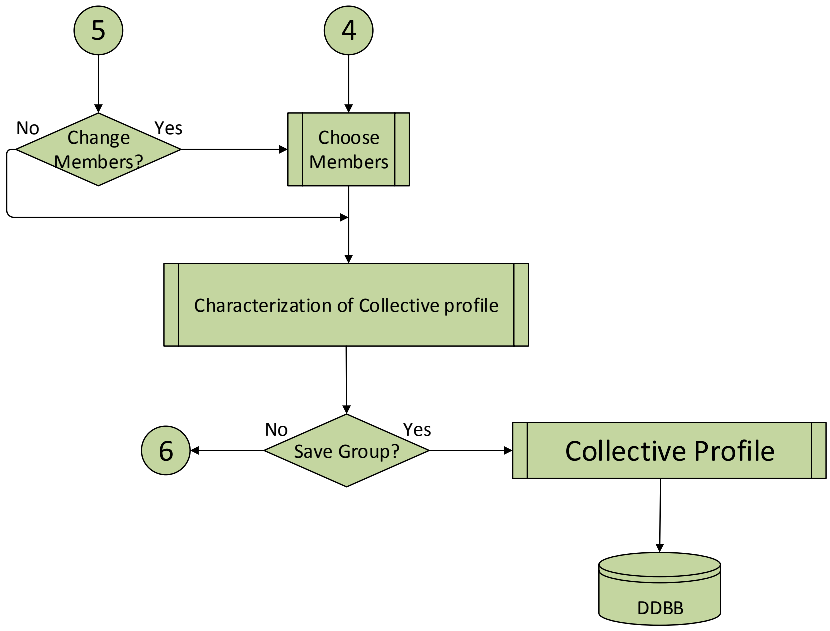

Figure 5.2-7 Group Process

5.3. Hybrid Recommendation Algorithm

The recommendation process begins with the selection of the group name, the group photo, and the remaining members. The recommendation process carries out the comparison between tags profile and collective profile by means of the similarity metric. The result of this comparison is the level of similarity of each tag with the collective profile. Tags are organized according to their level of similarity. After that, the tags with the highest level of similarity in the different subprofiles are selected by means of the proposed threshold filtering function (see Chapter 4). In this sense, the places with the highest rating are selected in proportion to the level of belonging to the group in each subprofile. This process is described in Figure 5.3-1. This process is not shown to the users.

The resulting recommendation is shown in the visualization process, Figure 5.3-2. The visualization process consists in showing the recommended places on the website. These places are shown on a dynamic map. The information of the recommended places also is shown in the information window of each marker. The markers are shown in different colors according to their belonging to the subprofiles.

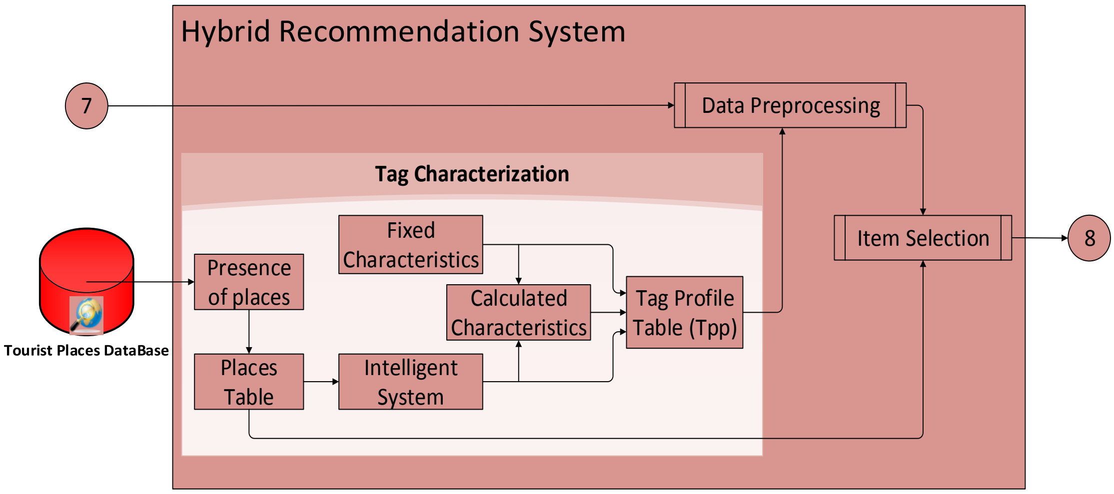

Figure 5.3-1 Recommendation Process

The proportion of recommended places is determined by the level of belonging to the collective profile in the different subprofiles, this proportion is calculated with the proposed item selection method (see Chapter 4). In the Figure 5.3-3 is shown an example of the resulting recommendation on the web site.

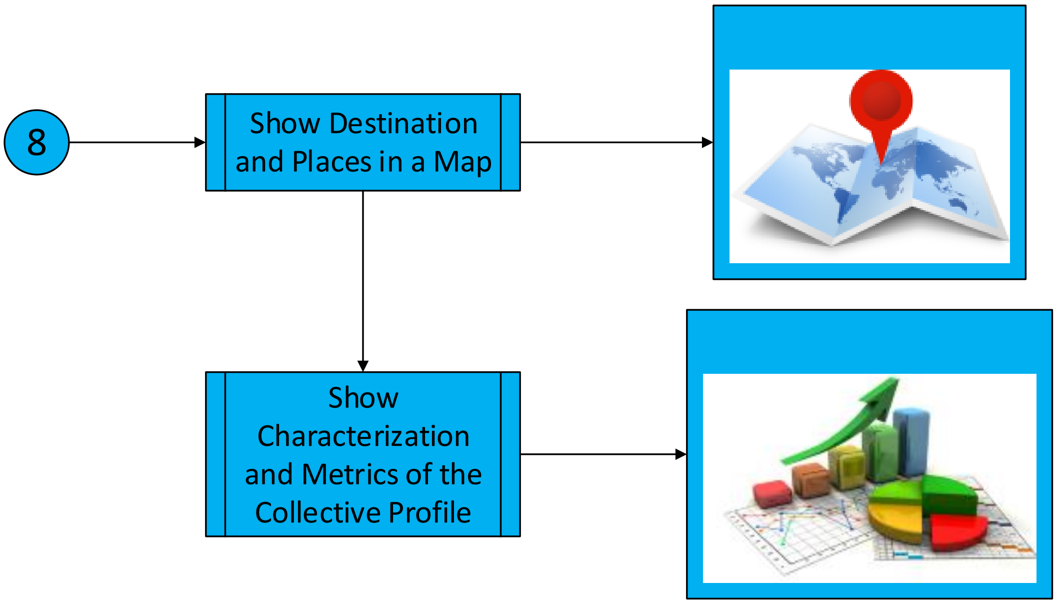

Figure 5.3-2 Visualization Process

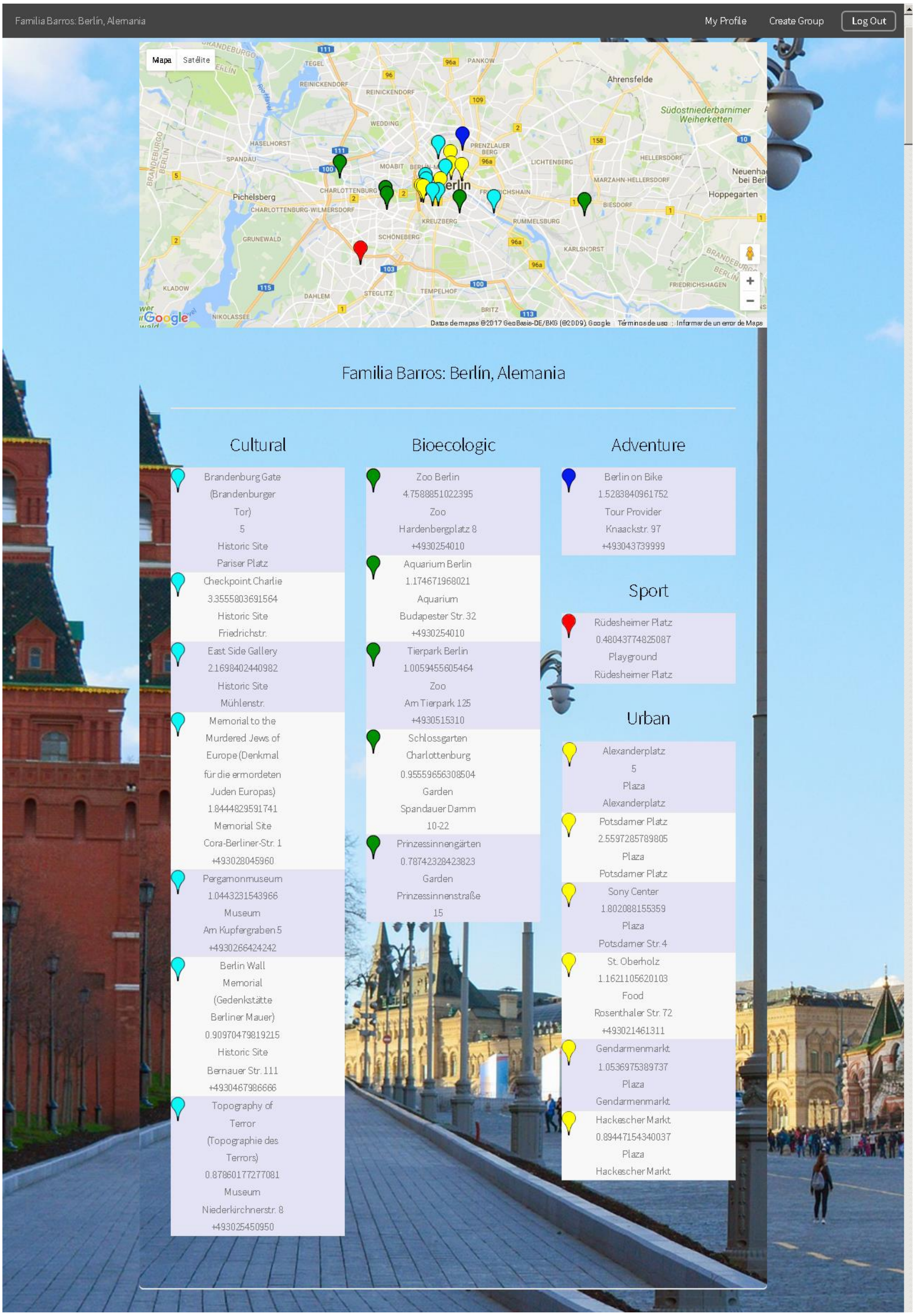

Figure 5.3-3 Recommendation

PART II

EXPERIMENTS RESULTS AND CONCLUSIONS

Chapter 6

Analysis of the Experimental Results

This chapter presents the discussion and analysis of the empirical experiments and testing that have been carried out for the proposed test beds. The results depicted in this chapter demonstrate the utility, feasibility, and reliability of the overall proposed approach presented in the previous chapters.

6.1. Experimental Design

The described approach has been implemented on the algorithm based on the GTravel system, Figure 6.1-1. This algorithm computes a series of simulation experiments to get a quantitative assessment of the proposed approach over the non-intelligent techniques.

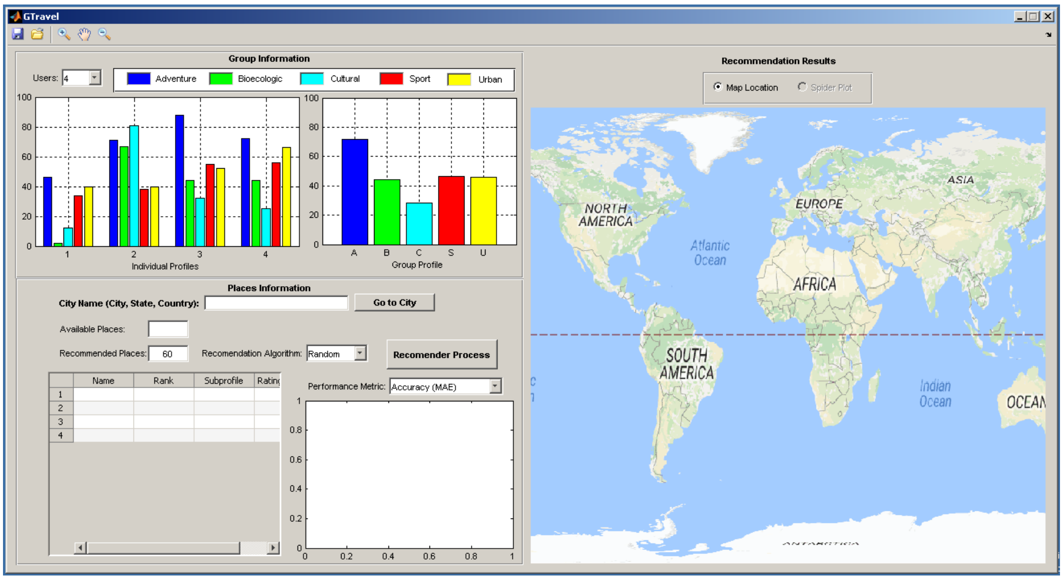

Figure 6.1-1 GTravel System

Each simulation experiment involves a trip to a destination city and a collective profile formed by groups among two or nine members. In addition, each experiment has been performed with a different and random configuration of the trips for making independent data collection. This includes different destination cities and group members. In Figure 6.1-2 is shown an example of a group with two (2) members and its collective profile.

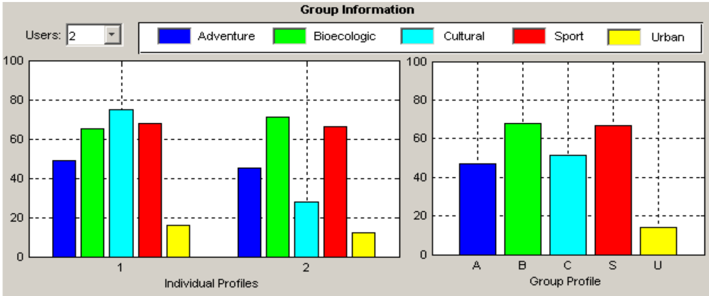

Figure 6.1-2 Example of a group with two (2) members

In Figure 6.1-3 is shown an example of a group with six (6) members and its collective profile.

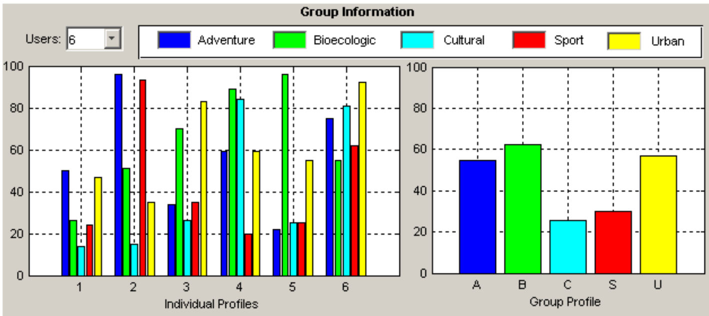

Figure 6.1-3 Example of a group with six (6) members

With this in mind, the experimental design is defined as follow:

<u>Hypothesis or presentation of the problem:</u>

What similarity metric should use to make effective recommendations for collective profiles, having in mind the destination city and the number of group members? This determines which the similarity metric to accomplish this goal is.

<u>Response Variables:</u>

1. Accuracy Rating Prediction (Root Main Square Error): this variable evaluates the accuracy rating prediction of the items. This variable determines the margin of error among the collective profile and the calculated recommendations. If the result of this metric is zero (0) there is not error.
2. Accuracy Rating Prediction (Main Absolute Error): this metric also determines the margin of error among the collective profile and the calculated recommendations. Different from the RMSE, this metric could

detect the big errors in a better way. If the result of this metric is zero (0) there is not error.

3. Geographical Dispersion: this is a proposed metric of this dissertation. This metric is calculated based on the recommendation area and the distance among the recommended places and the half-way point of the recommendation area. The median distance of the places to the half-way point is the dispersion. This metric has units in kilometers. If the result of this metric is zero (0) there is not dispersion. This metric is described in (6.1-1).

$$dp = median\left(\sqrt{(mx - lon)^2 + (my - lat)^2}\right) \quad (6.1 - 1)$$

Where:

dp: Geographical dispersion

mx: Median of the longitude vector

my: Median of the latitude vector

lon: Longitude vector

lat: Latitude vector

4. Accuracy Usage Prediction: this metric is used to evaluate the user preference of the items. This metric calculates the proximity among the result of the recommendation and the user preferences. The accuracy usage prediction and the next metrics are based on the confusion matrix shown in Table 6.1-1.

	Recommended	Non-recommended
Used	True-Positive(tp)	False-Negative(fn)
Non-used	False-Positive(fp)	True-Negative(tn)

Table 6.1-1 Confusion Matrix

The result of this metric change among zero (0) and one (1). If the result is zero the recommendation is cut off from the user preferences. If the result is one the recommendation is perfect.

5. Precision: this metric is defined as the fraction of the recovered items which are relevant to the user preferences. If the result is zero (0) the recommendation is cut off from the user preferences. If the result is one (1) the recommendation is perfect.

6. Sensitivity: this metric calculates the number of real positive which have been identified correctly. If the recommendation is cut off from the user preferences the result is zero (0). If the recommendation is suitable for the user preferences the result is one (1).
7. Mathew's Correlation Coefficient (MCC): this metric is the correlation coefficient among the observed and predicted binary classification. The result of this metric change among minus one (-1) and one (1).

Factors:

1. Random
2. Euclidean Distance
3. Manhattan/City Block Distance
4. Minkowski Distance
5. Pearson´s Correlation
6. Cosine Similarity
7. Tanimoto Coefficient

The number of possible trips (N) is calculated by means of the number of possible groups (g) and the number of the selected destination cities (cities) for the experiment. In this case, there are nine (9) possible groups, the number of group members varies among two (2) or ten (10). As destination cities, the most popular then tourist cities in all over the world has been selected.

1. London, United Kingdom
2. Istanbul, Turkey
3. Paris, France
4. Prague, Czech Republic
5. Rome, Italy
6. Hanoi, Vietnam
7. New York, United States of America
8. Ubud, Indonesia
9. Barcelona, Spain
10. Lisbon, Portugal

$$N = \ g * cities = 9 * 10 = 90 \qquad (6.1 - 2)$$

The totally population of possible trips is calculated in (6.1-2) and the result is ninety possible trips. A sample (n) of the population of possible trips could be taken by means of the confidence interval (Z_α), the standard deviation (σ) and the suitable boundary of the sample error (e). The size of the sample is calculate by means of (6.1-3) (Marquéz, Ibarra, & León, 2010).

$$n = \frac{N * \sigma^2 * {Z_\alpha}^2}{e * (N-1) + \sigma^2 * {Z_\alpha}^2} \qquad (6.1 - 3)$$

The selected margin of the confidence interval is 95% and the suitable boundary of the sampling error is 5%. The value of the standard deviation is determined with a test pilot of the experiment. This test pilot is carried out with ten samples. The test pilot is carried out with the same conditions of the experiment.

As a result of the test pilot it was calculate the next value for the standard deviation:

$$\sigma = 0{,}2533$$

The calculated value of the standard deviation is used to carry out the experiment.

$$n = \frac{N * \sigma^2 * {Z_\alpha}^2}{e * (N-1) + \sigma^2 * {Z_\alpha}^2} = \frac{(90) * (0{,}2533)^2 * (1{,}96)^2}{(0{,}05) * (90 - 1) + (0{,}2533)^2 * (1{,}96)^2} = 47{,}3042$$

$$n \approx 47$$

The result of the number of samples is forty-seven (47) possible trips.

For each response variable is carried out a factorial design where it is had one factor and seven levels. Each experiment has forty-seven (47) repetitions. For each repetition, it is selected randomly one destination city and the number of group members in order to form the possible trips.

6.2.Analysis of Results

Group recommendation systems need a suitable strategy to calculate suitable items for users. The method used to determine the suitable items is carrying out item-item, item-user or user-user comparisons by means of the similarity metrics. These similarity metrics calculates the level of similarity among the used factors according to the implemented recommendation system.

In the proposed approach is designed a collective profile-tag comparison, in order to determine the suitable tags for the collective profile (see Chapter 4). To start this stage of analysis and discussion of results, based on the proposed approach and the exposed implementation, in Chapter 5 have been exposed the designed group recommendation system. Based on this the random recommendation algorithm does not use any comparison method among collective profile and tags. The random recommendation algorithm takes tags randomly and later takes places randomly of the selected tags previously. The other algorithms use the similarity metrics exposed in Chapter 2 to compare the collective profile and the tags profile. After that, the items are selected according to their level of reputation. The system recommends items with a high level of reputation. In the Figure 6.2-1 is shown the result of the experiment.

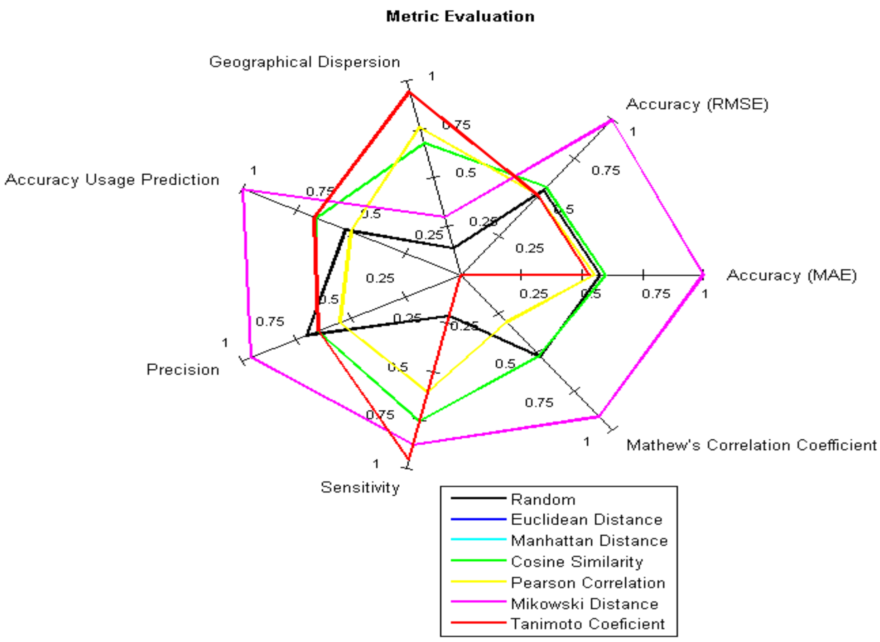

Figure 6.2-1 Result of the Experiment

In Figure 6.2-1 it is observed the Mikowski distance as the best similarity metric. But the Manhattan and Euclidean distances have the same performance. These cases are shown in Figure 6.2-2 and Figure 6.2-3.

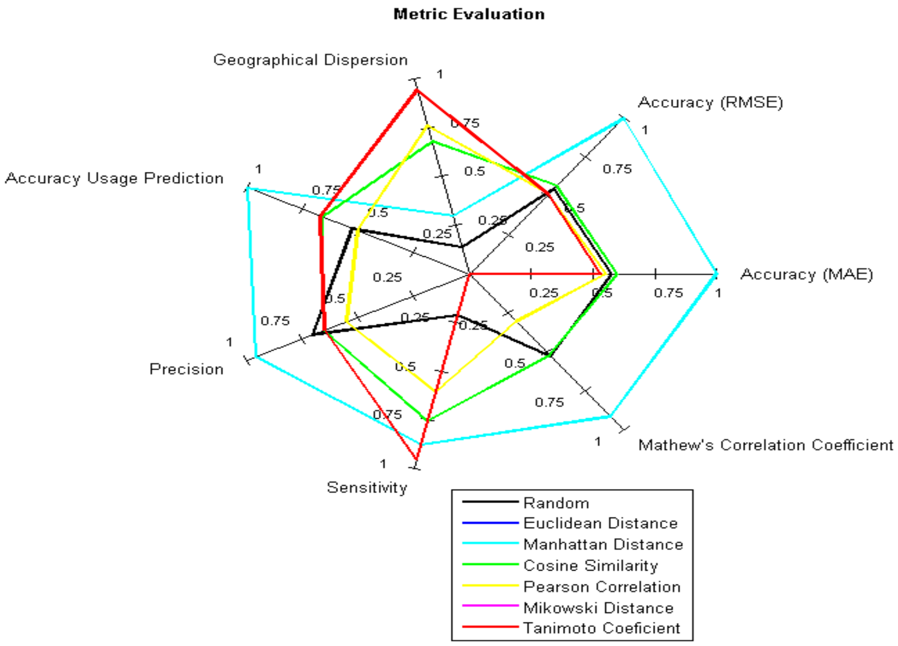

Figure 6.2-2 Manhattan Distance Behavior

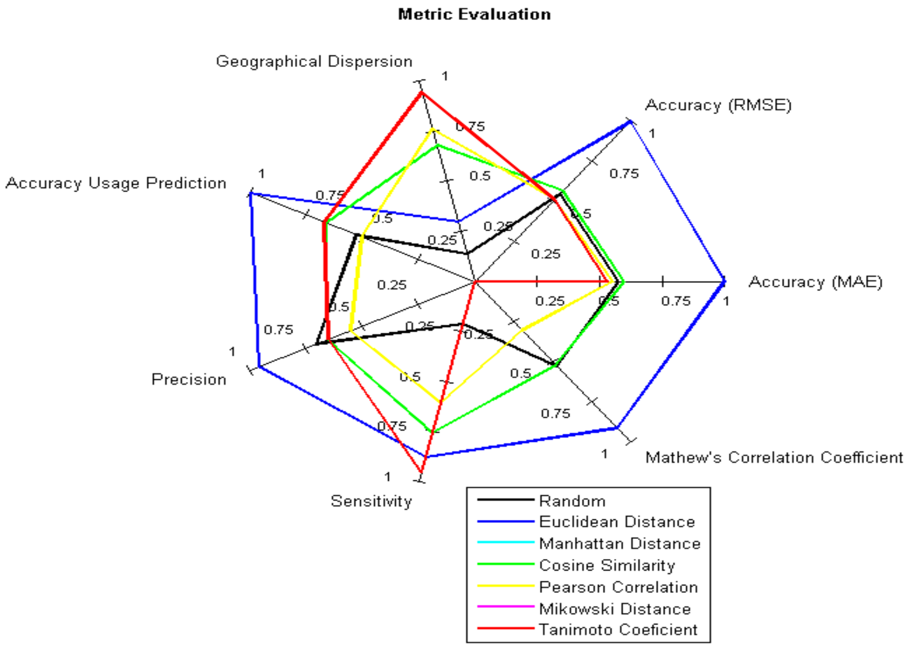

Figure 6.2-3 Euclidean Distance Behavior

The result shows the Euclidean, Manhattan and Minkowski distance like the similarity metrics with the best performance for the proposed approach. The Euclidean distance is used in the system to calculate the recommendation.

Now are shown three different cases study based on the same criterion of the experimental design.

As a case study of recommendations for collective profiles, we have a group with six (6) members. This group has the collective profile shown in Figure 6.2-4.

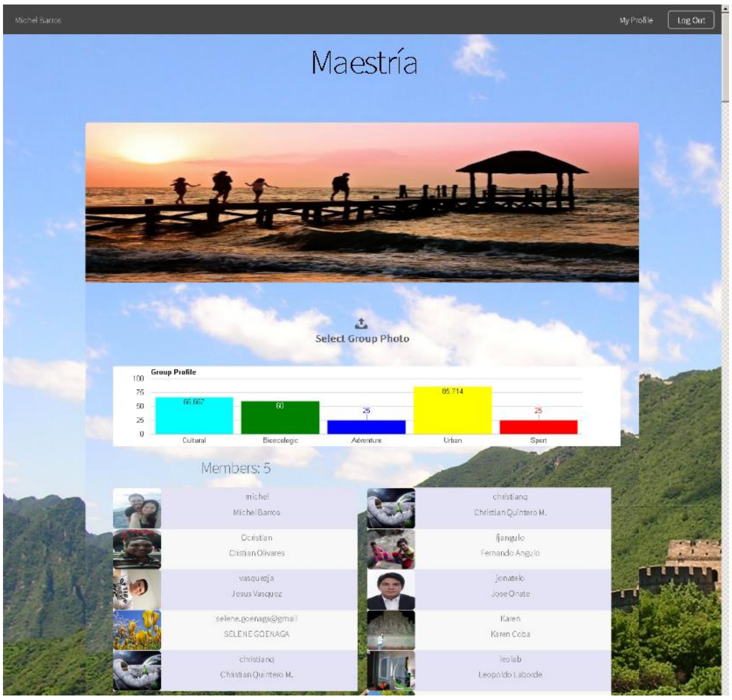

Figure 6.2-4 Example of a group with 6 members

This group selects New York as a destination city and twenty-five places to recommend, Figure 6.2-5.

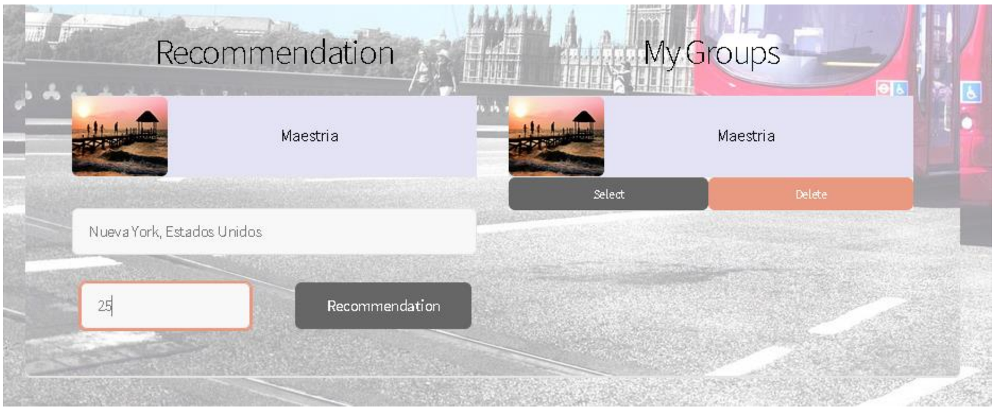

Figure 6.2-5 Selection of the Destination City and the Number of Places

The system calculates the recommendation for the collective profile, Figure 6.2-6.

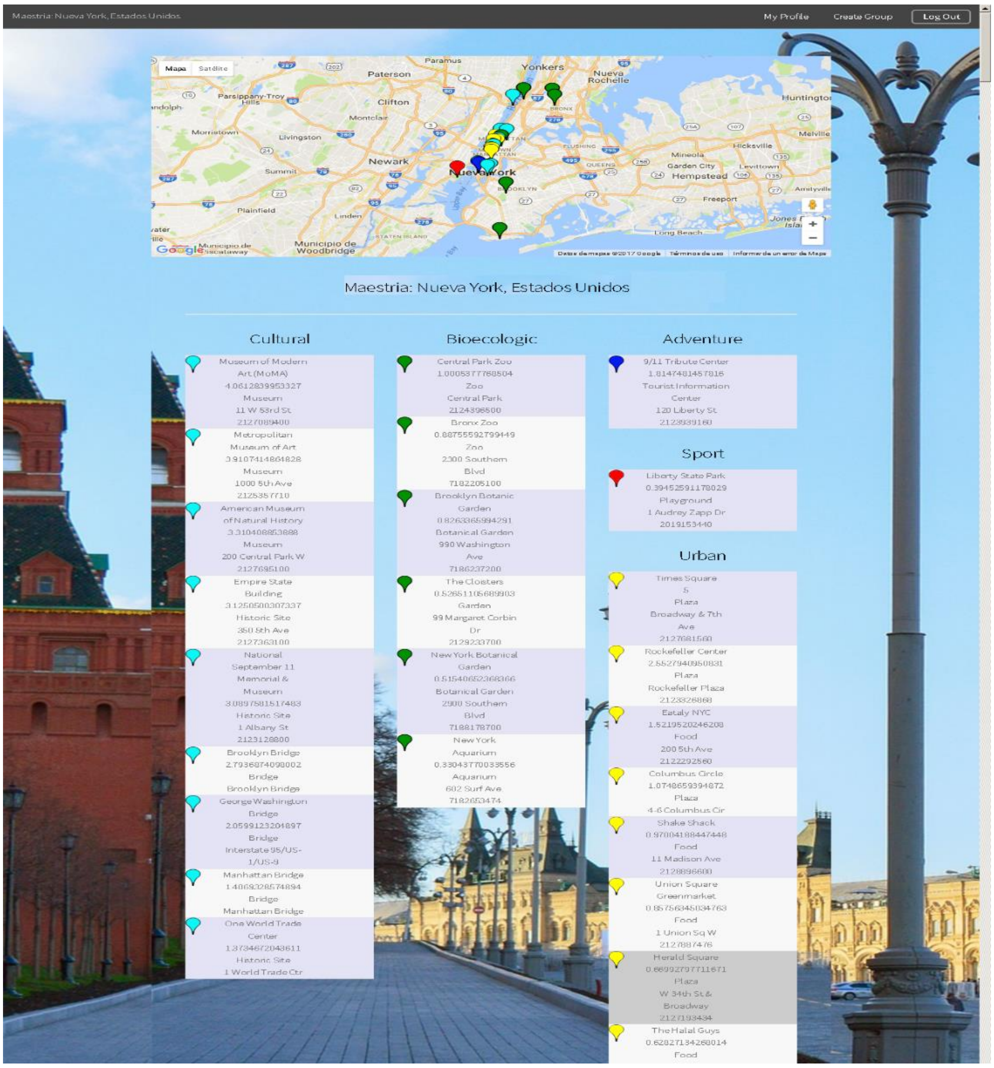

Figure 6.2-6 Recommendation for the Case Study

The next case study uses the same group, but it has another destination city, Figure 6.2-7.

Figure 6.2-7 Case Study with a Different Destination City

The recommendation of this case study is shown in Figure 6.2-8.

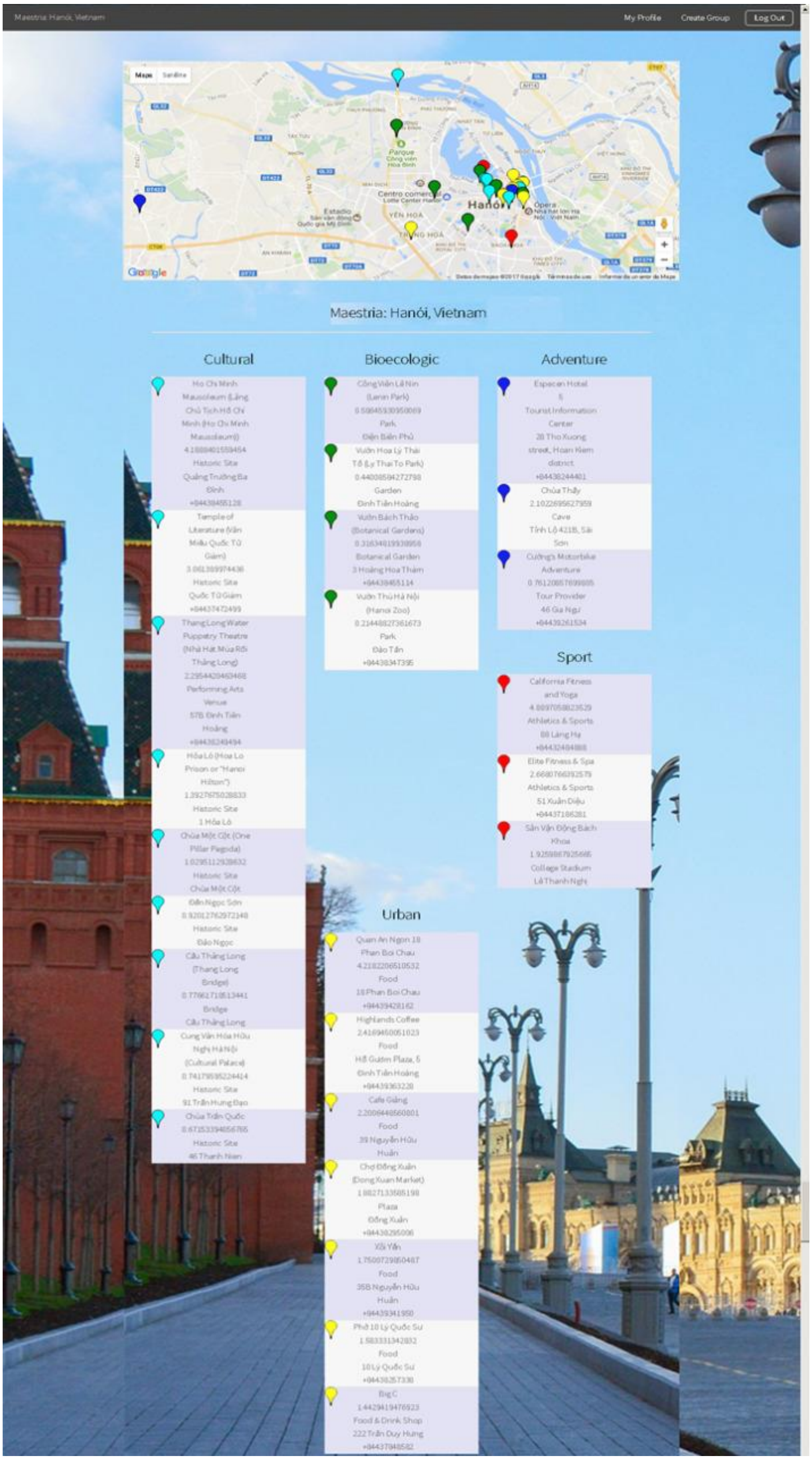

Figure 6.2-8 Recommendation for the Case Study with other Destination City

Next we have another example, but in this case, we use another group with a different collective profile, Figure 6.2-9.

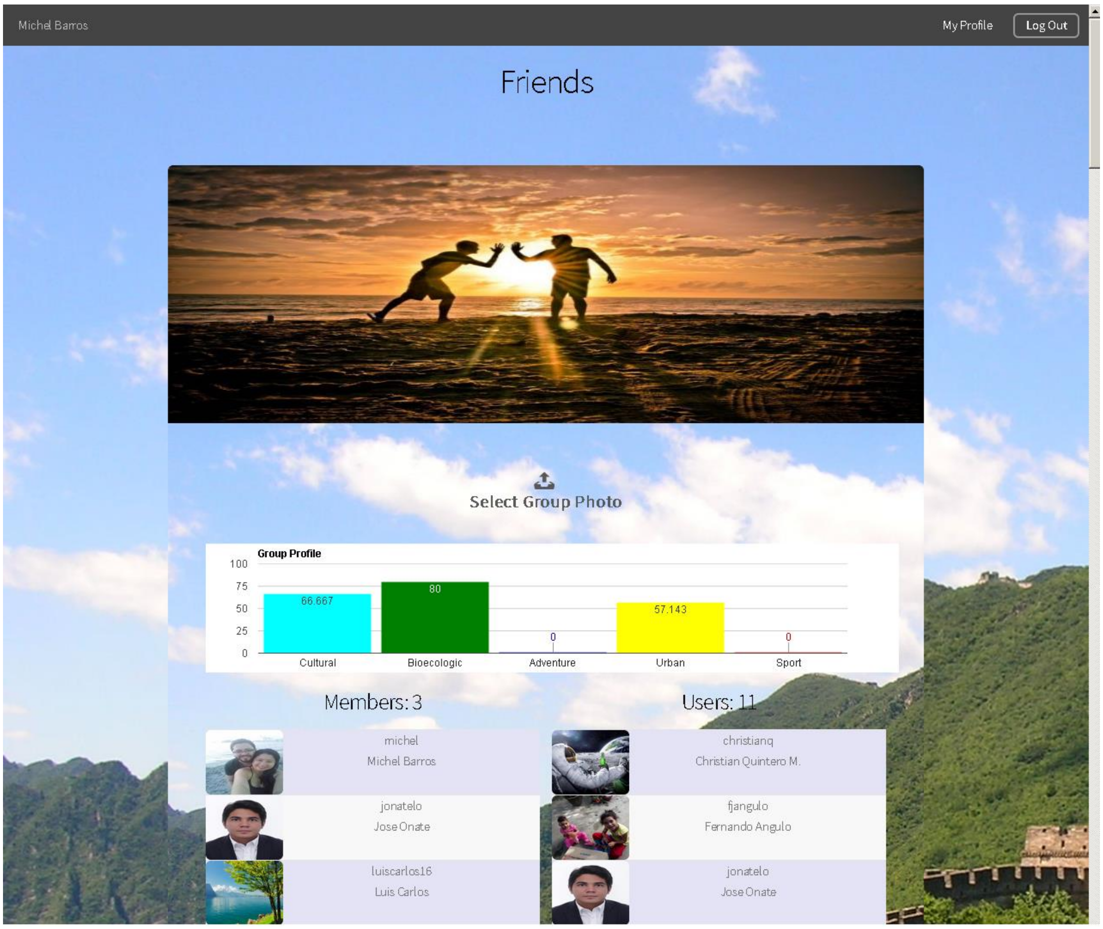

Figure 6.2-9 Case Study with a Different Collective Profile

This group selects other destination city and number of places, Figure 6.2-10.

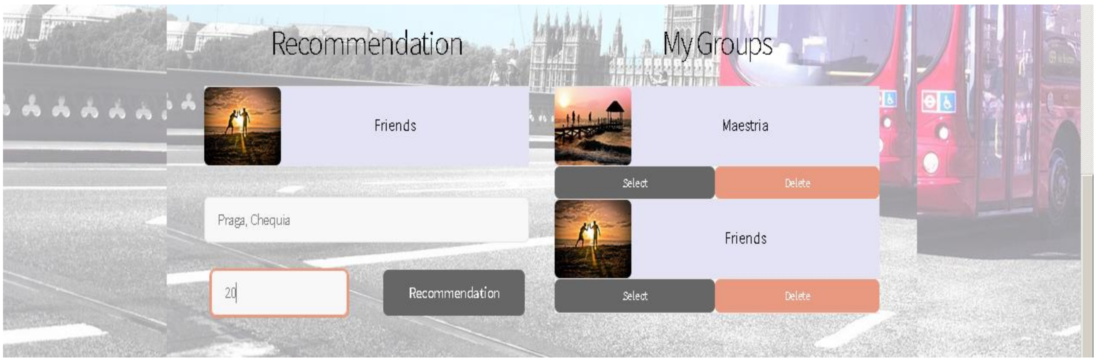

Figure 6.2-10 Number of Places and Destination City for the Case of Study

The figure 6.2-11 shows the calculated recommendation for this case study.

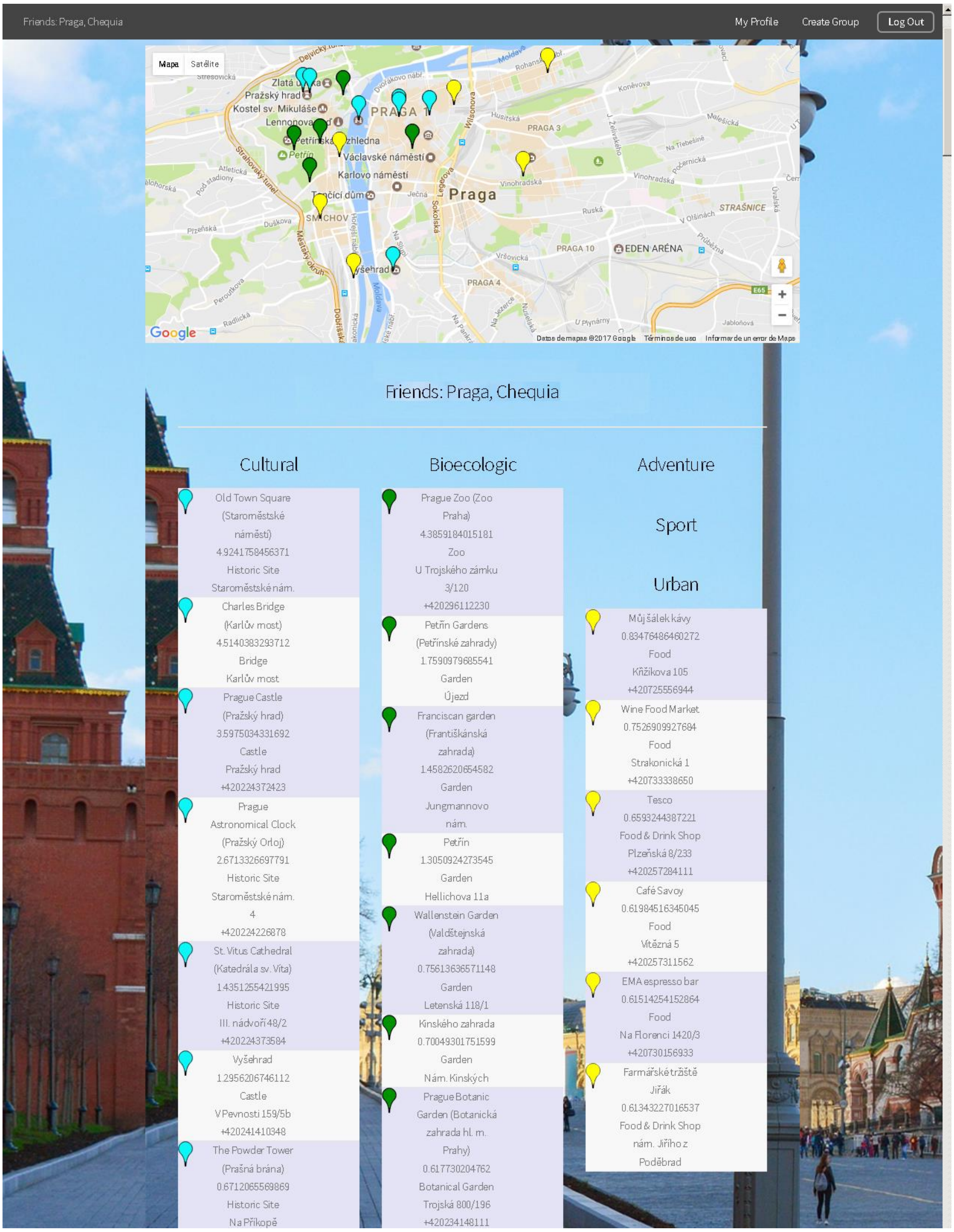

Figure 6.2-11 Recommendation for the Last Case Study

Chapter 7

Conclusions and Future Works

This chapter summarizes the main conclusions arisen of the analysis and discussion of the results reported in this work. The chapter also reviews the dissertation's scientific contributions and then discusses promising directions for future research and applications in certain topics in which the work of this thesis can continue. Finally, some concluding remarks are drawn.

7.1. Conclusions

The results and analysis carried out during this proposal has allowed highlighting the characteristics of each criterion keeping in mind the response variables discussed in section 6.1. However, the main goal in this proposal is calculating recommendations for collective profile and reaching the satisfaction of all group members; hence it is necessary to consider the performance of all proposed approach. According to the concepts exposed in Chapter 2, this proposal is suitable because it allows recommending tourist places to collectives profiles in order to satisfy the preferences of all group members.

The proposed group modelling strategy calculates a collective profile to represent the preferences of all group members. The group modelling strategy takes the preferences of all group members and these preferences are combined in order to calculate the collective profile. This approach contributes to satisfy the group preferences in order to minimize the bias of the consensus functions described in section 2.2.1.

The proposed hybrid recommendation system characterizes the tags related to tourist activities. This characterization allows calculate a profile for each tag and carrying out comparisons among them and the collective profile. The result of this comparison is the selection of the suitable tags for the collective profile. In this sense, the suitable activities are selected for the collective profile. After that, the places related to the suitable tags are selected according to their level of reputation.

This selection has the suitable places of the destination city related to the selected tags previously.

7.2 Main Contributions

This thesis presents a computational system oriented to recommend tourist places to collective profiles. Our approach presents a better performance in accuracy, precision, and sensitivity in comparison to the non-intelligent recommendation technique.

This thesis makes the following contributions in group recommendation system problems related to tourist places:

- *An intelligent approach for the places recommendation to collective profiles keeping in mind the preferences of all group members and the conventional problems of recommendation systems like the cold start, new user, new item, and ramp up among others.*
- *To improve the model of group preferences by means of the selected group modelling strategy and the proposed consensus function.*

7.3 Future Research and Directions

The development of an intelligent tourist recommender system focused on collective profiles in order to satisfy the group preferences represent an aspect interesting towards to the recommendation of items to groups. The developed system requires a methodology capable of characterizing and recommending tourist places in all over the world.

- Developing a method to recommend similar users in order to create groups.
- Incorporating new functions to characterize and recommending destination places to the collective profiles.
- Developing a fast method to download the places information.
- Implementing more sources of information.

References

Fernández, G., López, W., Olivera, F., Rienzi, B., & Rodríguez-Bocca, P. (2014). Let's go to the cinema! A movie recommender system for ephemeral groups of users. *2014 XL Latin American Computing Conference (CLEI)*. http://doi.org/10.1109/CLEI.2014.6965161

Alghamdi, H., Zhu, S., Saddik, A. El, & Systems, A. R. (2016). E-Tourism : Mobile Dynamic Trip Planner, 185–188. http://doi.org/10.1109/ISM.2016.76

Ashworth, G., & Page, S. J. (2011). Urban tourism research: Recent progress and current paradoxes. *Tourism Management*, *32*(1), 1–15. http://doi.org/10.1016/j.tourman.2010.02.002

Baraglia, R., Frattari, C., Muntean, C. I., Nardini, F. M., & Silvestri, F. (2012). RecTour: A Recommender System for Tourists. In *2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology* (pp. 92–96). IEEE. http://doi.org/10.1109/WI-IAT.2012.88

Bernier, C., Brun, A., Aghasaryan, A., Bouzid, M., Picault, J., Senot, C., & Boyer, A. (2010). Topology of communities for the collaborative recommendations to groups. *3rd International Conference on Information Systems and Economic Intelligence SIIE2010*, 1–6. Retrieved from http://hal.archives-ouvertes.fr/hal-00546932/

Bogdanov, D., Haro, M., Fuhrmann, F., Xambó, A., Gómez, E., & Herrera, P. (2011). A content-based system for music recommendation and visualization of user preferences working on semantic notions. *Proceedings - International Workshop on Content-Based Multimedia Indexing*, 249–252. http://doi.org/10.1109/CBMI.2011.5972554

Buckley, R. (2007). Adventure tourism products: Price, duration, size, skill, remoteness. *Tourism Management*, *28*(6), 1428–1433. http://doi.org/10.1016/j.tourman.2006.12.003

Buckley, R. (2012). Rush as a key motivation in skilled adventure tourism: Resolving the risk recreation paradox. *Tourism Management*, *33*(4), 961–970. http://doi.org/10.1016/j.tourman.2011.10.002

Buckley, R., McDonald, K., Duan, L., Sun, L., & Chen, L. X. (2014). Chinese model for mass adventure tourism. *Tourism Management*, *44*, 5–13. http://doi.org/10.1016/j.tourman.2014.01.021

Carlisle, S., Johansen, A., & Kunc, M. (2016). Strategic foresight for (coastal) urban tourism market complexity: The case of Bournemouth. *Tourism Management*, *54*, 81–95. http://doi.org/10.1016/j.tourman.2015.10.005

Carnicelli-Filho, S., Schwartz, G. M., & Tahara, A. K. (2010). Fear and adventure tourism in Brazil. *Tourism Management*, *31*(6), 953–956. http://doi.org/10.1016/j.tourman.2009.07.013

Cater, C. I. (2006). Playing with risk? Participant perceptions of risk and management implications in adventure tourism. *Tourism Management*, *27*(2), 317–325. http://doi.org/10.1016/j.tourman.2004.10.005

Chen, J. H. (2013). Hybrid recommendation system for tourism. *Proceedings - 2013 IEEE 10th International Conference on E-Business Engineering, ICEBE 2013*, 156–161. http://doi.org/10.1109/ICEBE.2013.24

Cibinskiene, A., & Snieskiene, G. (2015). Evaluation of City Tourism Competitiveness. *Procedia - Social and Behavioral Sciences*, *213*, 105–110. http://doi.org/10.1016/j.sbspro.2015.11.411

Cobbinah, P. B. (2015). Contextualising the meaning of ecotourism. *Tourism Management Perspectives*, *16*, 179–189. http://doi.org/10.1016/j.tmp.2015.07.015

Curice, J., Phillips, M., and C., & E. (2016). UNWTO Tourism Highlights 2016 Edition, 140. http://doi.org/10.18111/9789284418145

Dachary, A. C., & Arnaiz Burne, S. M. (n.d.). EL ESTUDIO DEL TURISMO. ¿Un paradigma en formación? *Estudios Y Perspectivas En Turismo*, *15*(2), 179–193.

Dai, Y., Ye, H., & Gong, S. (2009). Personalized Recommendation Algorithm Using User Demography Information. In *International Workshop on Knowledge Discovery and Data Mining* (pp. 100–103). IEEE. http://doi.org/10.1109/WKDD.2009.156

Das, M., & Chatterjee, B. (2015). Ecotourism: A panacea or a predicament? *Tourism Management Perspectives*, *14*, 3–16. http://doi.org/10.1016/j.tmp.2015.01.002

Dong, F., Luo, J., Zhu, X., Wang, Y., & Shen, J. (2013). A personalized hybrid recommendation system oriented to E-commerce mass data in the cloud. *Proceedings - 2013 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2013*, 1020–1025. http://doi.org/10.1109/SMC.2013.178

Dong, J., Li, X., & Fang, B. (2016). Community-Based Recommendations: A Solution to the Vulnerability Problem. In *2016 12th International Conference on Semantics, Knowledge and Grids (SKG)* (pp. 150–153). IEEE. http://doi.org/10.1109/SKG.2016.032

Duhme, L. (2013). *Cultural tourism: Case study Portugal*. (R. Raj, K. Griffin, & N.

Morpeth, Eds.)*Annals of Tourism Research* (Vol. 17). Wallingford: CABI. http://doi.org/10.1079/9781845939236.0000

Edwards, D., Griffin, T., & Hayllar, B. (2008). Urban Tourism Research. Developing an Agenda. *Annals of Tourism Research*, *35*(4), 1032–1052. http://doi.org/10.1016/j.annals.2008.09.002

Esfahani, M. H., & Alhan, F. K. (2013). New hybrid recommendation system based On C-Means clustering method. *The 5th Conference on Information and Knowledge Technology*, 145–149. http://doi.org/10.1109/IKT.2013.6620054

Fatemi, M., & Tokarchuk, L. (2013). A Community Based Social Recommender System for Individuals & Groups. In *2013 International Conference on Social Computing* (pp. 351–356). IEEE. http://doi.org/10.1109/SocialCom.2013.55

Fernandez, G., Lopez, W., Olivera, F., Rienzi, B., & Rodriguez-Bocca, P. (2014). Let's go to the cinema! A movie recommender system for ephemeral groups of users. In *2014 XL Latin American Computing Conference (CLEI)* (pp. 1–12). IEEE. http://doi.org/10.1109/CLEI.2014.6965161

Flanigan, S., Blackstock, K., & Hunter, C. (2014). Agritourism from the perspective of providers and visitors: A typology-based study. *Tourism Management*, *40*, 394–405. http://doi.org/10.1016/j.tourman.2013.07.004

Guerrero González, P. E., & Ramos Mendoza, J. R. (2014). *Introducción al turismo*. Larousse - Grupo Editorial Patria. Retrieved from http://site.ebrary.com/lib/unortesp/detail.action?docID=11013166&p00=Introducción+al+turismo

Hsu, C. H. C. (2006, July). Tourist Behaviour: Themes and Conceptual Schemes. *Annals of Tourism Research*. http://doi.org/10.1016/j.annals.2006.03.004

Hsu, C. H. C., & Huang, S. (2010). Formation of tourist behavioral intention and actual behavior. *2010 7th International Conference on Service Systems and Service Management, Proceedings of ICSSSM' 10*, 717–722. http://doi.org/10.1109/ICSSSM.2010.5530150

Ichimura, T., & Tachibana, I. (2014). Affective recommendation system for tourists by using emotion generating calculations. *2014 IEEE 7th International Workshop on Computational Intelligence and Applications, IWCIA 2014 - Proceedings*, 9–14. http://doi.org/10.1109/IWCIA.2014.6987727

Jiménez Bulla, L. H. (2008). *Ecoturismo oferta y desarrollo sistémico regional* (Vol. 6). Ecoe Ediciones. Retrieved from http://site.ebrary.com/lib/unortesp/detail.action?docID=10472953&p00=ecoturismo

Jiménez Bulla, L. H., & Jiménez Barbosa, W. G. (2013). Turismo tendencias globales y planificación estratégica. Retrieved from http://site.ebrary.com/id/10692891

Kavinkumar, V., Reddy, R. R., Balasubramanian, R., Sridhar, M., Sridharan, K., & Venkataraman, D. (2015). A hybrid approach for recommendation system with added feedback component. *2015 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2015*, 745–752. http://doi.org/10.1109/ICACCI.2015.7275700

Kembellec, G., Chartron, G., & Saleh, I. (2014). *Recommender Systems*. (Jean-Charles Pomerol, Ed.). London: Wiley-ISTE. Retrieved from http://site.ebrary.com/lib/unorte/detail.action?docID=10995110&p00=Recommender+Systems

Kennelly, M., & Toohey, K. (2014). Strategic alliances in sport tourism: National sport organisations and sport tour operators. *Sport Management Review*, *17*(4), 407–418. http://doi.org/10.1016/j.smr.2014.01.001

Kenteris, M., Gavalas, D., & Mpitziopoulos, A. (2010). A Mobile tourism recommender system. *Proceedings - IEEE Symposium on Computers and Communications*, 840–845. http://doi.org/10.1109/ISCC.2010.5546758

Kim, N., & Lee, J.-H. (2014). Group recommendation system: Focusing on home group user in TV domain. In *2014 Joint 7th International Conference on Soft Computing and Intelligent Systems (SCIS) and 15th International Symposium on Advanced Intelligent Systems (ISIS)* (pp. 985–988). IEEE. http://doi.org/10.1109/SCIS-ISIS.2014.7044866

Kumar, A., Gupta, S., Singh, S. K., & Shukla, K. K. (2015). Comparison of various metrics used in collaborative filtering for recommendation system. *2015 Eighth International Conference on Contemporary Computing (IC3)*, 150–154. http://doi.org/10.1109/IC3.2015.7346670

Li, X., Yuan, B., & Jin, P. (2012). Innovation of Tourism Information Services based on Internet Technology. *2012 9th International Conference on Fuzzy Systems and Knowledge Discovery*, (Fskd), 2178–2182. http://doi.org/10.1109/FSKD.2012.6233790

Lin, W. (2014). The optimization of weights in weighted hybrid recommendation algorithm, 14–17.

López-Guzmán, T., Vieira-Rodríguez, A., & Rodríguez-García, J. (2014). Profile and motivations of European tourists on the Sherry wine route of Spain. *Tourism Management Perspectives*, *11*, 63–68. http://doi.org/10.1016/j.tmp.2014.04.003

López Bonilla, J. M., & López Bonilla, L. M. (2012). *Investigación de mercados* *turísticos*. *Journal of Chemical Information and Modeling* (Vol. 53). Larousse - Ediciones Pirámide. Retrieved from http://site.ebrary.com/lib/unortesp/detail.action?docID=11059426&p00=Investigación+de+mercados+turísticos

Ma, M. Y., & Uchyigit, G. (2008a). *Personalization Techniques and Recommender Systems*. Hackensack, NJ: World Scientific Publishing Company. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=e000xww&AN=236130&lang=es&site=ehost-live

Ma, M. Y., & Uchyigit, G. (2008b). *Personalization Tecniques and Recommender Systems*. Hackensack, NJ: World Scientific Publishing Company. Retrieved from http://site.ebrary.com/lib/unorte/reader.action?docID=10255447

Mai, J., Fan, Y., & Shen, Y. (2009). A neural networks-based clustering collaborative filtering algorithm in E-commerce recommendation system. *2009 International Conference on Web Information Systems and Mining, WISM 2009*, 616–619. http://doi.org/10.1109/WISM.2009.129

Mark Piekarz, Ian Jenkins, P. M. (2016). Risk and safety management in the leisure, events, tourism and sports industries. *Tourism Management*, *54*, 296–297. http://doi.org/10.1016/j.tourman.2015.12.006

Marquéz, A. A., Ibarra, J. A., & León, O. G. (2010). *Introducción a la inferencia estadística*. (Carlos Mario Ramirez Torres, Alejandro Agustín Gómez Ruiz, & Juan José García Guzmán, Eds.) (PRIMERA ED). Naucalpan de Juárez, Estado de México: PEARSoN CUSToM PUblIShING.

Mehta, S. J., & Javia, J. (2015). Threshold based KNN for fast and more accurate recommendations. In *2015 IEEE 2nd International Conference on Recent Trends in Information Systems (ReTIS)* (pp. 109–113). IEEE. http://doi.org/10.1109/ReTIS.2015.7232862

Miyazawa, Y., Yamamoto, Y., & Kawabe, T. (2013). Context-Aware Recommendation System Using Content Based Image Retrieval with Dynamic Context Considered. *2013 International Conference on Signal-Image Technology & Internet-Based Systems*, 779–783. http://doi.org/10.1109/SITIS.2013.126

Mohamad Danial Md Sabri, Suratman, M. N., Mohd Salleh Daim, Abd Rahman Kassim, & Khamis, S. (2011). A preliminary study on the influences of ecotourism activities to the stand structure of protected forests of Taman Negara Pahang. In *2011 IEEE Symposium on Business, Engineering and Industrial Applications (ISBEIA)* (pp. 76–79). IEEE. http://doi.org/10.1109/ISBEIA.2011.6088890

Olcina Cantos, J. (2012). Turismo y cambio climático: una actividad vulnerable que debe adaptarse. *Revista Investigaciones Turísticas*, *4*(4), 1–34. http://doi.org/10.14198/INTURI2012.4.01

Pawar, K. R., Ghorpade, T., & Shedge, R. (2016). Constraint based recipe recommendation using forward checking algorithm. In *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 1474–1478). IEEE. http://doi.org/10.1109/ICACCI.2016.7732256

Pujahari, A., & Padmanabhan, V. (2015). Group Recommender Systems: Combining User-User and Item-Item Collaborative Filtering Techniques. In *2015 International Conference on Information Technology (ICIT)* (pp. 148–152). IEEE. http://doi.org/10.1109/ICIT.2015.36

Qi, S., Mamoulis, N., Pitoura, E., & Tsaparas, P. (2016). Recommending Packages to Groups. In *2016 IEEE 16th International Conference on Data Mining (ICDM)* (pp. 449–458). IEEE. http://doi.org/10.1109/ICDM.2016.0056

Quispe, L. C., & Luna, J. E. O. (2016). A content-based recommendation system using TrueSkill. *Proceedings - 14th Mexican International Conference on Artificial Intelligence: Advances in Artificial Intelligence, MICAI 2015*, 203–207. http://doi.org/10.1109/MICAI.2015.37

Rabbiosi, C. (2015). Renewing a historical legacy: Tourism, leisure shopping and urban branding in Paris. *Cities*, *42*(PB), 195–203. http://doi.org/10.1016/j.cities.2014.07.001

Ricci, F., Rokach, L., Shapira, B., Kantor, P. B., & Ricci, F. (2011). *Recommender Systems Handbook*. (F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor, Eds.). Boston, MA: Springer US. http://doi.org/10.1007/978-0-387-85820-3

Richards, G., & Munsters, W. (2010). *Cultural Tourism Research Methods*. *Cultural Tourism Research Methods*. http://doi.org/10.1016/j.annals.2011.07.014

Ritchie, B. W., Hall, C. M., & Cooper, C. (2004). *Sport Tourism : Interrelationships, Impacts and Issues*. Channel View Publications. Retrieved from http://site.ebrary.com/lib/unorte/detail.action?docID=10096131&p00=sport+tourism

Secall, R. E., Bernier, E. T., García, R. F., & Rojo, M. M. M. (2006). *Estructura de mercados turísticos*. Editorial UOC, S.L. Retrieved from https://books.google.com.co/books?id=WSkyO-x2EZAC

Shambour, Q., & Lu, J. (2010). A framework of hybrid recommendation system for government-to-business personalized e-services. *ITNG2010 - 7th International Conference on Information Technology: New Generations*, 592–597. http://doi.org/10.1109/ITNG.2010.114

Sharpley, R. (2006). *Travel and Tourism*. *Journal of Chemical Information and Modeling*. SAGE Publications Inc. (US). Retrieved from http://site.ebrary.com/lib/unorte/detail.action?docID=10218254&p00=Travel+and+Tourism

Shishmanova, M. V. (2015). Cultural Tourism in Cultural Corridors, Itineraries, Areas and Cores Networked. *Procedia - Social and Behavioral Sciences*, *188*, 246–254. http://doi.org/10.1016/j.sbspro.2015.03.382

Song, H., Lu, P., & Zhao, K. (2011). RETRACTED ARTICLE: Improving item-based collaborative filtering recommendation system with tag. *2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce, AIMSEC 2011 - Proceedings*, (70871029), 2142–2145. http://doi.org/10.1109/AIMSEC.2011.6011041

Tatli, I., & Birtürk, A. (2011). A tag-based hybrid music recommendation system using semantic relations and multi-domain information. *Proceedings - IEEE International Conference on Data Mining, ICDM*, (2), 548–554. http://doi.org/10.1109/ICDMW.2011.17

Thiengburanathum, P., Cang, S., & Yu, H. (2015). A decision tree based recommendation system for tourists. *2015 21st International Conference on Automation and Computing: Automation, Computing and Manufacturing for New Economic Growth, ICAC 2015*, (September), 11–12. http://doi.org/10.1109/IConAC.2015.7313958

Torres Bernier, E., Secall, R. E., & Fuentes García, R. (2006). *Estructura de mercados turísticos*. *Fundamentos de márketing*. Retrieved from http://site.ebrary.com/lib/unortesp/detail.action?docID=11126718&p00=Estructura+de+mercados+turísticos

Tsai, C. Y., & Chung, S. H. (2012). A personalized route recommendation service for theme parks using RFID information and tourist behavior. *Decision Support Systems*, *52*(2), 514–527. http://doi.org/10.1016/j.dss.2011.10.013

Valeri, B., Baez, M., & Casati, F. (2013). Come along: Understanding and motivating participation to social leisure activities. *Proceedings - 2013 IEEE 3rd International Conference on Cloud and Green Computing, CGC 2013 and 2013 IEEE 3rd International Conference on Social Computing and Its Applications, SCA 2013*, 211–218. http://doi.org/10.1109/CGC.2013.41

Vekariya, V., & Kulkarni, G. R. (2012). Hybrid recommender systems: Content-boosted collaborative filtering for improved recommendations. *Proceedings - International Conference on Communication Systems and Network Technologies, CSNT 2012*, *36*(4), 649–653. http://doi.org/10.1109/CSNT.2012.218

Wang, Q., Yuan, X., & Sun, M. (2010). Collaborative filtering recommendation algorithm based on hybrid user model. *Fuzzy Systems and Knowledge Discovery (FSKD), 2010 Seventh International Conference on*, *4*(Fskd), 1985–1990. http://doi.org/10.1109/FSKD.2010.5569479

Weaver, D. B. (2001). *The Encyclopedia of Ecotourism*. Retrieved from http://site.ebrary.com/lib/unorte/detail.action?docID=10066738&p00=the+encyclopedia+ecotourism

William Navidi. (2006). *Estadística para ingenieros y científicos*. (S. A. D. C. V. McGRAW-HILL/INTERAMERICANA EDITORES, Ed.) (Primera ed). Ciudad de México.

Xiang, Z., Magnini, V. P., & Fesenmaier, D. R. (2015). Information technology and consumer behavior in travel and tourism: Insights from travel planning using the internet. *Journal of Retailing and Consumer Services*, *22*, 244–249. http://doi.org/10.1016/j.jretconser.2014.08.005

Yoon, K., Lee, J., & Kim, M. U. (2012). Music recommendation system using emotion triggering low-level features. *IEEE Transactions on Consumer Electronics*, *58*(2), 612–618. http://doi.org/10.1109/TCE.2012.6227467

Zhang, L. (2012). Techniques for dynamic and diversified relaxation in constraint-based recommender systems. *Proceedings of the 2012 2nd International Conference on Business Computing and Global Informatization, BCGIN 2012*, 565–568. http://doi.org/10.1109/BCGIN.2012.153

Annex A

Relationship between Tags and Subprofiles

Each tag of the next list represents a type of places or activities available in the destination, these tags are taken of the Foursquare API and these are selected based on the activities described in the characterization profiles (see Section 4.2). These tags are characterized according to the proposed tag characterization (see Section 4.5.1). In the next tables the initials S.I. and C.A. mean defined by the intelligent system and calculated, respectively.

A.1. Tags related to the Adventure Subprofile

Tag Name	Foursquare API ID	Adventure (%)	Bioecologic (%)	Cultural (%)	Sport (%)	Urban (%)
Boat Rental	5744ccdfe4b0c0459246b4c1	100	S.I.	C.A.	0	S.I.
Boat or Ferry	4bf58dd8d48988d12d951735	100	S.I.	C.A.	0	S.I.
Cave	56aa371be4b08b9a8d573511	100	S.I.	C.A.	0	S.I.
Cruise	55077a22498e5e9248869ba2	100	S.I.	C.A.	0	S.I.
Dive Shop	52f2ab2ebcbc57f1066b8b1a	100	S.I.	C.A.	0	S.I.
Dive Spot	52e81612bcbc57f1066b7a12	100	S.I.	C.A.	0	S.I.
Fishing Store	52f2ab2ebcbc57f1066b8b16	100	S.I.	C.A.	0	S.I.
Go Kart Track	52e81612bcbc57f1066b79ea	100	S.I.	C.A.	0	S.I.
Indoor Play Area	5744ccdfe4b0c0459246b4b5	100	S.I.	C.A.	0	S.I.
Lighthouse	4bf58dd8d48988d15d941735	100	S.I.	C.A.	0	S.I.
RV Park	52f2ab2ebcbc57f1066b8b53	100	S.I.	C.A.	0	S.I.
Rafting	52e81612bcbc57f1066b7a29	100	S.I.	C.A.	0	S.I.
Reservoir	56aa371be4b08b9a8d573541	100	S.I.	C.A.	0	S.I.
Rock Climbing Spot	50328a4b91d4c4b30a586d6b	100	S.I.	C.A.	0	S.I.
Tour Provider	56aa371be4b08b9a8d573520	100	S.I.	C.A.	0	S.I.
Tourist Information Center	4f4530164b9074f6e4fb00ff	100	S.I.	C.A.	0	S.I.
Trailer Park	52f2ab2ebcbc57f1066b8b55	100	S.I.	C.A.	0	S.I.
Volcano	5032848691d4c4b30a586d61	100	S.I.	C.A.	0	S.I.
Waterfall	56aa371be4b08b9a8d573560	100	S.I.	C.A.	0	S.I.
Waterfront	56aa371be4b08b9a8d5734c3	100	S.I.	C.A.	0	S.I.

Table A.1-1 Adventure Tags

A.2. Tags related to the Bioecologic Subprofile

Name	Foursquare API ID	Adventure (%)	Bioecologic (%)	Cultural (%)	Sport (%)	Urban (%)
Aquarium	4fceea171983d5d06c3e9823	10	100	C.A.	0	S.I.
Botanical Garden	52e81612bcbc57f1066b7a22	10	100	C.A.	0	S.I.
Campground	4bf58dd8d48988d1e4941735	10	100	C.A.	0	S.I.
Farm	4bf58dd8d48988d15b941735	30	100	C.A.	0	S.I.
Field	4bf58dd8d48988d15f941735	10	100	C.A.	0	S.I.
Forest	52e81612bcbc57f1066b7a23	10	100	C.A.	0	S.I.
Garden	4bf58dd8d48988d15a941735	10	100	C.A.	0	S.I.
Island	50aaa4314b90af0d42d5de10	30	100	C.A.	0	S.I.
Lake	4bf58dd8d48988d161941735	30	100	C.A.	0	S.I.
Mountain	4eb1d4d54b900d56c88a45fc	10	100	C.A.	0	S.I.
Mountain Hut	55a5a1ebe4b013909087cb77	10	100	C.A.	0	S.I.
National Park	52e81612bcbc57f1066b7a21	30	100	C.A.	0	S.I.
Nature Preserve	52e81612bcbc57f1066b7a13	30	100	C.A.	0	S.I.
Other Great Outdoors	4bf58dd8d48988d162941735	30	100	C.A.	0	S.I.
Park	4bf58dd8d48988d163941735	10	100	C.A.	0	S.I.
River	4eb1d4dd4b900d56c88a45fd	30	100	C.A.	0	S.I.
Stables	4eb1baf03b7b2c5b1d4306ca	30	100	C.A.	0	S.I.
Summer Camp	52e81612bcbc57f1066b7a10	30	100	C.A.	0	S.I.
Trail	4bf58dd8d48988d159941735	30	100	C.A.	0	S.I.
Tree	52e81612bcbc57f1066b7a24	10	100	C.A.	0	S.I.
Zoo	4bf58dd8d48988d17b941735	10	100	C.A.	0	S.I.

Table A.1-1 Bioecologic Tags

A.3. Tags related to the Cultural Subprofile

Name	Foursquare API ID	Adventure (%)	Bioecologic (%)	Cultural (%)	Sport (%)	Urban (%)
Amphitheater	56aa371be4b08b9a8d5734db	10	S.I.	100	0	S.I.
Antique Shop	4bf58dd8d48988d116951735	10	S.I.	100	0	S.I.
Art Gallery	4bf58dd8d48988d1e2931735	10	S.I.	100	0	S.I.
Arts & Crafts Store	4bf58dd8d48988d127951735	10	S.I.	100	0	S.I.
Auditorium	4bf58dd8d48988d173941735	10	S.I.	100	0	S.I.
Ballroom	56aa371be4b08b9a8d5734cf	10	S.I.	100	0	S.I.
Bookstore	4bf58dd8d48988d114951735	10	S.I.	100	0	S.I.
Bridge	4bf58dd8d48988d1df941735	10	S.I.	100	0	S.I.
Canal	56aa371be4b08b9a8d573562	10	S.I.	100	0	S.I.
Canal Lock	56aa371be4b08b9a8d57353b	10	S.I.	100	0	S.I.
Castle	50aaa49e4b90af0d42d5de11	10	S.I.	100	0	S.I.
Cemetery	4bf58dd8d48988d15c941735	10	S.I.	100	0	S.I.
Christmas Market	52f2ab2ebcbc57f1066b8b3b	10	S.I.	100	0	S.I.
College Theater	4bf58dd8d48988d1ac941735	10	S.I.	100	0	S.I.
Concert Hall	5032792091d4c4b30a586d5c	10	S.I.	100	0	S.I.
Country Dance Club	52e81612bcbc57f1066b79ef	10	S.I.	100	0	S.I.
Cultural Center	52e81612bcbc57f1066b7a32	10	S.I.	100	0	S.I.
Exhibit	56aa371be4b08b9a8d573532	10	S.I.	100	0	S.I.
Festival	5267e4d9e4b0ec79466e48c7	10	S.I.	100	0	S.I.
Historic Site	4deefb944765f83613cdba6e	10	S.I.	100	0	S.I.
Library	4bf58dd8d48988d12f941735	10	S.I.	100	0	S.I.
Memorial Site	5642206c498e4bfca532186c	10	S.I.	100	0	S.I.
Movie Theater	4bf58dd8d48988d17f941735	10	S.I.	100	0	S.I.
Museum	4bf58dd8d48988d181941735	10	S.I.	100	0	S.I.
Music Festival	5267e4d9e4b0ec79466e48d1	10	S.I.	100	0	S.I.
Music Venue	4bf58dd8d48988d1e5931735	10	S.I.	100	0	S.I.
Palace	52e81612bcbc57f1066b7a14	10	S.I.	100	0	S.I.
Pedestrian Plaza	52e81612bcbc57f1066b7a25	10	S.I.	100	0	S.I.
Performing Arts Venue	4bf58dd8d48988d1f2931735	10	S.I.	100	0	S.I.
Public Art	507c8c4091d498d9fc8c67a9	10	S.I.	100	0	S.I.
Salsa Club	52e81612bcbc57f1066b79ec	10	S.I.	100	0	S.I.
Samba School	56aa371be4b08b9a8d5734f9	10	S.I.	100	0	S.I.
Scenic Lookout	4bf58dd8d48988d165941735	10	S.I.	100	0	S.I.
Sculpture Garden	4bf58dd8d48988d166941735	10	S.I.	100	0	S.I.
Spiritual Center	4bf58dd8d48988d131941735	10	S.I.	100	0	S.I.
Street Fair	5267e4d8e4b0ec79466e48c5	10	S.I.	100	0	S.I.
Theme Park	4bf58dd8d48988d182941735	10	S.I.	100	0	S.I.
Used Bookstore	52f2ab2ebcbc57f1066b8b30	10	S.I.	100	0	S.I.
Vineyard	4bf58dd8d48988d1de941735	10	S.I.	100	0	S.I.

Table A.2-1 Cultural Tags

A.4. Tags related to the Sport Subprofile

Name	Foursquare API ID	Adventure (%)	Bioecologic (%)	Cultural (%)	Sport (%)	Urban (%)
Athletics & Sports	4f4528bc4b90abdf24c9de85	10	S.I.	C.A.	100	S.I.
Bike Trail	56aa371be4b08b9a8d57355e	10	S.I.	C.A.	100	S.I.
Bowling Alley	4bf58dd8d48988d1e4931735	10	S.I.	C.A.	100	S.I.
College Stadium	4bf58dd8d48988d1b4941735	10	S.I.	C.A.	100	S.I.
Disc Golf	52e81612bcbc57f1066b79e8	10	S.I.	C.A.	100	S.I.
Fishing Spot	52e81612bcbc57f1066b7a0f	30	S.I.	C.A.	100	S.I.
Gun Range	52e81612bcbc57f1066b7a11	30	S.I.	C.A.	100	S.I.
Laser Tag	52e81612bcbc57f1066b79e6	10	S.I.	C.A.	100	S.I.
Mini Golf	52e81612bcbc57f1066b79eb	10	S.I.	C.A.	100	S.I.
Playground	4bf58dd8d48988d1e7941735	10	S.I.	C.A.	100	S.I.
Pool Hall	4bf58dd8d48988d1e3931735	10	S.I.	C.A.	100	S.I.
Racecourse	56aa371be4b08b9a8d573514	10	S.I.	C.A.	100	S.I.
Racetrack	4bf58dd8d48988d1f4931735	10	S.I.	C.A.	100	S.I.
Roller Rink	52e81612bcbc57f1066b79e9	10	S.I.	C.A.	100	S.I.
Ski Area	4bf58dd8d48988d1e9941735	30	S.I.	C.A.	100	S.I.
Ski Shop	56aa371be4b08b9a8d573566	10	S.I.	C.A.	100	S.I.
Sporting Goods Shop	4bf58dd8d48988d1f2941735	10	S.I.	C.A.	100	S.I.
Stadium	4bf58dd8d48988d184941735	10	S.I.	C.A.	100	S.I.
Water Park	4bf58dd8d48988d193941735	30	S.I.	C.A.	100	S.I.

Table A.3-1 Sport Tags

A.5. Tags related to the Urban Subprofile

Name	Foursquare API ID	Adventure (%)	Bioecologic (%)	Cultural (%)	Sport (%)	Urban (%)
Adult Boutique	5267e446e4b0ec79466e48c4	10	S.I.	C.A.	0	100
Bar	4bf58dd8d48988d116941735	10	S.I.	C.A.	0	100
Bay	56aa371be4b08b9a8d573544	10	S.I.	C.A.	0	100
Beach	4bf58dd8d48988d1e2941735	30	S.I.	C.A.	0	100
Betting Shop	52f2ab2ebcbc57f1066b8b40	10	S.I.	C.A.	0	100
Big Box Store	52f2ab2ebcbc57f1066b8b42	10	S.I.	C.A.	0	100
Bike Rental / Bike Share	4e4c9077bd41f78e849722f9	10	S.I.	C.A.	0	100
Brewery	50327c8591d4c4b30a586d5d	10	S.I.	C.A.	0	100
Business Center	56aa371be4b08b9a8d573517	10	S.I.	C.A.	0	100
Candy Store	4bf58dd8d48988d117951735	10	S.I.	C.A.	0	100
Casino	4bf58dd8d48988d17c941735	10	S.I.	C.A.	0	100
Chocolate Shop	52f2ab2ebcbc57f1066b8b31	10	S.I.	C.A.	0	100
Circus	52e81612bcbc57f1066b79e7	10	S.I.	C.A.	0	100
Clothing Store	4bf58dd8d48988d103951735	10	S.I.	C.A.	0	100
Club House	52e81612bcbc57f1066b7a35	10	S.I.	C.A.	0	100
Comedy Club	4bf58dd8d48988d18e941735	10	S.I.	C.A.	0	100
Comic Shop	52f2ab2ebcbc57f1066b8b18	10	S.I.	C.A.	0	100
Conference	5267e4d9e4b0ec79466e48c6	10	S.I.	C.A.	0	100
Convenience Store	4d954b0ea243a5684a65b473	10	S.I.	C.A.	0	100
Convention	5267e4d9e4b0ec79466e48c9	10	S.I.	C.A.	0	100
Convention Center	4bf58dd8d48988d1ff931735	10	S.I.	C.A.	0	100
Cosmetics Shop	4bf58dd8d48988d10c951735	10	S.I.	C.A.	0	100
Department Store	4bf58dd8d48988d1f6941735	10	S.I.	C.A.	0	100
Discount Store	52dea92d3cf9994f4e043dbb	10	S.I.	C.A.	0	100
Dog Run	4bf58dd8d48988d1e5941735	10	S.I.	C.A.	0	100
Entertainment Service	56aa371be4b08b9a8d573554	10	S.I.	C.A.	0	100
Food	4d4b7105d754a06374d81259	10	S.I.	C.A.	0	100
Food & Drink Shop	4bf58dd8d48988d1f9941735	10	S.I.	C.A.	0	100
Fountain	56aa371be4b08b9a8d573547	10	S.I.	C.A.	0	100
Fruit & Vegetable Store	52f2ab2ebcbc57f1066b8b1c	10	S.I.	C.A.	0	100
General Entertainment	4bf58dd8d48988d1f1931735	10	S.I.	C.A.	0	100

Name	Foursquare API ID	Adventure (%)	Bioecologic (%)	Cultural (%)	Sport (%)	Urban (%)
Harbor / Marina	4bf58dd8d48988d1e0941735	10	S.I.	C.A.	0	100
Hobby Shop	4bf58dd8d48988d1fb941735	10	S.I.	C.A.	0	100
Jewelry Store	4bf58dd8d48988d111951735	10	S.I.	C.A.	0	100
Karaoke Box	5744ccdfe4b0c0459246b4bb	10	S.I.	C.A.	0	100
Massage Studio	52f2ab2ebcbc57f1066b8b3c	10	S.I.	C.A.	0	100
Night Market	53e510b7498ebcb1801b55d4	10	S.I.	C.A.	0	100
Nightclub	4bf58dd8d48988d11f941735	10	S.I.	C.A.	0	100
Other Nightlife	4bf58dd8d48988d11a941735	10	S.I.	C.A.	0	100
Outdoor Supply Store	52f2ab2ebcbc57f1066b8b22	10	S.I.	C.A.	0	100
Outlet Mall	5744ccdfe4b0c0459246b4df	10	S.I.	C.A.	0	100
Outlet Store	52f2ab2ebcbc57f1066b8b35	10	S.I.	C.A.	0	100
Pachinko Parlor	5744ccdfe4b0c0459246b4b8	10	S.I.	C.A.	0	100
Parade	52741d85e4b0d5d1e3c6a6d9	10	S.I.	C.A.	0	100
Perfume Shop	52f2ab2ebcbc57f1066b8b23	10	S.I.	C.A.	0	100
Plaza	4bf58dd8d48988d164941735	10	S.I.	C.A.	0	100
Pool	4bf58dd8d48988d15e941735	10	S.I.	C.A.	0	100
Record Shop	4bf58dd8d48988d10d951735	10	S.I.	C.A.	0	100
Recreation Center	52e81612bcbc57f1066b7a26	10	S.I.	C.A.	0	100
Shopping Mall	4bf58dd8d48988d1fd941735	10	S.I.	C.A.	0	100
Shopping Plaza	5744ccdfe4b0c0459246b4dc	10	S.I.	C.A.	0	100
Smoothie Shop	52f2ab2ebcbc57f1066b8b41	10	S.I.	C.A.	0	100
Social Club	52e81612bcbc57f1066b7a33	10	S.I.	C.A.	0	100
Souvenir Shop	52f2ab2ebcbc57f1066b8b1b	10	S.I.	C.A.	0	100
Spa	4bf58dd8d48988d1ed941735	10	S.I.	C.A.	0	100
Speakeasy	4bf58dd8d48988d1d4941735	30	S.I.	C.A.	0	100
Stoop Sale	52f2ab2ebcbc57f1066b8b54	10	S.I.	C.A.	0	100
Thrift / Vintage Store	4bf58dd8d48988d101951735	10	S.I.	C.A.	0	100
Toy / Game Store	4bf58dd8d48988d1f3941735	10	S.I.	C.A.	0	100
Video Game Store	4bf58dd8d48988d10b951735	10	S.I.	C.A.	0	100
Video Store	4bf58dd8d48988d126951735	10	S.I.	C.A.	0	100
Watch Shop	52f2ab2ebcbc57f1066b8b2e	10	S.I.	C.A.	0	100

Table A.4-1 Urban Tags