An Intelligent Group Decision Support System for Urban Tourists

Development and evaluation of a group recommendation mechanism

Master Thesis

By K.J.A. Touw Ngie Tjouw

July 29, 2010
Abstract

A new Group Decision Support System (GDSS) has been developed that simplifies and improves group
decision-making for travel consumers. Prior studies have shown that various irrational factors influence
a group decision-making process. Moreover there is a lack of structured information that supports the
decision process. Presently, different studies have proposed group recommendations systems that
improve the decision-making process. However, they appear ineffective in leading a group towards a
satisfying outcome. The objective of this research is to develop and evaluate an intelligent group
recommendation mechanism that is integrated in a well-structured decision process. As a result, a
prototype of an online intelligent group decision support system for travel consumers has been built,
which we refer to as Trip.Easy. Subsequently, we set up a user experiment and invited over a 120
participants, divided into 30 groups. Each group was instructed to organize a city trip while using
Trip.Easy. After each session, the participants were asked to fill out a questionnaire. The results show
that users were satisfied with the outcome of the session and with the recommendations of Trip.Easy
during each session. Furthermore, analyses show that even groups with a relatively high measure of
conflicting preferences, are guided towards a satisfying outcome. Finally the results show that the
satisfaction with the recommendations significantly increases as the session continues. We conclude
that Trip.Easy is able to lead a group towards a satisfying outcome.
Preface
More than a year ago, me and my business partner Koh S. Ngai, came up with the idea to develop an online platform, where users can plan, search and book vacations together. Inspired by the research on artificial negotiation systems by Professor Catholijn Jonker and Koen Hindriks, we worked out a concept. Supervised by Koen Hindriks, we started on a long journey of designing, developing and evaluating a first prototype. The size and the complexity of the project required it to be divided into two complementary thesis projects.

Next to the scientific value, this project also has a huge commercial potential. Together with Koen Hindriks we therefore have applied for the ‘Valorization Grant’. This fund supports the commercialization of scientific concepts. The application included an application-document and a well-prepared presentation. This required Koen, Koh and me to put in a lot of time and effort and to be working as a team.

Therefore, I’d like to thank my supervisor Koen Hindriks for his enthusiasm and guidance during the thesis project and for the effort put in the funding application. Catholijn Jonker, I’d like to thank for the inspiring research and her advice during the funding application phase. Furthermore, I’d like to thank Maarten van Dijk, founder of ‘vakantiepunten.nl’, for supporting our experiment and for believing in the commercial potential of this concept.

Many thanks to my girlfriend Mirthe Boersbroek for her support on statistical analysis. Thanks also to my parents Kenneth sr. and Alice Touw Ngie Tjouw. Thanks to Ingrid Schierbeek for her support and advice. Also thanks to Shari Shield for arranging the meetings with travel marketing specialists. Thanks to the more than 120 participants of the experiment.

Finally, I’d like to thank Koh S. Ngai, for being a great friend and business partner. Together we have initiated this project and we will turn it into a commercial success. Therefore we traveled to the United States of America to orientate on possible investors and market trends. It is here in a hotel room on the 19th floor with a great view, where the last sentences of this thesis were written.

Kenneth Touw Ngie Tjouw
San Francisco, USA
13th of June 2010
Table of Contents

Abstract .......................................................................................................................... 3
Preface ........................................................................................................................... 4
Table of Contents ......................................................................................................... 6
Table of Figures ........................................................................................................... 9
Table of tables ............................................................................................................. 9
1 Introduction ............................................................................................................. 10
  1.1 Research objectives .......................................................................................... 11
  1.2 Trip.Easy Prototype ......................................................................................... 12
  1.3 Thesis organization ......................................................................................... 13
2 Background theory .................................................................................................. 14
  2.1 Group decision process ................................................................................... 14
    2.1.1 Group of travel consumers ................................................................. 15
  2.2 Related work .................................................................................................... 16
    2.2.1 MusicFX ............................................................................................... 17
    2.2.2 Let’s Browse ........................................................................................ 18
    2.2.3 PolyLens ............................................................................................... 18
    2.2.4 Travel Decision Forum ......................................................................... 19
    2.2.5 Limitations ............................................................................................ 20
    2.2.6 Conclusion ............................................................................................ 21
3 Models and Algorithms .......................................................................................... 22
  3.1 Introduction ..................................................................................................... 22
  3.2 Domain Model .................................................................................................. 23
  3.3 Preference Elicitation and Aggregation ............................................................ 24
    3.3.1 Classic voting ......................................................................................... 25
    3.3.2 Intensity of preferences ........................................................................ 25
    3.3.3 Cp-nets .................................................................................................. 27
    3.3.4 Kemeny rule and Spearman’s Footrule ............................................... 29
    3.3.5 Conclusion ............................................................................................ 31
  3.4 Preference model .............................................................................................. 32
3.5 Preference aggregation algorithm ................................................................. 34
3.6 Utility function .............................................................................................. 35
3.7 Well-structured decision process .................................................................. 36

4. Development and Design .............................................................................. 38
4.1 Functional design .......................................................................................... 38
4.1.1 Users ........................................................................................................ 38
4.2 Technical design ............................................................................................ 43
4.2.1 Web-application ......................................................................................... 43
4.2.2 Trip.Easy Recommendation Mechanism .................................................. 44
4.2.3 Database .................................................................................................... 47
4.2.4 Development approach ............................................................................. 49
4.2.5 Conclusion ................................................................................................ 49

5. Evaluation ......................................................................................................... 50
5.1 Introduction ..................................................................................................... 50
5.2 Research ......................................................................................................... 51
5.2.1 Evaluation of the outcome of the session ................................................... 51
5.2.2 The evaluation of the outcome of the session ............................................. 51
5.2.3 Similarity of preferences versus the satisfaction of the session outcome .... 51
5.2.4 Evaluation of the recommendations ............................................................ 52
5.3 Model .............................................................................................................. 57
5.4 Experimental setup ....................................................................................... 59
5.4.1 Protocol ...................................................................................................... 59
5.4.2 Participants ................................................................................................ 60
5.4.3 Equipment and setting .............................................................................. 60
5.4.4 Pilot test ..................................................................................................... 61
5.4.5 Reliability & Validity .................................................................................. 62
5.5 Results ............................................................................................................. 63
5.5.1 Satisfaction of the session outcome ............................................................ 63
5.5.2 Dissimilarity of preferences versus the satisfaction of the session outcome ........................................................................................................ 63
5.5.3 Evaluation of the recommendations ............................................................ 66

6. Discussion & Conclusion .................................................................................. 73
6.1 The evaluation of the outcome of the session ............................................... 73
Table of Figures

Figure 1: General consumer behavior model ................................................................. 15
Figure 2: Representation of a city .................................................................................. 23
Figure 3: CP net expressed in a graph G and the conditional preference tables ............. 28
Figure 4: The preference graph (left) and the majority graph (right) for the profile of the example ................................................................. 29
Figure 5: preference model .......................................................................................... 32
Figure 6: An abstract overview of the aggregation algorithm ....................................... 35
Figure 7: Overview of the decision-making process combined with the recommendation mechanism ................................................................. 37
Figure 8: user interface generate step .......................................................................... 39
Figure 9: user interface reduce step ............................................................................ 40
Figure 10: user interface clarify step .......................................................................... 41
Figure 11: user interface evaluate step ....................................................................... 42
Figure 12: web application in general ........................................................................ 43
Figure 13: Architecture of the recommendation mechanism ........................................ 44
Figure 14: Interaction diagram of recommendation mechanism and decision process .... 47
Figure 15: Global view of the Trip.Easy Database ......................................................... 48
Figure 16: Reinig’s goal attainment model .................................................................. 58
Figure 17: Schematic outline of the various steps of the experiment ............................. 59
Figure 18: test room arrangement .............................................................................. 60
Figure 19: Pilot test ..................................................................................................... 61
Figure 20: distances between group 4 members ............................................................ 64
Figure 21: overall score of Satisfaction of proposed cities (recommendations) ............... 70
Figure 22: the overall score of the satisfaction of the amount of the proposed cities ......... 71

Table of tables

Table 1: overview of aggregation methods vs. our criteria ............................................ 31
Table 2 : ANOVA table for process step vs. SATC ....................................................... 72
1 Introduction

“The last time you organized a holiday with a couple of friends, it is likely that you or some of your friends were unhappy about some parts of the holiday. Such a reaction is possible despite the fact that you have tried to take their preferences into account by gathering information and discussing it with them. After days of searching through all kinds of travel information about several destinations, you try to select some good alternatives. Subsequently you discuss these alternatives with your friends while using basic communication resources such as e-mail or telephone. Now that your friends are getting a bit more involved, they set new requirements and preferences, making it harder for you to come up with good alternatives. As a result, the group appears to be in conflict. In addition to that, some of your friends appear to be very dominant by imposing their opinions and ideas. Finally, you all decide to choose something which is probably a suboptimal alternative for the whole group.”

Group decision processes of tourists for destination selection, as described in the example above, are known to be suboptimal. In (Mansfeld, 1992), (Nichols, 1988) and (R.Gitelson, 1995) it is stated that members are often not very satisfied with certain parts of the vacation that were chosen by the group. This can be explained by the fact that the process of travel group decision making is typically ineffective. Even though not much research has been done into the decision making of groups of tourists, observations show that a lot of irrational influences exist within a group decision process. An example is the presence of a dominant member that pushes his or her preferences. Another aspect is the large amount of information online, which may lead to a phenomenon called information overload. As the amount of (unstructured) information increases beyond a certain threshold, consumers’ subjective states towards decisions might adversely become worse (Chen, 2007). All of the information is generated by travel agents or by independent websites, which function as so-called decision support systems (DSS’s). The current functioning online DSS’s for tourists are often limited to just sharing general information combined with user generated content. Above all they do not have any support for group decision processes.

In the literature we found several studies in different domains that claimed to have developed an intelligent group decision support system. However, these systems appeared to be solely group recommendation systems. The challenge remained of how to be able to support a group throughout the decision process and lead them towards a satisfying outcome. This research project shows that it can be achieved by combining an intelligent group recommendation mechanism with a well-structured decision process. We have developed and evaluated a prototype of an intelligent group decision support system, that we call Trip.Easy. This prototype facilitates the decision support for a group of travel consumers who are deciding about the destination of their next city trip.

This has been a large and complex project, thus we divided it into different sub-projects. One sub-project contains the design, development and evaluation of the structured decision process facilitated by Trip.Easy, which is covered in (Ngai, 2010). This thesis covers the design, development and evaluation of the recommendation mechanism of Trip.Easy.
1.1 Research objectives

The first goal of this research is to design and implement a group recommendation mechanism for a group decision support system for groups of travel consumers. We will deal with several technical design challenges, such as the design and implementation of a domain model for city trips, a preference model for travel consumers, a preference aggregation procedure, a utility function and a recommendation function. Finally everything needs to be integrated in a well-structured decision process; from thereon we will establish a prototype of an intelligent group decision support system for travel consumers.

The second goal is to show that the Trip.Easy prototype is able to support groups to reach a satisfying outcome. Therefore we evaluate the Trip.Easy prototype by performing a user evaluation. This will be an indication of whether the user accepts the technology and if he or she will use it in the future. The evaluation is based on empirically validated methods. The ultimate goal of the evaluation is to answer the main question formulated below:

**Do group members believe that the outcome established using Trip.Easy GDSS represents a satisfying outcome for the group?**

To answer this question we observe two main aspects. The first aspect is user satisfaction of the outcome of Trip.Easy. This indicates whether the users are satisfied with the outcome established by the group. This aspect is formulated by the following question:

*Are users satisfied with the outcome?*

In addition we are interested in the effect of dissimilar preferences of group preferences on the satisfaction with the outcome. This is formulated by the question:

*Is there a qualitative difference in outcome between groups with members that have similar preference and groups with members that have less similar preferences?*

The second aspect is the user satisfaction of the recommendations. This indicates how well the recommendations support the group in reaching a satisfying outcome and is formulated by the following research question:

**Do group members believe that the recommendations of Trip.Easy support the group adequately to arrive at a satisfying outcome?**

To answer this question we observe the following specific sub-aspects:

First we observe the user satisfaction of the recommendations. The recommendations influence the outcome of a session a lot, simply by the fact that it selects the alternatives from which the group members must choose and decide. Therefore we want to monitor the user satisfaction level of the recommendations throughout the decision process. This aspect is formulated by the following sub-question:
Are users satisfied with the proposed cities during each process step?

Secondly, we are interested in the course of the satisfaction level of the recommendations throughout the decision-process. This aspect indicates whether or not the user satisfaction of the recommendations is converging to a satisfying outcome. This aspect is reflected by the question:

Does the user satisfaction with the proposed cities increase during the whole process?

Finally, we observe the satisfaction of the amount of proposed cities. Too many recommendations can lead to information overload, a poor amount of recommendations however, can be perceived as a scarcity of options. Therefore we want to answer the question:

Are users satisfied with the amount of proposed cities in each process step?

1.2 Trip.Easy Prototype

Trip.Easy is an online group decision support system that supports a group with organizing a city trip. The goal of the system is to facilitate a group decision process and intelligently generate suggestions. Eventually, the group is lead towards an outcome that can be considered a best-fit for the group. Functionally, the prototype provides group members the ability to:

- Invite friends and establish a group.
- Insert his or her preferences
- View different cities (recommended by Trip.Easy) that fit these preferences
- Share feedback on these cities
- Make a final decision

The ability of generating recommendations for a group requires the system to have a lot of knowledge of the city domain, such as: city attractions, activities, events and other city characteristics. We have collected data from more than 150 cities in our database. This data has been organized according to the domain model that we designed. This way the system can perform calculations and active reasoning on the data. To be operational Trip.Easy is dependent on the following components:

- An interactive User Interface that guides the users through the whole process
- Domain model
- Recommendation mechanism
- Well-structured decision process

These components are embedded in a web-based architecture that runs on a web server.
1.3 Thesis organization

In the next chapter the background theory is discussed. This includes the rationale of this research and general related work. Chapter 3 discusses the models and algorithms we apply, followed by the actual development of the Trip.Easy prototype in chapter 4. In chapter 5 we elaborate on the results of the evaluation. Thereafter we discuss these results and come to a conclusion in chapter 6. The limitations and problems are covered in chapter 7, followed by the Future work in chapter 8.
2 Background theory
This chapter discusses the background theory of the decision-making process of travel consumers. We will elaborate on the problem of group decision-making. Furthermore we will discuss several group recommendation systems proposed in prior studies.

2.1 Group decision process
Group decision making is the process of deciding on which action or item the group should choose, while taking into consideration the various preferences of different group members (Flyvbjerg, 2006). In (Mansfeld, 1992) it is shown that when a group of travel consumers must decide on a trip, they go through a group decision process. It has been widely held among researchers in the consumer behavior field, that the consumer’s decision to purchase is a multi-staged process, based on the so called information-processing theory outlined by (Siakaya, 2003) and (Solomon, 2006). This theory is central to the general consumer behavior model described by (Solomon, 2006).

The general consumer model (Figure 1) identified that consumers transit five different stages (see Figure 1). At first they start with the ‘need recognition’ stage, in which the consumer recognizes a problem or a need. Secondly, in the ‘information search’ stage, the consumer is indulged to search for more information. This state of indulgence may encourage the consumer to actively search for information or heighten their attention to relevant information sources including advertising.

During the ‘evaluation of alternatives’ stage, different solutions of the problem (need) are being evaluated. The evaluation process is complex, because it varies per situation. However, in general terms the consumer will examine the attributes of the product, assign different levels of importance to such attributes, determine the likely level of overall satisfaction with each alternative and derive an attitude toward the different solutions/brands. The next stage is called, the ‘choice of product or service’. The consumer purchases the product or service that it considers to be the best solution for its need (problem). The final stage considers the ‘post purchase evaluation’, where consumers rank brands and form purchase intentions for the future.
In (A. Pizar, 1999) it is claimed that the ‘tourist purchase decision’ is also based on the stages as presented in the consumer behavior model. According to (A. Pizar, 1999) all decision making goes through the same sequential process. The travel decision model therefore can be seen as an instantiation of the consumer behavior model (Figure 1).

2.1.1 Group of travel consumers

There are no models that describe the travel decision process of other social groups like friends and colleagues. Nevertheless, the level of communication, the type of decision making (consensus, bargaining, voting, dictatorship), and the result of confrontation (agreement versus conflict) are important determinants of these group decision processes (Mansfeld, 1992). The group decision-making for a trip was studied in (R. Gitelson, D. Kerstetter, 1995). The following influential categories of a participant were distinguished:

- Sole decision maker, i.e., 100% responsible for the decision
- Dominant influence, i.e., the greatest percentage allocation of any decision maker
- Shared influence, i.e., equal influence with at least one other decision maker, but more than the other decision makers
- Lesser role, i.e., a percentage of influence less than some other decision maker, or
- No influence.

Regarding the influential categories of a participant, we notice that in all categories the influences are not equal; one or some participants of a group have greater influence on the decision making process than others. These influences cause suboptimal outcomes from travel group decision making, because the influences are affecting all stages of the consumer behavior model (Figure 1). For instance the choice of product is made without taking all the participants’ preferences into account, but only those of some. In (Siakaya, 2003) it is analyzed that the travel consumer behavior model actually is a problem solving process. Thus travel group decision-making can be regarded as a joint problem solving process. Joint
problem solving can be optimized by facilitation of a well-structured collaborative process (G.L. Kolfschoten, 2007). For more background information please consult the literature study (Touw, 2009). In our approach for developing a group decision support system for travel consumers, we take the following essential issues into account:

1. Facilitating a fair and workable decision process that involves all members
2. Intelligently providing recommendations to the group during the process while taking all preferences into account.

Within this thesis, we focus on the second element. The first is discussed in (Ngai, 2010). In the next paragraph we discuss some related work on group decision support systems.

### 2.2 Related work

Systems that intelligently provide recommendations to a group are called group recommendation systems. In general, the group recommendation process has four phases (Jameson, 2004).

1. **Preference elicitation**: we also refer to this step as the preference elicitation step. During this step, the system tries to extract information about the users’ preferences.
2. **Preference aggregation**: after the preferences are gathered, the system aggregates the individual preferences into a representative group preference model. Subsequently, it generates recommendations based on the group preference model.
3. **Presentation of recommendations**: the recommendations are presented on a display. This can be done in combination with information about the preferences, to create awareness among the users.
4. **Final decision**: finally the members should choose their most preferred alternative.

In this paragraph we discuss four of the, according to (Jameson, 2004), best-known group recommendation systems that have been presented in the literature. Thereafter we will discuss some limitations. We conclude with a description of essential issues that we need to take into account for developing and evaluating the Trip.Easy prototype.
2.2.1 MusicFX

The MusicFX system (McCarthy J. F., 1998) is a music channel selector for music to be played in a fitness center. This is done on the basis of preferences that have been previously specified by the members of the fitness center who are working out at that moment. The system chooses one of the 91 possible music channels. The choice procedure also includes some randomness to ensure variety. The MusicFx system is one of the earliest group recommendation systems which were developed.

2.2.1.1 Procedure

Preference specification: MusicFX users will have to insert their preferences by insertion of a rating for each music genre that exists in the MusicFX database. The rating ranges from -2 (= I hate this music) to +2 (= I love this music) on a 5-point Likert-scale. Once the member has inserted the complete rating form, the MusicFX has the complete profile of a member. Finally the MusicFX system selects music based on a combination of all profiles.

Preference aggregation: MusicFX aggregates the preferences by averaging the ratings of all the users. The best scoring music genre will be selected. However the selection procedure contains some randomness to secure a certain level of variety in the selected music.

Presentation of recommendations: The MusicFX directly presents the recommendations by playing the music. Users can directly evaluate the result and therefore determine whether the recommendations are suitable or not.

Final decision: As explained above the final decision is made by MusicFX. If users disagree with the music that has been selected, they can update their preferences. Apart from that, there are no other involvements of the users during the process of selecting a music channel.

2.2.1.2 Evaluation

A poll was conducted to find out how well the members liked the music being played in the fitness center. Members were asked whether they thought the music being played was better, worse or the same, compared to before the system was installed. A total of 71 responses were received, which represents more than 25% of the active members of the fitness center. The majority (71%) said they like the music selected by MusicFX better than the music that was played before the installation of MusicFX. Only 7% thought the music is worse. Another 8% thought the music was about the same. However 14% was unable to answer this question, since they joined the fitness club after MusicFX was installed. More than 30% of the respondents said that they wanted to have more influence on the selection of music.
2.2.2  Let’s Browse
Let’s Browse (Lieberman, 1999) recommends web pages to a group of two or more users who are browsing the web together.

2.2.2.1  Procedure
Preference specification: The preferences of the users are determined by analyzing the identity of a user. Therefore Let’s Browse consults a database where all the information (such as interests, hobby’s, etc) exists. Based on that information the Let’s Browse system determines a profile of that user and thus his or her preferences and interests.

Preference Aggregation: The preference aggregation is done by a simple linear combination of the profiles of each user. Consequently a combined profile was generated which is used to search through a set of websites and calculate a score. The score is based on the amount of content matching with the combined profile. The websites with the highest score is then recommended to the group.

Presentation of recommendations: A set of links is simply presented on the screen. In addition it also explains why the particular websites have been chosen.

Final decision: By clicking on one of the links presented by Let’s Browse the members decide which recommendation is accepted.

2.2.2.2  Evaluation
No controlled experiments were conducted (Lieberman, 1999).
This way of displaying the recommendations provide maximum information while minimizing the load on users. Recommendations are sorted in order of decreasing group prediction.

**Final decision:** By selecting one of the recommendations presented by PolyLens the members decide which recommendation is accepted. However, the final decision is made by one member.

### 2.2.3.2 Evaluation
The user satisfaction was measured by a survey that focused on how well the users liked the group recommendation features. The survey question was: “I found group recommendations more helpful than individual recommendations when deciding on a moving to see”. The results show that 22% of the users ‘strongly agreed’ with this. Furthermore 55% ‘agreed’, and 23% disagreed. Another observation was that 80% of the users find that the final decision should be made by the group instead of just one group member.

### 2.2.4 Travel Decision Forum
The Travel Decision Forum was build to help a group of users to agree on desired attributes of a vacation that they are planning to take together (Anthony Jameson, Stephan Baldes, and Thomas Kleinbauer, 2004). The system proposes a preference model. The preference model exists of attributes such as hotel, transportation and facilities (sauna, sun studio, etc.) and users tune this model by changing certain values (ratings). Based on those adjustments, the system again proposes a new preference model. No concrete recommendation is generated. The system supports users that are not at the same location and therefore not able to engage in face-to-face discussions. (The evaluation of the last part is the main objective of the research.)

#### 2.2.4.1 Procedure
**Preference specification:** Each group member specifies his or her preferences regarding the joint vacation (at least partially) by filling in a preference specification form. Novel aspects of this form include the possibilities of (a) optionally viewing, copying, and post-editing the preferences that have already been entered by other group members and (b) adding verbal arguments (which can be viewed by other members) to explain particular specifications. The preferences are specified by an importance level of each aspect of a vacation. The user must specify the importance level by inserting a rating value based on a 4-point scale that represents the importance level (which can also be copied from other members as point (a) explains).

**Preference aggregation:** Within the Travel Decision Forum different aggregation methods have been implemented. The first method is an averaging function. The average group ratings are calculated and then shown to the users. The second function is based on the median. This means that the group member with the median value (rating) with respect to the other group members is seen as the group preference value. So based on that the group preferences are determined and shown to the users. Another method is just making a random choice. One of the objectives within this project was to create a non-manipulable aggregation method. This means that users cannot manipulate the recommendation mechanism (for instance by inserting low ratings on purpose). An averaging method for instance is manipulable, while the median and random choice mechanisms are non-manipulable (Jameson, 2004).
**Presentation of the recommendations:** The Travel Decision Forum does not give any advice on what vacation packages to choose. Its objective is to have the users reach consensus about the joint preference model that the Travel Group Forum proposes. So the group preference model itself is important.

**Final decision:** The final decision is determined by accepting or rejecting the proposal of the group preference model.

2.2.4.2 **Evaluation**

No experiments were conducted.

2.2.5 **Limitations**

In the previous sections we have described four group recommendation systems. These systems only offer a substantial solution for the problem of group decision making. The MusicFX for instance does nothing more than averaging the users’ preferences. It is not focused on helping users reach a consensus based outcome that is acceptable for the group. The system decides what music is played and it does so autonomously without requesting any content of the user. Observing the Let’s Browse system however, we notice that more attention has been given to consensus building, by motivating how the recommendations are established. Still, the performance of supporting the group towards a consensus is very poor. For instance, it is assumed that the final decision making is the responsibility of one person instead of the whole group. The same can be said about the PolyLens system, which appears to presuppose that only one group leader will make the decision. However, the evaluation revealed that this was a non-desirable feature of their system; more than 80% of their users wanted to be able to be involved in the final selection of the group recommendation. The Travel Decision Forum however elicits formal feedback about the comprehensibility and acceptability of the proposals generated. Based on that feedback, new proposals are generated. Users can also adjust their preferences in order to arrive at a more agreeable outcome. Finally the Travel Decision Forum does not solve the problem of a group of travel consumers, because the group preference model is not translated to a concrete solution (i.e. a real vacation).

In general we notice that the approach of the design of the systems is very much derived from individual recommendation systems. MusicFX, PolyLense and Let’s Browse are all derived from individual recommendation systems. That is why they fail in supporting a group sufficiently in establishing a satisfying outcome. The Travel Decision Forum however pays more attention to that aspect. It offers more possibilities for involvement of users with the decision-making process. Thus users have more influence on the recommendations. However it seems that users can endlessly adapt their preferences and share feedback, on which the system responds with a new recommendation. A step in this process where users can make their final decision is nonexistent. The Travel Decision Forum only focuses on having the users arrive at an agreeable group preference model instead of a concrete solution (i.e. a vacation to Rome for 8 days).
2.2.6 Conclusion

In this section we make some concluding remarks about the group recommendation systems described in the previous sections. We will discuss these remarks by relating them to the four phases we introduced at the beginning of this paragraph.

Non-manipulability and fairness are important aspects of preference elicitation and preference aggregation (respectively phase 1 and phase 2). The Travel Decision Forum it appeared that it is very easy to manipulate an averaging aggregation method for instance. This means that one user can control the outcome in a way that is favorable for him, but very unfavorable for the rest of the group and thus the recommendations of the system become suboptimal with respect to the group.

During presentation of recommendations (phase 3), it appears to be appreciated when users have insight in each other’s preferences. This creates awareness of each others’ interest, which has a positive effect on the decision making.

Finally the fourth aspect, (phase 4) making the decision of which recommendation to accept, is still an issue for all group recommendation systems described above. In the previous section we showed that none of the proposed systems have an answer to this. Therefore we decided to take a different approach for designing and developing Trip.Easy in the sense that Trip.Easy should be able to facilitate the decision making process.

As discussed in previous section, the travel consumer behavior actually is a problem-solving process (Siakaya, 2003). Thus a group of travel consumers needs to solve that problem together. The decision making process is in fact a collaborative process. It is this aspect that also needs to be taken into account when designing a group recommendation system. A collaborative process can be designed using Collaborative Engineering (G.L. Kolfschoten, 2007), (Briggs, 2006) and (Bostrom, 2002). For a full elaboration of how this theory is applied to the design and development of Trip.Easy please check with (Ngai, 2010).

Within this thesis we will show that a group decision support system (or group recommendation system): (i) is able to generate useful recommendations for the group by using well designed preference elicitation and aggregation methods, and (ii) facilitates a well-structured collaborative process; is able to have group members arrive together at an outcome that is satisfying for the group.

In the next chapter we will discuss solutions for the design of the preference elicitation and aggregation methods. We will also focus on more technical issues like algorithms, the domain model and the utility function. The design of the collaborative process is further discussed in (Ngai, 2010). However in chapter 4 we will discuss the integration of both the collaborative process and the intelligent recommendation engine.
3 Models and Algorithms

3.1 Introduction
Within this chapter we will describe the theoretical design of the Trip.Easy Group Decision Support System with respect to the group recommendation components. First we elaborate on the different theoretical preference elicitation and preference aggregation methods in order to choose an approach that we can apply for the design of our aggregation models and aggregation algorithms. Another important part is the domain model, which identifies and describes the structure of a city. Based the domain model, (i) users are able to specify their preferences to Trip.Easy, which from thereon is able to (ii) aggregate those preferences and finally (iii) able to return recommendations that support the group adequately to arrive at a satisfying outcome. In the next section the domain model and its design is treated. In paragraph 3.3 we elaborate on the preference elicitation and aggregation methods followed by the preference model in paragraph 3.4. Subsequently, paragraph 3.5 discusses the design of the aggregation algorithm. In paragraph 3.6 we discuss the utility function. Finally, in paragraph 3.7 briefly describes the decision-making process.
3.2 Domain Model

As explained during the introduction of this thesis the aim of Trip.Easy is to facilitate the decision making process of a group of travel consumers. Trip.Easy will be operational in the so-called city trips domain. Its aim is to support travel consumers to decide and agree on which city they will visit. To be able to extract and aggregate the preferences and subsequently generate recommendations, it is essential to have a consistent general representation (model) of a city. To be able to generate a recommendation, the system will need to be able to recognize the preferences and match it with a certain city. Therefore the preferences should be based on certain features of a city, by which a city can be described. Examples of city features are; museums, ancient monuments, restaurants, parks, etc. Users often find certain features more important than others, thus prefer one feature over another. Trip.Easy uses such information to select cities that match these preferences. In Figure 2 the complete domain model is shown. The model is based on the study described in (Ruetsche, 2006).

Figure 2: Representation of a city

In (Ruetsche, 2006) it is explained that a city can be described in terms of certain elements (shown in Figure 2). These elements (features) are categorized by categories represented by the rectangles in Figure 2. Users will express their preferences according to this model. The system has a database with cities that are described according to this model. Subsequently, a weight (between 0 and 1) is assigned...
to each feature, expressing the presence of a certain element (i.e. New York is a very popular
destination for shopping, thus it gets a high weight value). The weights are based on information from
content providers (such as TripAdvisory.com). The weights were assigned manually via the
administration system of Trip.Easy (see next chapter). Consequently the system is able to interpret the
preferences of the user by matching it with the cities in the database. The interpretation is based on the
utility function described in 3.6. In the next paragraph we will discuss some preference aggregation
methods from which we derive the method that we will use.

3.3 Preference Elicitation and Aggregation

Preference elicitation is viewed as a process of extracting information about user preferences to the
extent that it is necessary to make good or even optimal decisions on behalf of that user (Braziunus,
2006). Within group decision support systems, preference elicitation is often succeeded by preference
aggregation. Preference aggregation is the process of combining the preferences of several individuals in
order to get a representation of the preferences of the group of these individuals (Thomas L. Saaty, Jen
S.Shang, 2005).

Every aggregation approach demands an aggregation function, which is a function that outputs a
collective preference relation or utility function (Thomas L. Saaty, Jen S.Shang, 2005) based on the
individual preferences. In the travel domain we find that we are dealing with cardinal as well as ordinal
preferences. After discussing the different approaches, we will select an aggregation method on which
we base our aggregation method design. In order to make a selection we have defined certain criteria:

- **Non-manipulability**: This means that users are not able to manipulate the recommendation
  mechanism. It prevents certain members from having more influence on the outcome than others.
- **Usability**: Preference aggregation methods are effective if they have the right preference
  information about users. However such information can get very complex and cumbersome.
  This leads to a preference elicitation process that is too exhaustive for the users and thus
  unusable.
- **Feasible aggregation**: Preference aggregation is an optimization problem. Solving such
  problems can be quite difficult and depends on the preference representation.
- **Empirical validation**: The method must be tested in real world situations. Based on that data
  we can predict the performance of the method and thus of our aggregation mechanism.
- **Representation of preferences**: The method must allow an effective representation of the
  preferences of users. It is essential that users can entirely express their preferences.

The next sections give an overview of some prominent preference aggregation solutions

---

1 The strength of the preferences
2 The ranking of preferences
3.3.1 Classic voting
In many situations, we need to represent and reason about the simultaneous preferences of several users over the same objects. To aggregate the users’ preferences, we can query each user in turn and collect the results.

We can see this as each user “voting” whether an outcome dominates another. We can thus obtain different semantics by collecting these votes together in different ways. However, according to Arrow’s theorem, a fair voting system with two or more users and which in total orders three or more candidates does not exist (McCarthy K., 2007).

Aggregation function
\[ \sum_{i=0}^{n-1} x_{ik}, \] where \( x_i \) is a vote for an item \( k \). The item with the highest amount of votes is chosen.

Voting is very effective when fast decision making is needed. When divergence appears during negotiation or decision making processes, voting can be a tool to come to a decision quickly.

Evaluation
- **Usability**: Effective for decision making about complex issues with a lot of participants. The information needed is very simple and can easily be queried from the users.
- **Feasible aggregation function**: Since the optimum is represented by the item that received most votes, it is quite simple to determine the group optimum.
- **Non-manipulability**: Though strategic voting is possible, voting is quite solid as it comes to manipulation. It is quite difficult to have a lot of influence on the outcome.
- **Empirical validation**: Voting is a very old method and has been used for a long time.
- **Representation of preferences**: Within classic voting, assumptions exist about the nature of preferences. Namely, preferences are assumed to be binary. There is no representation of the intensity of certain preferences.

3.3.2 Intensity of preferences
The problem of combining individual preferences to form a set of group preferences can be solved in different ways. Decision making with classic voting is based on counting heads and the option that gains the highest amount of votes wins. When we use voting, we ignore a lot of user- information that is actually very important. For instance, a person that prefers one issue over its opposite only by a proportion of 51 to 49 percent. Yet, when the person votes it is recorded as definitely for or definitely against. If a group votes with mixed feelings then the outcome of that voting indicates a stronger win than is justified by reality. This is especially a problem in contexts such as the travel group decision support system, where success depends on the quality of the consensus that a group has reached. The next flaw of the voting system is that the decision derived from a majority vote may result in an outcome that is the opposite of what the collective wants. Another problem with the traditional voting method is that it does not take dependency between different issues into account. It is possible that certain issues influence other issues with respect to the attractiveness of that issue. If issues are bound together (dependent) to some extent, it can happen that an earlier issue with bearing on what follows is
voted out, killing the chance to successfully influence the others, unless it is brought back again for reconsideration. Such yes-no voting often prevents the decision-makers from following a comprehensive view of issues as a whole. It can lead to a chain of decisions that are hard to realize. This opens a floodgate for unsettling and paradoxical results. A better way would be to discuss all the relevant issues simultaneously and make decisions by ranking these issues (Thomas L. Saaty, Jen S. Shang, 2005).

The Analytical Hierarchy Process: Instead of classic voting we can use a different type of preference aggregation, by taking into account the intensity of the preferences. The AHP (Analytic Hierarchy Approach) is a cardinal approach for aggregating individual preferences (Thomas L. Saaty, Jen S. Shang, 2005). AHP provides a method for aggregating individual cardinal preferences into a unique group preference. AHP facilitates the group process to capture preference intensities of individuals and incorporates them into a final group decision. It ensures validity of the outcome as it relates to the real world, because we now have insight in how much people like or dislike a certain outcome.

In (Saaty, 2008) the analytical hierarchy process is described as follows:

To make a decision in an organized way in order to generate priorities we need to decompose the decision into the following steps.

1. Define the problem and determine the kind of knowledge sought.
2. Structure the decision hierarchy from the top with the goal of the decision, then the objectives from a broad perspective, through the intermediate levels (criteria on which subsequent elements depend) to the lowest level which usually is a set of the alternatives).
3. Construct a set of pair wise comparison matrices. Each element in an upper level is used to compare the elements in the level immediately below with respect to it.
4. Use the priorities obtained from the comparisons to weigh the priorities in the level immediately below. Do this for every element. Then for each element in the level below add its weighed values and obtain its overall or global priority. Continue this process of weighing and adding until the final priorities of the alternatives in the bottom most level are obtained.

To make comparisons, we need a scale of numbers that indicates how many times more important or dominant one element is over another element with respect to the criterion or property with respect to which they are compared. This can be done by assigning values to preferred features based on a 0 to 9 scale. Based on those assignments we can create a matrix that shows the relative values of the preferences with respect to a pair of criteria. Based on this we can calculate the weight of each criterion (feature).
**Aggregation function**

Once we have calculated the weight of each feature per user we need to aggregate these values in order to get a group preference expression. The way to do that is to calculate the geometric mean (Venable, 2005) of the weights:

Given N decision-makers, the elements $a_{ij}$ of the preference matrix aggregated by the geometric mean is:

$$\left[ \prod_{k=1}^{N} a_{ij}^{N} \right]^\frac{1}{N},$$

where $a_{ij}^{N}$ is the judgment of the $k^{th}$ voter when comparing item i with item j.

Next the individual values are replaced by the geometric mean. Finally based on these values the preferred features can be selected and after some data processing a decision can be made. The results in (Thomas L. Saaty, Jen S.Shang, 2005) show that it is an improvement of the conventional voting based group decision process.

**Evaluation**

- **Usability:** The pair-wise comparisons that need to be extracted from the users can become a time-consuming process as the amount of features increases. This has a negative effect on usability. The (future) size of our domain model would cause the AHP procedure to be or become time-consuming. This can finally have a negative effect on the usability of the procedure and thus a negative effect on the overall user satisfaction.

- **Feasible aggregation function:** the aggregation function is relatively simple.

- **Non-manipulability:** it is difficult to manipulate the outcome.

- **Empirical validation:** AHP is used a lot in the real world and has proven to be very effective.

- **Representation of preferences:** Preferences are well represented, since the intensity of the preferences is taken into account.

**3.3.3 Cp-nets**

CP-nets are an attractive model for representing individual preferences, in part because they allow us to find the best outcome for a user in time that is proportional to just the number of features in an outcome. CP-nets (Conditional Preference Networks) are a graphical model compactly representing conditional and qualitative preference relations. They exploit conditional preferential independence by structuring an agent’s preferences under the *ceteris paribus* assumption. Informally, CP nets are sets of *conditional ceteris paribus (CP)* preference statements (F.Rossi). Formally, according to (Lang), a CP-net is a pair consisting of a directed graph $G$ whose vertices are the variables and a collection of conditional preference tables, one for each variable. Formally, a CP-net is a pair consisting of a *directed graph* $G$ whose vertices are the variables and a collection of *conditional preference tables*, one for each variable. The table for variable $X$ contains, for each instantiation of its parent variables in $G$, a preference relation on the value domain of $X$. Consider the following CP-net over the two binary variables $X$ and $Y$, with possible values $x, \bar{x}$ for $X$ and $y, \bar{y}$ for $Y$ (Figure 3).
Figure 3: CP net expressed in a graph G and the conditional preference tables

The directed graph G means that the agent’s preference over the values of X is unconditional, and that her preference over the values of Y (resp. Z) is fully determined given the value of X (resp. the values of X and Y). The local preference tables express preferences between the values of a variable, everything else being equal (*ceteris paribus*). For instance, in the table for Y, the item $x > y$ means that when x is true, then y is preferred to $\bar{y}$ for any fixed value of Z. Therefore, in the preference relation expressed by the CP-net, we have $xyz > \bar{x}yz$ and $xyz > \bar{x}\bar{y}z$. Instantiations of X, Y and Z could be respectively Destination = New York, Price = 1000 euro and Hotel = 4 star rating.

**Aggregation function**

By creating partial CP-nets, we are able to build a mCP-Net, where m stands for the m partial CP-nets of m users. In partial CP-nets non-ranked features can occur. The presence of non ranked features is needed since a user’s preferences may depend on another user’s preferences. Different partial CP nets can be aggregated into one mCP-Net. Next voting is used to reason about the mCP-net. Different semantics can be used such as, Pareto, Majority, Max, Lex and Rank (Lang).

**Evaluation**

More research is needed before one can apply this approach in an application like the group travel decision support system.

- **Usability**: The size of our domain model causes the establishment of a preference model based on conditional preferences to become a cumbersome process.
- **Feasible aggregation function**: It is difficult to combine conditional preference models from different users to generate a group preference model.
- **Non-manipulability**: it is difficult to manipulate the outcome.
- **Empirical validation**: No best practices (yet) for real-world domains (as far as we know). This makes it difficult for us to predict the performance. (in future research it might be interesting to test approach)
- **Representation of preferences**: The intensity of the preferences is not taken into account.
3.3.4 Kemeny rule and Spearman’s Footrule

The Kemeny rule has been proposed as a way of seeking a compromise ranking in the majority vote when there are cycles present in the majority preference relation. The Kemeny rule satisfies the Condorcet criterion\(^3\) and a weaker version of local independence of irrelevant alternatives\(^4\). When a group of people express their preferences over a set of attributes and we try to manufacture an order that represents the preferences of the group, we can use the Kemeny order to find the consensus order.

Example (from (Davenport)):

Consider the profile of rankings (A,B,C), (B,C,A), (C,A,B). There is a majority of 1 vote for \(A > B\), a majority of 1 vote for \(B > C\) and a majority of 1 vote for \(C > A\). This profile defines a cycle in the majority relation of \(A > B\), \(B > C\) and \(C > A\).

![Figure 4: The preference graph (left) and the majority graph (right) for the profile of the example](image)

The example shows us that we cannot get a profile that represents the group preferences. The majority graph shows a cycle and therefore we do not have a Condorcet winner, which is a ranking that satisfies the Condorcet criterion. Kemeny proposed a way of breaking such cycles, using a notion of distance for orders. Given a total order \(r\), a weak order \(r^j\) and two alternatives \(s\) and \(t\), we define:

\[
\delta_{st}(r, r^j) = \begin{cases} 
1 & \text{if } r_s < r_t \text{ and } r_s^j \leq r_t^j \\
0 & \text{otherwise}
\end{cases}
\]

And

\[
\Delta(r, r^j) = \sum_{s \in X} \sum_{t \in X} \delta_{st}(r, r^j)
\]

The value of \(\delta_{st}(r, r^j)\) indicates whether there is a disagreement in the relative ranking of \(s\) and \(t\) between \(r\) and \(r^j\). \(\Delta(r, r^j)\)\(^5\) measures the total number of disagreements between \(r\) and \(r^j\) which is the

---

\(^3\) If some alternative is ranked ahead of all other alternatives by an absolute majority of voters, then it should be ranked first in the consensus ranking

\(^4\) If some alternative is ranked ahead of another alternative then it should remain unchanged after introduction of a new alternative

\(^5\) Kendall’s tau distance; the number of pair-wise of adjacent transportations to transform one order into another
defined as the Kendall Tau’s distance: The Kendall tau distance between two ranked lists is represented by the number of pair-wise disagreements in the relative rankings of items in the two lists. That is, 
\[ K(r_1, r_2) = \left| \{(i, j) \mid r_1(i) < r_1(j) \& r_2(i) > r_2(j)\} \right| \]

Another way to measure the total number of disagreements between \( r \) and \( r^j \) in the relative rankings of items between the two lists is Spearman’s Footrule which is defined by (Beg, 2003) as:

The Spearman Footrule Distance measures the distance between two ranked lists by summing up the differences in the rankings of each item. That is, given two complete rankings, where \( r^1 = r \) and \( r^2 = r^j \).

\[ F(r_1, r_2) = \sum_{i \in U} |r_1(i) - r_2(i)| \] (Sculley)

**Aggregation function**

The distance between a complete order \( r \) and a profile of rankings \( R \) is given by \( d(r, R) = \sum_{j=1 \ldots n} \Delta(r, r^j) \).

Now finally a Kemeny order for a profile \( R \) is an order \( r \) which minimizes the distance \( d(r, R) \). The Kemeny order is the order with a minimum number of disagreements with the pair-wise rankings in the profile \( R \). According to (Lin Padgham, Michael Winikoff, 2003 ) preference aggregation based on the Kemeny rule leads to less cycles in the majority ranking of alternatives. When we use the Spearman’s Footrule for rank aggregation, the order \( r \) that minimizes \( d(r, R) \) for \( d(r, R) = \sum F(r_1, r_2) \) is the optimal ranking between the different rankings of \( R \).

When we compare Kemeny and Spearman’s Footrule optimization, then we learn that Kemeny is computationally much more expensive then Spearman’s Footrule. Kemeny will be NP – hard from as few as four lists (Beg, 2003). While Spearman’s Footrule optimization can be done in polynomial time (Li, 2009).

**Evaluation**

Ordered lists are one of the most common used representations of preferences. We noticed that a lot of preference elicitation techniques are based on ordered lists (rankings). We learned that in order to aggregate preferences representing rankings, we need distance measures. Kendall’s Tau and the related and less complex Spearman’s distance (Diaconis, 1977) are mostly used for this purpose.

- **Usability**: It is relatively easy for users to give rankings, therefore this method is relatively usable
- **Feasible aggregation function**: Depends on the distance measure. Kendall’s Tau can be computationally hard to find an optimum. Spearman’s Footrule however is much easier.
- **Non-manipulability**: it is difficult to manipulate the outcome.
- **Empirical validation**: rankings have proven to be quite effective for preference aggregation; both Kendall’s Tau and Spearman’s Footrule have been widely used.
- **Representation of preferences**: Quite well, however the intensity has not been taken into account.
3.3.5 Conclusion

In the previous paragraphs we have elaborated different preference aggregation methods. In order to make a choice we have evaluated them against the criteria that we set in paragraph 3.3.

<table>
<thead>
<tr>
<th></th>
<th>Usability</th>
<th>Feasible aggregation function</th>
<th>Non-manipulability</th>
<th>Empirically validated</th>
<th>Representation of preferences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classic voting</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>AHP</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>CP-NETS</td>
<td>✗</td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Kemeny rule</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Spearman’s Footrule</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 1: overview of aggregation methods vs. our criteria

Spearman’s Footrule and Kemeny rule both meet all the defined criteria. However, in comparison with the Kemeny rule (rank aggregation by using Kendall’s Tau), Spearman is computationally less complex, which is why we prefer Spearman’s Footrule over the Kemeny rule. Furthermore, in (Dinu, 2006) it is shown that the Kemeny rule has a strong connection with Spearman’s Footrule and that Spearman’s Footrule in fact approximates the Kemeny rule. “If \( \sigma \) is a Kemeny optimal aggregation for the multi-set of rankings \( \{\tau_1 \ldots \tau_k\} \) and \( \sigma' \) is a Footrule aggregation for the same multi-set, then \( K(\sigma,\{\tau_1 \ldots \tau_k\}) \leq 2K(\sigma',\{\tau_1 \ldots \tau_k\}) \) where \( K \) is the Kemeny distance of the optimal aggregation for a multi-set of rankings.” (Dinu, 2006) This is why the Spearman Footrule is very useful for aggregation of full rankings.

A direct consequence is that the preferences will be represented by rankings of city categories and features. The preference model is discussed in the next paragraph.
3.4 Preference model

Until now we have discussed the domain model and the preference aggregation methods. Subsequently we chose an aggregation method that we will apply within the system (further explained in 3.5). First we establish the preference model.

A preference model can be seen as a representation of the preferences of users concerning a set of items. The preference model is dependent on the domain model, since the domain model determines the items (features) that users use to establish their preferences. The aggregation method determines how the preferences are represented within the system. As explained in 3.3.4, the Spearman Footrule determines the distance between two rankings. Thus it is required that a preference model exists of a ranking to able to apply the Spearman Footrule. Therefore we propose that users will express their preference by ranking the different features (described by the domain model) based on its importance. In Figure 5 an abstract representation of a preference model is displayed.

![Preference Model Diagram]

Figure 5: preference model

The grey elements in Figure 5 represent the categories of the city model. Below each category the blue elements show the features of the city model. The user can rank the categories as well as the features belonging to that category. The numbers represent the rank of a category or a feature. The highest rank is 1 and the lowest is N or K (where N in case of categories; is equal to the amount of categories while in case of features; K is equal to the amount of features belonging to a category).
The preference model requires that the user specifies his or her preferences by ranking the (N) categories and (K) features per category according to what he or she finds important. Every user is requested to establish such a preference model.

The next step is to aggregate the preference models. In the next section we will discuss the preference aggregation algorithm, which is based on the aggregation method (Spearman’s Footrule) that we selected in the previous section.
3.5 Preference aggregation algorithm

In the previous sections we have described the aggregation method and the preference model that we use to represent the preferences of the users. Within this paragraph we discuss the algorithm that we can use to aggregate the preference model by using the Spearman Footrule. As explained earlier the Spearman Footrule is a measure of distance between two rankings. It is calculated by the following function:

\[
F(r_1, r_2) = \sum_{i \in U} |r_1(i) - r_2(i)| \quad \text{(Sculley)}
\]

Where \( r_1(i) \) is considered as the index of element \( i \) of ranking \( r_1 \), while \( r_2(i) \) the index of element \( i \) in \( r_2 \). The distance between ranking \( r_1 \) and ranking \( r_2 \) is the summation of all the differences between the indexes of the elements in one ranking and the corresponding element in the other. This distance is called the Spearman Footrule distance. During the aggregation process we will compare; (i) the rankings of the categories from one user to another and (ii) per category the ranking of the features from one user to another. So the preference model is actually divided into 8 rankings. The first of the categories (from 1 to 7) and the other 7 rankings of the features per category (all having different sizes with the largest size being 7) The 8 rankings of a group member’s preference model are all independently compared with the corresponding rankings of the other group members’ preference models. Now the challenge is to find a preference model that has a minimum Spearman Footrule distance to all al the user’s preference models. It is this preference model that can be considered as the optimal group preference model. The distance is of course based on the distance between the individual rankings. Thus the question is how to generate an optimal ranking. Let’s again look at the case where we have two rankings, \( r_1 \) and \( r_2 \). Now we have to find a ranking \( r_{opt} \) that has a minimum Spearman Footrule distance to the rankings; \( r_1 \) and \( r_2 \). Ranking \( r_{opt} \) is then regarded as the optimal ranking between \( r_1 \) and \( r_2 \).

Finding \( r_{opt} \) concerns a search problem that can be solved by different algorithms, such as Cross-entropy (CE) Monte Carlo algorithm (Pihur Vasyl, Susmita Datta, Somnath Datta, 2007) or a genetic algorithm such as explained in (Pihur, 2009). However in (Pihur Vasyl, Susmita Datta, Somnath Datta, 2007) it is said that when we deal with a small set of items, a brute force approach is certainly feasible. In our case the rankings of the categories have a maximum size of 7 (as there are 7 categories). When we look at the amount of features within each category, we notice that the largest set also has a size of 7 items. This is a relatively small amount of items and thus we use a brute force approach in order to solve the optimization problem.

The brute force approach means that we have the system generate all possible rankings. Subsequently we calculate the Spearman’s Footrule of each generate ranking to the rankings specified by the users and determine the ranking with the minimum spearman distance. We do this for each ranking (thus the categories and the features belonging to these categories) and after all optimal rankings have been determined, the group preference model can finally be constructed. The rankings are generated by a simple recursive permutation algorithm (Heap). An abstract overview is given of the aggregation algorithm (Figure 6).
is now possible to match the preference model with city information that Trip.Easy has in its database.

The utility function translates the preference models into a relative measure of value based on the rankings of the categories or features (Norvig, 2003). Based on those values the importance of the city categories and city features can be determined by Trip.Easy. In the previous sections we explained that the preference model is constructed from rankings of ‘categories’ or ‘features’. Thus to be able to interpret the preference model, Trip.Easy must translate these rankings into a value that represents the importance of the items ranked within such ranking. The value is what we call a weight. The weights are calculated by the utility function. The input of the utility function is a ranking and the output is weight related to each item of the ranking. The utility function is based on an interview that we had with a travel consumer marketing expert (Appendix E). However, during the development of the prototype, we had it fine-tuned based on tests. The utility function \( U(i) \) where \( i \) is the index of an item in a ranking, works as follows:

\[
U(i) = \begin{cases} 
  \frac{1}{\sqrt{i}} \times 0.3 & \text{if } i \leq \text{median}(r) \\
  \frac{1}{\sqrt{i}} \times 5 & \text{if } i > \text{median}(r)
\end{cases}
\]

Subsequently the result of \( U(i) \) is normalized. As a result we get weights between 0 and 1, where the items with the higher rank have more weight than the items with lower ranks. Based on these weights it is now possible to match the preference model with city information that Trip.Easy has in its database. Based on the matching procedure, the top twenty set of cities is selected from the database and presented to the user. (The actual code that is responsible for this can be found in Appendix E)

Effectively, the matching procedure is a multiplication of the weights of each city category or feature and the corresponding utility value of the city category and feature from the preference model of a user. Given; category \( c \) and feature \( f \), then the utility values are \( U_c \) and the feature belonging to this category \( U_{cf} \). The weights are represented as \( W_c \) and \( W_{cf} \). The procedure is represented by the following expression:

\[
S(p) = \sum_{i=k}^{i=0} (W_{c,i} \times U_{c,i}) \sum_{j=m}^{j=0} W_{cf,m} \times U_{cf,m}
\]

\( S(p) \) is the score of a city based on the preference model \( p \). The top twenty scoring cities are presented to the user. In the next section we will briefly discuss the decision-making process that we will apply in the prototype.
3.7 Well-structured decision process

Finally, the recommendation mechanism is integrated with the well-structured decision process established by (Ngai, 2010). The decision process exists of four main steps; the ‘Generate’, ‘Reduce’, ‘Clarify’ and ‘Evaluate’ step. These steps are derived from Collaboration Engineering Theory (Briggs, 2006).

**Generate:** In (Briggs, 2006) it is explained as the process of collecting and share known concepts. A concept can be regarded as a possible solution the problem of a problem-solving process. Thus, in our domain it is regarded as a possible city to visit, given the preferences of the individual users. Therefore in this step the preference models of the users are determined, as described in paragraph 3.4. Subsequently, the individual preferences are aggregated into a group preference model as discussed in paragraph 3.5. Furthermore, the group preference model is matched with the city characteristics, as discussed in the previous section. As a result, a list of cities is generated and shared with all users.

**Reduce:** This step is explained as the process of reducing the generated concepts to a set that the group deems worthy of further attention (Briggs, 2006). Therefore users will select a set of cities that they find most interesting. In effect, they rank the cities according to their preferences. The average top six ranked cities are chosen for the next step.

**Clarify:** To create awareness and understanding among group members, the clarify step facilitates users to share opinions and intensify their preferred concepts (Briggs, 2006). Therefore users are able to rate the cities that are still left. Furthermore they are able to share comments and opinions in order to make other group members aware of their interests. The top four average highest rated cities are selected for the next step. The comments and/or opinions are used in the next step to ease the making of the final decision.

**Evaluate:** Move from less to more understanding of the relative value of the concepts under consideration (Briggs, 2006). This is effectuated by a debate among the group members. As a result, the final choice is made.

We conclude this chapter with an abstract overview of the whole decision process combined with the recommendation mechanism (Figure 7). In the next chapter we discuss the process in somewhat detail as an implementation of steps discussed above.
Introduction GSS process and Goal

Start debating process

Evaluate: debate

Final choice

Display end result

Evaluate: choice

Generate: Gathering

Ranking

Reduce: Filtering

Display cities recommended cities

Display cities + user feedback

Clarify: Building shared understanding

Rate the recommended cities + give feedback

Get all the user feedbacks

Display cities + user feedback

Display cities recommended cities

Get all cities selected by all members

Insert preferences over general city features

Aggregate individual model into group model

Map groupmodel to ordered list of cities based on score function s(x)

Display 20 best matching cities

Generate: Gathering

Reduce: Filtering

Clarify: Building shared understanding

Evaluate: debate

Evaluate: choice

Figure 7: Overview of the decision-making process combined with the recommendation mechanism
4 Development and Design

In this chapter we discuss the realization of the prototype. We first start with a functional design, by which we describe the most essential functionalities divided over two stakeholders: user and administrator. Subsequently we discuss the technical design.

4.1 Functional design

An overview is given of the functionalities for the users and the administrators.

4.1.1 Users

- **Request for friendship**: Users are able to request other users to be their friend. If the other user accepts the request, the two are connected to each other.
- **Initiate a Trip.Easy Session**: A Trip.Easy Session facilitates the decision-making process for a group of users. Every user of Trip.Easy can start a session.
- **Invite friend**: as soon as a user has initiated a Trip.Easy Session, they can invite their friends to join in.
- **Registration**: new users need to fill-out the registration form.
- **Login**: when users are registered they can login by using their account.

In the subsequent sections we will elaborate on the functionalities per process step, which were introduced in the previous section.

4.1.1.1 Generate step

**Insert preferences**: During this step the user can insert his or her preferences concerning the features defined by the domain model in paragraph 3.2, which the preference model (paragraph 3.4) is determined. The recommendation mechanism will aggregate all individual preference models and generate a group preference model. Finally, recommendations are generated as explained in the previous chapter. Figure 8 shows the user interface.
The colored blocks (Figure 8) in the middle show the categories. Below each category the features are represented as blocks in the same color as the category. The user can drag the features from important (left) to less-important (right). The same can be done for a whole category. A detailed elaboration of the user interface and interaction design can be found in (Ngai, 2010).
4.1.1.2  Reduce step

**View group recommendations:** Based on the specified preferences, the system recommends cities that fit best to the group. These cities are listed according to their match with the group preference model.

**Reduce the alternatives:** Furthermore, users are able to rank the cities to select their top six most favorite cities.

---

**Figure 9: user interface reduce step**

Figure 9 shows the user interface of the reduce step. The recommended cities are represented on the left. Effectively, users can rank the cities by dragging them according to their preferred order. A detailed elaboration of the user interface and interaction design of this step can be found in (Ngai, 2010).
4.1.1.3 **Clarify step**

Specify intensity of preference: By rating the cities the user can assign a level of importance to the cities. The user can specify this, by rating the cities. Based on these ratings, the top four highest rated cities are selected for the next step.

Share opinions: Another important functionality is the possibility to share opinions or other remarks with the other group members.

---

**Figure 10: user interface clarify step**

Figure 10 shows the user interface of the clarify step. The user can insert the ratings by using the clickable ‘stars’ represent the rating (form 1 to 5). Opinions and remarks can be shared by using the
input-boxes (middle) and the text-box (right). A detailed elaboration of the user interface and interaction design of this step can be found in (Ngai, 2010).

4.1.1.4 **Evaluate step**
During this step the users will have a debate by which they will make the final decision. The opinions, remarks and the ratings are shown to support the users during their debate.

![User Interface Evaluate Step](image)

After the debate the final decision is confirmed by using the selection button (middle). A detailed elaboration of the user interface and interaction design of this step can be found in (Ngai, 2010)
4.1.1.5  **Administrators**

The administrator can manage and configure the Trip.Easy prototype.

- **City management**: Administrators can add/remove/edit the cities.
- **Weight management**: Administrators can configure the weights related to categories and features of a city (pre-defined by the domain model).
- **Domain model management**: Administrators can add/remove/edit the categories and features of the system domain model.
- **Management of online questionnaire**: The administrators can add/remove/edit the questions of the online questionnaire.

Screen impressions of the functionalities named above can be found in Appendix I.

4.2  **Technical design**

Within the technical design we explain how the recommendation engine is implemented and integrated with the decision process. Furthermore we discuss the essential components of the system and explain how the components interact with each other.

4.2.1  **Web-application**

The Trip.Easy prototype is a web application (Figure 12) that runs on a web-server. The application is accessed over the internet. This setup is inherent to the type of software we are developing, because a lot of complex information is retrieved from each user and needs to be processed centrally; next the results are broadcasted to each group member. Secondly, the application can be maintained and updated on the server, without having to distribute and install it on all the different client computers. This saves a lot of time during the testing phase and experimental phase.

Finally, the support for cross-platform compatibility makes it possible to access the application from different devices, such as pad-computers, notebooks and desktops. As long as a device has a supported web browser installed.

![Figure 12: web application in general](image-url)
4.2.2 Trip.Easy Recommendation Mechanism

The recommendation mechanism has several key components that are depicted in Figure 13.

Preference Model Constructor: This component is responsible for constructing a preference model (paragraph 3.4) out of the ranked categories and their ranked features. After all preference models are constructed they are sent to the recommendation engine. The recommendation engine exists of several components. Each step in the decision process is supported by a single or several component(s) of the recommendation engine.

Preference Aggregator: To generate a group preference model the individual preference model are aggregated by the Preference Aggregator. Within this component we have implemented the preference aggregation algorithm as described in paragraph 3.5. Below we show the three steps: *generate all possible rankings, calculate distances and get optimal ranking* in pseudo code.
Generate all preference models:

```java
generateRankings(int n) {
    if(n == 1) {
        generatedRankings.add(foundation)
    } else{
        for(int i=0; i<n; i++) {
            permut(n-1)
            if(n%2 == 1)swap(foundation,0,n-1)
            else swap(foundation,i,n-1)
        }
    }
}
```

This code is responsible for generating permutations. These permutations are all saved in the list called generatedRankings. The system uses this procedure to generate all possible rankings (permutations) for each ranking of the preference models of the individual group members.

Calculate distances:

```java
private def spearmanDistance(def generatedRanking, def ranking) {
    def distance= 0
    for (int i=0; i<inputOrder.size();i++) {
        distance += (generatedRanking.indexOf(inputOrder[i]) - i).abs()
    }
    return distance
}
```

For each generated ranking the system calculates Spearman’s distance with the corresponding rankings of the user preference models. Based on this distance the rankings are ordered and saved in a list called orderedRankings.

Get optimal ranking:

```java
getOptimalRanking(def orderedRankings) {
    return orderRankins[0]
}
```

Finally the ranking with the smallest distance to the corresponding rankings in the preference models of the group members is selected. This ranking then is the optimal aggregation of the corresponding group rankings.

This process is executed for each ranking of the preference model. Thus, the ranked categories are aggregated into an optimal group ranking and the ranked features for each category are aggregated into an optimal group ranking. All the rankings together, form the group preference model. The group preference model has exactly the same structure as Figure 5 in paragraph 3.4. Subsequently, the obtained group preference model is further processed by the Utility Calculator.
Utility Calculator: This component determines the utility values of the categories and features of the preference model. Within this component, we have implemented the procedure discussed in paragraph 3.6. As a result the preference model now carries utility information. Eventually, to be able to generate recommendation the preference models are matched with city information from the database. The pseudo code is shown in Appendix E.

City Matcher: this component matches the cities with the preference models. Based on the category and feature information of both, the City matcher is able to compare a city with a preference model and calculate how well it fit to a preference model. Within this component we have implemented the procedure described in paragraph 3.6. As a result a list of best-fit cities is generated and presented to the user. The pseudo code is shown in Appendix E.

Ranked City Aggregator: This component aggregates the ranked cities of each group member into a list of ranked cities. This list is ordered by the average ranking of the cities by the group. The average top six ranked cities of this list are selected and returned. The procedure is as follows: given a set of lists of ranked cities $R = \{r_1, r_2, r_3, ..., r_k\}$ where $k$ the amount of group members, then the score is calculated by:

$$S_{\text{rank}}(c) = \frac{1}{k} \sum_{k}^{1} \frac{1}{\sqrt{j}}$$

Where $S_c$ is the average score of a city is $c$ based on the ranks of that city is the lists of $R$ and $j$ is the rank index of $c$ in the $k$ rankings in set $R$. Subsequently, the top six cities, based on the $S_c$, are selected and returned to the group.

Rated City Aggregator: This component aggregates the list of rated cities of all group members into one list of rated cities. This list is ordered by the average rating of the cities by the group. The average top four rated cities are selected and returned. The procedure is as follows: given a set of lists of rated cities $R = \{r_1, r_2, r_3, ..., r_k\}$ where $k$ the amount of group members, then the score is calculated by:

$$S_{\text{rate}}(c) = \frac{1}{k} \sum_{k}^{1} v$$

Where $S_{\text{rate}}(c)$ is the average score of a city $c$ based on the rates of that city in the lists of $R$. Variable $v$ is the value of the rating. Subsequently, the top four cities, based on the $S_{\text{rate}}(c)$, are selected and returned to the group.

In Figure 14 an interaction diagram of the components described above and the structured group decision-making process facilitated in Trip.Easy is shown.
During the ‘Generate’ step users insert their preferences and a preference model is constructed for each group member. The preference models are then aggregated thereafter a group preference model is generated by the ‘Preference Aggregator’. The ‘Utility Calculator’ determines the utility values of the items (categories and features paragraph 3.4) of the group preference model. Finally the ‘City Matcher’, matches the cities with the group preference model and returns the top twenty best-fit cities to the group. Subsequently, the group members rank this list of cities individually. These ranked lists are then aggregated by the ‘Ranked City Aggregator’ into one list containing the average top six ranked cities. Thereafter the users will rate the list of top six ranked cities individually. These lists of rated cities are then aggregated by the Rated City Aggregator into one list containing the average top four rated cities. Based on these results the group has to make their final choice during the ‘Evaluate’ step.

4.2.3 Database

The database of Trip.Easy can globally be divided into three types of data, starting with cities. Trip.Easy contains information about 150 cities world-wide (Appendix H). Each city is described by the categories and features defined by the domain model (paragraph 3.2). Additional information such as name, country and price has been added. Furthermore, the database has information about the users. It has data about individual preference models and general information such as name, age, username, password, etc. Subsequently, the database contains information about the Trip.Easy sessions conducted by a group of users. The data exists of group preference models, the recommendations generated by Trip.Easy per process step and the final outcomes of groups. (Effectively, almost all decisions and actions of the group during a session is saved in the database.) Finally the database contains experimental data, such as questions of the questionnaire (Appendix A) and the results of the experiment (the answers of the participants). Figure 15 gives an abstract overview of the Trip.Easy database.
Figure 15: Global view of the Trip.Easy Database
4.2.4 Development approach
The development of the Trip.Easy prototype required an agile approach. This means that we designed and developed parts of the system. Starting with the design, implementation and testing of the algorithms and models described in chapter 3. Based on that integrated these components with a web framework. Thereafter we designed and build the database. Finally we integrated the user interfaces and the database with the framework and the components. This established the full-functioning Trip.Easy prototype.

4.2.5 Conclusion
In this research our first goal was to design and implement a group recommendation mechanism for a group decision support system for groups of travel consumers. In the past two chapters, we have shown the development and design of a recommendation mechanism for a group decision-making process for groups of travel consumers. This resulted in Trip.Easy, an intelligent group decision support system for group of travel consumers. In the next chapters we will discuss the evaluation of this system.
5 Evaluation

5.1 Introduction
The previous chapter described the design and implementation of the prototype of the Trip.Easy GDSS. The goal of the Trip.Easy GDSS is to support a group of travel consumers that are organizing a city trip that represents a best-fit for the group. To this end a well-structured decision process has been designed in combination with intelligent technology that provides smart recommendations. The recommendations are based on the preferences of the group, represented by a group preference model. The group preference model is established by preference aggregation of individual preference models. The individual preference models were extracted from the users via preference elicitation technology. During each step of the process, the recommendations done by Trip.Easy are fine-tuned and adapted to the actual preferences of the group. Consequently, the goal is to lead the group towards a satisfying outcome. We have formulated the following main research question:

Do group members believe that the outcome established using Trip.Easy GDSS represents a satisfying outcome for the group?

The research question is divided into sub-questions that focus on (i) the evaluation of the outcome of the sessions and (ii) the evaluation of the recommendations. We apply Reinig’s Goal Attainment Model to conduct the evaluation of the outcome of the sessions which we further discuss in 5.3. This model evaluates the satisfaction of a technology by measuring the satisfaction within, in this case, a group decision session that is supported by this technology. This is an important aspect to start with, because research has shown that users that find their experiences with a technology dissatisfying tend not to use that technology in the future (Reinig, 2001). Therefore we first want to determine whether users are satisfied with the session outcome, as defined by Reinig’s Goal Attainment Model. The session outcome is strongly influenced by the recommendations generated by Trip.Easy during each session. Therefore we have defined research questions that focus on the evaluation of the recommendations. Consequently we will evaluate the following key aspects per process step:

- The satisfaction of the proposed cities
- The development of the satisfaction of the proposed cities throughout the process
- The satisfaction of the amount of proposed cities

In paragraph 5.2 we describe the research in more detail. Thereafter we discuss the Goal Attainment Model and the User Satisfaction Analysis in paragraph 5.3. Subsequently we will discuss the Experimental Setup in 5.4. Finally the results and additional analyses are presented in 6.
5.2 Research

5.2.1 Evaluation of the outcome of the session
In order to answer the main question we will evaluate the satisfaction of the session outcome (SO) as explained in section 5.3. This variable will be evaluated at the end of each session by a group of participants. It indicates whether the session has lead to a satisfactory outcome. We expect that the final city that the group has selected will be considered a satisfying outcome. In addition, it is also interesting to look for further results of the experiments. As we already explained in chapter 3, we use the Spearman Footrule distance to measure the distance between two rankings during the first step of the decision process. At this point we determine the preferences of the city features and thus the preference model for each user. Based on those preference models the system generates a group preference model by applying the Spearman Footrule distance. We are interested in the influence of the distance between preference models (the dissimilarity), on the satisfaction level of the final outcome of the session.

5.2.2 The evaluation of the outcome of the session
We want to evaluate the outcome of the session by finding an answer to the question: Are users satisfied with the outcome? The goal of the system is to lead users towards a satisfying outcome, and we expect users to be satisfied with this outcome. This question gives rise to hypothesis H. 1:

H. 1: Group members find the session outcome satisfying.

To measure the satisfaction of the outcome, we have derived and adapted the following questions from the questionnaire from (Robert O. briggs, Bruce A. Reinig, Gert-Jan de Vreeds, 2006):

1. Do users like the outcome of the session?
2. Are users satisfied with the selected city?
3. When the session was finally over, did users evaluate the selected city positively?
4. Are users happy with the selected city?

5.2.3 Similarity of preferences versus the satisfaction of the session outcome
To compare the similarity of preferences with the satisfaction of the outcome, we want to answer the question: Is there a qualitative difference in outcome between groups with members that have similar preference and groups with members that have less similar preferences?

This question gives rise to H. 2:

H. 2: Trip.Easy leads users to a satisfactory outcome, despite the dissimilarity of preferences within a group.

Statistical analysis of the data will show if a relation exists between the dissimilarity of preference models within a group and the satisfaction level of the outcome. The results can be found in chapter 6. The next section elaborates on the hypothesis of the evaluation of the recommendations per process step.
5.2.4 Evaluation of the recommendations

In this section we explain the evaluation of the satisfaction of the cities that are proposed by Trip.Easy at the end of each process step. As we explained in chapter 4, Trip.Easy has four main process steps. Respectively the ‘generate’ step, the ‘reduce’ step, the ‘clarify’ step and the ‘evaluation’ step. During ‘generate’ step Trip.Easy elicits the preferences of the user. Subsequently, the individual preference models are aggregated into a group preference model. Finally, Trip.Easy recommends a set of cities that are considered to be a best-fit for the group. Furthermore, the users select the most preferred recommended cities during the ‘reduce’ step. Consequently, Trip.Easy proposes a reduced set of cities matching the group’s preferences. In the next step, users can share ratings and comments about the resulting cities. As a result, Trip.Easy proposes a further reduced set of cities based on the user-ratings. Finally during the ‘evaluation’ step, the users have a debate about these cities and make the final decision (the city they will visit).

The recommendations during the session have a lot of influence on the final outcome. Therefore we want to evaluate these recommendations per process step. Ideally, the recommendations of the Trip.Easy GDSS should be satisfactory for each group member and therefore lead to a satisfactory group-decision outcome. To get an indication of the performance of the Trip.Easy GDSS concerning the quality of the recommendations, we want to answer the following question: Do group members believe that the recommendations of Trip.Easy support the group adequately to arrive at a satisfying outcome?

To answer this question we observe the following aspects; (i) satisfaction of the recommendations, (ii) the course of the satisfaction of the recommendations throughout the decision-process and (iii) the satisfaction of the amount of proposed cities. These aspects are respectively reflected by the following questions:

- **Are users satisfied with the proposed cities during each process step?**
  This question is reflected by hypotheses: H. 3, H. 5, H. 7 and H. 9

- **Does the user satisfaction with the proposed cities increase during the whole process?**
  This question is reflected by H. 10

- **Are users satisfied with the amount of proposed cities in each process step?**
  This question is reflected by: H. 4, H. 6 and H. 8
5.2.4.1 Generate

During the ‘generate’ step, all individual preferences are merged together into a group preference model. The goal of this step is to present a reasonably satisfying first set of cities. To this end Trip.Easy proposes a list of cities that fits best with the group. We expect that users will perceive the presented cities as a reasonable first proposal. This means that we expect them to be at least somewhat satisfied about the proposed cities.

H. 3: Users perceive the cities presented during the ‘generate’ step as a reasonable first proposal.

The following questions measure the user satisfaction level of the proposed cities:

1. Do users perceive the cities that are proposed during the ‘generate’ step as satisfying?
2. Do users perceive the cities that are proposed during the ‘generate’ step as matching their expectations?
3. Do users perceive the proposed cities during the ‘generate’ step as matching their preferences?

The aim of the ‘generate’ step is to propose a balanced amount of cities. This means enough alternatives need to be proposed to facilitate a user to make a selection without being overloaded with information. Therefore, we are interested in whether the users are satisfied with the amount of cities that are proposed during this step. We expect that the amount will be sufficient, which is reflected by H. 4.

H. 4: Users are satisfied with the amount of cities that are proposed during the ‘generate’ step.

The following question measures the user satisfaction level of the amount of the proposed cities:

1. Do users perceive the amount of cities that are proposed during the ‘generate’ step as sufficient in order to make a selection?

The user will make a selection out of the proposed set of cities by ranking them according to his or her preferences. This happens during the ‘reduce’ step which is explained in the next section.
5.2.4.2  Reduce

The ‘reduce’ step is used to further reduce the set of cities resulting from the ‘generate’ step. The function of the ‘reduce’ step is to fine-tune the recommendations, based on rankings of the proposed cities. The top six cities with the average highest ranks remain and represent the cities that the group members prefer most. Therefore we expect users to be satisfied with the results.

**H. 5: Users are satisfied with the cities that are proposed during the ‘reduce’ step.**

The following questions measure the user satisfaction level associated with the proposed cities:

1. Do users perceive the cities that are proposed during the ‘reduce’ step as satisfying?
2. Do users perceive the cities that are proposed during the ‘reduce’ step as matching their expectations?
3. Do users perceive the proposed cities during the ‘reduce’ step as matching their preferences?

The proposed set of cities is smaller now. Still, the amount of cities that is left should be sufficient enough for the user to proceed with the decision process. This gives rise to hypothesis H. 6.

**H. 6 : Users are satisfied with the amount of cities being proposed during the ‘reduce’ step.**

The following question measures the user satisfaction level of the amount of the proposed cities:

1. Do users perceive that the amount of cities that are proposed during the ‘reduce’ step is sufficient in order to make a selection?
5.2.4.3  Clarify

The goal of the ‘clarify’ step is to create awareness among the members of each others’ preferences. Within this step, the members can openly share their opinions on the proposed cities. Based on the ratings that users assign to the cities, the system again selects a sub-set of cities. Only the top four cities with the highest ratings are chosen. Therefore we expect the satisfaction of this proposed set of cities to be relatively high. This is reflected by H. 7:

**H. 7: Users are satisfied with the cities that are proposed during the ‘clarify’ step.**

The following questions measure the user satisfaction level of the proposed cities:

1. Do users perceive the cities that are proposed during the ‘clarify’ step as satisfying?
2. Do users perceive the cities that are proposed during the ‘clarify’ step as matching their expectations?
3. Do users perceive the proposed cities during the ‘clarify’ step as matching their preferences?

The amount of cities being proposed is important. The intention is to provide the right amount of cities to make sure users converge on a solution. This gives rise to hypothesis H. 8:

**H. 8: Users will be satisfied with the amount of cities being proposed during the ‘clarify’ step.**

The following question measures the user satisfaction level of the amount of the proposed cities:

1. Do users perceive that the amount of cities that are proposed during the ‘clarify’ step is sufficient in order to make a selection?

The set of cities that results from this step is further discussed by the group during the ‘evaluate’ step, which eventually results in the final decision.
5.2.4.4 **Evaluate**
During the ‘evaluate’ step the users will discuss the top four rated cities resulting from the ‘clarify’ step. The goal is that they make their final choice. Therefore we are interested in the satisfaction level of the participants with the final result.

**H. 9: Users are satisfied with the final city that is chosen during the ‘evaluate’ step.**

The following questions measure the user satisfaction level associated with the proposed cities:

1. Do users perceive the cities that are proposed during the ‘evaluate’ step as satisfying?
2. Do users perceive the cities that are proposed during the ‘evaluate’ step as matching their expectations?
3. Do users perceive the proposed cities during the ‘evaluate’ step as matching their preferences?

The users have now decided which city they will visit. In the next section we will further analyze the satisfaction of the cities that are proposed by Trip.Easy. During that analysis we observe the development of the satisfaction level throughout the decision process.

5.2.4.5 **Evaluation of the satisfaction of the proposed cities between steps**
Until now we have discussed the satisfaction of the proposed cities per step independently. The goal of Trip.Easy is to lead a group towards a satisfying outcome. Starting at the ‘generate’ step, Trip.Easy proposes a reasonable set of cities, and via the ‘reduce’ - and ‘clarify’ step, it should result in a satisfying outcome at the ‘evaluate’ step. We expect that Trip.Easy will lead the group towards a satisfying outcome, as reflected by hypothesis H. 10.

**H. 10: Trip.Easy leads the group towards a satisfying outcome.**

To evaluate H. 10 we will compare the user satisfaction of the cities that were proposed by Trip.Easy during the ‘generate’, ‘reduce’, ‘clarify’ and ‘evaluate’ step.

To sum up, in this paragraph we have discussed the research questions and translated them into measurable hypotheses. In the next section we will further motivate the concepts of the research by explaining the Goal Attainment Model (Reinig, 2001). From this model we have derived the concept ‘satisfaction with the session outcome’.
5.3 Model

We use two methods to evaluate the hypotheses described in the previous paragraph. The first method is based on the Goal Attainment Model (Reinig, 2001). This model centralizes the concept of technology acceptance of users. Our goal is to find out whether users have positive experiences with the technology. In case they do, it is more likely that they are willing to use it in the future (Reinig, 2001).

The Goal Attainment Model evaluates the so-called meeting satisfaction of a group that went through a decision process while being supported by GDSS technology. According to (Reinig, 2001) the meeting satisfaction consists of two dimensions, namely, the satisfaction with the so-called meeting outcome (SO) and the satisfaction with the meeting process (SP). In case of Trip.Easy, the SO evaluates the satisfaction of the Trip.Easy users with respect to the city they have chosen at the end of the session. The SP evaluates the satisfaction with the process embedded in Trip.Easy. This is further evaluated in (Ngai, 2010).

The Goal Attainment Model is based on the assumption that individuals hold multiple goals. During the course of a meeting, some goals may be advanced (resulting in positive value appraisal) while others may be hindered (resulting in negative value appraisal). The model posits a cognitive mechanism that automatically and subconsciously aggregates the advances and hindrances associated with an object of satisfaction to arrive at a net value. A positive net value gives rise to satisfaction with an object. A negative net value gives rise to dissatisfaction with an object. The model also recognizes that one important goal people usually hold for a meeting process is to produce a satisfactory outcome. Meeting processes that produce satisfactory outcomes are more likely to be satisfying than processes that give rise to dissatisfactory outcomes. Thus, SO should also account for some portion of variance in SP. Nevertheless, this relationship would be expected to be weaker than the relationship between NPGA and SP (see Figure 16), because SO would be only one of many possible goals (Reinig, 2001). In distinguishing between SO and SP, the model is consistent with procedural justice research. These studies have shown that people care about the process by which they reach a decision, in addition to the decision outcome itself (W.C. Kim, 1995), (M.A.Korsgaard, 1995). Team members value process fairness, such as having their input considered and having influence over the final decision (M.A.Korsgaard, 1995). Thus, if a team member perceives a process as fair, then, ceteris paribus, such a perception should instantiate positive goal attainment and cause satisfaction with the meeting process. Furthermore, if a team member believes that his or her interests have been taken into consideration, then he or she is more likely to believe that those interests are reflected in the final outcome. Consequently he or she should exhibit more positive attitudes toward the final decision as well (M.A.Korsgaard, 1995). (Robert O. briggs, Bruce A. Reinig, Gert-Jan de Vreeds, 2006) In figure 3 the relations between the NPGA, SP and SO are shown. The NPGA has a positive effect on both the SO and SP. According to the model, the SO also has a positive effect on the SP (though it is just a weak effect).
In this research we only evaluate the Satisfaction of the meeting outcome. (In Trip.Easy a meeting is called a session, so therefore we will use the term session rather than meeting). The satisfaction with meeting process is evaluated in

Next to the evaluation of the Satisfaction of the session outcome, we will also evaluate the recommendations of Trip.Easy during each session. We will conduct a User Satisfaction Analysis to evaluate the recommendations generated by Trip.Easy. The method is partially derived from (Liang, 2007), where they have defined questions to measure the user satisfaction with system recommendations. We have adapted the questions in order to measure the **Satisfaction of the recommendations**. The questions are presented in Appendix A.

In the following paragraph the experimental setup is explained, in which we show how we measure the **Satisfaction of the session outcome (SO)** and the **Satisfaction of the recommendations**.
5.4 Experimental setup
In the introduction of this chapter it is explained that this research is focused on (i) the evaluation of the outcome of the session and (ii) the evaluation of the recommendations (proposed cities) by Trip.Easy. In paragraph 5.2 we have defined concrete hypotheses that we will test. In paragraph 5.3 we explained the methods from which we derived the variables that we need to test the hypotheses. In this paragraph we will explain the experimental setup that is needed to retrieve the required data by which we can finally answer our main question:

*Do group members believe that the outcome established using Trip.Easy GDSS represents a satisfying outcome for the group?*

5.4.1 Protocol
The whole process of the experiment is schematically depicted in Figure 17.

![Figure 17: Schematic outline of the various steps of the experiment](image)

**Introduction:** During the introduction the participants are informed about the aim of the study and what they can expect from the experiment. Moreover they were told that they could win a city trip. Another stimulus is that they will get the possibility to join “vakantiepunten.nl” and receive “vakantiepunten”. These are points that can be exchanged into discounts on future holiday trips.

**Assign roles:** Like in a real situation someone has to suggest going on a trip. Therefore, one of the participants is assigned to the role of initiator during each experimental session. This person has to invite other participants to join the session that is setup in the Trip.Easy GDSS.

**Start session:** Once the participants have accepted the invitation, the session is started. The group has to complete the whole decision process in order to decide which city they will visit (which is the session outcome). During the decision process users are not allowed to communicate with each other, except for the ‘evaluate’ step, where the debate takes place.

**Survey:** After the participants have successfully finished the session, an individual online survey is taken to evaluate each step of the Trip.Easy GDSS. The survey exists of eight sections that are completed sequentially: four sections that evaluate the satisfaction of the recommendations provided by Trip.Easy during the process steps and one section for the evaluation of the satisfaction of the session outcome. In addition we have two sections that measure the possible acceptance (future usage) and a final section for further improvements. Each section contains an open question where the participants are asked to provide remarks. The survey can be found in Appendix A. All quantitative and qualitative data are collected by our custom developed online survey tool. Subsequently the survey tool is able to

---

6 Sponsored by a private investor (Maarten van Dijk)
7 Sponsored by vakantiepunten.nl
automatically parse and export the data to SPSS. In SPSS we can execute the statistical analyses that are needed to test our hypotheses and finally answer our main question.

5.4.2 Participants
To gain representative results we took two issues into account: (i) the selection of the participants and (ii) the sample size. The selection of the participants is vital to the experiment. Subjects should be chosen in such way that they match the expected user population as closely as possible. Our target audiences for the Trip.Easy GDSS are users who are well known with internet applications and social media. This means that the demographic diversity may vary a lot. As mentioned earlier, another important factor is that a travel party mostly consists of members that know each other, like a group of friends, colleagues or a family. Therefore participants of our experiments were asked to invite three other persons to join their travel party. To secure the comparability of the results within our experiment, each travel party had the same size of four members. We used statistics to determine the necessary size of the sample. According to (Bartlett, 1977) a sample size of minimal n=96 is representative. However, our questionnaire is based on a 7-point Likert scale. Consequently, according to (Brinkman), a minimal sample size of n= 119 is acceptable. Our sample size equals n=120 participants, based on a population of 6.400.000\(^8\), with a margin error of 10%, a confidence level of 95% and a spread of 50%. Thus, to get the necessary sample size of 120 participants, we randomly choose 30 subjects. Each subject had to 3 more participants to form a travel party of 4 members in total.

5.4.3 Equipment and setting
The experiment took place at the office of Studio Kenneth&Koh. The room can host 1 group at the time. The duration of the test was approximately 40 minutes.

![Test room arrangement](image)

Figure 18: Test room arrangement
To conduct the research we needed 4 computers. Each computer has the same configuration and LCD monitor with a screen resolution of 1280x1024. Internet and a XHTML compatible browser were needed.

---

\(^8\) Volume and intensity of short holidays of the Dutch (length 2-8 days) in 2008: CBS Toerisme en recreatie in cijfers 2008
to run the Trip.Easy GDSS application. For this experiment, the participants used the Google Chrome browser.

5.4.4 Pilot test
Pilot tests were performed to evaluate and fine-tune the experimental protocol. Next we also wanted to make sure that there were no critical malfunctions or bugs in the system. In this section the process, findings and results of this pilot test are briefly described.

Sixteen subjects have participated in the pilot test. We selected 4 participants and asked them to bring 3 more friends to join in order to form a travel party of 4 participants each. These 4 groups did not participate in the real experiment in order to prevent biases. The subjects went through the experimental protocol and filled out the questionnaire. The pilot test indicated that several modifications to the Trip.Easy GDSS were needed, due to malfunctioning or performance problems. For instance wrong data was presented in the GUI during several steps in the decision process. Furthermore the execution of the preference aggregation component failed when multiple groups tried to execute it at the same time. This was caused by the limited computational power of our test machines. Another issue was that participants were communicating with each other throughout the session. As a result the duration of the experiment became unacceptably long. Moreover it caused the participants to influence each other’s preferences, which is undesirable with this experiment.

![Figure 19: Pilot test](image)

Based on the findings of the pilot tests, we resolved the technical issues and bugs and fine-tuned experimental instruments, such as improving the readability of the questionnaire. Additionally, we enhanced the ways of organizing the experiment by providing an instruction to our participants. For instance we explicitly set the rule that participants are not allowed to speak during the session except for the ‘evaluation’ step. The duration of the experiment was set to a 40 minutes maximum for the session with the Trip.Easy GDSS and approximately 30 minutes for the survey. After the adjustments of the system, the questionnaire and the experiment protocol we were ready for conducting the real experiment.
5.4.5 Reliability & Validity

Reliability analyses were conducted for the self-reported measures regarding the satisfaction with the session outcome and the satisfaction with the proposed cities. For the validation of the satisfaction with the session outcome, an expert on group decision-making measurement conducted an initial review to establish face validity. Furthermore, this instrument is derived from the Goal Attainment Model. The reliability and validity of this instrument has been shown in (Robert O. Briggs, Bruce A. Reinig, Gert-Jan de Vreeds, 2006). Cronbach’s $\alpha$ coefficients were calculated. In view of the exploratory nature of this research, a cut-off value of 0.70 was considered acceptable (Field, 2009). The reliability analyses indicated that the measurements of the satisfaction of the session outcome are reliable. (Cronbach’s $\alpha = .935$). To determine the reliability of the satisfaction of the recommendations, Cronbach’s $\alpha$ was calculated for each step in the decision-making process:

- ‘generate’ step: Cronbach’s $\alpha = .79$
- ‘reduce’ step: Cronbach’s $\alpha = .84$
- ‘clarify’ step: Cronbach’s $\alpha = .76$
- ‘evaluate’ step: Cronbach’s $\alpha = .82$

The instrument for measuring the satisfaction associated with the proposed cities is derived from (Liang, 2007). However, an expert on artificial intelligence has conducted an initial review to establish face-validity.
5.5 Results

Here we will discuss the results. In the first part we focus on the satisfaction of the outcome. Next we will treat the results of the evaluation of the recommendations during the process steps. Finally we will discuss some extra observations. In Chapter 6 we present a conclusion based on these results.

We have conducted the experiment with 120 users divided over 30 groups, where each group exists of 4 participants. The whole group of participants exists of 72 males and 48 females, between the ages of 18 and 55 years old ($M = 28, SD = 7.75$).

5.5.1 Satisfaction of the session outcome

As explained earlier, the satisfaction of the session outcome indicates whether users are satisfied with the end result of their session facilitated by Trip.Easy. The goal is to find an answer to the question:

*Are users satisfied with the session outcome?*

This question gives rise to hypothesis H. 1. In order to find out whether we can accept hypothesis H. 1, we perform a ‘one sample t-test’. We have set the test value at 5.0, which is the break-even point between ‘satisfied’ and ‘somewhat satisfied’. The results of the ‘one sample t-test’ show that on average, participants perceived a significantly greater satisfaction with the session outcome $M_{SO} = 6.0$ ($SE_{SO} = 0.74$) than the break-even point of 5.0 ($t(115) = 7.27, p < .05$). Thus we accept hypothesis H. 1:

**Group members find the session outcome satisfying.**

The results show that users are satisfied with the final decision. This can be seen as a first indication that Trip.Easy does lead the group of users to a satisfying outcome. In support of our hypothesis, the user comments indicate that the outcome of the session in most cases reflects the preferences of all users. Some users explicitly said that their interests were taken into account. According to (Reinig, 2001) this causes users to have a positive attitude towards the final decision. Some users commented that they had totally different preferences than the others in his or her group. This lead to an outcome that was somewhat less satisfying for these users.

5.5.2 Dissimilarity of preferences versus the satisfaction of the session outcome

In this section we discuss the effect of dissimilarity of preferences on the satisfaction of the session outcome. To this end we compare:

1. groups whose individuals have relatively more similar preferences, and
2. groups whose individuals have relatively more dissimilar preferences.

Before we are able to answer this question, we need to know how to determine the level of conflict between preferences. As explained earlier, the users have inserted their preference model by ranking the city features in the ‘generate’ step. Next Trip.Easy has created a preference model for each user that is based on ranked city features. A measure for similarity is represented by Spearman’s Distance (Li,
Therefore we have calculated the Spearman’s Distance from each user to every other user of that group (Figure 20).

Figure 20: distances between group 4 members

Finally we calculated the total Spearman’s distance of the group by summing up all the distances.

\[ d_{S_{\text{total}}} = dS(a,b) + dS(c,b) + dS(d,c) + dS(a,d) + dS(a,c) + dS(b,d) \]

Next we divided \( d_{S_{\text{total}}} \) by the amount of group members, which are 4 for each group in our case. As a result, we have for each group the average Spearman’s distance for each user to all other members of his or her group. In short we call this \( \overline{dS}_{\text{user}} \). The higher \( \overline{dS}_{\text{user}} \) the more conflict exists between users’ preferences in a group. Next we calculate the overall mean of \( \overline{dS}_{\text{user}} \) with respect to the whole dataset. We call this \( M_{\text{distance}} \). All groups that have a \( \overline{dS}_{\text{user}} \) greater then \( M_{\text{distance}} \), form one segment which we call segment A. All users that have a \( \overline{dS}_{\text{user}} \) smaller than \( M_{\text{distance}} \), will form another segment which we call segment B. Segment A consists of groups that have relatively more similar preferences, while segment B exists of groups that have relatively more conflicting preferences. Eventually we want to compare segment A and B with respect to the satisfaction of the session outcome (SO).

To find out whether the similarity of the preferences within a group influences the SO, we will perform an ‘independent t-test’ and look at the overall correlation between \( \overline{dS}_{\text{user}} \) and SO. The ‘independent t-test’ tests the difference between the average value of SO of segment A with that of segment B. The segment type (A or B) is an independent variable. Before the independent t-test can be conducted, certain assumptions concerning the data (cf. sampling distribution) have to be met.

---

9 Another measure for group conflict is: the maximum of the Spearman’s distance per group. Using this measure results in the same distribution of segment A and B.
5.5.2.1 Assumptions of independent t-test

To perform the ‘independent t-test’ we need to meet the following assumptions; normality, independent scores and homogeneity of variances. Below we describe these assumptions.

Normality: According to (Field, 2009) we have to make sure that the data is normally distributed by performing several normality checks:

- Shapiro-Wilk’s test
- Histograms
- QQ-Plots
- Skewness and kurtosis of the data

In Appendix D the results of these tests are presented. Shapiro – Wilk’s test is significant for both SO and $dS_{user}$, which indicates that the scores are significantly different from a normal distribution. However, because we are dealing with a large sample set (N = 116), this test can be significant even when the scores are only slightly different from a normal distribution. Thus we will need to conduct other tests to check the normality like histograms, Q-Q plots and the values of skewness and kurtosis. (Field, 2009)

The histograms of both SO and $dS_{user}$ (Appendix E) show that the scores are not very different from a normal distribution, however the distribution seems to look a bit jazzy. Moreover, the Q-Q plots of both SO and average distance look like a straight line, which indicates normality. The values of skewness for SO and $dS_{user}$ are both close to zero (-.312 respectively -.102), which means that the distribution is approximately normal. The values of kurtosis for SO and $dS_{user}$ lie both between 0 and -1 (-.579 respectively -.880), which indicates a slightly flat and light-tailed distribution. However, the values are still between 0 and 1 and thus we can conclude that the deviation from normality is too small to be of any importance. Finally, the fact that we are dealing with a relatively large sample (size is 100+) according to (Field, 2009) allows us to assume that the sample distribution is normal.

Independent scores: The scores of this test are independent, since each score is from a unique participant.

Homogeneity of variances: To find out whether the variances in the different data segments (A and B) are different, we use Levene’s test (Field, 2009). Levene’s test evaluates the null hypothesis that the variances in the two groups are equal (i.e. the difference between the variances is zero). Therefore, if Levene’s test is significant ($p \leq .05$), the assumption of homogeneity has been violated. If Levene’s test is non–significant ($p > .05$) we can assume homogeneity of variances. For these data, Levene’s test is non-significant ($p = .186$). Thus we can conclude that the assumption of homogeneity of variances is met.

5.5.2.2 The independent t-test

As explained earlier we want to find out whether the similarity of preferences between group members has any influence on the satisfaction of the session outcome (SO). Therefore we have divided the dataset into two segments. The first segment considers groups with a relatively high similarity in preferences (segment A), while the second segment considers groups with a relatively low similarity in
preferences (segment B). This segmentation is based on the \( dS_{user} \) of each user. The mean \( M_{distance} \) of the dataset was used as a threshold to distinguish the two segments.

The results of the ‘independent t-test’ show that the segment condition has a significant effect on the SO \( (t(114) = 2.10, p < .05) \). On average, participants from segment A show a higher SO \( (M_{SO} = 6.15, SE_{SO} = .09) \) than participants from segment B \( (M_{SO} = 5.86, SE_{SO} = .11) \). However, the effect is small-sized \( (r = .19) \). Thus this significant result is, according to (Field, 2009), less meaningful.

### 5.5.2.3 Correlation between SO and \( dS_{user} \)

To determine the correlation between the SO and \( dS_{user} \), the Pearson correlation coefficient is calculated. The usage of Pearson’s correlation requires that the assumption of a normally distributed sampling data is met. In section 5.5.2.1 it is already shown that the data is normal. Spearman’s correlation coefficient between the two variables SO and \( dS_{user} \) is -.171. This implies a negative relation between SO and \( dS_{user} \). However, the correlation between SO and \( dS_{user} \) is non-significant \( (p = .066) \).

### 5.5.2.4 Conclusion

In this section we want to evaluate the hypothesis that Trip.Easy leads users toward a satisfactory outcome, despite the dissimilarity of preferences within a group. Both analyses that have been conducted support hypothesis 2. Firstly, the results from the independent t-test show a small-sized and thus almost meaningless significant effect. Secondly, there is a non-significant correlation between the SO and \( dS_{user} \). In other words, a higher \( dS_{user} \) (thus relatively more conflicting preferences among group members) does not result into a significant lower SO. Accordingly, H. 2 can be accepted.

**Trip.Easy leads users to a satisfactory outcome, despite the dissimilarity of preferences within a group.**

### 5.5.3 Evaluation of the recommendations

In this paragraph we describe the observations concerning the satisfaction of the recommendations of the system during each process step. We start with the ‘generate’ step, followed by the ‘reduce’ and ‘clarify’ step and finally the ‘evaluate’ step. As explained before, during each step we have measured the satisfaction level about the following variables:

- Satisfaction of the proposed cities (recommendations) (SATC)
- Satisfaction of the amount of proposed cities (AMOC)

#### 5.5.3.1 Normality

To be able to carry out statistical analysis on the data we have to make sure the data is normal. Therefore we conducted several normality tests:

- Shapiro-Wilk’s test
- Histograms
- QQ-Plots

In Appendix J the results of these tests are presented. Shapiro – Wilk’s test is significant for both SATC and AMOC in every step, which indicates that the scores are significantly different from a normal
distribution. However, because we are dealing with a large sample set (N = 116), this test can be significant even when the scores are only slightly different from a normal distribution. When we observe the QQ-Plots we notice that it shows jazzy behavior. However, the histograms show that the data is negatively skewed. There are some negative outliers that are responsible for this. Without these outliers the data is much closer to the normal distribution. Moreover, the negative skewness leads to a conservative representation of the data, thus the real data should be more positive in fact.

Subsequently, according to (Field, 2009) a sample size greater than 100 is assumed to be normal. The sample size of the SATC and AMOC are both larger than 100. Consequently, we assume that the data of SATC and AMOC is normal.

5.5.3.2 Generate
The ‘generate’ step stands for collecting and aggregating all individual preferences over the general city features defined by the city model. Consequently, a group preference model is established. Based on this group preference model, Trip.Easy proposes a list of cities that should be a best-fit to the group. As a result, the outcome of this step will not be optimal for each individual. Therefore we expect the participants to be ‘somewhat satisfied’ at least, as reflected by hypothesis H. 1. This hypothesis states that users perceive the cities presented during the ‘generate’ step as a reasonable first proposal.

In order to find out whether we can accept hypothesis H. 1 we perform a ‘one sample t-test’. We use a test-value of 4.5, which is the break-even point between ‘somewhat satisfied’ with the proposed cities and not ‘satisfied’ with the proposed cities. The results of the ‘one sample t-test’ show that on average, the SATC ($M_{SATC} = 5.02, SE_{SATC} = 1.16$) is significantly greater than the test-value of 4.5 ($t(115) = 4.78, p < .05$). Based on these results we can accept H. 3: Users perceive the cities presented during the ‘generate’ step as a reasonable first proposal.

The acceptance of H. 3 means that users perceive the cities presented by Trip.Easy as a reasonable first proposal. This has everything to do with the fact that Trip.Easy does its recommendations based on the group preference model. This model represents the preferences of the whole group and is therefore by definition not optimal for an individual user. This is confirmed by the user comments concerning this step. For instance, some users expected different cities to be proposed by Trip.Easy. However, they also indicate that Trip.Easy presented enough interesting alternatives, which explains that the participants are on average ‘somewhat satisfied’ and not ‘dissatisfied’.

As explained above, Trip.Easy proposes a list of cities that should be a best-fit to the group. Subsequently, the participants are able to select their most favorite cities. This is done by ranking the cities which is part of the ‘reduce’ step. However, a sufficient amount of cities is needed to be able to make a decent selection. This is an important factor of group decision processes, as explained earlier. In order to find out whether we can accept hypothesis H. 4, which indicates that users are satisfied with the amount of cities proposed during the ‘generate’ step, we perform a ‘one sample t-test’. We use a test-value of 5.0, which is the break-even point between ‘not satisfied’ with the amount of the proposed cities and ‘satisfied’ with the amount of proposed cities. The results of the ‘one sample t-test’ show that
on average, participants have a significantly greater score on the AMOC with $M_{AMOC} = 5.47 \ SE_{AMOC} = 1.19$) than the break-even point of 5.0 ($t (115) =4.29, p < .05$). This implies that we can accept H. 4:

**Users are satisfied with the amount of cities that are proposed during the ‘generate’ step.**

The acceptance of H. 4 means that Trip.Easy generates enough recommendations (cities) for users, in order to make a selection. This is partially confirmed by the user comments. Several users indicated that they appreciated the amount of cities presented and that they were sufficient. However, some users also indicated that there were too much cities presented.

5.5.3.3 Reduce

In the ‘reduce’ step the users point out their most preferred cities from the list of cities that have been proposed by Trip.Easy as a result of the ‘generate’ step. This is done by ranking the cities. Trip.Easy then selects the most preferred set of cities from each user. Next the selected sets are aggregated to a set that is considered to be the most preferred by the group. In order to find out whether we can accept hypothesis H. 1, which indicates that users are satisfied with the proposed cities during the ‘reduce’ step, we perform a ‘one sample t-test’. We use a test-value of 5.0, which is the break-even point between ‘not satisfied’ with the proposed cities and ‘satisfied’ with the proposed cities. The results of the ‘one sample t-test’ show that on average, participants have a significantly greater score on the SATC with $M_{SATC} = 5.62 \ SE_{SATC} = 1.10$) than the break-even point of 5.0 ($t (115) =6.08, p < .05$). Based on these results we can accept H. 5:

**Users are satisfied with the cities that are proposed during the ‘reduce’ step.**

The acceptance of H. 5 implies that users are satisfied with the cities that were proposed during the ‘reduce’ step. Some users commented that they were a bit disappointed about the fact that their most preferred city did not appear on the proposed list of cities that Trip.Easy eventually generated. This occurred, due to the fact that the city was less popular with other group members.

Subsequently, we are interested in whether the amount of cities proposed by Trip.Easy (AMOC) is sufficient, as reflected in H. 6. With the AMOC we will evaluate hypothesis H. 6. In order to find out whether we can accept this hypothesis we perform a ‘one sample t-test’ with a test-value of 5.0. This value is the break-even point between ‘not satisfied’ with the amount of the proposed cities and ‘satisfied’ with the amount of proposed cities. The results of the ‘one sample t-test’ show that on average, participants have a significantly higher score on the AMOC with $M_{AMOC} = 5.62 \ SE_{AMOC} = 1.09$) than the break-even point of 5.0 ($t (115) =6.16, p < .05$). Based on these results we can accept H. 6:

**Users are satisfied with the amount of cities being proposed during the ‘reduce’ step.**

The acceptance of H. 6 indicates that users find the amount of cities, generated during the ‘reduce’ step, sufficient. However some user comments indicated that users wanted more proposed cities during this step. These users pointed out that the transition from 20 cities in the ‘generate’ step, to just 6 cities in this step was too large. Others appreciated this reduction and indicated that it simplified the decision process.
5.5.3.4  Clarify
During the ‘clarify’ step the users can rate the cities that were proposed by Trip.Easy as a result of the ‘reduce’ step. In addition users can share comments related to these cities (i.e. opinions about cities, motivations to visit a certain city, etc.) Next, Trip.Easy selects a subset of cities that contain the top-rated cities from each individual user. Finally, Trip.Easy establishes an aggregation of the cities that can be considered as a best-fit to the group. In order to find out whether we can accept hypothesis H. 1, which indicates that users are satisfied with the proposed cities during the ‘clarify’ step, we perform a ‘one sample t-test’. We use a test-value of 5.0, which is the break-even point between ‘not satisfied’ with the proposed cities and ‘satisfied’ with the proposed cities. The results of the ‘one sample t-test’ show that on average, participants have a significantly higher score on the SATC with $M_{SATC} = 5.47$ ($SE_{SATC} = 1.20$) than the break-even point of 5.0 ($t (115) = 7.27, p < .05$). This implies that the users are ‘satisfied’ with the proposed cities. Thus we can accept H. 7:

Users are satisfied with the cities that are proposed during the ‘clarify’ step.

The acceptance of H. 7 indicates that users are satisfied with the proposed cities in the ‘clarify’ step. The user comments confirm this result. Several users point out that the selected cities are acceptable for all group members. Secondly, they mentioned that the rating values and the possibility to share comments (opinions, motivations, etc) help the group in making the final choice (which is done in the next step).

Furthermore we are interested in the satisfaction of the amount of proposed cities (AMOC). Therefore we measure the AMOC to evaluate H. 8, which indicates that users are satisfied with the amount of proposed cities during the ‘clarify’ step. In order to find out whether we can accept hypothesis H. 1 we perform a ‘one sample t-test’ with a test-value of 5.0. This value is the break-even point between ‘not satisfied’ with the amount of the proposed cities and ‘satisfied’ with the amount of proposed cities. The results of the ‘one sample t-test’ show that on average, participants have a significantly higher score on the AMOC with $M_{AMOC} = 5.42$ ($SE_{AMOC} = 1.22$) than the break-even point of 5.0 ($t (115) = 3.72, p < .05$). This implies that we can accept H. 8:

Users are satisfied with the amount of cities being proposed during the ‘clarify’ step.

We conclude that the users find the amount of the proposed cities is sufficient to make a final decision. The final decision will be made during the ‘evaluation’ step (see next section). During this step users will have a debate in order to make the final decision.

5.5.3.5  Evaluation
The system has aggregated the ratings of each user resulting in the top four of highest ratings of the group. During the ‘evaluate’ step the users have a debate with each other about these 4 cities in order to make their final decision. At this moment the users make their final choice and thus establish a consensus. Therefore we measure the SATC to evaluate H. 9, which indicates that users are satisfied with the final city chosen during the ‘evaluate’ step.
In order to find out whether we can accept hypothesis H. 9, we perform a ‘one sample t-test’ with a test-value of 5.0. This value is the break-even point between ‘not satisfied’ with the proposed cities and ‘satisfied’ with the proposed cities. The results of the ‘one sample t-test’ show that on average, participants score significantly higher on the SATC ($M_{\text{SATC}}= 6.034$, $SE_{\text{SATC}}= 0.67$) than the break-even point of 5.0 ($t (115) =14.95, p < .05$). This implies that users are ‘satisfied’ with the final proposed city (cf. the final outcome of the session). Consequently we can accept H. 9:

**Users are satisfied with the final city that is chosen during the ‘evaluate’ step.**

All the ‘one sample t-test’ results from all the steps can be found in Appendix F.

5.5.3.6 **Summary**

The previous sections discussed the results of (i) the satisfaction of the proposed cities (recommendations) (SATC) and (ii) the satisfaction of the amount of proposed cities (AMOC). The variables we measured during each process step.

Figure 21 presents the scores of the overall satisfaction of the proposed cities (SATC) for each process step. The result show that the overall SATC score is: $5 \leq M_{\text{SATC}} \leq 6$, which is between ‘somewhat satisfied’ and ‘satisfied’. Furthermore, Figure 21 shows a positive trend of SATC with respect to the process steps, though with a small decrease during the transition from the ‘reduce’ to the ‘clarify’ step.

![Figure 21: overall score of Satisfaction of proposed cities (recommendations)](image-url)
Figure 22 shows the overall results of the average satisfaction of the amount of proposed cities (AMOC) for each process step. The result show that the average AMOC is: \(5.4 \leq M_{AMOC} \leq 5.6\), which is between ‘somewhat satisfied’ and ‘satisfied’. As we can see the scores of \(M_{AMOC}\) vary little throughout the decision process, thus it is relatively stable.

![Graph showing satisfaction of the amount of proposed cities](image)

**Figure 22: the overall score of the satisfaction of the amount of the proposed cities**

**5.5.3.7 Satisfaction with the recommendations during the process**

Figure 21 shows that the SATC score increases as the decision-making process proceeds. In this section we want to find out whether this concerns a significant increase of the SATC.

In order to test H. 10, we use the statistical analysis technique ‘one-way repeated measures ANOVA’. According to (Field, 2009), this analysis technique is used when there is one quantitative dependent variable and one qualitative within subject factor. The within subject factor is a variable that is repeatedly measured from the same participants. In our case, the process step (‘generate’, ‘reduce’, ‘clarify’ and ‘evaluate’) is the within subject factor and the SATC is the quantitative dependent variable. Due to the fact that the observations of the SATC in each process step are not independent, we need to test the data on the assumption of sphericity. Sphericity means that the data from the different measurements of each process step are univariate. If the data appears to be spherical, than a univariate approach is chosen for the calculation of the F-ratio. Otherwise, a multivariate approach is taken. To test the data for sphericity, we use Mauchly’s test (Field, 2009). Mauchly’s test shows that the assumption of sphericity has been violated \(X^2(2) = 22.25, p < .05\). Therefore a corrected F-value, namely, Green-House Geisser \((\varepsilon = .947)\) must be applied to the analysis. Finally we tested the data for normality and the results showed that we can assume that the data is normal.

The results of the ‘one-way repeated measures ANOVA’ show that the process step condition has a significant influence on the SATC \((F(1.90, 217.85) = 11.67, p < .05)\). The partial \(\eta^2\) is 0.09, which means that the effect is ‘medium’ sized. The results of the ANOVA are shown in Table 2.
Table 2: ANOVA table for process step vs. SATC

<table>
<thead>
<tr>
<th>Source</th>
<th>Df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
<th>partial eta²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>2.70</td>
<td>61.40</td>
<td>22.89</td>
<td>23.95</td>
<td>.000</td>
<td>0.17</td>
</tr>
<tr>
<td>Residue</td>
<td>308.70</td>
<td>295.00</td>
<td>.96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>315.70</td>
<td>356.40</td>
<td>23.85</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results of the ANOVA indicate that the process step significantly affects the SATC. However, this result does not tell us which differences are significant. Therefore, a ‘repeated contrast’ test (Field, 2009) has been conducted. This analysis is based on the comparison between ‘generate’ and ‘reduce’, ‘reduce’ and ‘clarify’ and finally ‘clarify’ and ‘evaluate’.

The results of the ‘repeated contrast’ test show a significant difference between the ‘generate’ and ‘reduce’ step with respect to the SATC ($F(1,115) = 22.11, p < .001$). The participant appointed a higher score to the ‘reduce’ step compared to the ‘generate’ step (respectively $M_{REDUCE} = 5.62$ and $M_{GENERATE} = 5.02$). This effect is ‘medium’ sized (partial eta² = 0.17). The difference in the SATC scores between the ‘reduce’ and the ‘clarify’ step is non-significant ($F(1,115) = 1.77, ns$). Finally, the results show a ‘medium’ significant effect, when the ‘clarify’ and ‘evaluate’ step are compared with respect to the SATC ($F(1,115) = 21.33, p < .001$, partial eta² = .17). The participants appointed a higher score to the SATC during the ‘evaluate’ step than during the ‘clarify’ step (respectively $M_{EVALUATE} = 6.034$ and $M_{CLARIFY} = 5.5$). Based on these results we can conclude that the SATC score significantly increases as the session proceeds, apart from the non-significant transition from the ‘reduce’ to the ‘clarify’ step. Overall, the process thus converges towards a satisfactory outcome. Consequently we can accept H. 10:

Trip.Easy leads the group towards a satisfying outcome.
6 Discussion & Conclusion

During this thesis project we have designed, developed and evaluated a new decision support system for groups of travel consumers. First we realized the design and development of the intelligent recommendation mechanism. These recommendations are intended to support the group to reach a satisfying outcome.

Secondly we tested the system for user acceptance, for which we conducted user evaluation experiments. These experiments were focused on the evaluation of the outcome of the session and the evaluation of the recommendations during the session. Therefore we formulated certain research questions that are answered in the following sections. Eventually we will answer the main research question in the last paragraph.

6.1 The evaluation of the outcome of the session

For the evaluation of the outcome of the session we observed two aspects; (i) the satisfaction of the session outcome (SO) and (ii) whether the dissimilarity of individual preferences has any influence on the satisfaction level of the outcome. These aspects are respectively measured to answer following research questions:

Are users satisfied with the outcome?

The results showed that hypothesis H. 1 can be accepted; therefore we conclude that users are satisfied with the outcome. This means that the final decision, the city they choose to visit, is satisfying. This is an indication that Trip.Easy leads groups towards a satisfying outcome and thus, according to (Reinig, 2001), is a first indication that users will accept this technology, which is promising for future use of Trip.Easy.

Is there a qualitative difference in outcome between groups with members that have similar preference and groups with members that have less similar preferences?

The results showed that hypothesis H. 2 is confirmed, thus we can conclude that Trip.Easy will lead users to a satisfying outcome, despite any dissimilarity of preferences. This aspect is an indication of the effectiveness of the recommendation mechanism. The first reason can be that the recommendation mechanism takes all users’ preferences into account, by generating a group preference model. When a group exists of group members with relatively more dissimilar preferences, than the generated group preference model also becomes less similar to the individual preference models of the group members. Therefore the recommendation mechanism proposes 20 cities in the first step. The fact that users can choose from this amount of options, there is always an option that satisfies the group members. This, in combination with a well-structured decision process, in which the recommendations are adapted and fine-tuned based on user feedback, increases the robustness of Trip.Easy with respect to dissimilar preferences.
6.2 The evaluation of the recommendations by Trip.Easy

The evaluation of the recommendations during the session is reflected by the research question: *Do group members believe that the recommendations of Trip.Easy support the group adequately to arrive at a satisfying outcome?* To answer this question we observed the following aspects; (i) satisfaction of the recommendations, (ii) the development of the satisfaction of the recommendations throughout the decision-process and (iii) the satisfaction of the amount of proposed cities. These aspects were measured to answer the following questions:

*Are users satisfied with the proposed cities during each process step?*

The results showed that hypotheses H. 3, H. 5, H. 7 and H. 9 can be accepted and thus users are satisfied with the proposed cities during each process step. This indicates that the recommendations generated by Trip.Easy do match the preferences of the users of the group. Thus the Trip.Easy is able to generate recommendations that fit the whole group. This is an indication that the recommendation mechanism performs well.

*Does the user satisfaction with the proposed cities increase during the whole process?*

The results showed that hypothesis H. 10 can be accepted and thus there is a significant increase of the user satisfaction with the proposed cities during the whole process. This is an indication that the recommendations contribute in supporting the user towards a satisfying outcome. During the process users build consensus, which is supported by the fine-tuning and adaptation of the recommendations, based on user feedback. This can lead to an increase of the satisfaction with the recommendations as the session proceeds.

*Are users satisfied with the amount of proposed cities in each process step?*

The results showed that hypotheses H. 4, H. 6 and H. 8 can be accepted and thus users are satisfied with the proposed cities during each process step. This means that there were enough options for the users to choose from and indicates that it had no influence on the overall satisfaction level. Earlier we explained that too many options can lead to an information overload, however too few options can lead to unsatisfactory outcomes, because users find that they do not have enough options to choose from.

Based on these results we can conclude that the recommendations of Trip.Easy support the group adequately to arrive at a satisfying outcome.
6.3 Conclusion

Based on the answers of the questions belonging to the (i) the evaluation of the outcome of the session (paragraph 6.1) and the questions belonging to (ii) the evaluation of the recommendations (paragraph 6.2) we are able to answer the main question of this research:

*Do group members believe that the outcome established using Trip.Easy GDSS represents a satisfying outcome for the group?*

Based on the results of this experiment we can conclude that group members believe that the outcome established by using Trip.Easy GDSS represents an acceptable outcome for the group.
7 Problems & Limitations

7.1 Research
The results of this research are hard to compare with the related work explained in paragraph 2.2. In those studies dedicated methods and models were used to measure the satisfaction of the user for instance. These studies were often explorative. In this experiment we used two methods, one based on the Goals Attainment Model, which is a general model by which group decision support systems can be evaluated.

7.2 Experimental
Within this experiment, the participants were instructed to simulate the situation that they wanted to plan a trip together. Thus the outcome did not have any consequences. This could be regarded as a weakness of this experiment; however we have no reason to believe that it had any influence. We noticed that the participants took it very seriously.

7.3 Technical
The experiments were conducted group by group. The architecture of the system did not allow multiple group sessions. Concurrency problems occur when multiple groups make use of the system at the same time. The amount of members of a group however does not have a significant effect on the system performance.

Another issue is that the aggregation algorithm can only be applied to a preference model with a limited size. The preference model is based on the domain model. A large domain model will lead to unacceptable computation-time that is needed to find an optimal group preference model. This already happens when the amount of city features per category, becomes larger than 9. This can be resolved by a different algorithm design, for instance genetic algorithm designs, as we explained in chapter 3. However, for our experiments this was not an issue.
8 Future Work

8.1 Experiment
Another interesting setup to evaluate Trip.Easy is to use real-world data powered by a travel organization and real travel consumers. The real satisfaction of the outcome actually is the satisfaction one measures after the trip took place. The results can be compared with current satisfaction studies of tourists.

8.2 Model
The evaluation of the user satisfaction with the session outcome is based on Reinig’s Goal Attainment Model (Reinig, 2001). In 5.3 we explained that this is a model for evaluation of group decision support systems. It evaluates the decision process and the outcome separately. In this thesis we only focused on the session outcome. We also indicated that the outcome is strongly influenced by the recommendations generated by Trip.Easy. These recommendations were evaluated by conducting a User Satisfaction Analysis. In future research it is interesting to integrate the evaluation of the recommendations with the evaluation of the outcome. The analysis then can give a better understanding and explanation of the results. This might improve the analysis of intelligent group decision support systems.

8.3 Technical
The recommendation mechanism needs to be more scalable to be able to handle a more complex domain model. This is necessary to be able to use real-world data from for instance a travel organization.

Another aspect is to design and implement an aggregation procedure that is able to handle these large and complex domain models. The current implemented algorithms will not be able to operate within a reasonable period of time. Therefore further research is necessary to come up with better solutions.
Bibliography


Appendix A

Questionnaire

<table>
<thead>
<tr>
<th>Likert value</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Zeer mee oneens</td>
</tr>
<tr>
<td>2</td>
<td>Mee oneens</td>
</tr>
<tr>
<td>3</td>
<td>Iets mee oneens</td>
</tr>
<tr>
<td>4</td>
<td>Neutraal</td>
</tr>
<tr>
<td>5</td>
<td>Iets mee eens</td>
</tr>
<tr>
<td>6</td>
<td>Mee eens</td>
</tr>
<tr>
<td>7</td>
<td>Zeer mee eens</td>
</tr>
</tbody>
</table>

Gathering

<table>
<thead>
<tr>
<th>Stelling</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>Q2</td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>Q3</td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>Q4</td>
<td>1 2 3 4 5 6 7</td>
</tr>
</tbody>
</table>
**Reduce**

<table>
<thead>
<tr>
<th>Stelling</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q5  Ik ben tevreden met de steden die na het ordenen zijn overgebleven</td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>Q6  De steden die gepresenteerd zijn door het systeem komen overeen met mijn verwachtingen</td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>Q7  Ik vind dat er voor het maken van een keuze voldoende steden overgebleven zijn</td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>Q8  Ik heb het gevoel dat de steden aansluiten bij mijn opgegeven ordening</td>
<td>1 2 3 4 5 6 7</td>
</tr>
</tbody>
</table>

**Clarify**

<table>
<thead>
<tr>
<th>Stelling</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q9  Ik kan me goed vinden in de steden die overgebleven zijn</td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>Q10 Ik vind dat er voldoende steden overgebleven zijn om over te debatteren</td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>Q11 Ik heb het gevoel dat de steden aansluiten bij mijn opgegeven waardeing (rating)</td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>Q12 De steden die overgebleven zijn komen overeen met mijn verwachtingen</td>
<td>1 2 3 4 5 6 7</td>
</tr>
</tbody>
</table>
Evaluate

<table>
<thead>
<tr>
<th>Stelling</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q13 Ik kan me goed vinden in de steden die overgebleven zijn</td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>Q14 De steden die overgebleven zijn komen overeen met mijn verwachtingen</td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>Q15 Ik heb het gevoel dat de steden aansluiten bij mijn belangen</td>
<td>1 2 3 4 5 6 7</td>
</tr>
</tbody>
</table>

Satisfaction with meeting outcome

<table>
<thead>
<tr>
<th>Stelling</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q16 Ik vond de uitkomst van de sessie van vandaag goed</td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>Q17 Ik ben tevreden over de uiteindelijke stad die we gaan bezoeken</td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>Q18 Ik vind dat het GDSS voldoende rekening heeft gehouden met elk van onze preferenties</td>
<td>1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>Q19 Ik ben tevreden met de stad die mijn groepje als bestemming heeft gekozen</td>
<td>1 2 3 4 5 6 7</td>
</tr>
</tbody>
</table>
### Appendix B

#### Independent t-test results

#### Group Statistics

<table>
<thead>
<tr>
<th>AvDistanceGroup</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>SO low distance</td>
<td>60</td>
<td>6.1458</td>
<td>.66748</td>
<td>.08617</td>
</tr>
<tr>
<td>high distance</td>
<td>56</td>
<td>5.8571</td>
<td>.80884</td>
<td>.10809</td>
</tr>
</tbody>
</table>

#### Independent Samples Test

<table>
<thead>
<tr>
<th>SO</th>
<th>Levene's Test for Equality of Variances</th>
<th>t-test for Equality of Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>Sig.</td>
</tr>
<tr>
<td>Equal variances</td>
<td>1.772</td>
<td>.186</td>
</tr>
<tr>
<td>assumed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equal variances</td>
<td>2.088</td>
<td>.039</td>
</tr>
<tr>
<td>not assumed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Correlation analysis between the Satisfaction of the session outcome (SO) and the average Spearman Distance within a group (averageDistance)

### Correlations

<table>
<thead>
<tr>
<th></th>
<th>averageDistance</th>
<th>SO</th>
</tr>
</thead>
<tbody>
<tr>
<td>averageDistance</td>
<td>Pearson Correlation</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>116</td>
</tr>
<tr>
<td>SO</td>
<td>Pearson Correlation</td>
<td>-.171</td>
</tr>
<tr>
<td></td>
<td>Sig. (2-tailed)</td>
<td>.066</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>116</td>
</tr>
</tbody>
</table>
## Appendix D

### Normality

Satisfaction with the session outcome (SO) and the average Spearman distance (averageDistance)

### Tests of Normality

<table>
<thead>
<tr>
<th></th>
<th>Kolmogorov-Smirnov&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Shapiro-Wilk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistic</td>
<td>Df</td>
</tr>
<tr>
<td>SOII</td>
<td>.176</td>
<td>116</td>
</tr>
<tr>
<td>averageDistance</td>
<td>.104</td>
<td>116</td>
</tr>
</tbody>
</table>

<sup>a</sup> Lilliefors Significance Correction
Histograms and QQ-plots

Skewness and Kurtosis

Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Skewness Statistic</th>
<th>Skewness Std. Error</th>
<th>Kurtosis Statistic</th>
<th>Kurtosis Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>SO</td>
<td>116</td>
<td>-.312</td>
<td>.225</td>
<td>-.579</td>
<td>.446</td>
</tr>
<tr>
<td>averageDistance</td>
<td>116</td>
<td>-.102</td>
<td>.225</td>
<td>-.880</td>
<td>.446</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>116</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```java
private def calculateWeight(double rank, double median) {
    double weight = 0.0
    if(rank <= median) {
        weight = (double)(Math.floor((1.0/Math.sqrt(rank * 0.3)))))
    } else {
        weight = (double)(Math.floor((1.0/Math.sqrt(rank * 5.0))))
    }
    return weight
}

private def normalizePreferenceModel(PreferenceModel preferenceModel) {
    def catRawTotalValue = 0
    preferenceModel.weightedCategories.each{catRawTotalValue += it.weight}
    preferenceModel.weightedCategories.each{
        it.weight = (it.weight/catRawTotalValue)
    }
}
```

This code is responsible for the translation of the preferences into utility values. First the weights are calculated by `calculateWeight` and subsequently all the weights are normalized by `normalizePreferenceModel`. These values will later be used to match the preference model with the cities.
This code is responsible for calculating the score that Trip.Easy uses as a measure to determine how well a city matches with a preference model. The utility values are multiplied by the associated weights of the city. The city with the highest score is considered to be the best-fit to the group.
Appendix F

Results of the one sample t-tests

The Satisfaction with the session outcome (SO)

One-Sample Statistics

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>SO</td>
<td>116</td>
<td>6.0065</td>
<td>.74997</td>
<td>.06963</td>
</tr>
</tbody>
</table>

One-Sample Test

<table>
<thead>
<tr>
<th></th>
<th>T</th>
<th>Df</th>
<th>Sig. (2-tailed)</th>
<th>Mean Difference</th>
<th>95% Confidence Interval of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>SO</td>
<td>14.454</td>
<td>115</td>
<td>.000</td>
<td>1.00647</td>
<td>.8685 - 1.1444</td>
</tr>
</tbody>
</table>

The Satisfaction with the proposed cities (SATC)

One-Sample Statistics

<table>
<thead>
<tr>
<th>SATC</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>generate</td>
<td>116</td>
<td>5.02</td>
<td>1.165</td>
<td>.108</td>
</tr>
<tr>
<td>reduce</td>
<td>116</td>
<td>5.62</td>
<td>1.101</td>
<td>.102</td>
</tr>
<tr>
<td>clarify</td>
<td>116</td>
<td>5.47</td>
<td>1.197</td>
<td>.111</td>
</tr>
<tr>
<td>evaluate</td>
<td>116</td>
<td>6.03</td>
<td>.745</td>
<td>.069</td>
</tr>
</tbody>
</table>
### One-Sample Test

Test Value = 5.0

<table>
<thead>
<tr>
<th>SATC</th>
<th>T</th>
<th>Df</th>
<th>Sig. (2-tailed)</th>
<th>Mean Difference</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>generate</td>
<td>.159</td>
<td>115</td>
<td>.874</td>
<td>.017</td>
<td>-.20</td>
<td>.23</td>
</tr>
<tr>
<td>reduce</td>
<td>6.074</td>
<td>115</td>
<td>.000</td>
<td>.621</td>
<td>.42</td>
<td>.82</td>
</tr>
<tr>
<td>clarify</td>
<td>4.187</td>
<td>115</td>
<td>.000</td>
<td>.466</td>
<td>.25</td>
<td>.69</td>
</tr>
<tr>
<td>evaluate</td>
<td>14.951</td>
<td>115</td>
<td>.000</td>
<td>1.034</td>
<td>.90</td>
<td>1.17</td>
</tr>
</tbody>
</table>

95% Confidence Interval of the Difference

### The Satisfaction with the amount of the proposed cities (AMOC)

### One-Sample Statistics

<table>
<thead>
<tr>
<th>AMOC</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>generate</td>
<td>116</td>
<td>5.47</td>
<td>1.190</td>
<td>.111</td>
</tr>
<tr>
<td>reduce</td>
<td>116</td>
<td>5.62</td>
<td>1.085</td>
<td>.101</td>
</tr>
<tr>
<td>clarify</td>
<td>116</td>
<td>5.42</td>
<td>1.224</td>
<td>.114</td>
</tr>
</tbody>
</table>
## One-Sample Test

Test Value = 5.0

<table>
<thead>
<tr>
<th>T</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
<th>Mean Difference</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMOC_generate</td>
<td>4.290</td>
<td>115</td>
<td>.000</td>
<td>.474</td>
<td>.26</td>
</tr>
<tr>
<td>AMOC_reduce</td>
<td>6.163</td>
<td>115</td>
<td>.000</td>
<td>.621</td>
<td>.42</td>
</tr>
<tr>
<td>AMOC_clarify</td>
<td>3.717</td>
<td>115</td>
<td>.000</td>
<td>.422</td>
<td>.20</td>
</tr>
</tbody>
</table>
Appendix G
Field research – Interview with domain expert

Interview with:

Marjolein Visser
(Tourist) Marketing Specialist

Introduction
During our literature survey we found out that the information about tourist classification based on preferences is very limited. Though it is not very important for the group decision support system, it could be very useful for preference elicitation (predicting preferences from other preferences) and for preference aggregation (distance function and difference measuring). Therefore we decided to consult a domain specialist to find out:

- Whether there exists a preference based model.
- The current methods used.
- Advice about how to develop such a model.
  - How to conduct field research
- Where we can get feature information of travel plans.

Subject:
Tourists’ preference modeling and classification

Current situation:
Consumer analysis of tourists is very basic and general. Most organizations focus on pushing products to consumers instead of customized offers based on their preferences. There are no preference models and thus classification or profiling based on preferences is not possible at the moment. Exceptions are the niche markets where travel organizations only focus on one type of travelling. These organizations are more focused on very specific target audiences for example safari tourists or adventure tourists. They require the tourists to have a certain level of travelling experience and knowledge, so they can customize a tour by elicitation of very specific preferences.

Transition:
At the moment a process of differentiation is going on within the travel market. Specialized tour operators and travel organizations that focus on specific travel products or services, are serving an increasing part of the market. The offers of these organizations are also more preference based. An example of such an organization is vakantie-wijzer-van.nl. This is an independent decision support website, which advises an older audience based on preferences. Another example is singleplus.nl which is a decision support website that advises singles, based on the preferences of this group. The difference between general holiday websites and these examples is that these specialized decision support
websites focuses on just one audience and therefore does recommendations that are much more optimal and complete. They take into account the most important and very specific preferences that only (or mostly) exist within their audience.

**Recommendation for group decision support for tourists:**

The interview with Mrs. Visser was focused on preference elicitation and classification of tourists. As said before, there is no existing preference based classification method whatsoever. In order to use classification, Mrs. Visser recommends three options:

1. Analyzing search behavior of tourists by mining over tourists search data and retrieve relations that can be used for classification. (Currently researched by Auwke Pot, researcher at the Vrije Universiteit van Amsterdam)

2. Developing a preference based model based on the means & chain model (Figure 3). This can be done by a research that focuses on question: “What’s important for tourists?” This can be conducted through interviews with tourists.

3. Conduct preference elicitation based on forms with questions. Preferences are often represented as rankings of essential elements holidays. (Attractions, activities, prices, culture, etc). These rankings are then translated into quantitative values. This process is called preference-rank translation. (Explained below)

Another issue is where to get the feature information for travel plans. Mrs. Visser told us that services exist, which provides feeds containing such information in different formats. These feeds can be read and used by our system. Examples of such services are:

- Tradetracker.nl
- Daisycon.com

**Conclusion**

Finally we can conclude that classification based on preferences is a difficult challenge, because of the fact that a tourist preference model does not exist at the moment. In order to get such a model, we will have to develop it ourselves by interviewing tourists. Finally it is recommended that we narrow down our domain to the level of a niche market or activity based holiday domain (e.g. skiing vacations, or safari, etc). This way we can take into account very specific preferences. This way we can our group decision support system for tourist can generate more customized travel plans, which raise the added value of the system and is also more feasible to develop.

**Rank translation**

To translate rankings into quantitative values, a weighing scheme that defines weights for each rank is used. For instance:
A typical weighting scheme is:

- first choice = 75%
- second choice = 17%
- third choice = 6%
- fourth choice = 2%
- fifth choice = 0%

Our utility function is derived from this idea. However, in marketing these schemes are static. In our case we needed a dynamic scheme, since our preference models contain rankings of with varying sizes. Therefore we have designed function (paragraph 3.6 ) that we have fine-tuned with several test sessions.
## Appendix H
### City Database

<table>
<thead>
<tr>
<th>Name</th>
<th>price</th>
<th>Region</th>
<th>Name</th>
<th>price</th>
<th>region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Innsbruck</td>
<td>450</td>
<td>Austria</td>
<td>Hangzhou</td>
<td>1100</td>
<td>China</td>
</tr>
<tr>
<td>Vienna</td>
<td>500</td>
<td>Austria</td>
<td>Hong Kong</td>
<td>925</td>
<td>China</td>
</tr>
<tr>
<td>Antwerp</td>
<td>120</td>
<td>Belgium</td>
<td>Split</td>
<td>630</td>
<td>Croatia</td>
</tr>
<tr>
<td>Bruges</td>
<td>120</td>
<td>Belgium</td>
<td>Zagreb</td>
<td>500</td>
<td>Croatia</td>
</tr>
<tr>
<td>Brussels</td>
<td>130</td>
<td>Belgium</td>
<td>Pula</td>
<td>460</td>
<td>Croatia</td>
</tr>
<tr>
<td>Mostar</td>
<td>850</td>
<td>Bosnia &amp; Herzegovina</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sarajevo</td>
<td>650</td>
<td>Bosnia &amp; Herzegovina</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sofia</td>
<td>330</td>
<td>Bulgaria</td>
<td>Prague</td>
<td>310</td>
<td>Czech Republic</td>
</tr>
<tr>
<td>Toronto</td>
<td>700</td>
<td>Canada</td>
<td>Copenhagen</td>
<td>280</td>
<td>Denmark</td>
</tr>
<tr>
<td>Vancouver</td>
<td>820</td>
<td>Canada</td>
<td>Helsinki</td>
<td>570</td>
<td>Finland</td>
</tr>
<tr>
<td>Ottowa</td>
<td>850</td>
<td>Canada</td>
<td>Toulouse</td>
<td>630</td>
<td>France</td>
</tr>
<tr>
<td>Montreal</td>
<td>700</td>
<td>Canada</td>
<td>Bordeaux</td>
<td>670</td>
<td>France</td>
</tr>
<tr>
<td>Edmonton</td>
<td>1000</td>
<td>Canada</td>
<td>Marseille</td>
<td>500</td>
<td>France</td>
</tr>
<tr>
<td>Whistler</td>
<td>900</td>
<td>Canada</td>
<td>Paris</td>
<td>200</td>
<td>France</td>
</tr>
<tr>
<td>Banff</td>
<td>800</td>
<td>Canada</td>
<td>Nice</td>
<td>750</td>
<td>France</td>
</tr>
<tr>
<td>Victoria</td>
<td>1000</td>
<td>Canada</td>
<td>Lourdes</td>
<td>340</td>
<td>France</td>
</tr>
<tr>
<td>Calgary</td>
<td>1050</td>
<td>Canada</td>
<td>Cannes</td>
<td>250</td>
<td>France</td>
</tr>
<tr>
<td>Quebec City</td>
<td>870</td>
<td>Canada</td>
<td>Lyon</td>
<td>150</td>
<td>France</td>
</tr>
<tr>
<td>Chengdu</td>
<td>1000</td>
<td>China</td>
<td>Cologne</td>
<td>250</td>
<td>Germany</td>
</tr>
<tr>
<td>Guilin</td>
<td>1050</td>
<td>China</td>
<td>Berlin</td>
<td>220</td>
<td>Germany</td>
</tr>
<tr>
<td>Beijing</td>
<td>1000</td>
<td>China</td>
<td>Munich</td>
<td>150</td>
<td>Germany</td>
</tr>
<tr>
<td>Shanghai</td>
<td>1000</td>
<td>China</td>
<td>Hamburg</td>
<td>200</td>
<td>Germany</td>
</tr>
<tr>
<td>Suzhou</td>
<td>950</td>
<td>China</td>
<td>London</td>
<td>450</td>
<td>Great Britain</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Edinburgh</td>
<td>450</td>
<td>Great Britain</td>
</tr>
<tr>
<td>Name</td>
<td>price</td>
<td>Region</td>
<td>name</td>
<td>price</td>
<td>region</td>
</tr>
<tr>
<td>--------------</td>
<td>-------</td>
<td>--------</td>
<td>-----------</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td>Shenzhen</td>
<td>1100</td>
<td>China</td>
<td>Santorini</td>
<td>540</td>
<td>Greece</td>
</tr>
<tr>
<td>Urumqi</td>
<td>1200</td>
<td>China</td>
<td>Athens</td>
<td>450</td>
<td>Greece</td>
</tr>
<tr>
<td>Xian</td>
<td>1100</td>
<td>China</td>
<td>Nuuk</td>
<td>1500</td>
<td>Iceland</td>
</tr>
<tr>
<td>Kunming</td>
<td>1100</td>
<td>China</td>
<td>Reykjavik</td>
<td>680</td>
<td>Iceland</td>
</tr>
<tr>
<td>Lhasa</td>
<td>1600</td>
<td>China</td>
<td>Budapest</td>
<td>340</td>
<td>Hungary</td>
</tr>
<tr>
<td>Mumbai</td>
<td>890</td>
<td>India</td>
<td>Puerto</td>
<td>1200</td>
<td>Mexico</td>
</tr>
<tr>
<td>Goa</td>
<td>900</td>
<td>India</td>
<td>Vallarta</td>
<td>1000</td>
<td>Mexico</td>
</tr>
<tr>
<td>Udaipur</td>
<td>900</td>
<td>India</td>
<td>Tijuana</td>
<td>865</td>
<td>Mexico</td>
</tr>
<tr>
<td>Delhi</td>
<td>850</td>
<td>India</td>
<td>Cancun</td>
<td>500</td>
<td>Mexico</td>
</tr>
<tr>
<td>Jakarta</td>
<td>1100</td>
<td>Indonesia</td>
<td>Monte Carlo</td>
<td>950</td>
<td>Monaco</td>
</tr>
<tr>
<td>Bali</td>
<td>1150</td>
<td>Indonesia</td>
<td>Budva</td>
<td>530</td>
<td>Montenegro</td>
</tr>
<tr>
<td>Ubud</td>
<td>1150</td>
<td>Indonesia</td>
<td>Oslo</td>
<td>325</td>
<td>Norway</td>
</tr>
<tr>
<td>Dublin</td>
<td>570</td>
<td>Ireland</td>
<td>Warsaw</td>
<td>300</td>
<td>Poland</td>
</tr>
<tr>
<td>Cork</td>
<td>880</td>
<td>Ireland</td>
<td>Krakow</td>
<td>500</td>
<td>Poland</td>
</tr>
<tr>
<td>Rome</td>
<td>250</td>
<td>Italy</td>
<td>Albufeira</td>
<td>500</td>
<td>Portugal</td>
</tr>
<tr>
<td>Naples</td>
<td>390</td>
<td>Italy</td>
<td>Porto</td>
<td>400</td>
<td>Portugal</td>
</tr>
<tr>
<td>Milan</td>
<td>350</td>
<td>Italy</td>
<td>Lisbon</td>
<td>400</td>
<td>Portugal</td>
</tr>
<tr>
<td>Venice</td>
<td>450</td>
<td>Italy</td>
<td>Bucharest</td>
<td>350</td>
<td>Romania</td>
</tr>
<tr>
<td>Tokyo</td>
<td>1000</td>
<td>Japan</td>
<td>Irkutsk</td>
<td>750</td>
<td>Russia</td>
</tr>
<tr>
<td>Kyoto</td>
<td>1050</td>
<td>Japan</td>
<td>Saint Petersburg</td>
<td>450</td>
<td>Russia</td>
</tr>
<tr>
<td>Fukuoka</td>
<td>1100</td>
<td>Japan</td>
<td>Moscow</td>
<td>550</td>
<td>Russia</td>
</tr>
<tr>
<td>Osaka</td>
<td>1000</td>
<td>Japan</td>
<td>Sochi</td>
<td>710</td>
<td>Russia</td>
</tr>
<tr>
<td>Sapporo</td>
<td>1200</td>
<td>Japan</td>
<td>San Marino</td>
<td>850</td>
<td>San Marino</td>
</tr>
<tr>
<td>Riga</td>
<td>330</td>
<td>Japan</td>
<td>Belgrade</td>
<td>470</td>
<td>Serbia</td>
</tr>
<tr>
<td>Luxembourg City</td>
<td>200</td>
<td>Luxemburg</td>
<td>Bratislava</td>
<td>370</td>
<td>Slovakia</td>
</tr>
<tr>
<td>Skopje</td>
<td>620</td>
<td>Macedonia</td>
<td>Ljubljana</td>
<td>600</td>
<td>Slovenia</td>
</tr>
<tr>
<td>Veracruz</td>
<td>1000</td>
<td>Mexico</td>
<td>Bled</td>
<td>600</td>
<td>Slovenia</td>
</tr>
<tr>
<td>Name</td>
<td>price</td>
<td>Region</td>
<td>name</td>
<td>price</td>
<td>region</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------</td>
<td>---------</td>
<td>--------------</td>
<td>-------</td>
<td>----------</td>
</tr>
<tr>
<td>Mexico city</td>
<td>735</td>
<td>Mexico</td>
<td>Busan</td>
<td>1350</td>
<td>South Korea</td>
</tr>
<tr>
<td>Acapulco</td>
<td>1000</td>
<td>Mexico</td>
<td>Seoul</td>
<td>1250</td>
<td>South Korea</td>
</tr>
<tr>
<td>Playa del Carmen</td>
<td>875</td>
<td>Mexico</td>
<td>Barcelona</td>
<td>330</td>
<td>Spain</td>
</tr>
<tr>
<td>Oaxaca</td>
<td>835</td>
<td>Mexico</td>
<td>Madrid</td>
<td>230</td>
<td>Spain</td>
</tr>
<tr>
<td>Cabo San Lucas</td>
<td>1030</td>
<td>Mexico</td>
<td>Granada</td>
<td>600</td>
<td>Spain</td>
</tr>
<tr>
<td>Guadalajara</td>
<td>1035</td>
<td>Mexico</td>
<td>Malaga</td>
<td>470</td>
<td>Spain</td>
</tr>
<tr>
<td>Isla Cozumel</td>
<td>925</td>
<td>Mexico</td>
<td>Cordoba</td>
<td>400</td>
<td>Spain</td>
</tr>
<tr>
<td>Mazatlan</td>
<td>1030</td>
<td>Mexico</td>
<td>Valencia</td>
<td>450</td>
<td>Spain</td>
</tr>
<tr>
<td>Salamanca</td>
<td>400</td>
<td>Spain</td>
<td>San Francisco</td>
<td>880</td>
<td>USA</td>
</tr>
<tr>
<td>San Sebastian</td>
<td>650</td>
<td>Spain</td>
<td>New York</td>
<td>900</td>
<td>USA</td>
</tr>
<tr>
<td>Girona</td>
<td>450</td>
<td>Spain</td>
<td>Memphis</td>
<td>850</td>
<td>USA</td>
</tr>
<tr>
<td>Santiago de Compostella</td>
<td>640</td>
<td>Spain</td>
<td>Nashville</td>
<td>900</td>
<td>USA</td>
</tr>
<tr>
<td>Sevilla</td>
<td>360</td>
<td>Spain</td>
<td>Jackson (Mississippi)</td>
<td>1000</td>
<td>USA</td>
</tr>
<tr>
<td>Stockholm</td>
<td>350</td>
<td>Sweden</td>
<td>Indianapolis</td>
<td>900</td>
<td>USA</td>
</tr>
<tr>
<td>Hualien</td>
<td>2600</td>
<td>Taiwan</td>
<td>Houston</td>
<td>800</td>
<td>USA</td>
</tr>
<tr>
<td>Taipei</td>
<td>1200</td>
<td>Taiwan</td>
<td>Detroit</td>
<td>900</td>
<td>USA</td>
</tr>
<tr>
<td>Chiang mai</td>
<td>1250</td>
<td>Thailand</td>
<td>Dallas (Texas)</td>
<td>900</td>
<td>USA</td>
</tr>
<tr>
<td>Phuket town</td>
<td>1075</td>
<td>Thailand</td>
<td>Hanoi</td>
<td>1025</td>
<td>Vietnam</td>
</tr>
<tr>
<td>Bangkok</td>
<td>1150</td>
<td>Thailand</td>
<td>Ho Chi Minh city</td>
<td>1000</td>
<td>Vietnam</td>
</tr>
<tr>
<td>Rotterdam</td>
<td>70</td>
<td>The Netherlands</td>
<td>Zurich</td>
<td>550</td>
<td>Switzerland</td>
</tr>
<tr>
<td>Amsterdam</td>
<td>70</td>
<td>The Netherlands</td>
<td>Geneve</td>
<td>550</td>
<td>Switzerland</td>
</tr>
<tr>
<td>Ankara</td>
<td>570</td>
<td>Turkey</td>
<td>Chicago</td>
<td>800</td>
<td>USA</td>
</tr>
<tr>
<td>City</td>
<td>Distance</td>
<td>Country</td>
<td>City</td>
<td>Distance</td>
<td>Country</td>
</tr>
<tr>
<td>----------------------</td>
<td>----------</td>
<td>---------</td>
<td>----------------------</td>
<td>----------</td>
<td>---------</td>
</tr>
<tr>
<td>Istanbul</td>
<td>430</td>
<td>Turkey</td>
<td>Miami</td>
<td>950</td>
<td>USA</td>
</tr>
<tr>
<td>Chicago</td>
<td>800</td>
<td>USA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miami</td>
<td>950</td>
<td>USA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Las Vegas</td>
<td>1000</td>
<td>USA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kansas City</td>
<td>1000</td>
<td>USA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boston</td>
<td>850</td>
<td>USA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Austin</td>
<td>1200</td>
<td>USA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cleveland (Ohio)</td>
<td>1000</td>
<td>USA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Los Angeles</td>
<td>900</td>
<td>USA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seattle</td>
<td>1000</td>
<td>USA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denver (Colorado)</td>
<td>800</td>
<td>USA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atlanta</td>
<td>850</td>
<td>USA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspen</td>
<td>1500</td>
<td>USA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Honolulu</td>
<td>1300</td>
<td>USA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Philadelphia (Pennsylvania)</td>
<td>950</td>
<td>USA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Appendix I

### Management of cities

Create/edit/delete cities

<table>
<thead>
<tr>
<th>ID</th>
<th>City Name</th>
<th>Type</th>
<th>Price</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Berlín</td>
<td>metropolis</td>
<td>250.0</td>
<td>Germany</td>
</tr>
<tr>
<td>2</td>
<td>Istanbul</td>
<td>metropolis</td>
<td>430.0</td>
<td>Turkey</td>
</tr>
<tr>
<td>3</td>
<td>London</td>
<td>metropolis</td>
<td>200.0</td>
<td>Great Britain</td>
</tr>
<tr>
<td>4</td>
<td>Madrid</td>
<td>metropolis</td>
<td>230.0</td>
<td>Spain</td>
</tr>
<tr>
<td>5</td>
<td>Paris</td>
<td>metropolis</td>
<td>200.0</td>
<td>France</td>
</tr>
<tr>
<td>6</td>
<td>Rome</td>
<td>metropolis</td>
<td>250.0</td>
<td>Italy</td>
</tr>
<tr>
<td>7</td>
<td>Amsterdan</td>
<td>city</td>
<td>70.0</td>
<td>The Netherlands</td>
</tr>
<tr>
<td>8</td>
<td>Barcelona</td>
<td>city</td>
<td>330.0</td>
<td>Spain</td>
</tr>
<tr>
<td>9</td>
<td>Brussels</td>
<td>city</td>
<td>130.0</td>
<td>Belgium</td>
</tr>
<tr>
<td>10</td>
<td>Budapest</td>
<td>city</td>
<td>940.0</td>
<td>Hungary</td>
</tr>
<tr>
<td>11</td>
<td>Copenhagen</td>
<td>city</td>
<td>280.0</td>
<td>Denmark</td>
</tr>
<tr>
<td>12</td>
<td>Dublin</td>
<td>city</td>
<td>570.0</td>
<td>Ireland</td>
</tr>
<tr>
<td>13</td>
<td>Lisbon</td>
<td>city</td>
<td>400.0</td>
<td>Portugal</td>
</tr>
<tr>
<td>14</td>
<td>Lyon</td>
<td>city</td>
<td>250.0</td>
<td>France</td>
</tr>
<tr>
<td>15</td>
<td>Milan</td>
<td>city</td>
<td>350.0</td>
<td>Italy</td>
</tr>
<tr>
<td>16</td>
<td>Munich</td>
<td>city</td>
<td>220.0</td>
<td>Germany</td>
</tr>
<tr>
<td>17</td>
<td>Naples</td>
<td>city</td>
<td>390.0</td>
<td>Italy</td>
</tr>
<tr>
<td>18</td>
<td>Stockholm</td>
<td>city</td>
<td>350.0</td>
<td>Sweden</td>
</tr>
<tr>
<td>19</td>
<td>Vienna</td>
<td>city</td>
<td>500.0</td>
<td>Austria</td>
</tr>
<tr>
<td>20</td>
<td>Hongkong</td>
<td>metropolis</td>
<td>925.0</td>
<td>China</td>
</tr>
<tr>
<td>21</td>
<td>Barri</td>
<td>town</td>
<td>800.0</td>
<td>Canada</td>
</tr>
<tr>
<td>22</td>
<td>Toronto</td>
<td>metropolis</td>
<td>700.0</td>
<td>Canada</td>
</tr>
<tr>
<td>23</td>
<td>Vancouver</td>
<td>metropolis</td>
<td>820.0</td>
<td>Canada</td>
</tr>
<tr>
<td>24</td>
<td>Ottawa</td>
<td>town</td>
<td>650.0</td>
<td>Canada</td>
</tr>
<tr>
<td>25</td>
<td>Montreal</td>
<td>city</td>
<td>700.0</td>
<td>Canada</td>
</tr>
</tbody>
</table>
Adjust weights of the city categories and features

Management of domain model

Create/edit/delete categories and features of the domain model

<table>
<thead>
<tr>
<th>Id</th>
<th>Category</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Culinar</td>
<td>Lokale keuken</td>
</tr>
<tr>
<td>3</td>
<td>Culinar</td>
<td>Internationale keuken</td>
</tr>
<tr>
<td>4</td>
<td>Culinar</td>
<td>Lunchfeesten</td>
</tr>
<tr>
<td>5</td>
<td>Culinar</td>
<td>Koek workshops en evenementen</td>
</tr>
<tr>
<td>6</td>
<td>Culinar</td>
<td>Musea</td>
</tr>
<tr>
<td>7</td>
<td>Culinar</td>
<td>Theater</td>
</tr>
<tr>
<td>8</td>
<td>Culinar</td>
<td>Congressenarth</td>
</tr>
<tr>
<td>9</td>
<td>Sportief</td>
<td>Wandeltouren</td>
</tr>
<tr>
<td>10</td>
<td>Sportief</td>
<td>Fietsburen</td>
</tr>
<tr>
<td>11</td>
<td>Amusement</td>
<td>Nachtclub</td>
</tr>
</tbody>
</table>
Questionnaire management

The online experiment for users

Starting screen of the questionnaire

Bedankt voor het deelnemen aan dit onderzoek. Als beloning krijg je 400 vakantiepunten en maak je kans op een D-Reizen waardebon van 150 euro! Zorg wel ervoor dat je e-mail overeenkomt met het opgegeven e-mail bij vakantiepunten.nl.

Ik heb de informatie hierboven geleezen betreffende
• het doel van dit onderzoek en mijn rechten;
• procedures, aparatuur, en regionale dat tijdens het onderzoek gebruikt zullen worden.

Ik ben op de hoogte dat
• er verdere wet zal optreden met alle informatie dat tijdens het onderzoek verkregen is;
• alle informatie dat tijdens het onderzoek verkregen is veilig wordt opgehouden;
• de verzamelde informatie voor onderzoek- en commenteerdelen beter zullen worden gebruikt;
• de verzamelde resultaten verwerkt kunnen worden in interne verslagen en wetenschappelijke publicaties;
• mijn identiteit geheel anonime blijft in alle publicaties;
• ik het recht heb om al mijn rechten toe te kennen en medewerking aan het onderzoek in te trekken zonder enige consequenties.
Example of the questions of the questionnaire

Categorie 7

Legenda

<table>
<thead>
<tr>
<th>Score</th>
<th>1</th>
<th>Zuur-heel onvoldoende</th>
<th>4</th>
<th>Neutraal</th>
<th>7</th>
<th>Zeer-heel voldoende</th>
</tr>
</thead>
</table>

Vraag 1
Ik vind de gekozen stad leuk

Vraag 2
Ik ben bovenop over de uitverkochte stad die we gaan bezoeken

Vraag 3
Toen de sessie was afgelopen, was ik tevreden over de stad die we gekozen hebben

Vraag 4
Ik ben blij met de stad waarop we zitten uitgekomen vandaag

Vraag 5
Ik heb een goed gevoel over de sessie van vandaag

Vraag 6
Ik wil mijn leden over de wijze waarop de sessie van vandaag is gevoerd

Vraag 7
Ik wil mijn leden over de gebruikelijke stappen in de sessie van vandaag

Vraag 8
Ik wil mijn leden over de wijze waarop we de stappen in de sessie van vandaag hebben uitgevoerd

Ga naar de volgende stap
Appendix J

Normality check for the satisfaction of the recommendations (SATC)

Generate step:

Reduce step:

Clarify step:
Evaluate step:

Normal Q-Q Plot of Satisfaction of the recommendations: clarify step

Normal Q-Q Plot of Satisfaction of the recommendations: evaluate step
Tests of Normality for Satisfaction of the recommendations (SATC)

<table>
<thead>
<tr>
<th>SATC</th>
<th>Kolmogorov-Smirnov</th>
<th>Shapiro-Wilk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistic</td>
<td>df</td>
</tr>
<tr>
<td>SATCsgenerate</td>
<td>.192</td>
<td>116</td>
</tr>
<tr>
<td>SATCreduce</td>
<td>.238</td>
<td>116</td>
</tr>
<tr>
<td>SATCclarify</td>
<td>.267</td>
<td>116</td>
</tr>
<tr>
<td>SATCevaluate</td>
<td>.257</td>
<td>116</td>
</tr>
</tbody>
</table>

Normality check of Satisfaction of the amount of recommendations (AMOC)

Generate step:
Reduce step:

Clarity step:
Tests of Normality of Satisfaction of the amount of recommendations (AMOC)

<table>
<thead>
<tr>
<th></th>
<th>Kolmogorov-Smirnov</th>
<th>Shapiro-Wilk</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistic</td>
<td>df</td>
</tr>
<tr>
<td>AMOC generate</td>
<td>.248</td>
<td>116</td>
</tr>
<tr>
<td>AMOV reduce</td>
<td>.283</td>
<td>116</td>
</tr>
<tr>
<td>AMOC evaluate</td>
<td>.259</td>
<td>116</td>
</tr>
</tbody>
</table>