UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Knowledge aggregation in people recommender systems: matching skills to tasks

by

Jennifer Nguyen

Knowledge aggregation in people recommender systems: matching skills to tasks

doctoral thesis
by

Jennifer Nguyen

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy in Artificial Intelligence
Co-supervised by Dr. Cecilio Angulo Bahón and Dr. Núria Agell Jané

Universitat Politècnica de Catalunya
Barcelona, Spain
November 2019

For my parents.

Acknowledgments

This research was supported by the INVITE research project (TIN2016- 80049-C2-1-R and TIN2016-80049-C2-2-R), funded by the Spanish Ministry of Economy and Competitiveness.

To my supervisor Dr. Núria Agell Jané, without whom this thesis would not have been possible. Your patience and kindness has been unwavering during the countless hours devoted to the teaching and research behind this thesis. Thanks to you, I discovered the field of research. You have always encouraged dialogue and helped me to see things from different perspectives, making this endeavor less intimidating. Above all, thank you for instilling confidence in me and allowing me to be a part of your team.

To my supervisor Dr. Cecilio Angulo Bahón, thank you for your guidance and support. I appreciate your ability to see beyond just the idea itself but to what it could become.

To Dr. German Sánchez, thank you for teaching me to embrace coding, especially the power of R and for mentoring me through the maze called research.

To Dr. Albert Armisen, thank you for listening and keeping me motivated. I value our conversations on postdoctoral life.

I would like to express my gratitude for our team of five with whom I have had the pleasure of working. This experience is richer because of your warmth and enthusiasm.

To Dr. Francisco J. Ruiz, you make me see the beauty in the intricacies of math. Thank you for your patience. You make math enjoyable.

To Dr. Xari Rovira, it has been a pleasure learning from you the values of teaching.

To Dr. Jordi Montserrat, thank you for letting me lean on you and walking this journey with me.

Thank you to the members of the GREC research group for your invaluable input and feedback to each stage of these works.

My time in this program has been made enjoyable in large part because of my friends, Ryan Federo, Karla Trejo, Roy Mouawad, Asier Pereda, Juan Acevedo, Feranita, Francesca Ferre, Ben Ngoye, and Mohammed (Zack) Zakriya who shared in the experience with me. I'm thankful for our camaraderie and hope that we will start another chapter together soon.

Laura Bierley, thank you for listening to me all these years and keeping me grounded.

Rich Brauer, your energy and enthusiasm for all things science and new is contagious.

To my second family, Masdeu Margalef, I am grateful for your continuous support, friendship, and cultural nourishment.

To my Aunt Hoa Nguyen, for your love, generosity and support.

To my Uncle Sia Afshari, for always being there for me. I learned the foundations for this thesis from you.

To my parents, Nhung and Thé Nguyen, for always believing in me and encouraging me to follow my ambitions. Thank you for teaching me the value of hard work and dedication. Most of all, thank you for showing me that anything is possible.

To Maria Jose Masdeu Margalef, you inspired me to take this leap and supported me on this quest. With your encouragement, this thesis has come to fruition. This thesis is as much yours as it is mine.

Thank you all very much.

Estima la vellesa que t'envolta i que se't dona clara i generosa, i en tu i per tu estima també els altres, car d'aquests dos amors te'n pervindrà l'harmonia i, tal volta, la grandesa.

Miquel Martí i Pol

Abstract

People recommender systems (PRS) are a special type of RS. They are often adopted to identify people capable of performing a task. Recommending people poses several challenges not exhibited in traditional RS. Elements such as availability, overload, unresponsiveness, and bad recommendations can have adverse effects. This thesis explores how people's preferences can be elicited for single-event matchmaking under uncertainty and how to align them with appropriate tasks.

Different methodologies are introduced to profile people, each based on the nature of the information from which it was obtained. These methodologies are developed into three use cases to illustrate the challenges of PRS and the steps taken to address them. Each one emphasizes the priorities of the matching process and the constraints under which these recommendations are made. First, multi-criteria profiles are derived completely from heterogeneous sources in an implicit manner characterizing users from multiple perspectives and multi-dimensional points-of-view without influence from the user. The profiles are introduced to the conference reviewer assignment problem. Attention is given to distribute people across items in order reduce potential overloading of a person, and neglect or rejection of a task. Second, people's areas of interest are inferred from their resumes and expressed in terms of their uncertainty avoiding explicit elicitation from an individual or outsider. The profile is applied to a personnel selection problem where emphasis is placed on the preferences of the candidate leading to an asymmetric matching process. Third, profiles are created by integrating implicit information and explicitly stated attributes. A model is developed to classify citizens according

to their lifestyles which maintains the original information in the data set throughout the cluster formation. These use cases serve as pilot tests for generalization to real-life implementations. Areas for future application are discussed from new perspectives.

Contents

List of Tables

Listing of figures

Listing of Acronyms

CQA	Community Question Answering
CV	Curriculum vitaes
DM	Decision Maker
FHOWA	Fuzzy Heavy Ordered Weighted Averaging
FOWA	Fuzzy Ordered Weighted Averaging
GSD	Global Similarity Degree
HAC	Hierarchical Agglomerative Clustering
HFLTS	Hesitant Fuzzy Linguistic Term Set
KMS	Knowledge Management Systems
LASSO	Least Absolute Shrinkage and Selection Operator
LDA	Latent Dirichlet Allocation
MCDA	Multi-Criteria Decision-Aiding
MCDM	Multi-Criteria Decision Making
MEMCDM	Method Multi Experts Multi Criteria Decision Making
MSD	Marginal Similarity Degree

NLP	Natural Language Processing
OS	Overall Scores
OWA	Ordered Weighted Averaging
PRS	People Recommender Systems
QI	Quality Index
RIM	Regular Increasing Monotone (Quantifier)
RS	Recommender Systems
SNA	Social Network Analysis
SVM	Support Vector Machine
TF	Term Frequency
TF-IDF	Term Frequency - Inverse Document Frequency
TOPSIS	Technique for Order Preference by Similarity to an Ideal Solution

Seeing someone reading a book you love is seeing a book recommending a person.

Anonymous

1 Introduction

In a *knowledge society* information is transformed into resources which enable us to make decisions in our everyday lives [106]. Thanks to advances in information and communication technology people have access to immense amounts of data. With the explosion of social networks and online tools, it has become easier for people to contribute and share knowledge. People can promote awareness and offer their opinions about matters in order to give meaning to them. This in turn builds context around matters with which other people have limited or no familiarity but in which they may be interested. The ease with which people can impart their knowledge has led to an overwhelming amount of available content [131] which is intrinsically heterogeneous and unstructured [106].

Parsing through this information can hinder a person's ability to identify a solution to their matter at hand in a timely manner. For example, if a person has a broken water pipe in their home, it is expected that they would need to find a plumber quickly. However, not knowing a specific plumber would require them to search for one using online tools and sort through the deluge of information available to them. An alternative might be to ask for a recommendation from a neighbor. With this recommendation, the neighbor is not only providing a plumber's name but a sense of trustworthiness facilitating the decision to contract this plumber. Recommender systems (RS) have been developed to assist people in finding pertinent content or information [6, 104, 110, 130, 131]. They have been developed to tackle decisions regarding a diverse set of domains including e-learning [26, 54, 99, 166], movies [38, 165], and travel [27, 99]. They are capable of filtering large amounts of information in order to introduce people to items for the first time or suggest relevant items for the matter at hand [131].

RS are tools and techniques which augment this social process [130] guiding people toward interesting results in a personalized manner [30]. Thereby, RS assist people with these various decision-making processes. Formally, RS are tools and techniques which suggest items of relevance to users [131]. In general, they are directed towards individuals who do not have sufficient experience to evaluate the alternative items. Upon a user's request, a RS makes a recommendation based on data about the user, available items, and previous transactions. The user then decides whether or not to accept the recommendation. The user's response may be used to enhance future recommendations.

In order to make recommendations, RS gather information about *items*, *users*, and *transactions* [131]. Items are anything which are recommended. They may be represented by their complexity or relevance to a user. Complexity refers to the different aspects and features of a item like its format or sensitivity to time. Users are people to whom items are being suggested. They are seeking recommendations according to their individual criteria. Recommending items according to their preferences provides a personalized component which differentiates RS from information retrieval or search engines which assist users with searching for various forms of content but often neglect the preferences of users [131]. Criteria can include different aspects such as the features or ratings for a specific attribute of an item [8], context in which an item is selected [7], or preference variation over time [119, 131]. In RS, this information is structured according to the recommendation technique [131]. Transactions are relationships between the user and the item recorded from interaction that the user maintains with the RS. Examples of transactions may be ratings or items purchased. The goal of a RS is to predict which items are most relevant to the user according to her criteria [19].

Traditionally, approaches to the recommendation problem are described broadly as *content-based* or *collaborative filtering* [6, 50]. Content-based approaches identify items having the same characteristics as items for which a user has previously rated positively. These approaches assume there is a rich information profile describing the characteristics of each of the items [25, 50]. For example, if a user likes Movie A and rated it positively, the system may recommend Movie B because both movies A and B have the same actors and genre. In contrast, collaborative filtering approaches identify users of the system who provide similar ratings to the same items. Given that these users have rated past items in a similar way, it is expected that if one of these users is introduced to an item rated positively by some of the other users, he may like it. Collaborative filtering approaches can help to overcome limitations of content-based approaches due to over-specialization [25, 50]. Through ratings feedback, collaborative filtering can recommend items to users despite having limited content information about items. Moreover, collaborative filtering can reduce over-specialized recommendations by recommending novel items liked by another user who has been assigning similar ratings as the active user [19, 50]. For example, it has been determined that a user A likes action and comedy a lot and dislikes romance based on movies she previously rated. For a given movie, some users have rated it high while others have rated it low. If all of the high ratings are contributed by users who like action and low ratings by users who like romance, the system may recommend the movie to user A because it is similar to the movies she previously rated high.

1.1 Systems to recommend people

Traditionally, RS identify a list of items which match a user's preferences. The utility of the user is the primary criteria taken into account in the recommendation. However, the receiver of the recommendations may not always be the sole stakeholder in the system [183]. Where multiple individuals and organizations can benefit from recommendations a multi-stakeholder environment exists [3]. The objective of these environments is to generate an item recommendation which considers the utilities of multiple stakeholders [183] with limited loss to the accuracy of the recommendation [184]. For example, a plumbing association operates a site which recommends local plumbers. Its main objective is to reduce the transport expenses of the plumbers and to distribute work solely among members of its organization. For this service, plumbers pay a fee to be members of the association. In order for the platform to thrive, its needs to attract and retain participants. Customers expect valuable recommendations and plumbers want to be matched with users likely to purchase their services. Satisfaction can lead to repeat business or referrals on the customer's side and continued membership on the plumber's side. Dissatisfaction can lead to both sides abandoning the platform. Therefore, the RS operator has his own objectives which are a function of the utilities of the bilateral relationship between the plumber and the customer [31]. Research in multi-stakeholder recommendation include approaches in recommendation hybrids, multi-objective optimization, and multi-agent architectures [3].

A special case of multi-stakeholder RS is a reciprocal RS [183] which focuses on recommending people to people, whereby, the preferences of each stakeholder in the recommendation needs to be satisfied [91, 125, 183]. For example, identifying a match between two users of an online dating platform requires that the preferences of both users be satisfied. Reciprocal RS introduce concepts of reciprocity, limited availability, sparsity [91, 125], and passiveness [91]in addition to those of traditional recommenders. Reciprocal recommendation has been addressed in domains such as online recruiting [178], online dating [157, 167], and expertise management [83]. Recently, it has been a topic of interest with researchers and competitions [1].

Although reciprocity is an important aspect of people-to-people recommenders, many works [126, 149] in which RS are used to assign people do not mention reciprocity [124]. One possible reason according to [124] may be that these systems are focused on satisfying the proactive user. *Proactive users* are those who are actively searching for a recommendation and *reactive users* are those who are being recommended [125]. For example, a user searching for reviewers for papers will have certain preferences towards the characteristics of the expert such as his area and level of expertise. In comparison, the reviewer may have few to no preference towards the user. Returning to the example of the plumber, while it is important to the homeowner that the plumber is capable of fixing his broken pipe, the plumber may have no preferences towards the person hiring him. We will refer to RS which assign people without need of reciprocity as people recommender systems (PRS). Often PRS are adopted to identify people capable of fulfilling a task such as code reviewers and company experts [18, 49]. In these scenarios, the proactive user is a an activity looking to be matched with a person, the reactive user. Because the matching in PRS is the reverse of traditional user-item RS, it is characterized by different challenges from how to portray the reactive user to matching him with his respective activity. Some main differences between traditional RS and reciprocal RS are *reciprocity*, *limited availability*, *sparsity* [91, 124], and *passiveness* [91]. We discuss the last three concepts with respect to PRS in the following sections.

1.1.1 Building profiles for people to recommend

There has been considerable work in the area of user profiling for PRS, especially in expert finding. In general, these profiles define a person's area and level of knowledge or interest in order to recommend them. Some such RS are to identify a candidate for a job in a hiring process [95], determine the right reviewer for a paper in a conference [149], or find an expert to help with a problem [175]. Other reasons to define a person's knowledge are to detect which items may appeal to him. In the case of scholars, items may be scientific articles or academic papers [151]. Although the context of expert profiling could be extended beyond the scope of researchers, evidence of expertise is more readily available for them in the form of academic papers, books, published articles, and personal websites.

There is no unified method to define a stakeholder's utility [184]. Previous methods include utilizing previous interactions with items [94], explicitedly requesting preferences [96], extracting them from textual information [42], or gleaning them from social networks [10]. Finding individuals having knowledge in specific areas is highly dynamic, difficult to qualify, and varying in degree of knowledge [60, 109]. Knowledge can be categorized as tacit and explicit. Explicit knowledge can be articulated and codified. Tacit knowledge, on the other hand is difficult to codify. Managing tacit knowledge is a important to the core competencies of an organization for it is harder for competitors to copy this intellectual asset than explicit knowledge [92]. Tacit knowledge resides within a person, profiling them and constructing a topic directory is an effective way to manage knowledge and identify experts who can help others in the organization [175]. However, finding relevant experts in a directory is difficult because the information seeker is uncertain of his information needs [109] which may involve multiple topic areas. In addition, the growing amount of knowledge and associated taxonomy complicates the search.

Within the context of reciprocal RS, people expect to provide explicit profile information regarding both their preferences and personal characteristics [125]. In comparison, people prefer to define minimal information in traditional RS as it may pose a time or privacy imposition. Even though users may elect to give more information, often times explicitly stated preferences may differ from actual preferences. This difference may be due to a person's uncertainty in his preferences or a need to have a more attractive profile [125]. Implicit preferences may illuminate discrepancies between actual preferences and explicit preferences. PRS more closely resemble traditional RS in this regard as people may not have a need in providing an updated profile or may have incentive to exaggerate their knowledge. Therefore, people's knowledge is difficult to validate [60, 109]. RS have begun to integrate secondary sources of information to provide more comprehensive profiles [60]. Knowledge of user expertise may be spread across multiple sources of information. Different sources can add dimensions to profiles enabling them to be refined and provide a different point-of-view about a person's preferences [17]. Despite its benefits, extracting knowledge items from different sources illuminates inconsistency [126] and heterogeneity among taxonomy [11]. For example, reviewers may have multiple profiles on a website owing to how his name was written on the published article. Another example, one source could characterize a person as having very high expertise in one area while another says the opposite. This is possibly due to the publications, books, and other sources of information from which the website is extrapolating its data. Therefore, seeking information from multiple sources requires knowledge unification and transformation to reconcile these discrepancies.

Given that users on both sides of reciprocal RS are actively engaged, either content-based or collaborative filtering approaches are appropriate to identify users matching the proactive user's preference and vice versa [124]. Once the two sides of the matching have been determined the system can identify the overlapping users and recommend them to their respective others [105, 178]. In traditional RS, rich implicit consumption history can be obtained from repeated transactions to re-enforce explicit preferences or pin-point explicit preferences which are not reflective of actual preferences. However, the nature of people-to-people RS limits the ability to obtain repeated transaction information. For if the system performs well, a match between two people will be identified following a few transactions after which the users will leave the system. Due to this data sparsity [91, 124], implicit preferences cannot be leveraged to enhance user profiles nor can it be leveraged to enhance the profiles of those similar to them making the system more reliant on robust and reliable explicit preferences. Without these implicit preferences, inferring which explicit preferences have inherent uncertainty is a challenge. When users of the RS are not active or there are limited transactions then group generalizations may be required to obtain the preferences of the non-active user [124]. These preferences may be gleaned by relaxing success requirements and identifying intermediate interest. Rather than solely evaluating job candidate and employer preferences at the moment a job is fulfilled, success may be defined by a candidate applying for a position and a position receiving applications.

1.1.2 Challenges to recommending people

Recommending people poses several challenges not exhibited in traditional RS. Special consideration to control for limited availability, passiveness, *overspecialization*, *rejection*, and *neglect* may bare more relevant recommendations. Traditional RS do not necessarily limit the number of users recommended an item, rather they assume there is an abundant supply of an item. However, for reciprocal RS where the item is a user, there is limited availability [91, 124]. Overspecialization occurs when variability and diversity of recommendations is limited [25, 124]. A reactive user may be recommended very often and receive a lot of attention or expressions of interest and therefore, become overloaded with recommendations [124]. The overload may cause the reactive user to reject the proactive user. Rejections can also be a result of bad recommendations which do not align with a proactive user's preferences or the reactive user does not share mutual interest [125]. As expressions of interests are based on an expectation of reciprocity, rejection can leave people feeling disappointed. Consider an expert RS where an expert seeker is recommended an expert. After being contacted, the expert may reject the request to collaborate, wasting the seeker's time and effort. Furthermore, as expressions of interests are directed towards reactive users, these users may not respond making their engagement passive [91]. Neglect refers to users who are never recommended making them more likely to leave the site [125].

These challenges can be extended to PRS where people are being recommended for tasks. Let's assume the homeowner has been referred to a popular plumber, one who is frequently called upon to work in the community. He is now too busy to attend to any additional calls and the homeowner must look for another plumber. However, members of the community are unaware of other plumbers to recommnend. The popularity of the plumber may cause him to become overwhelmed leaving him to reject or ignore future matches and users to search for other recommendations. As discussed above, it is important in PRS to minimize rejections. Given a similar scenario in a traditional RS, a popular movie recommendation would not prevent a user from watching the movie unless he elected not to watch it. Pizzato et al. [124] proposed to balance the distribution of recommendations by aligning users along their popularity groups. Another strategy recommends items at the border of users' areas of interest rather than at the center [2]. Novelty, may not be an appropriate solution for all cases of popularity. In our example, a commercial plumber may be available to repair the homeowner's broken pipe. However, a commercial plumber's qualifications exceed those of a residential plumber due to the nature of commercial facilities such as size, structural complexity, and types of problems. Although, the plumber's knowledge and expertise makes him capable of the repair, it may also impose an unnecessary cost to the homeowner. Finding a expert with skills more closely related to the problem reduce undue burden to the homeowner. Moreover, another plumber who is paying for the services of a RS may not be as popular and therefore, is not being recommended to any customers may opt to leave the site imposing less cost to himself.

1.2 Objectives of the thesis

The overall aim of this thesis is to contribute to literature on PRS. Specifically, we are interested in two areas. First, we consider how to represent people's profiles in a manner more expressive of their preferences. Second, we look at how to define matching systems which consider priorities in order to assist recommender systems in obtaining their matching goal. These two objectives are addressed through three cases: 1) conference reviewer assignment, 2) personnel assignment problem, and 3) lifestyle classification.

The main objectives of this thesis are:

1. **Creating user profiles.** Although, it is expected that users provide more explicit profiles in people RS than in traditional RS, the task can be time consuming. Therefore, obtaining user preferences through more implicit methods is an opportunity to minimize users' time. However, given that people matching does not occur with the same frequency as traditional item to user matching, obtaining implicit preferences solely from transactional information is a challenge. Adding to the challenges of implicit preferences, information provided explicitly by users is likely to have uncertainty. Our first research objective is to define user profiles with respect to two aspects:

   (a) *How can people's preferences be elicited for single event matchmaking?* This question will be addressed in Chapter 3 by proposing to develop a profile derived from publicly available information. A profile represented by categorical and numerical characteristics is developed which resolves preferences from unstructured information containing conflicting elements. We demonstrate the applicability of the proposed method within a conference reviewer assignment problem in a real case example.

   (b) *How to capture a person's preferences when they are not explicitly defined by the person?* This question will be addressed in Chapter 4. Similar to the reviewer assignment problem, we develop a user profile for students looking for internships in a real case example. We propose to derive preferences for both the students and internships from unstructured information and express them in terms of their inherent imprecision and hesitance. The internship and candidate profiles are modeled as hesitant linguistic terms.

(c) *How to generalize profiles based on opinions?* This question will be addressed in Chapter 5. We consider profiles created from implicit and explicit information. Implicit preferences derived from text provided in customer reviews is integrated with explicitly stated attributes and generalized to a class of visitors. The segmentation technique administers an aggregation of these attributes.

2. **Balancing recommendations.**

(a) *How to allocate users evenly across items? How to avoid assigning the most popular users to the most popular items and neglecting less popular items?* To address this question, a proposed RS is developed in Chapter 3 which assigns reviewers to papers based on coverage. The proposed method considers that some paper topics are more popular than others requiring more reviewers with expertise in that particular topic than others. A *popular* topic is one about which many papers are written or many reviewers have experience. Similarly, the distribution of expertise topics among reviewers is not uniform as some topics can be more popular than others among reviewers at any given moment. In addition, a person whose expertise covers many topics and is therefore, likely to be a good candidate to review many papers may create a situation of overspecialization. Likewise, papers whose topics attract a lot of interest from reviewers may do the same. Specifically, the proposed method assigns reviewers to papers according to the topics of the paper which need to be covered by reviewers.

(b) *How to increase recommendation exposure to relevant items?* To address this question, a methodology is presented in Chapter 5 which applies a preference distribution based on both students and internships. This model applies a fuzzy order weighted averaging (FOWA) operator to sort internships and recommend a selection of most relevant internships. Given that some careers may be more interesting than others, competition for those internships may be tougher. As these internships are limited in quantity, we are inspired to expose students to other positions that may be of equal relevance to their interests in order to increase the opportunities available to them.

1.3 Structure of the thesis

This introductory chapter provides an overview of RS and the challenges related to PRS which are used to guide the objectives of this thesis. Chapter 2 formalizes the user profiling process and reviews related research in the areas of profile creation and matching. In Chapter 3, the problem of matching people and items is introduced in a conference reviewer assignment problem. A proposed method for profiling and assigning reviewers to papers is implemented on a real case and the results are evaluated from five different perspectives, and show the interpretability of the results. The chapter is developed from the perspective of ranking reviewers. Chapter 4 further builds on the matching problem by approaching the problem from the opposite reference point, that of assigning jobs to job candidates. In contrast to the previous chapters, this chapter seeks to sort positions by relevancy rather than rank them. Given the different scenarios, the conditions which must be met for each assignment is different. For example, in Chapter 3 one constraint is that many reviewers need to be assigned to a single paper in order for it to receive multiple feedback. However, each reviewer need not cover all the topics of the paper he is reviewing. In Chapter 4 a job candidate is interested in receiving suggestions for multiple job openings so that he may have options available to him. In contrast to the previous case, suggested positions need to cover as many topics as possible. A fuzzy matching approach is applied to assign internships to students in order to capture the inherent uncertainty related to the personnel selection problem. A comparison of the results with two alternatives suggests the viability of the proposed method. Chapter 5 classifies people's lifestyles based on attributes of previously frequented restaurants and their reported experiences. This chapter develops generalized profiles integrated from customer opinions and elicited information from their past transactions. In Chapter 6 a discussion of the work presented, its relationship to previous work, limitations, and areas of future work are reviewed.

1.4 Contribution of the thesis

This thesis takes a new perspective on the topic of PRS. In contrasts to previous studies which focused on obtaining the optimal match between people and tasks, we directly consider the issues related to people recommender systems: limited availability, overspecialization, rejection, and neglect. The specific methodologies and algorithms resulting from this thesis are driven in part by previous theoretical studies in recommender systems and research in human resources for practitioners.

The scientific contribution of this thesis can considered as an action research paradigm. Action research is an "intervention approach to diagnose and treat a problem of a specific client" [89, 159]. It has two distinct features: (1) a client experiences an applied problem and (2) the problem is addressed by engaging with the client and intervening in his setting [89]. The client participates in the problem solving through data collection, feedback, and action [122]. In human resource development, the General Method of Theory Building [103] has been proposed to integrate the paradigms of theory and practice combining elements of conceptual development and application [145]. In this scenario, theory is applied to a real world application where it can be evaluated for usefulness and refined through inputs from the client [103, 145]. Therefore, it is a recursive process which enables the theory to remain relevant in practice. In this direction the main contributions of the thesis can be considered in the following three lines.

Contribution to the reviewer assignment problem: This problem has been previously studied and at times, systems developed from these studies have been implemented in real-life conference situations. While large conferences, such as IEEE INFOCOM [90] and NIPS [42], have adopted these systems for assigning reviewers to conferences smaller ones are slow to adopt. One reason may be due to the ease of adoption. Smaller conferences may lack the resources to implement systems from previous research or find it less of a necessity due to the size of the conference. Chapter 3 of this thesis proposes a general method which can facilitate the adoption of automated assignment. The methodology rests on the ease with which conference organizers can readily assemble reviewer biographies from publicly available resources, assign weights according to the criteria of importance and make assignments within specified constraints. It is interpretable and implementable by those outside the research area. The methodology has been validated against a ground truth, reviewer specified knowledge areas, and an optimized solution to the same problem. As will be discussed in Chapter 3, previous studies developed solutions for specific conferences, however, these solutions were rarely implemented in other conference environments or generalized to other reviewer assignment problems. The significance of our proposed methodology is that it has been piloted within an organization for assigning reviewers to medical project proposals, a related but different contextual environment as explained in Section 6.5. Similar to the aforementioned research in practice, throughout the research we conferred with stakeholders in defining the problem and obtaining feedback. In addition, the methodology contributes to the assignment problem by addressing the criterion of topic coverage in a multi-criteria matching problem. Previous studies consider topic coverage as part of a constraint based optimization problem [83, 149]. However, situations in which criteria other than that of expertise play a role in assigning tasks to experts, can benefit from multi-criteria matching. Lastly, the methodology implements a variation to LDA which permits the automated labeling of topics, a known drawback to the automation of LDA [128]. The output of LDA is a set of concepts which can be distributed across various topics. The method aligns these concepts with predefined conference topics.

Contribution to the personnel assignment problem: As many candidates may apply for a single position, candidates are pre-screened are typically based on preliminary information. The process can be extremely time consuming and may not lead to a shortlist of candidates which meet the organization's placement goals (ex. diversity). A potential candidate may be excluded from the process if he is unaware of a job opportunity and a poor match between an employee and a corporation can result in business costs. PRS can connect organizations with expertise outside the organization to reduce the cost of search, enable search beyond the local geographical area, and reach more distant and diverse audiences. However, the following challenges still exist: validating expertise [5, 16, 36], determining responsiveness and accessibility [16, 55], and managing expert profiles. Several commercial offerings have tried to address these challenges by applying data science to develop and predict a candidate's capacity to perform a job [1] [2] [3]. These PRS focus on finding the person with the "right level of expertise". Yet, it has been noted that candidates may not be completely honest about their skills and interests in order to be attractive to the company [36]. One solution has been to engage candidates in game playing to assess their interest in a position or industry [36].

In Chapter 4, we propose an alternative methodology of eliciting interest from candidates. It is understandable that candidates will put forth the most favorable representation of themselves in submitting their resume to an organization. Therefore, the presented methodology considers a general resume submitted to a job bank. It considers the entire matching process of candidate to position from the point of view of the candidate, the opposite of what was studied in the reviewer assignment problem. This process enables the areas of interest to be put into the voice of the candidate and narrows the broad list of potential candidates to only those exhibiting an interest in the position. Different from prior research which have considered a match between a candidate's interests and an employer's requirements or focused solely on the interests of the employer, the methodology in this thesis places emphasis on the interests of the candidate.

Contribution to lifestyle classification: Cultural similarities are defined as tastes, experiences, leisure pursuits, and self-presentation styles [28]. These similarities are often the bases on which merit is evaluated [51] and serve as markers for inclusion or exclusion from social opportunities[133, 164]. A study, by Northwestern University [133], showed that once a candidate passed an initial screening, cultural fit was usually given more weight than experience or coursework in the hiring process. A survey given by The Rockefeller Foundation [4] to 200 C-suite professionals and Human Resource professionals, found that the most important metric to measuring success of entry-level employees was cultural fit. Cultural fit is a sub-

[1] https://angel.co/company/gild
[2] https://www.entelo.com
[3] https://www.gapjumpers.me
[4] https://www.rockefellerfoundation.org/blog/key-findings-on-the-state-of-entry-level-employment-in-the-us/

jective measure and difficult to define [5] [36], and until recently left to the personal interviews for assessment. Organizations are moving towards game-based assessments. However, these have not been validated against job performance [36]. Rivera [133] describes the phenomenon of placing fit over qualifications as a process of cultural matching between the candidate and organization and introduces interpersonal processes which evaluate candidate lifestyle markers during the hiring process. First, evaluators assess candidate's cultural similarity to the firm based on their extracurricular interests and self-presentation styles. Second, candidates are judged according to their similarity to the evaluator in terms of extracurricular or extraprofessional similarities. Furthermore, the author's study implied that cultural similarities assisted with greater comprehension and valuation of the candidate.

As one's cultural fit is a subjective opinion, we chose to assess it based on one aspect of a person's lifestyle marker via his dining-out behavior. We assess the motivational drivers behind his decisions to frequent some establishments and not others and consider the characteristics of the visited places. It is assumed that candidates are willing to provide access to non-personal social media sites as it may appear less intrusive than game playing or corporate snooping of personal social media sites shared with friends.

With the previous elements in mind, we present a methodology which clusters people according to their motivational drivers and the attributes of the places they visit. In addition, the clustering process considers the degree to which the attributes are of relevance to the individual. Understanding where a job candidate fits into one of these clusters, could assist with understanding his social style. For example, the candidate may be a "foodie" or "socializer" based on his dining selection. Clustering individuals according to their lifestyles can assist with recommending potential answerers in question and answer forums. For example, when posting a question about the ambience of an after work establishment, a response may come from sociable person who enjoys frequenting happy hours or someone who goes on occasion to network. The opinions may differ depending on the point of view. Therefore, the relevant response depends on a match with the lifestyle of the user or context in which he is searching.

[5] refer to footnote 4

The objective of the methodology presented is different from the state-of-the-art RS of Netflix [6] and Amazon [7]. Netflix concentrates on increasing engagement, the time users spend on Netflix [70]. Their RS has a variety of algorithms to recommended movies based on the customer's behavior, general reoccurring viewer trends, and similarity to movies watched [70]. Therefore, its focus is on the movie recommendation. Amazon creates recommendation from related items purchased or viewed, applying an advanced item-based collaborative filtering algorithm [144]. Both systems are dedicated to recommending items and benefit from repeated customer interaction. In contrast, the presented methodology is interested in describing a person who consumes this item. Given an item, the features surrounding the item provide some context as to its selection and the person who has chosen it. For example, if an item is a hotel, the destination may not be as important in the description of the customer as the frequency with which he elects to stay in elegant full service hotels.

During the development of this thesis, various elements have been presented at different conferences and workshops within the artificial intelligence, multi-criteria decision making, and qualitative reasoning communities. The presentations are listed in Table 1.4.2. Our work culminates in two publications one in *Pattern Recognition Letters* (Q2 - Computer Science, Artificial Intelligence) and another in *Applied Soft Computing* (Q1 - Computer Science, Artificial Intelligence) detailed in Table 1.4.1. In addition, two publications derived from the thesis are under review. One is under the second round of reviews in *Neural Computing and Applications* and another is under the first round of reviews in *Knowledge-Based Systems*.

[6] `https://www.netflix.com/es-en/`

[7] `https://www.amazon.com`

Table 1.4.1: Journal publications resulting from this thesis

Article Title	Authors	Journal	Article	Journal Metrics
A decision support tool using Order Weighted Averaging for conference review assignment	J. Nguyen, G. Sánchez-Hernández, N. Agell, X. Rovira, C. Angulo	Pattern Recognition Letters	https://doi.org/10.1016/j.patrec.2017.09.020	IF: 1.952
A linguistic multi-criteria decision-aiding system to support university career services	J. Nguyen, G. Sánchez-Hernández, A. Armisen, N. Agell, X. Rovira, C. Angulo	Applied Soft Computing	https://doi.org/10.1016/j.asoc.2017.06.052	IF: 3.907

Table 1.4.2: Conferences where parts of this thesis have been presented

Conference Name	Date	Location
30th European Conference on Operational Research (**Euro 2019**)	June 23, 2019 - June 26, 2019	Dublin, Ireland
Artificial Intelligence International Conference (**A2IC 2018**)	November 21, 2018 - November 23, 2018	Barcelona, Spain
18th Conference of the Spanish Association for Artificial Intelligence (**CAEPIA 2018**)	October 23, 2018 - October 26, 2018	Granada, Spain
21st International Conference of the Catalan Association for Artificial Intelligence (**CCIA 2018**)	October 8, 2018 - October 10, 2018	Roses, Spain
31st International Workshop on Qualitative Reasoning (**QR 2018**)	July 13, 2018 - July 19, 2018	Stockholm, Sweden
20th JARCA Workshop on Qualitative Systems and Applications in Diagnosis, Robotics and Ambient Intelligence (**JARCA 2018**)	June 23, 2018 - June 26, 2018	Alfas de Pi, Spain
IEEE International Conference on Fuzzy Systems (**FUZZ-IEEE 2017**)	July 9, 2017 - July 12, 2017	Naples, Italy
19th International Conference of the Catalan Association for Artificial (**CCIA 2016**)	October 19, 2016 - October 21, 2016	Barcelona, Spain
18th JARCA Workshop on Qualitative Systems and Applications in Diagnosis, Robotics and Ambient Intelligence (**JARCA 2016**)	June 23, 2016 - June 29, 2016	Almeria, Spain
18th Congreso Español sobre Tecnologías y Lógica Fuzzy (**ESTYLF 2016**)	May 25, 2016 - May 27, 2016	San Sebastián, Spain
83rd European Working Group on Multicriteria Decision Aiding (**83rd EWG-MCDA**)	March 31, 2016 - April 2, 2016	Barcelona, Spain
IEEE International Conference on Fuzzy Systems (**FUZZ-IEEE 2015**)	August 2, 2015 - August 5, 2015	Istanbul, Turkey
17th JARCA Workshop on Qualitative Systems and Applications in Diagnosis, Robotics and Ambient Intelligence (**JARCA 2015**)	June 23, 2015 - June 29, 2015	Vinaros, Spain
17th International Conference of the Catalan Association for Artificial (**CCIA 2014**)	October 22, 2014 - October 24, 2014	Barcelona, Spain

2
Related Work

The overall aim of this thesis is to contribute to literature on PRS. Specifically, our focus is to represent people's profiles in a manner more expressive of their preferences and define matching systems which consider priorities in order to assist recommender systems in obtaining their matching goal. This chapter attributes a brief discussion of related work to both directions within the framework of people recommendation.

2.1 Introduction

There has been considerable work in the area of PRS. In particular, systems for identifying experts to fulfill a task have been proposed for identifying a candidate for a job [95], determining the right reviewer for a paper in a conference [149], or finding an expert to help with a problem [175]. Personalization of the recommendation can assist with information overload associated with decision-making by customizing information for individuals [67]. Personal preferences can be inferred from information pertaining to a person's actions and captured in *user profiles*. Depending on the domain of the people RS, a user's profile may consist of preferences, interests, skills, or areas of expertise. People are recommended to tasks based on their profiles and the matching objective.

To provide context to PRS and connect the concept of creating a profile with its end objective of defining a match, we have adapted the user profile framework introduced by Gauch et al. [67] to PRS. Figure 2.1.1 illustrates the adapted framework that is reviewed in this chapter. This framework is divided into five phases: 1) define the matching problem, 3) define the profile elements, 4) collect data, create the profile, and 5) perform the match. The first phase,"Define the matching problem", is understanding the matching problem to be solved, identifying the actors and tasks, and the requirements for an item and individual to be considered a match as defined by the user. The second phase, "Define profile elements", is about defining the features of the task which need to be satisfied and the preferences of the individual to be assigned in terms of the matching constraints set by the user. For example, if the amount of time an individual is required to have available to perform a task is a constraint, it is an additional feature for the individual's profile. In addition, based on the type and strictness of the matching defined by the user, different representations of the features and individual's preferences can be considered. It is apparent that the first and second phases happen together. The user in this sense may be the end user, owner of the system or both. The third phase, "Collect data" refers to techniques used to collect data about individuals. Elements of this phase include uniquely identifying the individual and information collection. The technique may be implicit, explicit or a hybrid of both. Data is collected with respect to the elements uncovered in the "Define profile elements" phase. The "Create a profile" phase refers to methods for defining and representing an individual's characteristics and preferences and the features of the task to be performed. Lastly, the "Perform a match" phase exploits the user profile to provide personalized services based on the requirements of the matching specified by the user. Therefore, it can be an iterative process between the user and the researcher. Where the outcome of the matching is not in accordance with the expectation of the user, adjustments to the requirements may be made. These systems assist people with finding experts for consulting [15, 59], for reviewing research projects

[143], and for collaborating within organizations [160].

In this thesis, we develop and test two solutions to real problems related to PRS to address the issues discussed in Chapter 1. Specifically, in Chapter 3 and 4, we follow the defined framework to match reviewers to papers in a conference environment and internships to students in a university environment, respectively. Each is a mirror image of the other's problem. Therefore, different requirements for a match are highlighted impacting the "Define the matching problem" and "Define profile elements" phases. Furthermore, the context around each problem influences the sources from which data may be collected and the information selected in the "Collect data" phase. Profiles in each case have been represented according to the problem. For instance, in Chapter 4 student feedback is requested regarding his auto-generated profile and thus should be displayed in an interpretable manner such as linguistic terms. Assignments are made based on the requirements of the problem which differ based on the perspective of the user. Emphasis was placed on the preferences of the user. Chapter 5 departs from the previous two matching problems and focuses on the development of profile elements. It focuses on defining an element of an individual's profile, lifestyle. Following the structure of the previous cases, it leverages publicly available data to minimize intruding on an individual in the "Collect data" phase. Similarly, features and preferences of each individual is determined and represented in the "Create profile" phase. Others sharing in the individual's lifestyle are identified in the "Perform match" phase. Understanding clusters derived from this lifestyle can enhance the profiles developed in the second case or be applied to its own application in which recommendations from people of similar lifestyles are preferred. For instance, when seeking travel advice, those traveling with families may place more value on suggestions given by others traveling with families than a single person who travels with friends. Therefore, identifying "answerers" with comparable lifestyles may be relevant for PRS.

In the following subsections we summarize the related work according the "Collect data", "Create Profile", and "Perform a Match" phases as shown in Table 2.1.1. As many of articles reviewed do not reference the first and second phases, they are not assessed here. Specific state-of-the-art literature pertaining to each use case

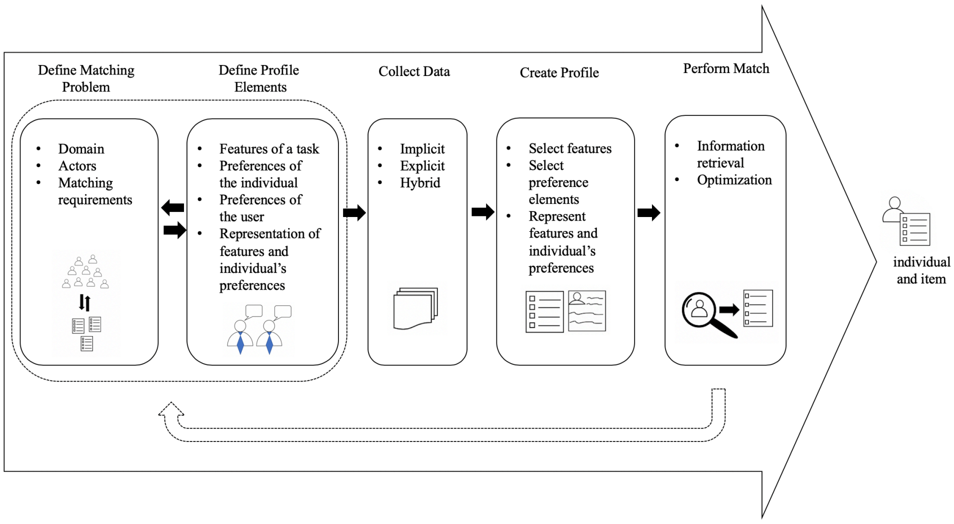

Figure 2.1.1: Personalization and assignment for people RS

addressed in the thesis will be discussed in its corresponding chapter.

Table 2.1.1: Summary of related work

Article	Year	Use case	Collect data	Create a profile	Match method
Yang and Huh [175]	2008	identify experts in organizations (KMS)	implicit	multi-criteria, prob. dist, keyword extract.	not defined
Liao et al. [94]	2009	recommend sources from digital library	implicit	concept, prob. dist., keyword extract.	heuristic
Kelemenis & Askounis [86]	2010	select job candidates	explicit	multi-criteria, linguistic or fuzzy, hetero. sources	aggregation-fuzzy
Li et al. [92]	2011	identify experts in organizations (KMS)	explicit	concept, linguistic or fuzzy, manual	aggregation-fuzzy
Dror et al. [52]	2011	routing questions	hybrid	multi-criteria, numerical, keyword extract.	machine learning
Tang et al. [149]	2012	assign reviewers to conference papers	implicit	concept, prob. dist., topic modeling	optimization
Charlin and Zemel [42]	2013	assign reviewers to conference papers	hybrid	concept, prob. dist., topic modeling	optimization
Li and Watanabe [93]	2013	assign reviewers to conference papers	not defined	multi-criteria, not defined, manual	optimization
Luukka and Collan's [101]	2013	select job candidates	explicit	multi-criteria, linguistic or fuzzy, hetero. sources	aggregation-fuzzy
Liu et al. [97]	2014	assign reviewers to conference papers	implicit	multi-criteria, prob. dist., topic modeling	optimization
Das and Goçken [48]	2014	assign reviewers to R&D proposals	explicit	concept, linguistic or fuzzy, manual	aggregation-fuzzy
Amini et al. [10]	2014	recommend scientific articles to scholars	implicit	concept, numerical, hetero. sources	Not defined
Tejeda et al. [151]		recommend sources from digital library	explicit	multi-criteria, linguistic or fuzzy, manual	aggregation-fuzzy
Gupta and Garg [73]	2014	select job candidates	hybrid	multi-criteria, numerical, manual	heuristic
Amini et al.[11]	2015	recommend scientific articles to scholars	explicit	concept, numerical, ontology	Not defined
Liu et al. [95]	2015	select job candidates	hybrid	keyword, numerical, hetero. sources	aggregation
Liu et al. [96]	2016	assign reviewers to R&D proposals	explicit	concept, linguistic or fuzzy, manual	heuristic
Protasiewicz et al. [126]	2016	assign reviewers to R&D proposals	hybrid	keyword, numerical, keyword extract.	heuristic
Shon et al. [141]	2017	assign reviewers to R&D proposals	hybrid	keyword, linguistic or fuzzy, keyword extract.	aggregation-fuzzy
Guo et al. [72]	2017	identify features of customer satisfactions	implicit	concept, prob. dist., topic modeling	not defined

Continued on next page

Table 2.1.1 – continued from previous page

Article	Year	Use case	Collect data	Create a profile	Match method
Rahimi et al. [127]	2017	classify diner's restaurant preferences	implicit	concept, numerical, NLP	machine learning
Martinez-Gil et al. [107]	2018	recommend jobs to candidates	not defined	multi-criteria, numerical, not defined	machine learning

2.2 Collect Data

Elements of the "collect data" phase include the source of information from which and manner in which personal knowledge data is collected. In general, RS profiles are assumed to be created for the system user and item in order to match an item according to the preferences of the user. However, it is important to note that, in some situations the user is not the individual being matched. For example in human resource management, a talent manager is looking to match a job candidate or an employee to a job. In this thesis, we name the person about whom the profile has been created and refer to the person requiring a recommendation as the user of the system.

Before discussing specific methods to learn expertise profiles, we would like to note that there has been significant work in the field of user modeling. Previous work has considered cross-platform modeling to address sparse user profiles. Specifically, one paper aggregated tags and form data on one set of social web sites [4] in order to fill in a user's profile for another site. Another paper, [61] proposed that user personality is available and can be used to better leverage cross platform data in order to provide recommendations. Although previous methods relied on user input through ratings, form completion, tagging, click-through data or consumption, ubiquity of personal technology such as smart mobile devices and wearables present the possibility of ubiquitous personalization, another area in user modeling [87]. Whether implicitly or explicitly collected, user modeling obtains much of its insight into user behavior through repeated system-user interaction. In contrast, the methodologies presented in this thesis refer to single-event matchmaking, where repeated interaction is limited. Moreover, the real case environments selected require knowledge of a user's expertise, an area not commonly exhibited on social networks. However, some articles evaluate code reviewers' [152] and community contributors' [52] abilities through repeated interaction and votes. Therefore, we reserve exploration and incorporation of user modeling for future research.

Methods to learn profiles have been categorized into *implicit* and *explicit* methods [32, 85, 130]. Information is collected explicitly when a person is directly asked to provide information about their preferences or the preferences of others.

Different types of explicit elicitation methods have been considered. Some systems asked people directly for their preferences. In Liu et al. [96] reviewers were asked to fill in a form related to their discipline areas and those of their published papers. Scholars were invited to provide articles related to their research interests in Amini et al. [11]. In other studies the input of outside decision makers was solicited to assess candidates according to their own area of expertise [86, 101]. Another technique was to ask users to evaluate an item's relevance with respect to their own preferences [48, 151]. Li et al. [92] combined two methods to recommend experts in an organization. Respondents evaluated candidate experts with respect to knowledge areas and the expert seeker rated documents deemed relevant to his problem at hand. Capturing knowledge with explicit techniques suffers from a knowledge acquisition bottleneck [46]. These methods require people's inputs appearing intrusive [112]. Thus, people may be unwilling to spend time answering questions regarding their profiles [32] or may not participate due to privacy concerns [67].

Implicit elicitation techniques can collect information about people's preferences, interests and tacit knowledge without their active participation providing a less intrusive method of knowledge acquisition. For example, there has been much work to improve scholars' profiles by engaging knowledge driven approaches which extract scholars' interests from textual content [10]. Topic modeling has been applied to learn the topics over published papers of reviewers and submissions [42, 97] and customer reviews [72, 127]. Guo et al. [72] applied topic modeling to identify dimensions of customer satisfaction in text reviews written by hotel visitors. Rahimi et al. [127] applied natural language processing techniques to restaurant reviews to infer which restaurant features were preferred by different stages of romantic relationships. Amini et al. [10] proposed to collect scholar's profiles and Tang et al. [149] proposed to collect reviewer's publications from the bibliographic database, Arnetminer [1]. The tool automatically identifies and extracts profiles from the Web by using social network analysis (SNA) and information extraction techniques. Liu et al. [97] extracted co-authorship information from Microsoft Academic Search system [2] to establish relationships between reviewers

[1] https://aminer.org
[2] https://academic.microsoft.com

and authors. Other sources rely on people to contribute artifacts to a database from which the system may subsequently extract information. Yang and Huh [175] proposed to identify expertise within an organization by analyzing knowledge artifacts contributed by employees. However, this technique is limited by the number of documents registered to a KMS [92]. Preferences can also be inferred from prior transactions using implicit methods [32]. Liao et al. [94] proposed to mine university library patron's interests from library loan records. Given that information seekers may be uncertain of their needs [109], implicit methods may be an appropriate means to glean these requirements. Nevertheless, transaction information may be insufficient given the nature of single-event matching in PRS. This gap is covered under the umbrella of objectives 1 and 2 which seek to elicit requirements implicitly and expose information seekers to relevant items beyond their search criteria.

Implicit methods may be preferred over explicit methods when people are only able to express their feedback in this manner [129]. In contrast, it may be preferable to collect information explicitly for new users to a platform to establish an initial profile of the user. Hybrid techniques benefit from both implicit and explicit elicitation. Information collected implicitly from databases or web pages can be supplemented with explicitly furnished information enabling the profile to be more refined and current. Protasiewicsz et al. [126] collected information about scientific publications from open access databases such as DBLP [3], personal web pages, and reviewer supplied documents. Charlin and Zemel [42] learned reviewers' topics of expertise from published papers submitted by the reviewer or crawled from Google Scholar[4] profile and supplemented it with reviewer self-assessments. Dror et al. [52] obtained user attributes from user interaction with questions on Yahoo! Answers[5], and explicit preferences specified such as keywords, categories, and people whom he was following. Shon et al. [141] extracted proposal and publication information along with user and reviewer supplied keywords from proposal and reviewer databases. Liu et al. [95] required faculty applicants to state their work experience, submit documents related to published research and col-

[3] `https://https://dblp.uni-trier.de`
[4] `https://scholar.google.com`
[5] `https://answers.yahoo.com`

lected areas of research from an expert database provided by the university and social network websites. Gupta and Garg [73] assumed either the candidate or the system would specify the preference and demographic data.

Finally, let us note that, in general, few papers focus on integrating multiple sources of information. The limited number of information sources considered can result in a profile with an incomplete view of a scholar [10]. Scholars partake in a variety of activities including formal education, studying articles, and authoring and exhibit knowledge in different forms such as homepages, blogs, and online communities. Different sources can help to complete and add dimensions to profiles enabling them to be refined without eliciting explicit input. Each source can provide a different point-of-view about a person's preferences [17]. Some systems extracted text from the corpus of publications and homepage content [10, 126], and curriculum [10]. Due to the challenges of obtaining publications from digital libraries like authorization and availability of full text, the authors engaged scholar profiles from digital libraries. These scholar profiles contain abstracts of the publication and reviewer information. Other systems included self-ascribed keywords [95] and self-provided documentation [95, 126]. Another technique employed multiple decision-makers to score candidates according to the criteria within their area of specialty[86, 101]. Despite its benefits, extracting knowledge items from different sources illuminates inconsistency and heterogeneity [126]. For example, reviewers may have multiple profiles on a website owing to how his name was written on the published article. Another example may be that one source characterizes a person as having very high expertise in one area while another says the opposite due to the different sources of information from which the website is making the determination. Therefore, seeking information from multiple sources requires knowledge unification and transformation to reconcile these discrepancies [10, 126]. Regarding the "collect data" phase there is a gap related to identifying knowledge from heterogeneous sources using implicit elicitation methods. This gap is covered under the umbrella of objective 1.

2.3 Creating a profile

A fundamental aspect of a RS is the source and type of information it will employ [32]. Ideally, RS should have information regarding the preferences of the active user and features describing items. Creating a profile encompasses the process of determining the preferences of the active user and features describing items and their representation. Representation includes descriptors used to define knowledge and how it is expressed to enable comparison to search requirements and/or preferences.

Keyword profiles contain keywords which represent the topics of interest to the user and weights to express the user's level of interest with respect to each topic [126, 141] or the recency of the interest [95]. Keyword profiles are the simplest to build. However, they have to capture and represent all the words that may be of interest to a person in the future [67]. To accomplish this task, considerable user feedback is required in order to learn the terminology by which a topic might be referred. Complicating matters, keyword-based systems have inherent challenges originating from natural language ambiguity. Polysemy [67, 98], the existence of multiple meanings for a word which can cause the wrong document to be deemed relevant [98]. Synonymy, a problem of multiple words having the same meaning which can cause documents to be missed if exact words are not used in both the document and the profile [98]. Furthermore, these systems are unable to capture the semantics of user interests because they primarily rely on string matching operations [98].

Semantic analysis, such as semantic networks, can solve these problems [67]. Keywords co-occurring within documents of interest to the user are linked to specific concepts and are associated with a weight which characterizes a person's interest in it. Because semantic profiles explicitly model the relationship between words and concepts, it can better manage the ambiguity of natural language implicit in polysemy.

Similar to semantic network profiles, concepts profiles are represented as conceptual nodes and the relationship between them. In concept profiles, these nodes are abstract concepts rather than keywords or sets of related words. A common method to represent a concept profile is as a vector of abstract concepts and their associated weights according to the user's interests. People describe their expertise as a combination of several topics [113] making modeling expertise with respect to the topics of published work critical to the assignment process. Different techniques have been proposed to represent expertise such as topics of research [10], areas of discipline [96], areas of expertise [42, 92, 149], and degree of relevance with a task [48, 92]. Liu et al. [96] characterized papers by areas of discipline and Li et al. [92] applied a degree of relevance to represent knowledge areas in which the user was seeking more information. People also describe items which appeal to them in multiple dimensions such as hotel [72] and restaurant features [127].

For some people assignment problems, criteria beyond area of knowledge is required in the selection of a person. Therefore, multi-criteria decision making (MCDM) analysis has been considered in these types of problems to determine the overall preference among alternative options [86]. A decision is made by evaluating each alternative based on a set of criteria. Each criterion is measured for each alternative and forms the basis for comparison by sorting or ranking of the alternatives. Criteria may refer to features of an item or to evaluation measures upon which an item is rated. Some systems considered relevance of an item plus quality of expertise [93, 151] quantity of expertise [93], recency [93, 175] frequency of contribution, and usefulness to other users [175]. Liu et al. [97] proposed to balance expertise, recognition from the scientific community, and diversity of reviewers' research interests. Other systems considered platform interaction. Dror et al. [52] employed question attributes consisting of the question title, body, best answer, other answers, topic category, and the user performing each role. User attributes contained the type of interaction in which the user engaged with a question.

Criteria explored also delved into more personal areas. Martinez-Gil et al. [107] predicted candidates interests according to their preference towards very high salary jobs, jobs located near home, high hourly rate, and big companies located in large cities. However, each interest was predicted individually, only the last one, which considered the size and location of the company, was predicted as a multi-attribute. Gupta and Garg [73] proposed to describe a job candidate by his age, gender, marital status, university major, degree, grade, experience, current location, and skills and a position by its company status, industry, position level, and pay scale. Although these elements provide a broad sense of the candidate, they may not be generalizable to environments where collection of this data is prohibited by equal opportunity employment laws.

For situations such as the human resource management problem, decision-makers face difficulties in assigning crisp values to criteria. Therefore, scholars tend to extend typical MCDM methods to fuzzy environments. Decision-makers have been allowed to select their own evaluation criteria, score candidates via fuzzy numbers [101], and select criteria weights [86].

Preferences for criteria are usually expressed as *measurable*, *ordinal*, *probabilistic* or *fuzzy* [8]. Measurable refers to a criterion which can be quantified on a scale. Expertise may be represented as a binary variable for absence and presence [73], weights [10, 11, 95, 126] or frequency, based on keywords or concepts [52, 127]. Attributes besides expertise may be considered. Martinez-Gil et al. [107] referenced number of employees, number of citizens, distance to home, salary, and work hours.

Probabilistic criterion are represented as probability distributions. Topic distribution of reviewers and papers [42, 97, 149], and customers [72] have been considered. Other papers have applied term or keyword distributions to create expert's profiles [175] or item profiles [94]. Frequency of keywords [94] or co-authorship [97] appearing in text representing interests have been adapted, as well.

Fuzzy is a criterion which is expressed in terms of its possibility to belong to a qualitative interval. The majority of developed systems consider measurable criteria. However, probabilistic and fuzzy criteria may better reflect the uncertainty in people's preferences. Liu et al. [96] expressed a reviewer's discipline as a binary variable but labeled the level of expertise in linguistic terms and assigned the labels values of 1, 2, and 3, respectively. Rather than representing linguistic terms in terms of crisp values, other methodologies expressed them as triangular fuzzy numbers [48, 86, 141], or 2-tuple linguistic values [92, 151]. Other methodologies directly applied triangular fuzzy numbers [101]. Although these papers reflect the expressed opinions in fuzzy terms, they were acquired through explicit or hybrid means. However, as previously mentioned, implicit elicitation techniques can provide a less intrusive method of knowledge acquisition. This gap is covered under the umbrella of objective 1 which seeks to define preferences in an implicit manner from unstructured information and express them in terms of their natural uncertainty.

Keyword-based profiles are created by extracting keywords from information collected from sources such as publications, resumes, and web pages [67]. Keyword weighting is performed to identify the most important keywords. Keywords have been identified using parts of speech tagging [52, 126] and TF-IDF [94, 141, 175]. TF-IDF has been applied to create the profiles for both sides of the matching in [94, 141]. Dror et al. [52] performed a process similar to TF-IDF to determine the distribution of topic categories over the terms. As previously discussed, there are inherent challenges to keyword-based profiles.

In recent years, ontology-based approaches have been applied to user profiling [11]. In general, profiling approaches leverage a domain ontology and learn scholars' preferences considering contextual information. Ontology, a conceptual framework containing concepts of a domain, their relationships and attributes [68] enables user interests to be inferred and applied to user profiles in recommender systems [112]. Amini et al. [11] proposed a method for profiling scholar's background knowledge by integrating multiple domain taxonomies into a reference ontology for the computer science domain in order to represent scholar's preferences. There were 747 topics represented in the final ontology and a case example built profiles for 25 scholars. Given a real implementation in a RS, the large number of topics may lead to a sparse matrix of topics to reviewers making it difficult to identify reviewers and items having high degrees of similarity. Moreover, creating a hierarchy is labor intensive and can be become costly [29].

A document typically encompasses multiple topics in different proportions. Topic modeling is a statistical model for discovering abstract topics occurring in a collection of documents. A common method of topic modeling in expert RS is Latent Dirichlet Allocation (LDA). Tang et al. [149], applied extensions of LDA to compute a matching score between each reviewer and paper employing a language model-based retrieval and Author-Conference-Topic modeling method. LDA has also been applied to learn the topics over the published papers of reviewers and submissions [42, 97]. Guo et al. [72] applied LDA to customer reviews and identified key dimensions of customer service expressed by hotel visitors. Rahimi et al. [127] applied natural language processing (NLP) techniques to reviews and inferred restaurant features from their associated nouns and adjectives.

Topic modeling algorithms [22] are able to discover a set of topics from a collection of documents. A topic is a distribution over a set of terms surrounding a central theme. They are an interpretable, low-dimensional representation of documents [40] and have been applied to corpus exploration, document classification, and information retrieval. Throughout the use cases in this thesis, we exploit the discovered topic structure of text to develop profiles, therefore, we review its concepts here.

Specifically, documents are represented as probability distributions over a mixture of topics and topics are probability distributions over a mixture of words. Let us assume K is the number of topics, each topic $\beta_k, k \in \{1, \ldots, K\}$ follows a Dirichlet distribution η. The generative process of LDA considers for each document $D_d\ ,\ d \in \{1, \ldots, M\}$ the following steps:

1. Draw the topic distribution θ_d for document D_d, considering α, a parameter of Dirichlet prior on the per document-topic distribution: θ_d ~Dirichlet(α)

2. For each word $w_n\ ,\ n \in \{1, \ldots N_d\}$, where N_d is the number of words in the document D_d:

   (a) Draw topic assignment $z_{d,n}$ for word w_n of document D_d from document's multinomial topic distribution: $z_{d,n}$ ~Multinomial(θ_d)

   (b) Draw word w_{n_d} from the topic's multinomial distribution: w_{n_d} ~Multinomial($\beta_{z_{d,n}}$)

Each document is a mixture of topics. The topic proportions are specific to each document. However, the set of topics are shared by all documents in the corpus. Each topic is a distribution over a fixed vocabulary and each word is drawn from a topic. A graphical model for LDA is presented in Figure 2.3.1.

Although LDA is the simplest topic model [23], it has been used in people RS to identify areas of expertise and to make recommendations, as discussed in the previous section. LDA has several advantages. First, it is an unsupervised topic modeling method used to learn underlying topics in a collection of textual documents [23]. This aspect is useful in RS to identify documents similar to ones the

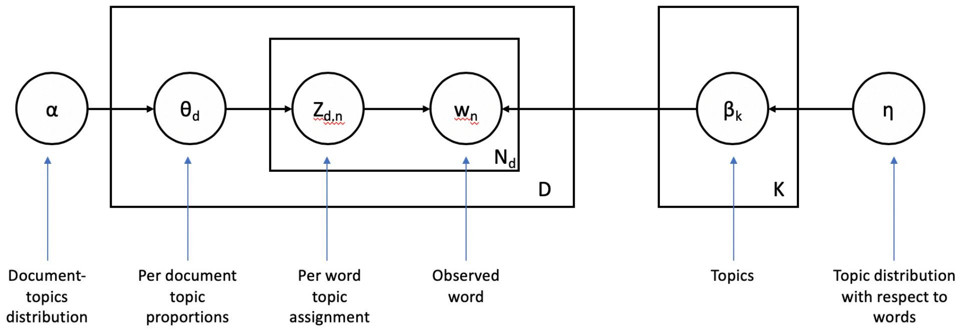

Figure 2.3.1: Graphical model representation for LDA [23]

user likes by generalizing unrated items [160]. Second, LDA does not limit a document to being described by a single topic, rather it allows for a mixed-membership between multiple topics leading to clearer estimates of word co-occurrance patterns.

2.4 Performing a match

Although there is some disagreement as to how to categorize existing methods for expertise matching, most studies suggest information retrieval and optimization as the two main groups [161, 169]. Information retrieval techniques compute the matching degree between a person and item [53, 134, 169]. Optimization techniques solve the problem from a mathematical or operational research perspective [149]. Within the optimization modeling group we can distinguish exact techniques and approximate ones. Exact optimization is obtained when it is possible to compute the optimal solution given by a fitness function. Approximate optimization is found when a solution is near to optimal and computed through iterations and aggregation functions. The methodologies that are developed in the subsequent chapters of this thesis focus on approximate optimization methods. Therefore, the papers reviewed in these related works concentrate in this area.

The assignment problem is a traditional problem in Operations Research and has been studied extensively [147]. Methods of exact optimization have been based on integer linear programming, [83], and minimum cost flow [74], and Hungarian algorithms [93]. Li and Watanabe [93] defined the matching problem as a combinatorial optimization problem and proposed an adaptation of the Hungarian algorithm to work with the constraints imposed on the matching problem. Liu et al. [97] formulated the reviewer assignment problem as a optimization framework that integrated random walk with restart and a sparsity constraint to obtain a balance of expertise, authority and diversity. Constraint-based optimization with the objective of maximizing the matching between reviewers and papers while satisfying general constraints has been employed in [42, 149]. In principle, since the set of possible solutions is finite, any combinatorial optimization problem could be solved exactly by enumerating all the outputs of the objective function and identifying the elements corresponding to the best value [62]. However, the number of feasible solutions can grow exponentially and may not be practical in all applications. In these cases, approximate optimization methods offer an alternative.

One manner to perform approximate optimization is machine learning. Martinez-Gil et al. [107] compared the results of random forest and Support Vector Machine (SVM) to recommend jobs to candidates. Dror et al.'s [52] methodology compared each user and question pair according to their attributes and evaluated their matching with Gradient Boosted Decision Trees. Rahimi et al. [127] experimented with principal component regression, partial least squares, least absolute shrinkage and selection operator (LASSO) regression, random forest, and regression trees to correlate restaurant features with stages of romance. These types of techniques may be beneficial in generalizing people when known personal preferences are limited. In addition, the latter two papers leveraged large data sets from commercial sites and incorporated information from multiple sources (perspectives) within them; restaurant ratings and customers reviews [127], and questions and answers [52]. On the other hand, these methodologies employed supervised learning, requiring a training set which may not always be available. Other methods of approximate optimization include keyword matching, latent semantic indexing [53], and topic modeling [84, 113]. As these methods are mainly for textual analysis, the multi-criteria nature of the people assignment problem is not a primary focus. This gap is covered under the umbrella of objective 1 where a methodology is developed which generalizes people's lifestyles according to their textual reviews and item attributes applying an unsupervised method of clustering.

Optimization methods are suited for well-defined problems [126, 148]. However, less strict algorithms may fare better for ill-defined problems. For these problems, not every constraint or variable may be known. In these cases heuristic algorithms or artificial intelligence methods have been found to be effective. Moreover, heuristic approaches may be helpful for solving optimization problems that are hard to approximate [62]. Examples are a greedy randomized search with genetic algorithm [148], a greedy and evolutionary algorithm [111], and fuzzy sets [92].

In heuristic methods, similarities are computed between users by aggregating similarity of an individual criterion or using multidimensional distance metrics [8]. Common techniques include correlation-based and cosine-based methods. These methods compute the adequacy of an item for a user based on observed data and in general, heuristic assumptions. Cosine similarity measure has been utilized to compare words in a search query to those describing a reviewer's expertise [126] and desired skills of a position and those of the candidate [73]. Other works applied rules such as assigning reviewers according to the primary discipline area of a paper [96] and computing the average weights for keywords shared between a book and those borrowed by a library patron to signify the patron's preference for the book [94].

Although some of the papers previously discussed considered multi-criteria profiles and by extension the matching process they did not consider topic coverage. It is criterion which has been implemented within the context of conference reviewer assignment and refers to the aspect that several topics may be discussed in a paper [83, 149]. A set of reviewers assigned to a paper should complement each other such that they are able to cover as many topics as possible within the constraints of the assignment problem. Given its importance, some papers focused on the criterion of topic coverage [83, 149] but did not expand it to a multi-criteria matching problem. This criterion is not limited to the assignment of reviewers but could be considered in other domains such as team formation where distributed expertise is advantageous, setting up the inverse criterion. In this direction, objective 2 focuses on the allocation of reviewers combining heuristic and aggregation methods. An heuristic method is developed to support coverage need while an aggregation method is employed to avoid eliminating candidate reviewers who do not match topics in totality prematurely.

Aggregation of preferences, criteria or similarities can happen at different stages in RS. It can occur at interim stages to match elements of profiles with items or during the final stages to match entire profiles. These types of functions take multiple variables as inputs and fuse them into a representative output [21]. The main families of aggregation include generalized means, Choquet and Sugeno integrals, Ordered Weighted Averaging (OWA), triangular norms and conorms (t-norms and t-conorms), and bipolar aggregation functions. Aggregation functions are generally implemented in collaborative filtering RS to aggregate ratings or preferences of similar users and ascertain user similarity. Content-based filtering may use aggregation functions in item score computation, similarity computation, and construction of profiles.

Liu et al. [95] applied a cosine similarity to measure the relevancy of a reviewer and university application based on previous project and publication. Then, a comprehensive score was computed multiplying the conflict of interest, relevancy, and quality scores. Li et al. [92] employed a linguistic weighted average operator to calculate the similarity between an expert and information seeker. Kelemenis and Askounis [86] set a veto threshold for each criterion and decision maker. The distance between each candidate and the vetoes of all criteria was calculated based on the steps of fuzzy technique for order preference by similarity to an ideal solution (TOPSIS). The candidate with the greatest positive distance from the vetoes was preferred. Tejeda et al. [151] proposed a switching hybrid approach. When a new source was added to the library, a content-based recommendation approach was executed which computed the cosine similarity of the linguistic values between the source's disciplines and patron's preferences. However, when a new patron was registered, a collaborative-based recommendation was made based on a nearest-neighbor algorithm. Next, the quality rating and relevance rating of a source were aggregated into a single score via a fuzzy linguistic operator and the recommended sources were re-ranked. Shon et al. [141] computed the similarity between a proposal and categories, and a reviewer and categories to derive the similarity between a proposal and reviewer. Das and Goçken [48] implemented the signed distance method and ranking of fuzzy numbers with integral value to match reviewers and papers. Luukka and Collan [101] introduced Fuzzy Heavy Ordered

Weighted Averaging (FHOWA) to aggregate the scores of the the criteria given by the decision-makers. Given that the aggregate score was a fuzzy number, a fuzzy similarity measure was proposed to compare each candidate to the ideal solution. Afterwards, candidates were ranked, accordingly.

Finally in this section, we provide a review of OWA operators and the implementation of quantifier guided aggregation methods. These methods are useful when the solution to a problem does need not satisfy all the criteria, rather only a portion of the criteria needs to be satisfied. In addition, the criteria are ordered differently for each occasion. According to Yager [172], the OWA operator has been considered an important aggregator in MCDM primarily for its ability to represent linguistic quantifiers. The author explains that its ability enables the use of OWA operators in quantifier guided aggregation allowing decision makers to express their criteria in natural language. This is the context in which the first two use cases are presented for the conference reviewer assignment problem and personnel assignment problem. Therefore, OWA plays an important role in the matching processes for this thesis and is discussed here.

OWA functions associate a weight with the relative order of the input in comparison to the other inputs. Yager introduced the family of aggregation operators called OWA operators in [170]. In general, there are three steps in the OWA aggregation process.

1. Reorder the input arguments. In this way, "the weights are not associated with a particular argument but with the ordered position of the arguments" [171]. This operation introduces a non-linearity into the integration process [172] differentiating it from the weighted averaging operator.

2. Determine a weighting vector for the operator

3. Use the weights to aggregate the reordered arguments to evaluate each alternative

The result of an OWA operator is a weighted aggregation such that the weights associated with each criterion depends on the order of the values of the criteria.

There are different methods to obtain the weight vector W. For our purpose we will use a linguistic quantifier guided aggregation, a process in which the decision maker selects a quantifier representing the proportion of criteria necessary for a good solution [171]. In Zadeh [180], it is proposed to represent linguistic quantifiers as fuzzy sets. If the fuzzy set Q satisfies $Q(0) = 0$, $Q(1) = 1$ and Q is increasing, then Q is defined as a Regular Increasing Monotone (RIM) quantifier. Included in these quantifiers are “all”, “most of”, and “many”.

Given a regular increasing monotone (RIM) quantifier, $Q : \mathbb{R} \to \mathbb{R}$, the vector of weights $W = (w_1, \ldots, w_p)$ can be defined as follows [170, 171]:

$$w_h = Q\left(\frac{h}{p}\right) - Q\left(\frac{h-1}{p}\right), \; h \in \{1, \ldots, p\}. \tag{2.1}$$

Example 2.4.1 *Let us assume that there is set of criteria {0.80, 0.093, 0.028, 0.576, 0.777}. A decision maker prefers that “most of” the criteria are satisfied. Applying an OWA operator guided by the RIM quantifier “most of”, consider $Q(r) = r^{1/2}$. The corresponding weight vector from Equation 2.1 is (0.447, 0.185, 0.142, 0.120, 0.106). The aggregation of the criteria values gives:*

$\varphi_{mostof}(0.80, 0.093, 0.028, 0.576, 0.777) = 0.80 \cdot 0.447 + 0.093 \cdot 0.185 + 0.028 \cdot 0.142 + 0.576 \cdot 0.120 + 0.777 \cdot 0.106 = 0.598$

Figure 2.4.1 depicts some examples of RIM functions of $Q(r) = r^a$ on the top, and their corresponding vector of weights on the bottom. These RIM functions guarantee that all the criteria contribute to the final aggregated value because they are strictly increasing functions. Note that the concave (convex) property of Q provides decreasing (increasing) weights. The graph corresponding to $a = 1$, gives us equally-valued weights and therefore, represents the mean operator.

The OWA operator has several desirable traits for an aggregation operator. As shown in [170] these functions are idempotent, symmetric, and strictly monotone assuming all weights are greater than zero.

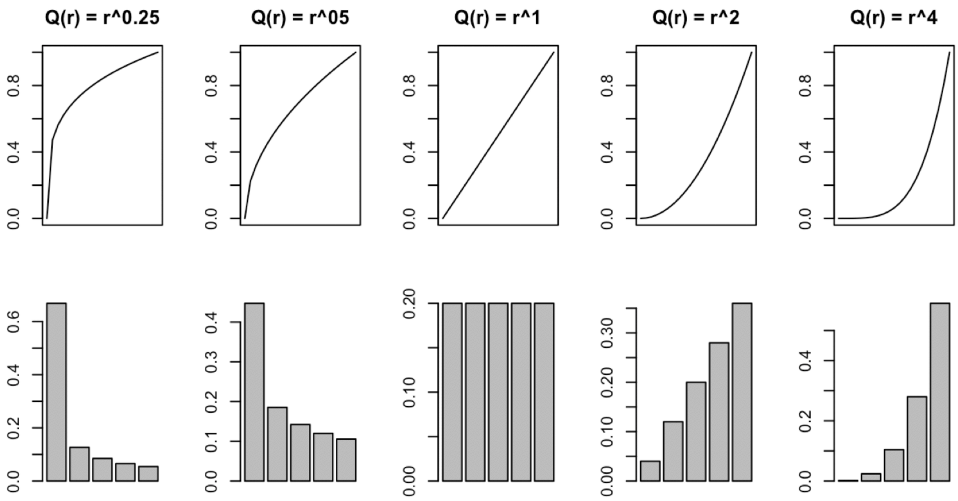

Figure 2.4.1: RIM functions and their corresponding weights [171]

3

A decision support tool using Order Weighted Averaging for conference review assignment[1]

3.1 Introduction

Assigning papers to reviewers is a non-trivial task for conference chairs and scientific committees. The task requires an optimal matching between reviewers and papers. To accurately accomplish this task, knowledge of reviewer expertise and paper topics are required. Often these assignments must be made within days after a submission deadline creating a huge burden on the conference chairs. This article goes in the direction of assisting conference chairs with matching a paper

[1]Parts of this chapter contributed to the article published in Pattern Recognition Letters by J. Nguyen, G. Sanchez-Hernandez, N. Agell, X. Rovira, and C. Angulo (https://doi.org/10.1016/j.patrec.2017.09.020)

and reviewers.

A number of academic research and commercial software have tried to address the automation of the reviewer assignment problem [42, 47, 132]. Reviewer preferences or bids are used to represent reviewer research interests. However, some shortcomings can be associated with the bidding process. A reviewer may bid on papers for their novelty rather than their alignment with his/her research interests [17]. In addition, reviewers may search for papers using keywords and bid on papers returned in their search rather than considering all the papers in the conference [42].

Some systems determine reviewer expertise from their publications or webpages [17, 42, 123]. This approach could help avoid the shortcomings from the bidding process. Websites like the ORCID[2] registry allow researchers to create a research profile with employment history, funding, publications, bibliography, and links to external websites. In addition, scientific indicators like ORCID facilitate the process of identifying reviewer's work. Finally, other systems obtain reviewer expertise by directly asking reviewers to select their areas of expertise from a predefined list of topics [74, 81, 148]. A pre-condition to our proposed method is to leverage publicly available information, like those mentioned above, to create reviewer profiles.

In order to provide conference chairs with an overall view of a reviewer's expertise, we propose to build a profile for each reviewer consisting of seven features. Five of these dimensions are aggregated into a single quality score representing a reviewer's publishing accomplishments. The sixth variable corresponds to a second score representing a reviewer's areas of research and the third score, the seventh variable, *recency*, refers to papers published in recent years. To reduce the amount of time required from reviewers, we propose to create profiles from information extracted from public web pages. As this process can be completed at any time, conferences can develop and update profiles in advance of the conference paper assignment process.

As argued in [150], there exists imprecision associated with reviewer expertise levels. However, often in prior studies, reviewer expertise across different domains has been considered as a crisp set. As our information comes from multiple

[2] https://orcid.org/

sources, an additional natural uncertainty exists. Therefore, we consider an Ordered Weighted Averaging (OWA) aggregation function [170] to summarize the information coming from different sources. According to Torra [153], an OWA enables each source of information to contribute equally to the final solution. It weights the values rather than the sources of information because each weight is attached to an ordered position and the values in each of the positions are determined, in our case, by their decreasing order regardless of their originating source of information. The OWA places emphasis on the most exhibited variable. This aspect of the OWA is in contrast to the weighted mean which assigns a weight to the value obtained from a source of information and can adjust for the reliability of each source. Candidate reviewers are ranked according to an overall score which is determined by an OWA operator applied to features in their profile and an availablity indicator. These features are limited to the topics of the paper being assigned, the recency, and quality scores. The candidate with the highest overall score is assigned to review the paper.

The rest of the chapter is organized as follows. First, in Section 3.2 we provide a review of related work. Next, in Section 3.3 we summarize the OWA operator and its associated weight function. In Section 3.4 we explain the proposed methodology for defining paper and reviewer profiles, and matching papers to reviewers. Then, in Section 3.5 we provide a simulated case example using data from some conferences. In Section 3.6, we evaluate our results from a real case example. Finally, in Section 3.7 we discuss our conclusions and future work.

3.2 Related Work

In this section, we review and compare related research on the *reviewer assignment problem*. Specifically, we characterize the existing literature according to four dimensions: Reviewer profile, Paper profile, Matching method, and Case implementation.

The first dimension, *Reviewer profile*, considers the elements which make up a reviewer profile and how they are determined. For example, a reviewer's area of expertise may be gathered by asking reviewers to select from a set of previously defined keywords specific to the conference. The second dimension, *Paper profile*, considers elements that make up a paper's profile and how they are determined. For example, authors of each paper may be asked to enter or select from a set of keywords which best describes their paper. The third dimension, *Matching method* refers to the algorithmic approach used to assign reviewers to papers. Lastly, the fourth dimension, *Case implementation* refers to how the methodology was implemented. If the method was implemented in a real case scenario, the environment is also considered.

As can be seen in Table 3.2.1, variables in both the reviewer and paper profiles were collected explicitly and/or implicitly. Information acquired explicitly requires input from the reviewer, whereas information acquired implicitly entails eliciting information in an automated way [81]. Most of the papers consider a set of predetermined keywords either for profiling reviewers or papers where a conference provides a set of keywords from which authors and reviewers select to represent their papers or expertise, respectively. However, the range of approaches considered for the matching method is very wide, varying from crisp to fuzzy methods. Regarding the types of applications, all of them are oriented towards either the conference reviewer assignment problem or assignment of experts to project proposals.

The first column of Table 3.2.1 describes how the reviewer profiles were generated. Most of the papers asked reviewers to select keywords which best represented their areas of expertise from a predefined list of words [69, 74, 81, 111, 142, 148]. Others asked reviewers for abstracts [53] or archived papers [42] which represented their areas of expertise, while others requested reviewers to evaluate the relevance of selected papers to themselves [48]. Lastly, Tayal et al. [150] assumed reviewer information to be previously available in an out of scope database. Different from these techniques, the presented methodology considers collected reviewer information from publicly available websites implicitly, without input from the reviewers, reducing the amount of effort required from them. Furthermore, with the exception of Tayal et al. [150], the papers reviewed considered keywords and bids to represent reviewers' areas of expertise. In contrast, Tayal et al. [150] and the presented methodology consider multiple elements to represent a reviewers' area and level of expertise.

The second column, paper profiles, describes how topics describing each conference paper were defined. Several papers reviewed, asked authors to select from a list of predetermined keywords the ones which best characterize their paper[69, 74, 81, 82, 111, 148]. One paper assumed that papers were previously grouped by keywords outside the scope of their methodology [48]. In another paper, the Technical Program Committee Chair assigned keyword to papers [142]. Two other papers applied topic modeling to conference papers [42, 53] and to representative abstracts or archived papers of candidate reviewers to obtain the similarity between the two. Although the presented methodology employs topic modeling to conference papers, it does not apply it to define the reviewer profile. Specifically, after employing topic modeling it seeks to align the topics identified in the papers with those of the conference itself. Likewise, topics of interest identified in reviewer profiles are also aligned to the same conference topics enabling a matching between reviewer, papers and conference topics. Furthermore, the LDA extension developed to define a paper's profile enables the methodology to name the concepts resulting from the LDA output automatically.

The third column refers to the matching method to assign reviewers to papers. Most of these methodologies implement an optimization method [69, 74, 111, 142, 148]. However, these methods may not be scalable to large conferences. Three papers consider the semantic similarity between reviewer and papers by applying topic modeling to reviewer abstracts [53] and archived papers [42], as previously mentioned or a taxonomy over the set of keywords [81, 82]. Unlike the methodology proposed by Tayal et al. [150], these methods do not deal with profile elements outside of keywords. Tayal et al. [150] compute a matching degree between reviewer and papers profiles employing multiple representations of reviewer expertise. The presented methodology in this chapter extends this concept to include topic coverage, that is, reviewers assigned to a paper should ideally cover all topics or, at least, most of them.

Lastly, nearly half of the papers reviewed did not demonstrate or demonstrated the applicability of the methodologies on simulated data [48, 69, 148, 150]. Our real case example is implemented on real conference data and has been introduced to a grant review environment exhibiting its ability to be generalizable. The second implementation is briefly explained in Section 6.5.

Our proposed method introduces two main advantages. First, it deals with information coming from several public sources to establish reviewer expertise and uses several variables to complete reviewer profiles. Second, an automated matching process, based on an aggregation function defined by an OWA operator, allows the simultaneous use of the relevant features without any filtering process.

3.3 Preliminaries

In this section we provide a summary of the OWA operator introduced by [170] which will be applied in the proposed methodology.

Definition 3.3.1 *An OWA operator of dimension n is a mapping of $\varphi : \mathbb{R}^p \to \mathbb{R}$ with an associated weighting vector W such that $w_h \in [0, 1]$ and $\sum_{h=1}^{p} w_h = 1$. The OWA operator is defined as:*

$$\varphi_{OWA}(a_1, \ldots, a_p) = \sum_{h=1}^{p} w_h \cdot a_{\sigma(h)} \tag{3.1}$$

where $(a_1, \ldots, a_p)$ is the vector of values associated with the set of criteria being aggregated and $\sigma : \{1, \ldots, p\} \to \{1, \ldots, p\}$ a permutation such that $a_{\sigma(h)} \geq a_{\sigma(h+1)}$, $\forall h \in \{1, \ldots, p\}$, i.e., $a_{\sigma(h)}$ is the h-th highest value in the set $\{a_1, \ldots, a_p\}$.

There are different methods to obtain the weight vector W. For our purpose we will use a linguistic quantifier guided aggregation as defined in Equation 2.1, in which the decision maker selects a quantifier representing the proportion of criteria necessary for a good solution [171].

Definition 3.3.2 *Given a regular increasing monotone (RIM) quantifier, $Q : \mathbb{R} \to \mathbb{R}$, we define the vector of weights $W = (w_1, \ldots, w_p)$ as follows:*

$$w_h = Q\left(\frac{h}{p}\right) - Q\left(\frac{h-1}{p}\right), \; h \in \{1, \ldots, p\}. \tag{3.2}$$

The use of the RIM quantifier in an OWA operator implies that the decision maker prefers that "most of" the criteria are satisfied.

Table 3.2.1: Comparison of different approaches to the reviewer assignment problem

Paper	Reviewer Profile	Paper Profile	Matching Method	Case Implementation
[53]	Abstracts provided by reviewers	Not applicable	Latent semantic Indexing	Conference Reviewer Assignment: Hypertext'91
[74]	Predetermined keywords	Predetermined keywords	Capacitated transhipment problem	Conference Reviewer Assignment: Decision Sciences Institute annual meeting 1998
[111]	Predetermined keywords	Predetermined keywords	1) Greedy algorithm 2) Evolutionary algorithm	Conference Reviewer Assignment: Parallel Problem Solving from Nature 2002
[69]	1) Predet. keywords 2) Bids	Predetermined keywords	1) Min. cost flow problem 2) Stable marriage problem	Conference Reviewer Assignment: No experimental results
[148]	Predetermined keywords	Predetermined keywords	1) Capacitated transportation problem 2) Heuristic	Assignment of experts to project proposals: Prototype - simulated data
[81] and [82]	Predetermined keywords	Predetermined keywords	Semantic similarity	Conference Reviewer Assignment: CompSysTech 2010 and 2011
[42]	1) Score from reviewer's papers 2) Bids	Not applicable	1) Latent Dirichlet Allocation 2) Supervised score prediction	Conference Reviewer Assignment: NIPS 2010, 2012, ICML, UAI, AISTATS, CVPR, ICCV, ECCV, ECML/PKDD, ACML, ICGVIP
[48]	Predetermined keywords	Not applicable	Fuzzy linear programming with fuzzy ranking	Assignment of experts to project proposals: Toy example
[150]	Expert quality measurement from indicators	Predetermined keywords	Fuzzy equality operator	Assignment of experts to project proposals: prototype - simulated data
[142]	Keywords determined by authors	Keywords determined by Tech Program Comm chair	Optimization method	Workshop Reviewer Assignment: SPAWC 2010

3.4 Proposed Methodology

The first three out of four defined dimensions, that is, Reviewer profile, Paper profile, and Matching, are described for our proposed method. The actual Case implementation performed is left for Section 5.

3.4.1 Defining the Reviewer Profile

We propose to represent a reviewer's profile using seven features which can be summarized into three measures related with his/ her research topic interests, recency, and quality. These seven features are determined from a cursory analysis of articles [34, 100, 138, 140, 156] on "good" journal reviews and reviewers some of which were written by editors of MIS Quarterly, Academy of Management, Journal of International Business Studies, International Review of Financial Analysis, and Journal of Behavioral and Experimental Finance. In total nine components were identified across all the articles. We proxy five of those components from information extracted from publicly available information. These are timeliness [100, 138, 140, 156], quality [156], diversity [138, 156], quantity [156], and developing others [34, 100, 140]. The remaining four components systems thinking [34, 140], positive voice [34, 100, 140], attention to detail [100, 140], and social objectivity [34, 140] more closely describe a well written review than a good reviewer and therefore, considered out of scope. The quality component referred to the H-index of a reviewer, diversity to the professional age and geographic location, quantity to the total number of publications, developing others to the ability to give constructive advice. The presented methodology considers quality, quantity, and developing others to be part of the quality measure, a representation of a reviewer's publishing accomplishments. Timeliness referred to the ability to submit a review within the time frame allotted and forms the availability part of a reviewer's profile. It is external to the information collected from public websites and therefore not part of these seven features and subsequent three measures. With respect to diversity, geographic location is not considered as the data set is from a local conference where all members are expected to be from the same region. Professional age includes young researchers on the frontier and the wisdom of more established scholars. The presented methodology considers professional age as the recency measure. As the recency score refers to the papers published in recent years for a reviewer, higher scores, these reviewers may be closer to the frontier. Reviewers with higher levels of expertise in a particular topic area may be more established scholars. The research topic interests vector represents a reviewer's area of expertise. Descriptions of the three measures, topics of interest, recency, and quality are described in the following paragraphs.

To gather information about each reviewer, we use global and local public sources. Global sources collect and aggregate information about researchers from around the world. The information can come from numerous sources. One example is Aminer[3]. In contrast, local sources collect information for a specific group of people such as faculty of a university. One example is TDX[4], a repository for theses defended in Catalunya, Spain. Prior research in [17] found that using more sources of information can lead to better performance.

Given a set of reviewers $\{Y_l\}_{l=1}^{L}$, all the possible research topics obtained from several websites for each reviewer are put into a common taxonomy using a dictionary of terms. The dictionary aligns common terms with the conference topics $\{T_k\}_{k=1}^{m}$. Automated alignment systems can be applied in this step, but it is vetted by an expert to ensure proper translation. Then, each research interest is translated to a conference topic. For each reviewer Y_l, the measure of his/her expertise in each conference topic is expressed as a vector $(y_{l_1}, \ldots, y_{lm})$.

The recency score is defined as a weighted average impact factor of the papers published by a reviewer in the past N years as defined by Aminer.

Regarding the quality score, the features considered in our methodology are: the number of PhDs supervised, books and book chapters written, papers published (both journals and conferences), and their H-index. Note that since we use several sources of information, data consistency is not warranted, each source of information can provide different values for the features considered in the quality score. Therefore, the maximum value of each feature from the different sources is selected. Using an Ordered Weighted Averaging (OWA) function these values are aggregated into the score called quality.

Besides the previously considered variables, a reviewer's profile also contains

[3] `aminer.org`
[4] `tdx.cat`

a list of previous co-authors and the reviewer's availability. The list of co-authors enables the system to avoid any conflicts of interest between the authors of conference papers and the proposed reviewer. The reviewer's availability is an indicator to be used in the matching procedure for assigning reviewers to papers.

3.4.2 Defining the Paper Profile

A paper profile consists of a paper's set of topic areas and a list of authors. The authors' names can be extracted from the paper submissions. To determine a paper's topic areas two steps are considered.

First, we determine a set of *concepts* from the entire set of paper submissions. The Latent Dirichlet Allocation (LDA) approach has been considered to generate this list. Originally introduced in [23], it is an unsupervised topic modeling method. LDA has been used in other reviewer assignment systems such as the Toronto Paper Matching System [42]. In our case, LDA considers the entire collection of n paper submissions $\{P_i\}_{i=1}^{n}$ and provides a set of concepts, $\{C_j\}_{j=1}^{s}$. Each concept is defined by a group of words. In addition, LDA calculates the proportion of each concept C_j represented in each paper P_i, α_{ij}, and satisfies,

$$\sum_{j=1}^{s} \alpha_{ij} = 1. \tag{3.3}$$

Second, to translate these concepts into the set of conference topics $\{T_k\}_{k=1}^{m}$, each set of words representing a concept is combined with the conference theme and a topic in a search using Google Scholar[5]. Then, the frequency that each concept appears with each conference topic is normalized by the frequency of the conference topic and theme, and collected in a matrix $\mathbf{G} = (g_{jk}) \in [0,1]^{s\times m}$, where each value of the matrix represents the frequency that the concept C_j appears with the conference topic T_k in the search of all papers received for the conference. It is worth noting that each concept represents a combination of several conference topics,

$$C_j = \sum_{k=1}^{m} g_{jk} T_k. \tag{3.4}$$

[5] scholar.google.com

Next, for each paper P_i, the vector of concept proportions provided by LDA, $(a_{i1}, \ldots, a_{ij}, \ldots, a_{is})$, is multiplied by the column in the matrix $\mathbf{G}$ representing the topic T_k to obtain the relationship r_{ik} between the paper P_i and the topic T_k,

$$r_{ik} = \sum_{j=1}^{s} a_{ij} \cdot g_{jk} \tag{3.5}$$

and avoid the use of 'intermediate' concepts. Hence, the matrix $\mathbf{R} = (r_{ik}) \in [0,1]^{n \times m}$ is considered, whose rows correspond to the proportions of the conference topics covered in each paper P_i.

3.4.3 Assigning Reviewers to Papers

In this use case, the assignment problem is considered from the perspective of identifying an appropriate reviewer for a paper. As reviewers' interests and skills continually evolve overtime, the methodology described here can be applied with each conference assignment. Reviewer profiles can be updated offline prior to paper submissions. However, it is the paper which needs to be reviewed and therefore, a reviewer's topics aligned with the papers. To this end, assigning reviewers to papers is the matching process to identify a set of reviewers who satisfy the needs of a paper. In the associated methodology, four types of indicators are used to evaluate a match between possible reviewers and papers. The first indicator is the match between topics covered in the paper and reviewers' expertise. The other three indicators are quality, recency, and availability. A diagram of the assignment process is shown in Figure 3.4.1.

The proposed matching methodology consists of five steps which are detailed below: compute paper coverage need, order papers by coverage need, assess reviewers per paper, rank reviewers by overall score, and assign reviewer and update availability. Once all of the paper submissions have been received, the process may begin. Prior to its start, the following parameters should be set up: *mp*, the maximum number of papers to be assigned to a reviewer, *nr*, the number of reviewers needed to fully cover a topic in a paper, and each reviewer's availability set to 1. Steps 1-5 are applied in an iterative loop until the criterion "all papers are assigned to the required number of reviewers" is met.

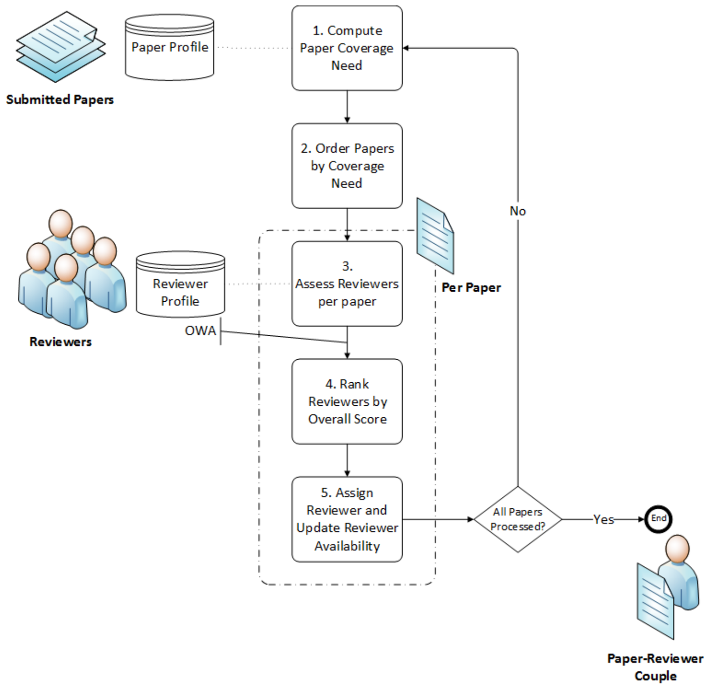

Figure 3.4.1: Process flow of the proposed matching methodology for assigning reviewers to papers

Step 1. Compute paper coverage need   In each iteration, coverage need for each paper is computed. Coverage need refers to the topics and number of topics included in a paper and the constraint that the *nr* parameter is fulfilled. Therefore, in the initial iteration, coverage need is calculated with all of the topics in the paper. However, in subsequent iterations, topics which have already been fulfilled (partially or fully) by a reviewer are taken into account.

Example 3.4.1 *Let us assume there are four papers and five conference topics. The parameter nr is set to one. The initial papers' Coverage Need are shown in the last column in Table 3.4.1.*

Step 2. Order papers by coverage need   Papers are ordered by the coverage need value obtained in Step 1 in decreasing order. Next, Steps 3 - 5 are performed for a single paper.

Table 3.4.1: Example of computed Coverage Need, CN

ID	Machine Learning	Computer Vision	AI App	Fuzzy Logic	Data Mining	CN
Paper1	1	0	1	1	1	4
Paper2	1	0	1	0	0	2
Paper3	1	1	1	0	0	3
Paper4	0	0	1	1	0	2

STEP 3. ASSESS REVIEWERS PER PAPER   Given the paper P_i with the greatest coverage need from Step 2, identify candidate reviewers:

(a) Filter out reviewers with an availability score of zero and any reviewer already assigned to the paper.

(b) For each selected reviewer, assess the partial scores for each indicator: 1) reviewer research topic expertise according to the topics coverage need of the paper, 2) quality, 3) recency, and 4) availability of the reviewer.

(c) Employ OWA from (3.1) to aggregate the p partial scores a_p into an *overall score* for each reviewer.

Weights w_h are computed by the RIM quantifier in (3.2) considering $Q(x) = x^{1/2}$.

Example 3.4.2 *Continuing with our example, we can determine that P_i is Paper1. For the available reviewers, we obtain the following partial and overall scores in Table 3.4.2 and Table 3.4.3, respectively:*

Table 3.4.2: Example of computed partial scores

ID	Mach Learn	AI App	Fuzzy Logic	Data Mining	Rec	Qual	Avail
Rev1	0	0	1	1	0.8	0.2	1
Rev2	1	0	1	1	0.8	0.8	1
Rev3	1	1	1	0	0.5	0.8	1
Rev4	1	1	0	0	0.5	0.5	1

Table 3.4.3: Example of computed Overall Scores, OS

ID	v1	v2	v3	v4	v5	v6	v7	OS
Rev1	.378	.157	.120	.081	.018	.000	.000	.754
Rev2	.378	.157	.120	.101	.071	.065	.000	.892
Rev3	.378	.157	.120	.101	.071	.040	.000	.868
Rev4	.378	.157	.120	.051	.045	.000	.000	.750

STEP 4. RANK REVIEWERS BY OVERALL SCORE Rank reviewers by overall score in descending order.

STEP 5. ASSIGN REVIEWER AND UPDATE AVAILABILITY Reviewer Y_l, with the highest score, is assigned to paper P_i and his/her availability, initially set to 1, is reduced as defined by,

$$Av(Y_l)_{new} = Av(Y_l) - \frac{1}{mp}, \tag{3.6}$$

where $Av(Y_l)$ is the current availability of reviewer Y_l.

Example 3.4.3 *Completing our assignment, Reviewer2 is assigned to Paper1. Reviewer2's availability is then adjusted by (3.6).*

In the case of ties between two or more reviewers with the highest score, the "exclusiveness" of the topics known by each reviewer (in terms of the number of reviewers knowing the same topics) is used, in order to choose the reviewer with the least exclusive knowledge.

Once a reviewer is assigned, the system checks if all papers have met the reviewer assignment criterion. If the criterion has not been met, the system completes another iteration beginning at Step 1. If the criterion has been met, the system exits the loop.

3.5 A SIMULATED CASE USING REAL DATA

To validate the reviewer assignment quality we generated a simulated case using real data from three consecutive editions of a small international conference.

3.5.1 Data Set

The data set consists of three consecutive conferences of the International Conference of the Catalan Association for Artificial Intelligence (CCIA 2014, 2015, and 2016). These conferences were combined into a simulated bigger one for two reasons. First, combining several conferences provided a larger number of paper submissions. Second, as these were the most recent conferences of CCIA, we were able to assume that the reviewer profile would be relatively similar for each year.

The papers and the reviewers from the three conferences were combined to form a single "conference". There were a total of 106 submitted papers and 96 Scientific Committee members. The Committee members' names are public on the conference web pages. We simulated the conference to take place in the current year. Therefore, the data collected to generate the reviewer profile is considered a representation of the current interests and activities of the reviewer.

To generate the reviewer profile we selected three global sources: Aminer, ResearchGate[6], and dblp[7] and one local source: TDX (Catalan database of PhD theses). Each reviewer was identified according to his/her name, organizational affiliation, and network, when necessary. For each website, we gathered all the available information for each reviewer. When there were multiple entries for a reviewer from a single source, we took the one containing the most recent publications with the assumption that it implied a more updated profile of the reviewer. If there were two records with articles published in the same year, we selected the one with the most profile information. We observed that the TDX website sometimes included the reviewer's own thesis in the collection of theses supervised and it was removed manually. All available information was translated to English. A dictionary of terms was created to translate terms representing reviewers' research interests from the different sources into the CCIA conference topics. Among the original 96 reviewers, only 51 had skills populated on their ResearchGate profiles. Therefore, as the main matching entities in a conference-reviewer environment are the expertise topics of the reviewers, we took into consideration only the subset of 51 reviewers for whom we could identify their skills.

[6] `researchgate.net`

[7] `dblp.uni-trier.de`

The rest of the case follows the methodology described above and it is implemented with the general constraints of a reviewer assignment problem. The number of reviewers per paper in this simulated case was set to 2, the maximum reviewer load (parameter *mp*) was initially set to 3 and the number of reviewers needed to fully cover a topic in a paper (parameter *nr*) was set to 2. However, the system automatically adjust to compensate for the ratio of reviewers per paper taking into consideration the parameter *mp*. Since 212 assignments are required (2 reviews per paper) from 51 reviewers, then the maximum number of papers assigned to a reviewer *mp* must be adjusted upward to 5.

3.6 Results and Evaluation

Many methods have been proposed to measure the performance of an automatic assignment system [42, 84, 113]. However, there is no standard method to our knowledge. We applied five different techniques to evaluate the performance of our method from the perspectives of the overall matching, reviewers, papers, an expert's opinion, and a baseline method.

3.6.1 The Overall Perspective

First, we assessed the overall output of the matching. Using our method, 106 papers were assigned to 46 reviewers, a ratio of 2.3 papers per reviewer. Considering the operation defined in (3.1) and performed in Section 3.4.3 Step 3, an overall score was assigned to each paper-reviewer couple. This score, which is an aggregation of the partial scores (topic interest coverage, availability, recency, and quality), gives us a grade about the adequacy of each selected couple. Globally, this overall score is in the range $[0, 1]$. Considering the 212 assignments (2 reviewers per paper), the average overall score was 0.789 with the minimum fixed to 0.650 and maximum equal to 0.933. The distribution of this overall score is depicted in Figure 3.6.2.

In order to evaluate the significance of this result, we compared the solution obtained using the presented methodology with one obtained by optimizing the paper-reviewer assignments. Specifically, the optimal assignment considered the topic coverage need of each paper and assigned two reviewers per paper such that the two reviewers covered all the topics of a paper. The reviewers were assigned with the same constraints imposed as in the presented methodology.

If β, the optimal solution, is represented by the number of topics of a paper where all topics are covered by the selected reviewers, then $\beta - 1$ represents the number of topics of a paper where all topics except one are covered by the selected reviewers. Likewise, $\beta - 2$ represents the number of topics of a paper where all topics except two are covered by the selected reviewers. The Wilcox Signed-Ranks test was applied to compare the results of the presented methodology with β, $\beta - 1$, and $\beta - 2$. This test was selected because neither the optimal solution nor the solution determined by the presented method followed a normal distribution. Using the Shapiro-Wilk test for normality, p-value = 1.619e-11 and p-value = 1.162e-09 were obtained for the solutions of the optimization and presented methodology, respectively. For the Wilcox Signed-Ranks test, the null hypothesis considered that the presented methodology obtained is worse than or equal to the results of the compared model. Table 3.6.1 shows the results of the Wilcox Signed-Ranks test for each comparison.

Table 3.6.1: Model comparison with Wilcox Signed-Ranks

model comparison	W	p-value
β	2405.5	1
$\beta - 1$	6242	0.1056
$\beta - 2$	9640.5	2.2e-16

As shown in Table 3.6.1, the results of the Wilcox Signed-Ranks test were not significant for β and $\beta - 1$. We fail to reject the null hypothesis that the solution of the presented methodology is worse than or equal to the solutions of β and $\beta - 1$. However, the results of the test are significant for $\beta - 2$. We can reject the null hypothesis. Therefore, the solution of the present methodology is between better than $\beta - 2$ and worse than $\beta - 1$. To refine the assessment, we defined a parameter $\lambda \in [0, 0.2]$. One hundred random samples of each value of λ were selected such that the compared model was $\beta - (1 + \lambda)$. P-values < 0.5 were consider significant. Figure 3.6.1 displays the results for each value of λ as the percentage of experiments where p-values < 0.5 were obtained. As can be seen in the figure, the presented methodology is able to cover all but one topic of a paper with the assigned reviewers' expertise in most of the matches and misses two or more topics on few occasions.

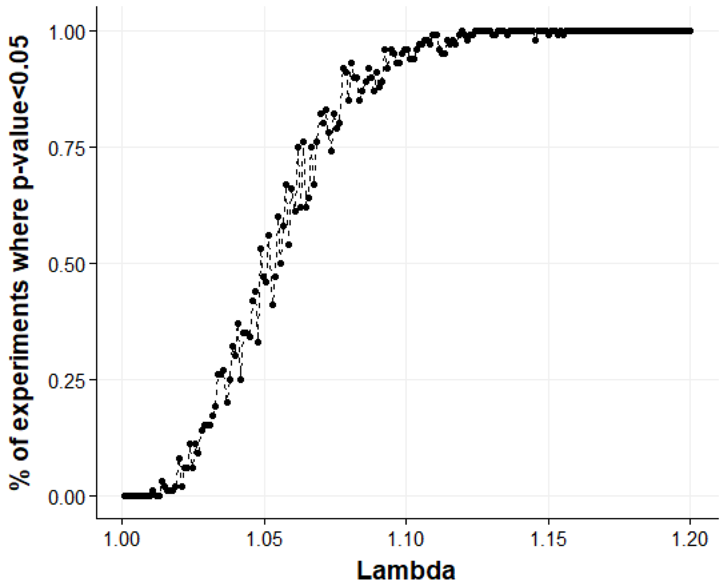

Figure 3.6.1: Results of 100 simulations with λ

3.6.2 The Reviewer's Perspective

Second, we compared the reviewer to paper assignments with the quality index (QI) defined in [142]. This measure represents, for each reviewer, the average percentage match between his/her topics and the topics of each paper to which he/she has been assigned. Out of the 46 assigned reviewers the mean and standard deviation of the QI for our method were 0.693 and 0.284, respectively.

3.6.3 The Paper's Perspective

Third, on a paper basis, we assessed the coverage of each paper's topics according to the assigned reviewers. We evaluated this measure in two parts. Using the assignments made by the system, we compared the topic coverage of each paper based on the paper and reviewer topics assigned by the system. Next, we compared the topic coverage with the paper and reviewer topics determined by an expert.

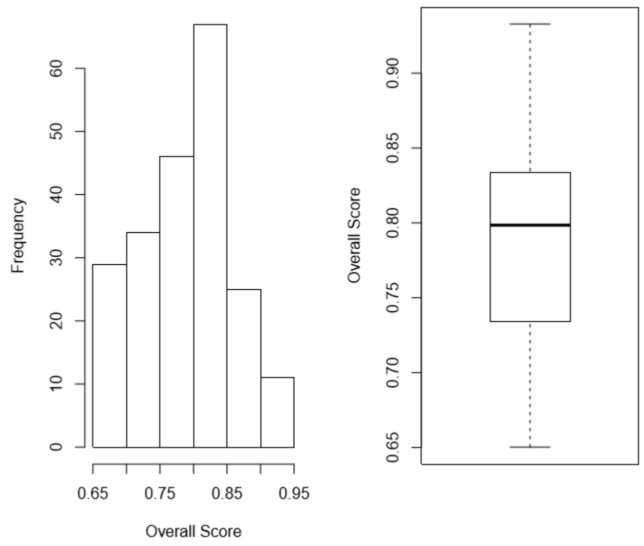

Figure 3.6.2: Distribution of the overall score with proposed method, with $nr = 2$

Applying the reviewer and paper profiles determined by the system we obtained the following results. Out of the 106 assigned papers, 104 had a complete match. We define a complete match as one where at least one topic of each reviewer assigned matches the topics of a paper. In addition, we observed that 2 of the papers had a partial match. We consider a partial match to be a paper having only one reviewer having topics that match the paper. There were no papers without a match. In other words, there were no papers where a reviewer assigned to a paper did not cover at least one topic of the paper.

3.6.4 The Expert's Opinion Perspective

In order to compare the results to that of an expert, a ground truth was created similar to [84]. An expert from the Artificial Intelligence community in Catalunya was consulted for the validation process. He assigned research topics from the CCIA conference to each of the Scientific Committee members. Then, he read the abstracts of each paper submitted to the conference and assigned relevant CCIA conference topics to each paper. This gave us a gold standard to evaluate our system.

Inter-rater reliability was computed for each of the topics assigned by the expert to a reviewer. During the period in which data was being collected, another project by the the University of Valencia [8], was asking CCIA conference attendees to enter their research topic areas into a database. We compared the inter-rater reliability of 28 members who provided their information with the expert's evaluation. Using percent agreement and Cohen's Kappa we obtained the following results for each topic. Cohen's Kappa is selected as it takes into account agreement by chance. Results are interpreted according to Landis and Koch [88]. As can be seen from Table 3.6.2, the expert and reviewers are in agreement for most of the topics with the exception of Cognitive Modeling. The low agreement may be due to only one reviewer being rated positively for the topic by the expert. We include the topic in the real case as the percent agreement is still high, although we acknowledge its limitation.

Table 3.6.2: Intr-rater reliability between expert and reviewer

	Applications of AI	Machine Learning	Fuzzy Logic & Reasoning	Computer Vision	Data Mining	Cognitive Modeling	Planning Optimization satisfiability	Agents & multi-agent systems	Natural Language	Constraint Programming	Robotics
% agreement	89.3%	92.9%	67.9%	100%	100%	71.4%	100%	92.9%	100%	100%	100%
Cohen's Kappa	0.788	0.858	0.4	1	1	-0.0667	1	0.826	1	1	1
Level of agreement	substantial	almost perfect	fair	perfect	perfect	slight	perfect	almost perfect	perfect	perfect	perfect

Applying the reviewer and paper profiles determined by the expert we obtained the following results. Out of the 100 assigned papers (6 papers were discarded by the expert due to their minimal relation with CCIA topics), 81 had a complete match. In addition, we observed that 17 of the papers had a partial match. There were 2 papers without a match. Results showed that with the expert opinion the matches between papers and reviewers slightly decreased. We attribute the decrease to the more accurate assignment of topics to reviewers and papers by the expert, thanks to his knowledge about the reviewers.

3.6.5 The Baseline Perspective

Lastly, we applied random assignment, imposing the constraints of the problem presented in the case, to match reviewers to papers. Thirty iterations of random as-

[8]https://nodes.acia.cat/modules.php?name=news&idnew=187&idissue=32

signment were ran and the mean of each evaluation measure was computed. These values are compared in Table 3.6.3 to the proposed method.

Table 3.6.3: Comparison of random and proposed methods based on matches

	QI	complete match	partial match	no match
Random	0.422	56.6%	34.8%	8.6%
Proposed	0.693	94.3%	5.7%	0

Each iteration of the random method assigned 106 papers to 51 reviewers, with an average ratio of 2.08 papers per reviewer. The number of papers assigned to an individual reviewer ranged from 1 to 5. As can be seen from Table 3.6.3, our method out performs the random method in both the evaluation of the matches and the QI.

3.7 Conclusion and Future Research

In this paper, a new method for assigning papers to reviewers for conferences has been introduced. This methodology improves existing systems because:

- It uses several sources of public information to define reviewers expertise profiles.
- It considers the whole set of papers submitted to the conference to define the most appropriate topics for each paper.
- The matching process is defined via the concept of coverage and uses an OWA operator, which allows us to avoid filtering but simultaneously consider several relevant variables for the process.

Furthermore, a novel method which leverages LDA enables conferences to readily label paper submissions within the context of the predefined conference topics. LDA topicing modeling was implemented in the methodology in order to extract the topics of the papers. Three known drawbacks to LDA are 1) assumption that the number of concepts is known prior to running the model [12], 2) concepts are learned without taking into consideration the labels to be predicted, therefore concepts are not very discriminative across document categories [168], and 3) how to model human interpretable labels while discovering unlabeled concepts [128]. These drawbacks became evident when we applied topic modeling to define the topics reflected in each paper. Given the conference environment, the topics for which papers were being solicited were known beforehand. These were the topics with which both the papers and reviewers needed to be aligned. The first challenge was to allow these topics to be used as labels in the topic modeling while allowing the topic model to discover the underlying concepts in the corpus of papers. The second challenge was that LDA returned a set of concepts related to one or more of the labeled topics.

In this chapter, we presented a method to address these challenges. Rather than obtaining the topics directly from LDA, we introduced a interim process which interprets the output of LDA as a set of concepts, each of which can contribute to one or more topics. LDA was applied to allow unlabeled concepts to be discovered. Then the conference topics were imposed as labels across multiple concepts considering the frequency with which concept keywords coincided with conference topics. This frequency was combined with the concept proportions given by LDA to determine the conference topics referenced in each paper.

This methodology is developed such that it can be implemented by organizations which can benefit from expert assignment without the desire to expend a large quantity of effort in obtaining it, as it is outside of their core business. Therefore, we simulate their environment and utilize data with which they would have access to define reviewers' profiles. As such, the methodology leverages publicly available data sets. These data sets are messy and although they may appear structured on their own, the aggregate information across all data sets is unstructured. Each web site has its own vernacular, structure, and sources of information leaving the summation of this information to the discretion of the user, the conference committee.

The presented methodology was applied to conference data from CCIA in order to demonstrate its applicability. As previously discussed, an optimization method was implemented for comparison. The small size of the conference allows an optimal solution which maximizes the topic coverage of each paper to be obtained with an exact optimization method. However, obtaining an optimal solution for a conference with thousands of papers is very costly. In this case, our methodology can approximate an assignment regardless the size of the conference and obtain good results.

We are considering different lines for future research. First, we would like to apply the method to a larger conference environment. Second, we would also like to explore the inclusion of inputs from authors regarding suggested reviewers to help refine assignments. Lastly, we aim at applying a similar methodology to the human resources problem that considers the assignment of candidates for a job position.

4

A linguistic multi-criteria decision-making system to support university career services[1]

4.1 Introduction

Organizations are challenged daily to make complex decisions. These decisions can be subjective, uncertain, and imprecise [114]. As data becomes continually available, these decisions become increasingly more complex, making the role of decision support tools more important. Specifically, this notion can be observed within human resource personnel selection. In general, personnel selection depends on a firm's specific targets, and the preferences of the hiring managers [86]

[1] This chapter is mainly based on the article published in Applied Soft Computing by J. Nguyen, G. Sanchez-Hernandez, A. Armisen, N. Agell, X. Rovira, and C. Angulo, (`https://doi.org/10.1016/j.asoc.2017.06.052`)

and candidates.

For global organizations, human resources personnel selection can be challenging as candidates are disperse and vary in level of knowledge of a topic. Their knowledge is difficult to qualify and changes frequently [109]. Personnel selection is subjective in nature with regards to assigning crisp values to the job requirements and evaluating candidate qualifications. Previous studies have extended Multi-Criteria Decision-Aiding (MCDA) methods to this problem to address its fuzziness [14, 43, 86].

Within universities, obtaining an internship is a specific personnel selection process. It may be the first time a student is applying for a position. Therefore, the terms used to describe the desired position may be unfamiliar making the job search process overwhelming. Students may not know which terms to use when searching for a specific position or for which position their skills are most relevant. Hence, the positions obtained in their search results may not be the best match for them. There are two different perspectives to personnel selection. The hiring company is looking for the best candidate to fill a position. On the other hand, the candidate is looking for a position which satisfies their interests. Knowing on which positions to focus their time is key to both the student and the hiring company.

The aim of this paper is to introduce a practical decision support system to assist students with identifying positions most related to their interests. A real case example is implemented with student and job information provided by a university's career services office. In terms of feature representation, the novelty of the application is two-fold. First, the requirements of a position are extracted in an implicit manner and represented via linguistic terms. Second, linguistic terms are also considered to represent students' interests. The model considered for linguistic descriptions is the hesitant fuzzy linguistic model. This model was introduced by Rodriguez et al. in [135] and further developed in [136].

The rest of the paper is structured as follows. First, a review of current MCDA applications to the personnel selection problem is presented. Next, we discuss tools used in the design of a linguistic MCDA system which include linguistic descriptors, and fuzzy matching and aggregation. These tools are applied to a decision support system to help students in the selection of their internship, presented in Section 4. Following the explanation of the methodology, a real case is provided with the implementation of the proposed method. Lastly, conclusions are presented and future research directions are proposed in Section 6.

4.2 State-of-the-Art in Personnel Selection

The personnel selection problem has been studied quite extensively [14, 57, 80]. In this section, we review and compare related research in personnel selection with specific attention to applications of MCDA to the problem. Nearly all of the papers reviewed assess candidates with respect to a position's requirement. As personnel selection is a two-sided problem, our study proposes to address the problem from the less studied point of view. Therefore, we define a support system for students to choose among a set of alternative internships. However, both sides of the problem share the main characteristics of defining applicant and job profiles, and an assignment process. We characterize the existing literature according to three dimensions that consider the ranking method, feature weights, and case implementation. The first dimension, ranking method, refers to the method by which the candidates for a position are ranked according to their qualifications. The second dimension, feature weights, considers how the importance of each feature for a position is assigned. The third dimension, case implementation has four components: a) environment, b) number of positions, c) number of candidates, and d) number of features. Environment refers to how the methodology was executed, number of positions refers to the number of jobs to which the case attempted to assign candidates, number of candidates refers to the number of candidates each case tried to assign to a position, and the number of features refers to the number of evaluation criteria assessed.

As can be seen in Table 1, most of the papers analyzed in the literature review implemented an illustrative case while only two papers had use cases. In the first group of papers, the authors selected positions, candidates, and features to suit their illustrative example. The features selected were estimated based upon their specific positions. Regarding the number of candidates considered in each paper, only two papers had 100 candidates while the others had six or fewer. The lower number of candidates may be to facilitate the illustrative example while the papers with 100 candidates had a full web implementation.

Table 4.2.1: Applications of MCDA to personnel selection

Paper	Ranking Method	Weights	Case Implementation			
			Environment	# Po	# Ca	# Fe
Canós and Liern, 2008 [35]	OWA and parametric aggregation	Learned weights and FWA	Illustrative example	1	5	6
Güngör et al., 2009 [71]	Comparison of fuzzy AHP and Yager's weighted method	Predetermined by recruiter	Illustrative example	1	6	17
Faliagka et al., 2012 [57]	AHP	Predetermined by recruiter	Use case	3	100	4
Kabak et al., 2012 [80]	Fuzzy TOPSIS and fuzzy ELECTRE	Fuzzy ANP computations to determine weights	Illustrative example	1	6	10
Baležentis et al., 2012 [14]	MULTIMOORA for group decision making using FWA operator	Predetermined by recruiter	Illustrative example	1	4	8
Yu et al., 2013 [177]	GHFPWA and GHFPWG operator used to aggregate hesitant fuzzy elements (HFE)	Prioritized average (PA) operator	Illustrative example	5	5	4
Faliagka et al., 2014 [58]	Learning to rank	N/A	Use Case	3	100	4

Our proposed method differs from existing methods for several reasons. As students, rather than positions, are the main focus of our method we propose to elicit the features from the students. We incorporate an existing automatic topic modeling technique to extract these features. Therefore, the number of features considered is determined through a process defined in [56] and is tailored to the students. Next, the required features are identified for each job description applying a posterior distribution based on the previously defined features. Lastly, an automated matching process, based on an aggregation function defined by a FOWA operator, allows the simultaneous use of the relevant features without any filtering process. Specifically, each component of a student's interests and position's features are compared by a fuzzy matching operator and aggregated with an ordered weighted averaging operator (OWA), introduced by Yager and Kacprzyk [174], to obtain a fuzzy linguistic label.

We present a real case study with 275 students. These students were the actual internship candidates for a business school in 2016. Given that these students were from the same college with similar backgrounds it is expected that they would compete for the same positions. Therefore, this scenario is analogous to the personnel selection problem, which human resource managers face, with many candidates for a single position.

4.3 Preliminaries

In this section, we briefly present the necessary tools to design a linguistic multi-criteria decision-aiding system, that is, the concept of fuzzy matching for linguistic descriptions and fuzzy aggregation operators for the selection of alternatives.

4.3.1 Linguistic descriptions and Fuzzy matching

To introduce a decision support system which proposes available positions to college students, there are some uncertainties that should be considered in evaluating the students' interests. The uncertainty is inherent in students' abilities to communicate their affinity for specific features of a position. Having had little experience with these features, it may be difficult to express their preferences as a single label. Given this uncertainty, as mentioned in the introduction, we propose the application of Hesitant Fuzzy Linguistic Term Set (HFLTS) [136] to manage the need for several labels to define preferences.

Other linguistic modeling techniques could have been considered such as multi-granular linguistic modeling [115], computing with words based on discrete fuzzy numbers [108], 2-tuple linguistic modeling [77], or linguistic modeling based on ordinal symbolic information [66]. In fact, our method could be considered a multi-granular linguistic model as it considers different levels of granularity in the linguistic assessments. However, in general, multi-granular linguistic modeling methods aggregate the opinions of experts across all of the alternatives prior to ranking them. In contrast, we propose to use a matching operator which enables matching student preferences to position features on an individual attribute and student level, and then computes an overall score. Secondly, with respect to computing with words based on discrete fuzzy numbers and 2-tuple linguistic modeling, experts would be required to provide additional information regarding the grade of the value contained in the semantic support as part of their qualitative or linguistic evaluation. In our method, we require less information from the participants because specifying a grade of a value would be difficult as students may not have this information. Finally, if we consider linguistic modeling based on ordinal symbolic information, as defined in [66], experts would be asked to pairwise compare features. In the context of this real case, students would not have the flexibility to express their preferences as "I don't know" which may be the case if they have had no experience with a feature.

The approach proposed in this paper relies on the use of linguistic terms based on a qualitative absolute order-of-magnitude model [9, 155] that allows us to deal with the imprecision and hesitance involved in decision processes. We will express this model by means of HFLTS introduced by Rodriguez et al. [136].

Let $\mathbb{S}_n$ be a finite set of totally ordered basic terms, $\mathbb{S}_n = \{B_0, \ldots, B_n\}$, with $B_0 < \ldots < B_n$ and the hesitant fuzzy linguistic terms set, $H_{\mathbb{S}_n}$, be the set of all consecutive linguistic basic terms of $\mathbb{S}_n$, i.e. $B_{ij} = \{x \in \mathbb{S}_n \,|\, B_i \leq x \leq B_j\}$ $\forall i, j \in \{0, \ldots, n\}$, with $i \leq j$. In general, each term corresponds to a linguistic label, with B_0 being the term "*None*". For simplicity, we will denote the singleton $B_{ii} = B_i$. The total order in the set of basic terms, $\mathbb{S}_n$, allows us to define a total order in $H_{\mathbb{S}_n}$ based on the lexicographic order such that: given two linguistic terms, $B_{ij}, B_{i'j'} \in H_{\mathbb{S}_n}, B_{ij} \leq_L B_{i'j'}$, iff $i < i'$ or $i = i'$ and $j \leq j'$.

For instance, let us consider $n = 3$ and B_0 = None $<$ B_1 = Low $<$ B_2 = Medium $<$ B_3 = High, then, terms B_{12} and B_{03} will represent the linguistic labels *Low or Medium* and *Unknown* (None, Low, Medium, or High), respectively. From the lexicographic order, we get $B_0 \leq_L B_{03} \leq_L B_1 \leq_L B_{12} \leq_L B_2$.

From this point forward, we consider $H_{\mathbb{S}_n^*}$, a subset of $H_{\mathbb{S}_n}$, which corresponds to the HFLTS obtained when the set of basic elements is $\mathbb{S}_n^* = \{B_1, \ldots, B_n\}$. In addition, in $H_{\mathbb{S}_n}$ we consider the subset inclusion to define the relation *"to be more precise or equal to"*. We say that B_{ij} *is more precise or equal to* $B_{i'j'}$, $B_{ij} \preceq B_{i'j'}$, if and only if, $B_{ij} \subseteq B_{i'j'}$, i.e, $i' \leq i$ and $j \leq j'$. For instance, in the previous example, we have $B_1 \preceq B_{02}$ and $B_{12} \preceq B_{13}$. Finally, the connected union operator

, $\sqcup$, is considered in $H_{\mathbb{S}_n}$ defined as $B_{ij} \sqcup B_{i'j'} = B_{kl}$ where $k = \min(i, i')$ and $l = \max(j, j')$. Following the previous example, $H_{\mathbb{S}_3}$, $B_{01} \sqcup B_3 = B_{03}$.

HFLTS can be used to compare individual's preferences to object's attributes to capture imprecision in decision processes. To this end, we will define an operator matching two basic terms and extend it to the entire set of HFLTS catching all possible combinations of hesitancy in both descriptions.

Definition 4.3.1 *The fuzzy matching operator is the map*

$$* : H_{\mathbb{S}_n} \times H_{\mathbb{S}_n^*} \to H_{\mathbb{S}_n}$$

such that:

1. $\forall B_i \in \mathbb{S}_n$ *and* $\forall B_j \in \mathbb{S}_n^*$, $B_i * B_j = B_{\min(n, n-(j-i))}$,
2. $\forall B_{ij} \in H_{\mathbb{S}_n}$ *and* $\forall B_{i'j'} \in H_{\mathbb{S}_n^*}$,
   $B_{ij} * B_{i'j'} = \bigsqcup\{B_k * B_l, i \leq k \leq j \text{ and } i' \leq l \leq j'\}$.

Note that, 2. coincides with 1. $\forall B_i \in \mathbb{S}_n$ *and* $\forall B_j \in \mathbb{S}_n^*$.

Example 4.3.1 *Let us consider that a candidate's preferences are represented by $H_{\mathbb{S}_n^*}$ and the features of each position are represented by $H_{\mathbb{S}_n}$, then given the previously considered HFLTS, $H_{\mathbb{S}_n^*}$, with $n = 3$, the results of the fuzzy matching operator for the basic terms are shown in Table 4.3.1.*

Table 4.3.1: Fuzzy matching operator *

*	**Low** (B_1)	**Medium** (B_2)	**High** (B_3)
None (B_0)	Medium (B_2)	Low (B_1)	None (B_0)
Low (B_1)	High (B_3)	Medium (B_2)	Low (B_1)
Medium (B_2)	High (B_3)	High (B_3)	Medium (B_2)
High (B_3)	High (B_3)	High (B_3)	High (B_3)

Interpreting the table, it can be seen that when the candidate has a "Low" preference for a feature, a position with the same value or higher for the feature is a "High" match. It is considered that the candidate's preference has been met or exceeded. A position having *a value one step lower than the candidate's preference is considered a "Medium" match as the feature partially meets the candidate's preference. A value two steps lower is a "Low" match because the preference of the candidate is barely met. Looking at the far right side of the table, when the candidate's preference is "High" but the position value is "None", the difference is three steps lower and the position does not contain this feature resulting in a "None" match leaving the preference unmet for this feature. Continuing with the candidate's preference of "High", a position with a "Medium" value partially meets and a "Low" value barely meets the candidate's preference. Therefore, the match qualities are "Medium" and "Low", respectively.*

From Example 4.3.1, it can be seen that the fuzzy matching operator deliberately returns the value *"High"* in half of the situations in order to capture the positions with features which meet or exceed student preferences.

Example 4.3.2 *To demonstrate how the $*$ operator works with non-basic labels let us consider, B_{02} and B_{12} along with Table 4.3.1. $B_{02} * B_{12} = \bigsqcup\{B_0 * B_1, B_0 * B_2, B_1 * B_1, B_1 * B_2, B_2 * B_1, B_2 * B_2\} = \bigsqcup\{B_2, B_1, B_3, B_2, B_3, B_3\} = B_{13}$.*

Proposition 4.3.1 *The fuzzy matching operator $*$ fulfills the following properties:*

1. *$\forall B_{ij}, B_{i'j'} \in H_{\mathbb{S}_n^*}$, then $B_{ij} * B_{i'j'} \neq B_{i'j'} * B_{ij}$.*
2. *$\forall B_{ij} \in H_{\mathbb{S}_n}$ and $\forall B_{i'j'} \in H_{\mathbb{S}_n^*}$, with $B_{i'j'} \leq_L B_{ij}$, then, $B_n \preceq B_{ij} * B_{i'j'}$.*
3. *$\forall B_{ij} \in H_{\mathbb{S}_n}$, $B_{ij} * B_n = B_{ij}$.*

From Property 1 we can infer that the order always matters when matching two different terms in $H_{\mathbb{S}_n^*}$. If the first one is greater than or equal to the second one, the result is less precise than B_n. In addition, whenever the first label, B_{ij} is matched with a second label of B_n, the result is always B_{ij}. It follows that the element B_n is neutral with respect to B_{ij}.

4.3.2 Fuzzy aggregation and alternatives selection

Given two k-dimensional different vectors, $X = (X_1, ..., X_k) \in (H_{\mathbb{S}_n})^k$ and $Y = (Y_1, ..., Y_k) \in (H_{\mathbb{S}_n^*})^k$, we analyze the existing matching between these vectors, comparing each component, by means of the fuzzy matching operator $*$, and a FOWA (fuzzy ordered weighted average).

Definition 4.3.2 *Given $X \in (H_{\mathbb{S}_n})^k$ and $Y \in (H_{\mathbb{S}_n^*})^k$, the fuzzy matching between X and Y is defined as:*

$$X * Y = (X_1 * Y_1, ..., X_k * Y_k) \in (H_{\mathbb{S}_n})^k$$

Example 4.3.3 *Continuing with Example 4.3.1, given the vectors $X \in (H_{\mathbb{S}_3})^5$ and $Y \in (H_{\mathbb{S}_3^*})^5$, $X = (B_2, B_1, B_3, B_0, B_2)$, and $Y = (B_2, B_2, B_1, B_3, B_1)$, the match is $X * Y = (B_3, B_2, B_3, B_0, B_3)$. In the same way, if $X = (B_{02}, B_{12}, B_1, B_{13}, B_0)$, and $Y = (B_2, B_1, B_{13}, B_{12}, B_2)$ the match is $X * Y = (B_{13}, B_3, B_{13}, B_{23}, B_1)$.*

As previously mentioned, we apply an OWA introduced by Yager and Kacprzyk [174] to our specific context, to obtain a fuzzy linguistic label from a vector of $(H_{\mathbb{S}_n})^k$.

Definition 4.3.3 *Given $Z = (Z_1, \ldots, Z_k) \in (H_{\mathbb{S}_n})^k$ we define its weighted average index as:*

$$\mu^Z = \sum_{i=1}^{k} w_i \cdot \phi(Z_{(i)})$$

with: $Z_{(i)}$ having the same terms as Z_i ordered from the largest to the smallest by means of the total order $\leq_L$, a set of decreasing weights, w_i, such that $w_i \in [0,1]$ and $\sum_{i=1}^{k} w_i = 1$, and an increasing function with respect to $\leq_L$, $\phi : H_{\mathbb{S}_n} \to \mathbb{R}$, such that $\phi(B_s) = s, \forall s \in \{0, \ldots, n\}$.

For our purpose, we consider the regular increasing monotone (RIM) function, introduced by Yager [171], guided by the linguistic quantifier 'most of', as defined in Equation 2.1. Note that a RIM function must be used to obtain positive weights w_i, and $Q(x) = x^a$ should be defined with $a \in [0,1]$ to obtain a concave operator able to model those aggregations with importance associated with them.

Definition 4.3.4 *Given $Z = (Z_1, \ldots, Z_k) \in (H_{\mathbb{S}_n})^k$ we define the fuzzy ordered weighted average operator $\Phi : (H_{\mathbb{S}_n})^k \to H_{\mathbb{S}_n}$ is defined as follows:*

$$\Phi(Z_1, \ldots, Z_k) = B_{\mu_1^Z \mu_2^Z}$$

where μ_1^Z and μ_2^Z are the rounded and ceiling values, respectively.

Definition 4.3.5 *Given $X \in (H_{\mathbb{S}_n})^k$, $Y \in (H_{\mathbb{S}_n^*})^k$, we define the degree of fitness of X to Y by means of the composition between the operator * and the function Φ defined previously, i.e.: $\varphi_Y(X) = \Phi(X_1 * Y_1, ..., X_k * Y_k)$.*

Example 4.3.4 *Continuing with Example 4.3.3, we can consider the increasing function: $\phi(B_{sl}) = s + \frac{l-s}{3+1-s}$, $\forall s, l \in \{0, 1, 2, 3\}$ for our example. The function chosen for $\phi(B_{sl})$ could be defined differently in other contexts. In addition, to define the set of weights, w_i, we consider the RIM function, guided by the linguistic quantifier 'most of', expressed as:*

$$w_i = Q\left(\frac{i}{5}\right) - Q\left(\frac{i-1}{5}\right), \quad i = \{1, \ldots, 5\}, \tag{4.1}$$

where $Q(x) = x^{\frac{1}{2}}$.

Then, given the matching vector, $X * Y = (B_{13}, B_3, B_{13}, B_{23}, B_1)$, between vectors X and Y, and applying Definitions 4.3.3, 4.3.4, and 4.3.5, the degree of fitness of X to Y is $\varphi_Y(X) = \Phi(X_1 * Y_1, ..., X_5 * Y_5) = \Phi(B_{13}, B_3, B_{13}, B_{23}, B_1) = B_{23}$.
This result comes from the fact that: $B_3 \geq_L B_{23} \geq_L B_{13} \geq_L B_{13} \geq_L B_1$, $\phi(B_3) = 3$, $\phi(B_{23}) = \frac{5}{2}$, $\phi(B_{13}) = \frac{5}{3}$, $\phi(B_1) = 1$, and $w_1 = \sqrt{\frac{1}{5}}$, $w_2 = \sqrt{\frac{2}{5}} - \sqrt{\frac{1}{5}}$, $w_3 = \sqrt{\frac{3}{5}} - \sqrt{\frac{2}{5}}$, $w_4 = \sqrt{\frac{4}{5}} - \sqrt{\frac{3}{5}}$, $w_5 = 1 - \sqrt{\frac{4}{5}}$.

4.4 Proposed multi-criteria decision-aiding system to support university career services

Multi-criteria decision-aiding systems are designed to help users in situations where there are several decision factors that may cause controversy or complexity in decision processes [63, 158]. When these factors are related to user preferences but not easily measurable, the introduction of fuzzy and linguistic descriptions brings an appropriate framework [33, 37]. Multi-criteria decision support systems are comprised of several steps. First, the set of alternatives to be considered are introduced into the system. Second, the user or decision maker (DM) introduces his/her preferences with regards to different criteria. Finally, the system ranks or selects the alternatives that are closest to the user preferences. In this section we introduce a MCDA system to support college students with the internship job market application process.

4.4.1 System Description

Much like online job boards, university career services have a database of available positions. Companies post internship offerings for the upcoming year that can be reviewed by students online. Each internship has a record with information about the position such as its title, organization, and requirements, all of which are qualitative values. Each piece of information is provided in a free text field making the information unstructured. Therefore, it is difficult for a student to search for any position by keyword alone.

The proposed system caters to the interests of students rather than the requirements of a position. Specifically, the system is intended to help students identify internship offerings which best match their individual interests. To accomplish this task profiles are created for each student and position to represent preferences and features of each, respectively. Preferences are student interests elicited from each student and features are requirements determined from each position. Student's preferences are compared with each position's features. The outputs of the decision-making model are internship positions sorted in a manner which represents student's preferences. A diagram of the process follows in Figure 4.4.1 and detailed descriptions of the individual steps are given below.

4.4.2 Determine Features from All CVs

Before the process begins all of the curriculum vitaes (CV) of the participating students for the internship cycle are collected. From these CVs a set of features are determined to represent the main interests of the student body and define features for positions. Although there may be small changes in the features selected from year to year, extracting features specific to the student participants enables the system to better discern between positions. This is particularly important if in a given year all the positions are closely related, e.g. being in a single type of position.

To obtain these features, Latent Dirichlet allocation (LDA) is applied to the entire set of CVs. Originally developed by Blei et al. [23], LDA is an unsupervised topic modeling method. It is a generative probabilistic model of a collection of

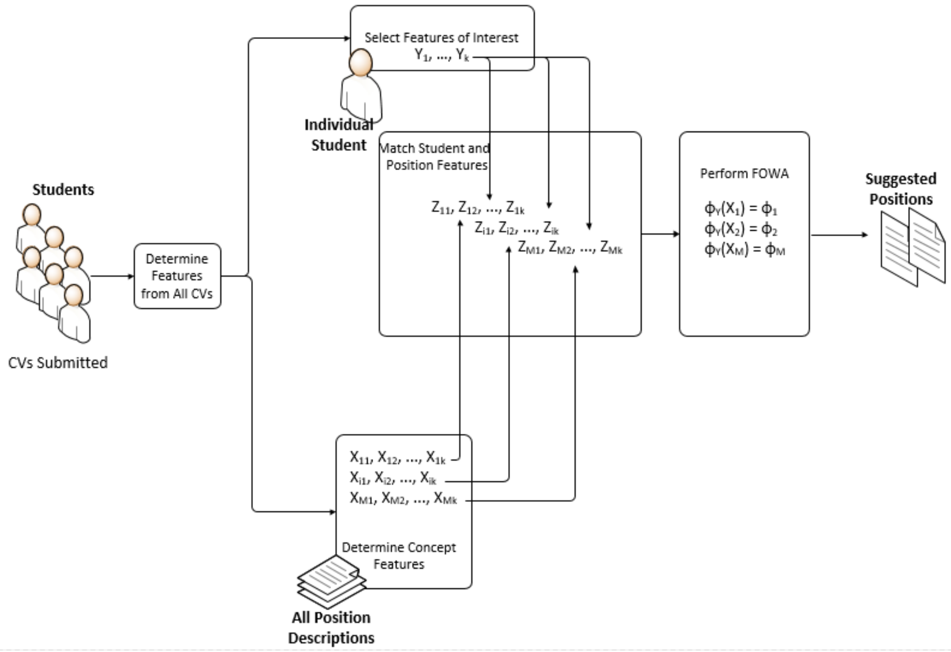

Figure 4.4.1: System process flow

documents. Each document is represented as a mixture of latent features based on keywords. The number of features, *K*, is determined using a qualitative approach, following [56]. This method consists of varying the number of features until an expert can recognize each feature from the keywords representing it (e.g. the keywords finance, economical and model are associated with finance). When two experts concur on the recognizable features the number of features is determined. Once the features have been determined, the student user interface is updated to reflect the considered features and the decision process begins.

4.4.3 Determine Features for Positions and Student

Initially the entire collection of internship postings are possible alternatives for every student. In order to be able to match these positions with the preferences of each student, the features of each position needs to be determined. One output of the LDA performed in Section 4.4.2 is a set of keywords related to each feature as shown in Table 4.5.1. For each student CV, there is a probability distribution of all possible features determined. Because this method seeks to provide equal value to all features, it normalizes each feature according to its distribution across students and jobs, respectively, by applying cutoff values. These cutoff values are then translated into a linguistic term (ie. "*None*" when $Feature_i < 10\%$, "*Low*" when $10\% \leq Feature_i < 50\%$, "*Medium*" when $50\% \leq Feature_i < 90\%$, "*High*" when $Feature_i \geq 90\%$).

When students enter the system, they will see the features determined in Section 4.4.2 available to them. The user interface is personalized to reflect each student's preference and level of preference based on the results of LDA. As LDA was applied to the entire collection of CVs to obtain the underlying features overall, it also computes the probability distribution of these features for each document. Each student may then adjust the feature preferences and levels presented to them, as necessary (e.g. change a feature preference from *"Low"* to *"Medium"-"High"*). Therefore, for each student, Y_j, the vector $Y_j = (Y_{j1}, \ldots, Y_{jk}) \in (H_{\mathbb{S}_n^*})^k$, with $k \leq K$, is set corresponding to his/her selected preferences expressed in hesitant fuzzy linguistic terms, as introduced in Section 4.3. The following figures detail the system's user interface. Specifically, from Figure 4.4.2 the student selects his/her interests and corresponding level. Note that a student may select a level that corresponds to two label categories (e.g. Jenn's preference for Sales and Marketing is between *"Low"* and *"Medium"*), or at least or at most some level of interest (e.g. Jenn has *"at most"* a medium preference for Strategy).

4.4.4 Match Student Interests and Position Features to Propose Positions

Each internship opportunity is an alternative for a specific student, Y_j. Therefore, to perform a match, we need to create the position profile, expressing the relevance of each feature determined from the collection of CVs. Once the student and position profiles have been created, a matching is performed between the preferences of the student and the features of each position. The system performs the matching process of Section 4.3.2. The process concludes with a proposed list of positions which best match the interests of the student as shown in Figure 4.4.3.

The match is performed between the preferences of the student and the features of the position, where only the features of each position representing the

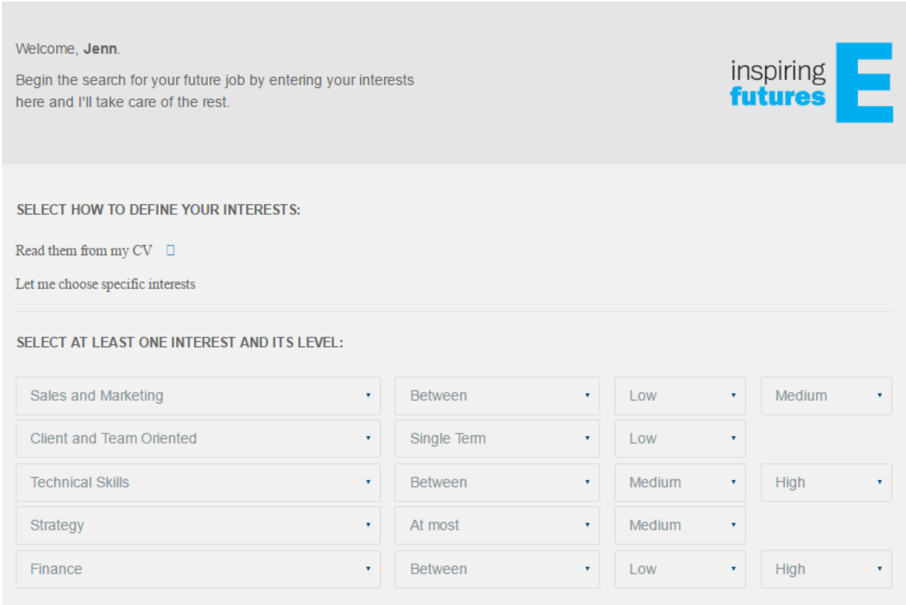

Figure 4.4.2: Interest selection user interface

preferences of the student are retained (i.e. the features that the individual identified as *"None"* are removed). Therefore, the position's vector is redefined as $X_i = (X_{i1}, \ldots, X_{ik})$ and is compared to the student's preferences, $Y_j = (Y_{j1}, \ldots, Y_{jk})$. The outcomes of the matching are linguistic labels, $H_{\mathbb{S}_n}$, that are assigned to a matching vector, $Z = (Z_1, \ldots, Z_k)$, based on the position's ability to satisfy the interests of the student.

Once the matching vector is obtained, a fuzzy order weighted average is computed. The FOWA, introduced in Section 4.3.2, is applied to aggregate the linguistic terms from the matching step in order to emphasize the features with the greatest match between students and positions. The resulting level of satisfaction is a fuzzy linguistic term set $\varphi_Y(X) = \Phi(X_1 * Y_1, ..., X_k * Y_k)$ obtained via the weighted average. Positions falling within the highest level of satisfaction are proposed. Note that the number of positions proposed can vary between students depending on the student preferencess and their match with each position's features.

Figure 4.4.3: Positions with highest level

4.5 A Real Case Example

In this real case example, the 2016 internship program for the Bachelor of Business Administration at ESADE Business School in Barcelona, Spain, was used to apply the proposed method. This program provides students with the opportunity to gain professional experience at an organisation. For some students, this may be their first-time working in their future profession.

4.5.1 Data Sets

The data set was composed of 275 student resumes and 1063 available internships. All resumes and internship descriptions in English were considered. The final data set consisted of 275 students and 549 internships. Student information was limited to the resumes provided for the purposes of the 2016 internship cycle. Internship positions included national and international postings.

4.5.2 Implementation and Results

We applied latent Dirichlet allocation (LDA) to extract features from the set of 275 CVs following the steps in Section 4.4.2. Five features were defined, as shown in Table 4.5.1, according to two experts as described in the method in [56].

Table 4.5.1: Features defined from collection of 275 CVs

Feature	Top 10 Keywords	Distribution
Client and Team Oriented	user, international, team, social, sales, sports, students, program, service, association	20%
Strategy	intern, project, consulting, strategy, competition, innovation, development, services, projects, case	20%
Sales and Marketing	marketing, sales, market, assistant, brand, managed, social, events, manager, collaborated	21%
Technical Skills	excel, word, office, access, powerpoint, marketing, spss, point, united, power	18%
Finance	financial, analysis, participated, team, companies, research, finance, investment, students, analyst	21%

With these features the system created the student profiles. The distribution of each feature was considered across all student resumes. Given this distribution, the percentiles 10th, 50th, and 90th were determined. For any student and feature, a value below the 10th percentile was discarded as it is assumed that the student would not have selected this feature. The linguistic terms "*Low*", "*Medium*", and "*High*" were assigned to the remaining features for each student. Therefore, the linguistic term set for this case includes the basic labels ("*Low*", "*Medium*", and "*High*") and its associated non-basic labels. Students are able to adjust the initial basic labels according to their preferences and apply basic or non-basic labels for each feature.

The rest of the case implementation follows the system description in Section 4.4.1. Finally, for each student, linguistic values are assigned expressing the fitness between the student and the position. The set of positions with a degree of satisfaction equal to *"High"*, according to the operator defined in Section 4.3.2, is shown to the student. Of the 549 positions, an average of 22 positions with a median of 13 were proposed to each student. The distribution of the variable, "number of positions selected for each student", is represented in Figure 4.5.1.

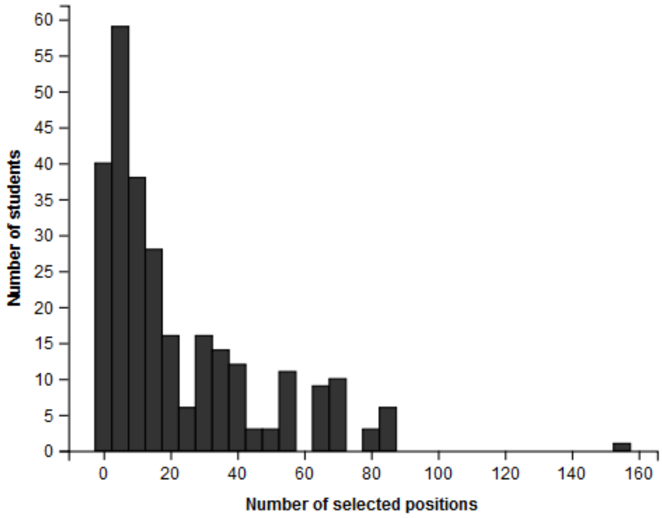

Figure 4.5.1: Distribution of number of positions selected per student using proposed method

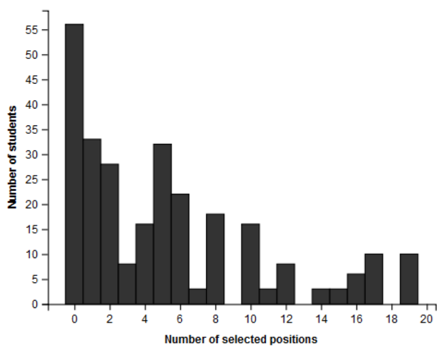

Figure 4.5.2: Distribution of number of positions selected per student using TOPSIS method

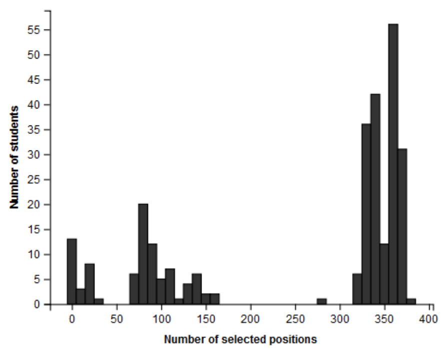

Figure 4.5.3: Distribution of number of positions selected per student using Hellinger distance

4.5.3 Evaluation of the method

As can be seen from Figure 4.5.1, using our proposed method, the number of positions for the student to review has been significantly reduced. By narrowing the focus for the student's internship search, he/she saves considerable time and can work more effectively with only the positions which match his/her interests. Overall, this efficiency is passed directly to the career services office. In a real life scenario, students would be able to refine their search by modifying their preference parameters, thus reducing the number of returned position results.

To evaluate the advantages and drawbacks of our proposed method, we will compare it to: 1) TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) and 2) a ranking method based on Hellinger distance. The first method is a ranking method based on a multi-granular fuzzy linguistic modeling that ranks alternatives based on comparing distances to a optimal alternative as defined in [20]. The second method is based on the classic Hellinger distance [24, 137] that does not convert attributes into linguistic terms but uses the frequency distribution of variables.

In order to compare these methods to ours, we used the same cut-off value for a recommendation to the user. In this case a "High" linguistic term (i.e. 90%). The results of the first comparison are depicted in Figure 4.5.2. This method recommended fewer jobs to individuals than our proposed method. In fact, our method recommended zero positions to at most 40 users while the TOPSIS method recommended zero positions to at most 55 users, demonstrating that more students received recommendations with our method. The results obtained with the second comparison method is based on the Hellinger distance used to determine the distance from a student's preferences to a position's features. As can be seen in Figure 4.5.3, this method recommended 65 or more positions to the majority of the students, while our method provided fewer than 40 recommendations to most students. The method presented recommends a number of positions between those of the Hellinger and TOPSIS methods. This number is closer to the one suggested than the other two methods. Sending 40-50 resumes to targeted companies is more effective than applying to every job on a job site [2]. Our method has a main advantage of an asymmetric matching of student preferences and position requirements that captures position features which meet or exceed student preferences. The main drawback is the loss of information due to the fact that the sorting method proposed is not symbolic and requires translation to numerical values to be computed and as the computation is with numerical values, the results need to

[2] `http://julliengordon.com/50-job-search-statistics-successful-job-seekers-need-know`

be translated to linguistic terms.

4.6 Conclusion

In this paper, a new method for sorting internship postings according to student interests has been introduced. This methodology improves existing methods in several ways. First, it proposes to perform a matching between students and internships from the perspective of the job candidate rather than the position. This is the reverse of the more popular matching to find the best candidate for a position. More specifically, the system is directed at students or new graduates with very little experience. Their interests may be a better representation of themselves since they have less relevant experience than seasoned professionals. In addition, as students may have had limited exposure to their fields of interest, they may not be aware of which keywords to use or they may not be aware of what types of available positions match their interests. A system such as this can facilitate the search process by narrowing the list of positions to the ones that best satisfy student interests. Second, the method considers a FOWA operator in the matching to capture the inherent uncertainty of personnel selection. Futhermore, the FOWA operator avoids filtering but simultaneously considers several relevant variables for the aggregation process. Lastly, the interests and features of the students and positions are represented as HFLTS, reflecting human tendency to opine with imprecision and hesitance in making decisions.

Our methodology can be extended to both sides of the general personnel assignment problem making the process more efficient. A position which is closely aligned with the interests of a job candidate may lead to better job loyalty. Therefore, as future research, we propose to adapt our methodology to other personnel selection environments like headhunting firms, online job boards, and industry human resources to uncover the interests of a job candidate prior to the interview process.

Regarding enhancements to the methodology, we plan to evaluate our method with a symmetric Sugeno Integral which is based only on min/max operations. The Sugeno Integral is useful to model situations where dependence of criteria are not certain [154]. In our specific problem context, features from which students select their interests are determined implicitly from their CVs. Therefore, the relationships between the features cannot be determined beforehand, making Sugeno Integral an interesting alternative. We would like to note that although the method is employing numeric operators with numeric weights, it does not match each of the labels to a numeric value, rather, it considers different levels of precision labels to be mapped to binary numerical values. The mapping of each of the labels is to a pair of numeric values in order to consider different levels of precision. The result obtained by applying a FOWA operator considers a lexicographic order among all labels. In this context, we expect label translations to be acceptable as the method is seeking to sort and group positions according to preferences rather than to identify the position having the best match. To that end, an extension of the method could include techniques which do not require label translation in order to better preserve human communicated preferences.

5

An OWA-based hierarchical clustering approach to understanding users' lifestyles

5.1 Introduction

When searching for items, such as destinations, attractions, restaurants, or travel arrangements people traditionally turn to recommender systems (RS) for assistance with finding items suited to their personal preferences. These RS leverage the ratings and profiles of individuals using the system in their search for an item. To identify items of interest to a user, RS use machine learning techniques such as clustering [163] and neural networks [120]. Clustering algorithms generally apply a distance measure based on similarity to aggregate individuals into groups. Individuals within the same group are more similar to each other than to those in other groups. User ratings, demographic or contextual attributes are commonly used to group users together and to predict a new user's satisfaction with an item. Demographic attributes about an individual, and contextual attributes about the situation under which a decision is made are considered in some cases. More recently, RS have been leveraging reviews to identify features [127], features and sentiment [65, 117], others have leveraged review metadata [120]. However, few have focused on grouping people according to a large number of attributes and content extracted from reviews to implicitly identify customer lifestyles.

With 4.9 million restaurant listings on TripAdvisor[1] and 19 million restaurant reservations a month on OpenTable [2], people's dining habits are changing. There were 177 million reviews posted on Yelp [3] last year of which 19% were about restaurants. There is a shift from spending less time cooking to dining out[2]. Food has assumed a prominent role in tourism due to its ability to bring people into its local culture [75]. The abundance of restaurant options makes the decision more challenging for the diner and attracting customers more difficult for the restaurateur[2]. Diners have many resources to which they may turn such as restaurant reviews and recommender systems. Community forums are recent additions [4] which aid diners with research when they cannot find an answer for a specific question. A main expectation of a poster to this forum is a timely and relevant response. However, due to large scale participation, meeting these expectations is a challenge [41] which has led to research on expertise identification [119, 162] and question routing [13, 41]. In general, these elements taken alone may bring about an appropriate and timely answer but lack personalization required for some questions.

In this paper, we propose to identify diners with shared lifestyles as compatible answerers and commenters to a question. This will reduce the set of possible advisers to those who share common interests. To this end, we propose to cluster individuals applying a new approach to measure the similarity between categorical variables. Specifically, the measure is a two-step approach. The first step considers the frequency distribution of categorical attributes and compares the attributes between two clusters by way of a marginal similarity degree. The second step employs an ordered weighted averaging aggregation (OWA) operator to the marginal similarity obtained for each couple. This similarity measure is implemented in a hi-

[1] http://ir.tripadvisor.com/static-files/6d4c71fd-3310-48c4-b4c5-d5ec04e69d5d

[2] https://openforbusiness.opentable.com/insider-information/how-diners-are-changing-and-what-it-means-for-restaurants/

[3] https://www.yelp.com/factsheet

[4] https://www.yelpblog.com/2017/02/qa

erarchical clustering framework. We demonstrate the applicability of the process to classify lifestyles based on text reviews and their associated restaurant attributes.

The rest of the paper is organized as follows. Section 5.2 is a review of related work. In Section 5.3 we describe the developed methodology. Next, the methodology is implemented in a real case example and the obtained results are discussed in Section 5.4. Finally, we provide concluding remarks and areas for future research in Section 5.5.

5.2 Related Work

5.2.1 Recommender systems for restaurants

Clustering is a common method which has been applied to group customers with similar preferences together [76] and customers with similar rating behaviors together [121]. The objective of the method is to divide a data set into groups such that the customers within them are homogeneous while the groups themselves are heterogeneous [118]. While these methods help to alleviate data sparsity, general aggregation of customer preferences can result in high overall customer group satisfaction but low individual satisfaction [44, 182]. Therefore, some works have considered that if restaurants providing similar services are frequented by customers with similar preferences, segmenting restaurants for different types of customers may balance customer group satisfaction and individual satisfaction [179, 182]. Different methods for recommending restaurants have been proposed. Zhang et al. [182] measure the correlation between customer and restaurant groups. Paradarami et al. [120] apply a neural network and combine implicit and explicit preference models. Others have considered restaurant recommendation in a mobile environment [181]. Fu et al. [64] develop a generative probabilistic model to exploit multi-aspect ratings of restaurant service quality. In this paper, customer lifestyles are inferred from the restaurants they have frequented and the reviews they have written about them. The methodology considers the explicit attributes of the restaurants and the inherent concerns of the reviewer in selecting a restaurant in determining the lifestyle clusters.

5.2.2 Profile Attributes

Multiple factors can influence restaurant selection. A study conducted by OpenTable[5] across three highly populated cities in the United States, found that occasions drive decisions around a meal. Depending on the occasion, different importance is placed on quality of food, price, location, ambience, and service. These are widely accepted attributes which explain restaurant patrons' behavior [79]. Food quality refers to patrons' perception of the food served such as its presentation, taste, or temperature. Price refers to whether or not the price charged for the visit is fair. Service quality is a contributor to satisfaction. Patrons' expect to dine in convenient locations. Ambience refers to environmental factors including lighting, space, and music.

Different methods have been proposed to identify elements influencing customers' decision to dine in a restaurant. Paradarami et al. [120] trained their restaurant recommendation model on a restaurant's rating and comments on users' reviews. Guo et al. [72] applied latent Dirichlet allocation to customer reviews and identified key dimensions of customer service expressed by hotel visitors. Rahimi et al. [127] applied natural language processing techniques to reviews and inferred restaurant features from their associated nouns and adjectives.

A common approach to representing user profiles in tourism is with a vector of numerical ratings. Each one corresponding to a user's interest towards an attribute of an item [27]. Also, as commented in [27], a user may be represented by a vector of preferred categories. Other works convert ratings into linguistic term sets [182]. Our paper follows a method similar to [127], the methodology developed in this paper uses natural language processing techniques and word frequency to infer restaurant features. Different from their study, this paper blends attributes obtained from reviews with those predefined about the restaurants blending both categorical and text-based elements. This frequency based approach can infer the degree of preference towards each attribute in the original decisions.

[5] https://openforbusiness.opentable.com/insider-information/how-diners-are-changing-and-what-it-means-for-restaurants/

5.2.3 Clustering Similarity Measures

Clustering is a method that groups a set of objects into undetermined number of classes [45]. The grouping is detected from similarities between characteristics found in the data with the objective of having high intra-cluster similarity and low inter-cluster similarity. The process is iterative until the clusters are stable.

Traditional methods use different measures (e.g. Pearson correlation, cosine, mean squared differences, and Euclidean distances) [25] to assess the cohesion of the clusters. However, recently, ordered weighted averaging (OWA) has been used frequently in classification [45, 102, 118]. OWA is useful in multi-criteria decision making problems which require the aggregation of distributed information [170]. These problems often require the inclusion of weights to signify the importance of different criteria [173]. Cheng et al. [45] applied an OWA operator to aggregate multi-attribute data into a "single attribute" in order to reduce the complexity of the clustering. Then, clustered this attribute following k-means. Nasibov and Kandemir-Cavas [118] considered an OWA-based linkage as a general form of the single, complete, and average linkage methods. They used it to find the distance between clusters for a hierarchical clustering scheme. Luukka and Kurama [102] applied OWA to a classification schema which first computes the similarity measure between each class vector and data vector, then, aggregates the similarity vector into a single value to determine to which class the data vector belongs. Two of these studies implemented their classification scheme with supervised learning technique and another two on medical data sets. In contrast, the methodology developed in this paper applies an unsupervised method to classify user lifestyles exhibited by their decision driver and attributes representative of their cumulative selected items. For each user, we have have an accumulation of observed purchases (restaurant choices). Each of these choices, are described by a set of attributes. Therefore, each user is represented by a set of distributions associated with the set of attributes. These attributes are qualitative in nature and are not converted to quantitative values for computation. The presented methodology implements a similarity measure which compares the frequency distribution of each variable and then aggregates these similarities using an OWA. The methodology enables the retention of the original information and therefore, is well suited for categorical data from popular search sites.

5.3 Proposed Methodology

In this section we introduce the formal framework needed for the cluster definition and the formal measures to compute and aggregate similarities. The process begins with a set of reviews and item descriptors. For ease of explanation we will refer to the item here as the restaurant as in our case example. All restaurants are described in terms of the same set of descriptors as defined on the social network platform. Each descriptor can have many dimensions where each dimension is a different attribute. The methodology developed is able to consider many attributes to preserve the original information in the clustering process. An example of possible descriptors and their related attributes are given Table 5.3.1.

Table 5.3.1: Sample descriptors and related attributes

Descriptor	Dimensions	Attributes
Food type	41	Comfort food, fast food, local flavor...
Dietary restrictions	7	Gluten-free, vegan, soy-free...
Special services	14	Catering, WiFi, delivery, take out...

Each customer is represented in terms of the descriptors of the restaurants he has frequented and descriptors driving his decision to dine at a restaurant. This decision driver was elicited from customers' reviews to understand concerns taken into consideration when choosing a restaurant. Customers are clustered according to their descriptors. Hierarchical agglomerative clustering (HAC) is implemented with a mixed similarity measure to determine the customer clusters. HAC has some advantages to identifying lifestyles which may be of interest. First, HAC is a method based solely on a similarity or distance measure. The algorithm does not require the computation of centroids in the process as do other clustering methods such as k-means. Second, HAC is less affected by outliers than other methods. Finally, HAC is adaptable to the similarity measure defined in this methodology because it maintains the content of the original information.

The presented methodology considers multiple descriptors to represent a user. Within each of these descriptors are numerous attributes. There are different quantities of attributes representing each descriptor. The proposed similarity measure allows the methodology to apply balanced importance to each of the descriptors. The similarity considers the relative frequencies of each attribute of each descriptor in a marginal similarity measure. Then the OWA operator is imposed to fuse marginal similarity of each descriptor into a single aggregated value measuring the similarity between the profiles of two individuals. The initial iteration is a comparison of individuals but subsequent steps occur between an individual and cluster or two clusters. The result of the process are groups of customers having shared lifestyles. For example, given a set of restaurants frequented, it can be surmised that a customer enjoys Korean and Mediterranean food. The places are best visited on Friday nights and weekends for groups of people. Reasons for selecting these places concern ambience and service. Therefore, we might infer that he enjoys visiting places which are social and entertaining in nature. Figure 5.3.1 illustrates the steps of the methodology.

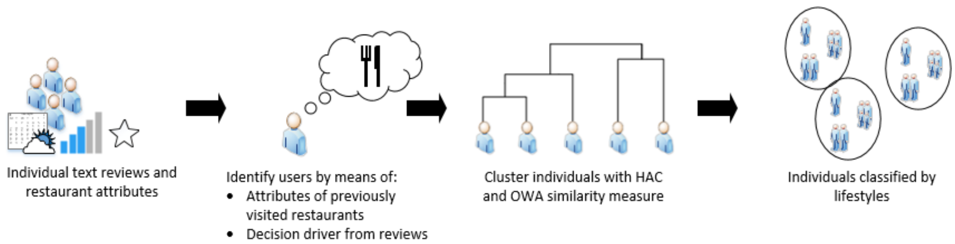

Figure 5.3.1: Proposed Methodology

5.3.1 Building customer profiles

The first part of a customer's profile contains all of the attributes about each of the restaurants he has frequented. These attributes are first categorized according to higher level descriptors. For example, "French", "Italian", and "Mexican", originally classified as different attributes may be combined into "Nationality". Similarly, "fast food", "buffet", and "drive-in" may be classified as "food type". For each customer, the frequency of each attribute exhibited by the restaurants he has reviewed are summed individually.

The second part of a customer's profile refers to his decision driver. The decision driver is derived directly from the restaurant patron's textual comments. They represent the concerns of a customer when dining out. A natural language processing technique is applied to lemmatize and tokenize the parts of speech of each patron's review text representation. The nouns related to the text reviews are retained. The top 100 nouns appearing across all the reviews are grouped according to four widely accepted attributes contributing to customer satisfaction: "food", "price", "location", "ambience", and "service" as discussed in Section 5.2. Each review text representation is then assessed for a frequency of term occurrence for each noun. For each customer the frequencies for each noun are summed individually. These attributes contribute to the descriptor "Decision Driver".

The attributes from the first and second parts are joined to give a vector relating each customer to the set of descriptors. Care is taken to review the descriptors for bias towards any one in particular. For instance, a high quantity of mentions about "service" may lead clusters to be highly stacked on the service variable misleading the interpretation of clusters. In other words, many customers will be clustered on the "service" variable making it difficult to separate customers into clusters clearly distinguishable by their lifestyles. Therefore, a normalization of the variables is imposed during the clustering process.

5.3.2 Similarity measure

Several similarity metrics are commonly used in literature with HAC to measure intercluster distancies: single, complete, average, weighted, and centroid linkage [116]. Inspired by [139], we propose to measure the similarity between clusters by a *marginal similarity degree* (MSD).

Given a partition $\mathcal{C}$ consisting of M clusters $\{C_1, \ldots, C_M\}$, μ_i^k is the marginal distribution of descriptor D_k, $k \in \{1, \ldots, P\}$, for each cluster C_i, $i \in \{1, \ldots, M\}$, that is defined by the frequencies of different attributes that the descriptor exhibits.

Given that $\{a_1, \ldots, a_{N_k}\}$ is a set of attributes for descriptor D_k, the MSD between each pair of clusters $C_i, C_j \in \mathcal{C}$ is computed for each descriptor D_k as:

$$Ms_{ij}^k = \frac{\sum_{m=1}^{N_k} f_{im}^k \cdot f_{jm}^k}{\sqrt{\sum_{m=1}^{N_k} (f_{im}^k)^2} \cdot \sqrt{\sum_{m=1}^{N_k} (f_{jm}^k)^2}}, \tag{5.1}$$

where f_{im}^k and f_{jm}^k are the normalized frequencies of values a_m of descriptor D_k in C_i and C_j, respectively.

Example 5.3.1 *Let us consider a set of customers divided into three clusters $C_1, C_2, C_3 \in \mathcal{C}$ and two descriptors $\{D_1, D_2\} = \{Nationality, Food\ type\}$. Table 5.3.2 provides the distribution of descriptors for each cluster.*

Table 5.3.2: Distribution of descriptors in each cluster

Cluster	D_1 **(Nationality)** {Korean, American}	D_2 **(Food type)** {Comfort food, Fast food, buffet, Local flavor}
C_1	(0.6, 0.4)	(0.3, 0.5, 0.2, 0.0)
C_2	(0.5, 0.5)	(0.2, 0.4, 0.3, 0.1)
C_3	(0.2, 0.8)	(0.6, 0.2, 0.1, 0.1)

The marginal similarity degree Ms_{ij}^k is computed for each pair of clusters for both descriptors using Eq. 5.1 (see Table 5.3.3).

Table 5.3.3: Marginal similarity degree (MSD) for cluster pairs

Ms_{ij}^k	C_1	C_2	C_3
C_1		(0.98, 0.95)	(0.86, 0.75)
C_2			(0.74, 0.68)
C_3			

Next, an OWA operator, introduced by [170], is applied to aggregate the marginal similarities between each pair of clusters to obtain a *global similarity degree* (GSD).

Definition 5.3.1 *An OWA operator of dimension n is a mapping $\varphi : \mathbb{R}^p \to \mathbb{R}$ with an associated weighting vector $W = (w_1, \ldots, w_h, \ldots, w_p) \in [0,1]^p$ such that*

$\sum_{h=1}^{p} w_h = 1$. *The OWA operator is defined as:*

$$\varphi_{OWA}(l_1, \ldots, l_p) = \sum_{h=1}^{p} w_h \cdot l_{\sigma(h)} \tag{5.2}$$

where $(l_1, \ldots, l_p)$ *is the vector of MSD associated with the set of descriptors being aggregated and* $\sigma : \{1, \ldots, p\} \to \{1, \ldots, p\}$ *a permutation such that* $l_{\sigma(h)} \geq l_{\sigma(h+1)}$, $\forall h \in \{1, \ldots, p\}$, *i.e.,* $l_{\sigma(h)}$ *is the h-th highest value in the set* $\{l_1, \ldots, l_p\}$.

There are different methods to obtain the weight vector W. For our purpose we will use a linguistic quantifier guided aggregation as defined in Equation 2.1 in which the decision maker selects a quantifier representing the proportion of criteria necessary for a good solution [171].

Note that a RIM function must be used to obtain positive weights w_i, and $Q(x) = x^a$ should be defined with $a \in [0, 1]$ to obtain a concave operator able to model those aggregations with importance associated with them.

The GSD for clusters C_1, C_2, C_3 are shown in Table 5.3.4. The pair of clusters with the greatest intercluster similarity are combined. In this example, C_1 and C_2 are joined together.

Table 5.3.4: Intercluster similarity of cluster pairs (matrix S)

Gs_{ij}	C_1	C_2	C_3
C_1		0.97	.83
C_2			0.72
C_3			

5.3.3 Hierarchical agglomerative clustering

Hierarchical agglomerative clustering (HAC) is a widely used approach to summarize data by grouping similar nodes together. It is a bottom-up approach where each individual starts as a singleton cluster. Then, in each iteration the two closest clusters are merged together until only one cluster remains containing all the nodes. Algorithm 1 describes our proposal for the HAC process.

Algorithm 1 Hierarchical Agglomerative Clustering process

1. Place each individual $X = (x_1, \ldots, x_P)$ into a singleton cluster. Each individual is defined by a set of qualitative descriptors $D_1, \ldots, D_P$ where x_k is the frequency that descriptor D_k occurs for the individual
2. Group identical individuals into clusters and consider the rest of the individuals as clusters
3. Compute intercluster similarities $Gs(C_i, C_j)$ into a squared matrix S
4. Using the matrix S, identify the two closest clusters
5. Merge the two closest clusters
6. Repeat steps three to five until only two clusters remain

The result of the process is a binary tree with each level representing a partition of the data. From these levels, a natural clustering is selected by satisfying interpretability requirements according to marketing experts.

5.3.4 Cluster selection

The usability of a clustering is based on its ability to inform and be interpreted. Therefore, examining partitions is advantageous to determine a number of clusters sufficient enough to generate new knowledge but small enough to produce an interpretable model. In marketing environments in which clustering is used to extract behavioral patterns to design market strategies, the number of clusters retained is usually between three and five [39]. This assumption does not imply that clusters outside of this range should be automatically discarded. It depends on the interpretability of the final partition retained. In this regard, clusters having at least a minimum number of individuals to be interesting enough to generate sufficient information are considered, in this methodology. This value varies with the domain and data. The first partition with at least five clusters satisfying this criteria is selected.

5.4 Real Case Example

In this section, we evaluate the performance of the proposed schema using a real-world data set of Yelp restaurant reviews [6]. First, we describe the data set and pre-processing steps taken. Then, we introduce the evaluation metric used in the real case example. Finally, we present the results of the proposed schema to show its interpretability in identifying customers' lifestyles.

5.4.1 Data set

To develop a real-case example, we used data from the Yelp 2017 Challenge [7]. The data set contained data objects for businesses, users, reviews, tips and check-ins. There were a total of 5,261,658 reviews and 174,567 businesses. We filtered for only restaurants to limit the scope of our implementation to one type of point of interest. There were 3,221,418 restaurant reviews.

To test our methodology, a pilot test has been conducted. As the proposed method applies hierarchical agglomerative clustering, a time intensive process, we sought to reduce the processing time by limiting the number of reviewers. In order to include a large volume of review text while limiting the number of reviewers, we selected the top 500 most prolific reviewers. The volume of review text assisted with the process of selecting the decision driver. All reviews in English were selected, leaving 499 reviewers. These reviewers generated 134,102 reviews for 31,562 restaurants. The number of reviews per person ranged from 169 to 2209.

To obtain the first part of a customer's profile from the restaurant attributes, we pre-processed the data in two sections. First, each restaurant was associated with a list of categories. From this list, there were 234 unique categories. Seven categories were removed which did not describe restaurants. One category that was only exhibited by restaurants visited by the previously removed reviewer was also removed. Two hundred twenty six categories were retained, each of which was grouped into four descriptors. These descriptors and the number of categories associated with each one are in the Table 5.4.1.

Nationality referred to the country of origin of the cuisine. Food type referred

[6] https://www.yelp.com

[7] https://www.yelp.com/dataset/challenge

Table 5.4.1: Distribution of categories per descriptor

Descriptor	Number of categories
Nationality	53
Food type	41
Specialty	46
Place	86
Total terms	226

to the type of cuisine such as barbecue or comfort food. Specialty referred to the entree in which the restaurant specializes. Place referred to the venue in which the restaurant is located. For example, a place might be a shopping mall.

The second section was obtained from the business features. There were 82 features from which we filter for those related to the restaurants our reviewers reviewed. Features containing only NA or missing values, and one not related to restaurants were removed. Features with different labels were expanded such that 58 features were retained, finally. These features were grouped into six descriptors. They are listed with their corresponding number of features in the Table 5.4.2. The two parts of the first component gave us ten descriptors. The number of reviews per restaurant is added as the eleventh descriptor. For ease of explanation, categories and features are referred to as attributes from this point forward. For each customer, the frequency of occurrences of each attribute were summed individually across all of the restaurants he reviewed.

Table 5.4.2: Distribution of attributes per descriptor

Descriptor	Number of attributes
Time	13
Parking	6
Dietary restrictions	7
Facilities	18
Special services	14
Total attributes	58

The second part of a customer's profile refers to his decision driver. This driver was obtained by applying natural language processing techniques to the reviews. Specifically, we leveraged the R package "UDPipe" to tag the parts-of-speech of each document. "UDPipe is an open-source pipeline performing tokenization, morphological analysis, part-of-speech tagging, lemmatization and dependency parsing" [146]. The nouns exhibited across all of the reviews were collected. These nouns were then grouped into five attributes: "food", "prices", "location", "ambience", and "service". For each customer the frequency of each noun exhibited in his review document is summed according to the attribute to which they belong. The vector of descriptors from the first and second parts are joined to represent a reviewer's profile. Elements contributing to the reviewer's profile are depicted in Figure 5.4.1.

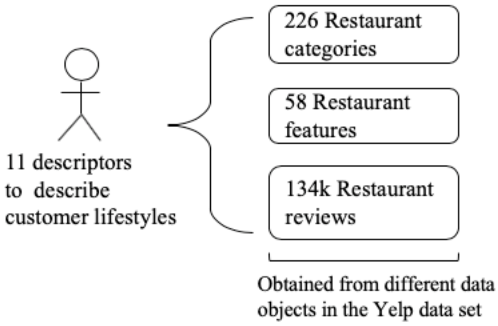

Figure 5.4.1: Reviewer profile elements

5.4.2 Implementation and Results

An iterative version of the HAC process described in Algorithm 1 is implemented in an R environment. For the OWA similarity measure, we defined the set of weights, w_i, considering the RIM function, guided by the linguistic quantifier "most of", where $Q(x) = x^{\frac{1}{2}}$. Once the clusters were generated by the HAC process, the partition containing the greatest number of interesting clusters was selected. Each of these clusters contained at least 30 individuals. The threshold 30 was used because it is the minimum number of samples to identify a pattern.

The selected partition defined a set of seven clusters. Note that there was one cluster that contained all the elements that were considered outliers. Each one represented a different lifestyle. Figures 5.4.2 and 5.4.3 depict some descriptors for the six interesting clusters in the final partition. The attributes in each descriptor which were exhibited in a cluster can be detected from these figures.

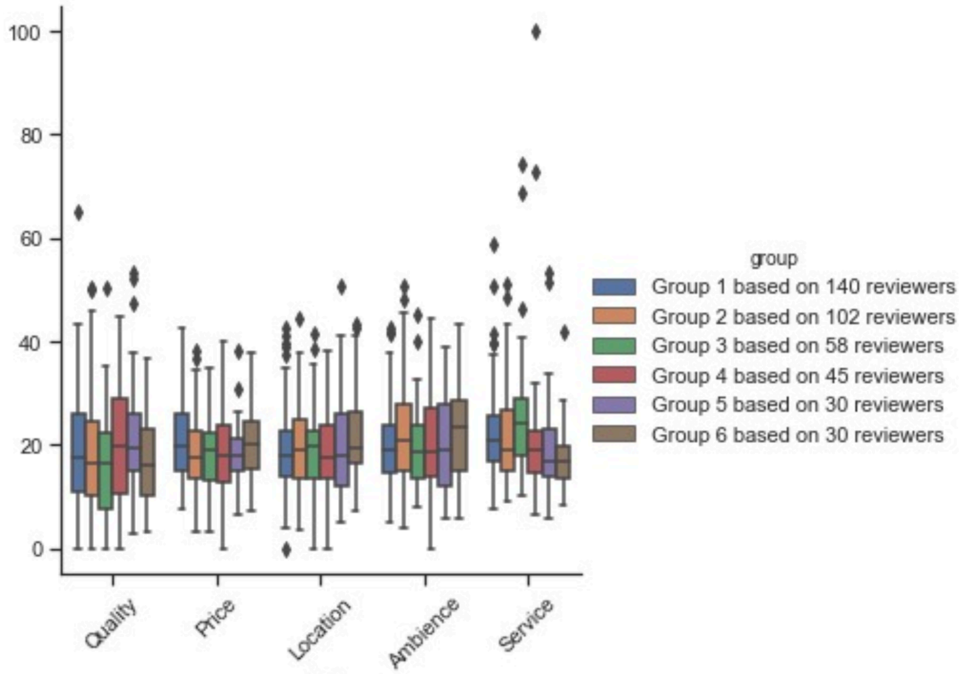

Figure 5.4.2: Decision driver considering the 6 clusters

Table 5.4.3 is a summary of the qualitative descriptions for each lifestyle as identified in each cluster. These descriptions are given in greater detail as follows:

- Cluster 1
  - The proportion of customers going to restaurants with a high number of reviews is very high
  - The proportion of customers going to restaurants with parking lots is low
  - The proportion of customers going to restaurants with coat check is high
  - The proportion of customers going to restaurants with outdoor seating is very high

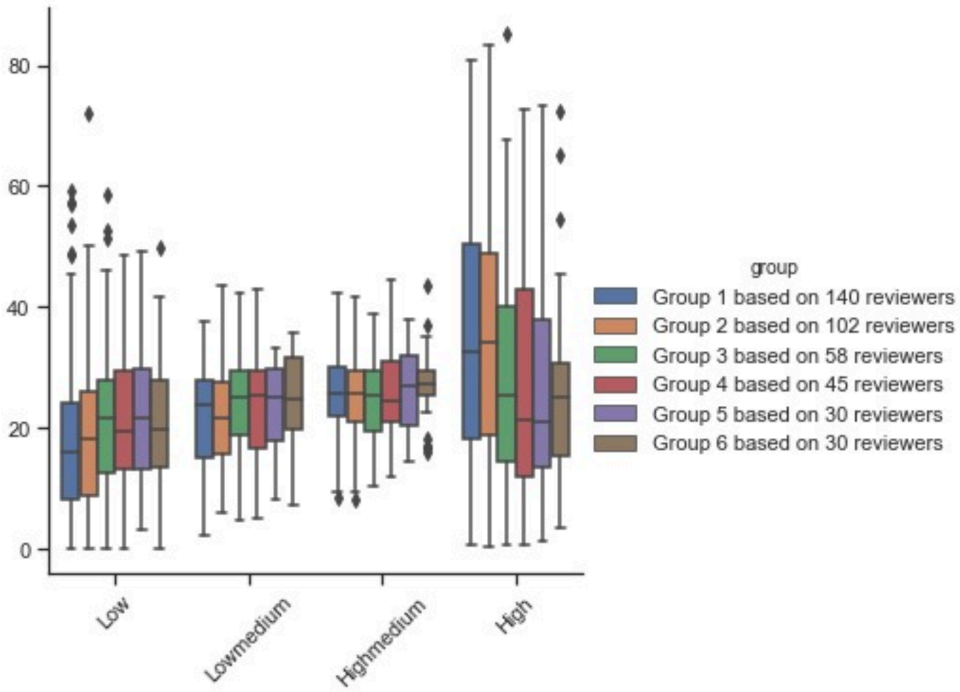

Figure 5.4.3: Number of reviews considering the 6 clusters

- The proportion of customers going to restaurants where dogs are allowed is high

- Cluster 2
  - The proportion of customers going to restaurants with a high number of reviews is very high
  - The proportion of customers going to restaurants with parking lots is low
  - The proportion of customers going to restaurants with valet parking is high
  - The proportion of customers going to restaurants with coatcheck is high
  - The proportion of customers going to restaurants during breakfast, brunch and dinner is high

  - The proportion of customers going to restaurants where dogs are allowed is high
  - The proportion of customers going to restaurants where music DJ is offered is high

- Cluster 3
  - The proportion of customers where service is a main concern of their decision driver is high
  - The proportion of customers going to restaurants with parking lots is low
  - The proportion of customers going to restaurants with wheelchair accessibility is high
  - The proportion of customers going to restaurants during happy hour is high
  - The proportion of customers going to restaurants good for kids is high

- Cluster 4
  - The proportion of customers where quality is a main concern of their decision driver is high
  - The proportion of customers going to restaurants where the best night is Wednesday is high
  - The proportion of customers going to restaurants good for kids is high

- Cluster 5
  - The proportion of customers going to restaurants serving alcohol is very high
  - The proportion of customers going to restaurants with parking lots is low
  - The proportion of customers going to restaurants during happy hour is very high

- Cluster 6
  - The proportion of customers going to restaurants serving alcohol is very high
  - The proportion of customers where ambience is a main concern of their decision driver is high
  - The proportion of customers going to restaurants with validated parking is high
  - The proportion of customers going to restaurants with wheelchair accessibility is high
  - The proportion of customers going to restaurants with outdoor drive-thrus is very high
  - The proportion of customers going to restaurants with outdoor seating is very high
  - The proportion of customers going to restaurants with TV and full bar is very high

Seven of the eleven variables revealed the most information about the lifestyles of the clusters retained from the HAC process. The qualitative descriptions distinguish the importance of different attributes for each lifestyle making the resulting clusters interpretable and informative. Thereby, supporting the selection of this cluster partition. The complete process took minutes on a 2.6 GHz Intel Core i7 MacBook Pro (2017) with 16 GB of memory demonstrating the efficiency of the methodology.

5.5 Conclusion

In this paper we proposed a methodology to cluster customers based on their inferred lifestyles. There are several benefits to our approach. First, the methodology considers all attributes from the original data set separately to obtain a more consistent assessment of customer considered descriptors. Second, these attributes

Table 5.4.3: Qualitative description of lifestyle for each cluster

Cluster	Diet	Driver	Rest. reviews	Parking	Services	Time	Facilities
#1			High(+)	Parking lot(-)	Coat-check(+)		Outdoor seating(+) Dog allowed(+)
#2			High(+)	Valet parking (+) Parking lot(-)	Coat-check(+)	Breakfast(+) Brunch(+) Dinner(+)	Music DJ casual(+) Dog allowed(+)
#3		Service(+)		Parking lot(-)	Wheelchair(+)	Happy hour(+)	Good for kids(+)
#4		Quality(+)				Best night Wednesday(+)	Good for kids(+)
#5	Alcohol(+)			Parking lot(-)		Happy hour(+)	
#6	Alcohol(+)	Ambience(+)		Validated parking(+)	Outdoor drive-thru(+) Wheelchair(+)		Has TV, full bar(+) Outdoor seating(+)

and the information contained within them are retained throughout the cluster formation until the cluster selection step. At which point only some of the clusters are retained and the original population of customers are reduced and the lifestyles generalized. The attributes retain their qualitative nature and the frequency of each occurrence allows a representation of each attribute's importance to each visitor to be expressed. The consistency of the first part can facilitate the interpretation of the lifestyles as it provides more information towards each descriptor enabling better discretization between them. Second, a marginal similarity degree is applied to measure the similarity between clusters simplifying the usual conversion to binary or dummy variables in order to be used in traditional methods. In essence, the proposed method maintains the number of dimensions, but does not incur a noticeable increase in computation. Lastly, an OWA similarity measure is applied to fuse multiple descriptors and provides the ability to include weights to signify the importance of different criteria.

As can be seen from the results of the real case example, the methodology leads to interpretable results. From the data set of reviewers, we were able to detect which reviewer exhibited which lifestyle attributes. Community forums can leverage this process to identify other diners on the platform with shared lifestyles who are suited to answer questions from a perspective similar to that of the individual asking the question.

Regarding enhancements to our methodology, we plan to build more flexibility into the model by allowing customers to be members of multiple clusters through methods such as fuzzy clustering. When deciding to which cluster a new customer belongs, there exists the same uncertainty. Further analysis and interpretation of clusters would be useful.

6 Conclusion and Future Research

People recommender systems (PRS) are a special type of recommender systems (RS). They are often adopted to identify people capable of fulfilling a task such as code reviewers and company experts [18, 49]. The matching in PRS is characterized by different challenges, from how to portray the reactive user to matching him with his respective activity.

The results obtained in this thesis are in two directions with respect to PRS. The first direction considers representation of people's profiles in a manner more expressive of their preferences. It explores how people's preferences could be elicited for single-event matchmaking environments. On other occasions, if a person were unable to convey her preferences, how might they be elicited? Given no certainty that previous matches were successful, can profiles be generalized to recommend future matches?

The second direction is centered around priorities in the matching processes of PRS. It considers how assigning a person too frequently might overextend her and not assigning a person leads to neglect. Additionally, it seeks to expose people to relevant items beyond those having the closest match to their preferences.

The main conclusions that are obtained in this thesis are addressed through the three use cases: 1) conference reviewer assignment, 2) personnel selection, and 3) lifestyle classification. Each use case is defined by a methodology which addresses its specific search problem and a real case example demonstrating its applicability.

This chapter is a discussion of the global contributions of this thesis. Theoretical contributions are highlighted from two directions: user profile creation and balanced recommendations. Afterwards, managerial implications are addressed, and limitations and areas for future research are presented. We conclude with some final remarks with respect to how these methodologies may be translated to practice.

6.1 Contribution to user profile creation

Three novel techniques for creating profiles are defined. These techniques allow us to obtain information implicitly. Generally, qualitative and linguistic opinions are explicitly provided by decision-makers [66, 77, 108]. Contrary to these approaches, a person's areas of interest are detected from his CV and expressed in terms of his uncertainty (Chapter 4). Therefore, reliance on a person's ability to explicitly state his preferences and outsider assessment can be avoided. Furthermore, profile features and respective preferences are generated together for all users, thus, their representation (level of interest and hesitancy) is measured on the same scale. This process precludes issues related to different people having different interpretations of the same linguistic term. Therefore, one user specifying a high interest in a particular area would not be different from another.

Some problems related to expert assignment necessitate criteria beyond expertise knowledge in aligning experts and tasks [93, 151, 175]. Different from the multi-criteria profiles reviewed, the profile attributes are derived wholly from external sources in an implicit manner (Chapter 3). This approach can characterize users from multiple perspectives and multi-dimensional points-of-view without influence from the user. Reconciliation of heterogeneous sources is simplified and has the potential to decrease the time needed to develop or update user profiles without diminishing the matching result benefiting specific environments (eg.

proposal, paper, project review) operating under time constraints.

A two-sided approach integrates explicitly stated attributes about items and those inferred from customer review text (Chapter 5). A result of the approach is a consistent method of extracting the level of preference from both implicitly and explicitly collected attributes. Thereby, all attributes represent customers in terms of the same scale and retain their qualitative nature. The similarity measure permits the original information contained in the data set to be retained throughout the cluster formation process and the importance of attributes to be considered.

6.2 Contribution to balanced recommendations

Different methodologies for assignment are developed in this thesis which contribute to PRS literature by taking initial steps to address overspecialization, rejection, and neglect. People who are frequently assigned may consume their availability while tasks in need of their knowledge remain unassigned. This situation can be disconcerting in PRS when there is limited resource availability and all items are of equal importance (ie. all items require a person to be assigned). In this thesis, we leverage coverage need (Chapter 3) to distribute people across items in order to prevent person overload, demonstrating the potential to limit overspecialization of a person, neglect of a task, leaving it without a relevant person assigned, and rejection of a task, a result of overspecialization or mismatch between knowledge and requirements. In addition, given a situation in which people are self-selected to tasks, we consider matching a user to tasks asymmetrically (Chapter 4) to capture tasks which meet or exceed user's preferences exposing him to tasks beyond those most comparable. It demonstrates that a reasonable number of tasks can be recommended and provides the variability lacking in overspecialized systems as in content-based methods [176]. From the opposite perspective, owners of tasks may have an expectation that there are candidates that can fulfill their tasks (Chapter 4). Exposure to tasks beyond those of an exact match can potentially lower the likelihood of neglect.

6.3 Managerial Implications

Our results offer several practical implications for PRS. The first implication is a methodology (as explained in Chapter 3) which creates user profiles from publicly available sources. The methodology does not impose on the candidate as all information is collected implicitly. In addition, it simplifies the reconciliation of heterogeneity inherent in information. Furthermore, managers can take advantage of this methodology to construct profiles of a candidate pool which can be maintained for future reference.

The second implication is a methodology (as seen in Chapter 4) which determines users' preferences in terms of their natural hesitancy. Building on the first methodology, candidates' underlying experience can be realized implicitly from textual information. Managers can assemble employee experience profiles from textual contributions made to corporate knowledge management systems (KMS). It can be useful to compare these profiles to objectives of the organization or employees for development purposes. The hesitancy in the profile is representative of the fluidity associated with continuous experience gain.

The third implication is a result of the methodology defined in Chapter 3. Consider an organization with several projects it needs to staff. This methodology identifies a person having satisfied most of the criteria; someone closely fitting the requirements. It incorporates constraints which prevent people from being over assigned reducing potential burnout which is a possible reason for rejecting future projects. Moreover, it can potentially manage human resource requests from multiple projects and fulfill the assignments without neglect.

Another implication relates to Chapter 4. This methodology is novel because of its consideration for the preferences of one side of the matching problem above another. Organizations can take advantage of this asymmetry. Candidates searching for a position would receive recommendations for which they are an exact match and those exceeding their requirements. Managers searching for a person to fulfill a position, will in turn, expand their candidate pool and introduce variability into the recommendation to include candidates whom they might not have previously considered.

The last implication is from Chapter 5 with respect to classifying people according to reviews and attributes. Managers seeking to assign employees to tasks can leverage customer reviews. If customer service is a primary concern, assigning a customer service representative who is both knowledgeable and compatible may be of interest. Analysis of past customer reviews of customer service experience is able to shed light on elements which are important to the customer in addition to the resolution of his problem. These elements may assist with matching customer service representatives at both a problem and compatibility level and make their service attendance more personalized.

6.4 Discussion, Limitations, and Future Research

In Chapter 3, expertise knowledge from heterogeneous sources is implicitly modelled. Terms that are extracted from heterogeneous sources are inherently conflicting. They can be ambiguous and of different granularity due to inconsistencies in sources, shortcomings of information retrieval techniques, and user supplied interests [10]. Each source may attribute different areas of interest, terms may not include contextual information with respect to the subject area, and authors may describe themselves in different granularities of interest areas. A limitation of the methodology presented is in the selection of the maximum value of each feature among all sources as the robustness of this process has not been tested. Extending the proposed methodology to account for uncertainty inherent in the sources themselves is a subject for future research.

In single-event matchmaking, level of interest cannot easily be inferred from repeated transactions. Methods to derive levels of interest from implicit information and reconcile them across sources can strengthen user profiles. Chapter 4 raises the topic of translating extracted levels of interest to linguistic values to represent users' uncertainty. However, further exploration and inclusion of the vagueness of the interpretation is interesting for future research.

The assignment process in Chapter 3 is iterative based on paper coverage need. Therefore, a limitation is the possibility that a paper having very few number of topics is not covered by a reviewer having relevant expertise. The paper's coverage need may be repeatedly fewer than other papers during each assignment iteration and at the time of this paper's assignment, reviewers with this expertise have exhausted their availability. Nevertheless, an OWA operator offers benefits over more traditional methods, like optimization, since it more closely resembles human reasoning and allows the inclusion of fuzzy approaches. As future research, a comparison of an optimization method and the method presented will be conducted.

In Chapter 4, preference labels are mapped to a pair of numeric values in order to consider different levels of precision during the matching process. Although, a reasonable number of items is recommended, techniques which do not require label translation may better preserve human communicated preferences. The methodology performs an asymmetric matching with primary attention given to satisfying the interests of the user. However, other methods which consider preference criteria for the level of satisfaction may be interesting. Maintaining the original preference label can provide flexibility to these matching techniques and is a subject interesting for future research. Furthermore, as features are determined implicitly, the relationships between the features cannot be determined. The Sugeno integral could be an alternative matching operator in future research.

Chapter 5 clusters people according to their lifestyle as determined from the integration of explicitly and implicitly obtained attributes taken from a large data set of transactions. Future application to single-event matchmaking, such as Community Question Answering (CQA) forums, can be beneficial to identify a group of potential answerers who, because of their shared lifestyles, can provide applicable advice within the desired context of the question.

The clustering process in Chapter 5 assigns each person to a unique cluster. As future research, introducing more flexibility, such as fuzzy descriptors, into the model will allow us to consider customers belonging to multiple clusters (i.e. obtaining a fuzzy segmentation). Moreover, further analysis and interpretation of clusters can assist with assigning a new customer to a cluster.

6.5 From theory to practice

The methodologies developed have been carried out in uses cases to demonstrate their applicability. In Section 6.3, managerial implications discussed how the presented methods may contribute to organizational operations. Building upon that discussion, these use cases serve as a pilot test for generalization to a real-life implementation.

Use cases one and two are from two different perspectives of identifying a candidate for a task. The first case considers the problem from the requirements of the position by finding a close match, whereas the second case emphasizes the preferences of the candidate based on asymmetrical matching. An extension of these use cases to a corporate career site may serve to identify both internal and external candidates for a position. Another extension of the first use case may be for the assignment of grant proposal reviewers. The third use case detects similar groups of people based on their past experiences which could be interesting for determining relevant items based on those sharing in the same lifestyle context. In addition, it may be useful in a CQA environment to identify people knowledgeable in the question at hand and who share similar lifestyles with the person asking the question. These aspects have the potential to increase the likelihood of a response and to be relevant to the context in which it was asked. Another extension may be to assist smart cities administrators with identifying activities to develop in the interest of its citizens.

The first use case has already been implemented in a real-life scenario. A private grant agency in Spain wanted to automate the task of matching reviewers to grants. Although, reviewer assignment is subject to constraints similar to those discussed in conference reviewer assignment [78], additional constraints were imposed by the grant agency. To support its first priority of optimizing reviewers' coverage of a proposal's topics, reviewers were required to cover the different topics of a multidisciplinary proposal collectively. The second priority mandated that the overall group of invited reviewers was a balanced representation of both genders. Third, as reviewers were paid to participate in each call, it was critical that each reviewer selected to review one proposal was assigned a minimum of five proposals to review. In order to make the assignments, the methodology from Chapter 3 of recommending reviewers to papers was adapted with the additional requirements. Two instances have been completed as part of this collaboration.

The third use case is currently being implemented as part of a project for the Spanish Ministry of Economy and Competitiveness. The project intends to discover tourist's lifestyles to personalize recommended destinations in terms of what and when to visit. Furthermore, understanding citizens' lifestyles can assist city planners with detecting areas of interests for potential growth and development.

Bibliography

[1] . *RecSys Challenge '17: Proceedings of the Recommender Systems Challenge 2017*, New York, NY, USA, 2017. ACM. ISBN 978-1-4503-5391-5.

[2] Zeinab Abbassi, Sihem Amer-Yahia, Laks VS Lakshmanan, Sergei Vassilvitskii, and Cong Yu. Getting recommender systems to think outside the box. In *Proceedings of the third ACM conference on Recommender systems*, pages 285–288. ACM, 2009.

[3] Himan Abdollahpouri, Robin Burke, and Bamshad Mobasher. Recommender systems as multistakeholder environments. In *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*, pages 347–348. ACM, 2017.

[4] Fabian Abel, Eelco Herder, Geert-Jan Houben, Nicola Henze, and Daniel Krause. Cross-system user modeling and personalization on the social web. *User Modeling and User-Adapted Interaction*, 23(2-3):169–209, 2013.

[5] Mark S Ackerman, Juri Dachtera, Volkmar Pipek, and Volker Wulf. Sharing knowledge and expertise: The cscw view of knowledge management. *Computer Supported Cooperative Work (CSCW)*, 22(4-6):531–573, 2013.

[6] Gediminas Adomavicius and Alexander Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE transactions on knowledge and data engineering*, 17(6): 734–749, 2005.

[7] Gediminas Adomavicius and Alexander Tuzhilin. Context-aware recommender systems. In *Recommender systems handbook*, pages 217–253. Springer, 2011.

[8] Gediminas Adomavicius, Nikos Manouselis, and Youngok Kwon. *Multi-Criteria Recommender Systems*, pages 769–803. Springer US, Boston, MA, 2011. ISBN 978-0-387-85820-3. doi: 10.1007/978-0-387-85820-3_24. URL https://doi.org/10.1007/978-0-387-85820-3_24.

[9] Núria Agell, Mónica Sánchez, Francesc Prats, and Llorenç Roselló. Ranking multi-attribute alternatives on the basis of linguistic labels in group decisions. *Information Sciences*, 209:49–60, 2012.

[10] Bahram Amini, Roliana Ibrahim, Mohd Shahizan Othman, and Ali Selamat. Capturing scholar's knowledge from heterogeneous resources for profiling in recommender systems. *Expert Systems with Applications*, 41(17): 7945–7957, 2014.

[11] Bahram Amini, Roliana Ibrahim, Mohd Shahizan Othman, and Mohammad Ali Nematbakhsh. A reference ontology for profiling scholar's background knowledge in recommender systems. *Expert Systems with Applications*, 42(2):913–928, 2015.

[12] Rajkumar Arun, Venkatasubramaniyan Suresh, CE Veni Madhavan, and MN Narasimha Murthy. On finding the natural number of topics with latent dirichlet allocation: Some observations. In *Pacific-Asia conference on knowledge discovery and data mining*, pages 391–402. Springer, 2010.

[13] Mossaab Bagdouri. Cross-platform question routing for better question answering. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1053–1053. ACM, 2015.

[14] Alvydas Baležentis, Tomas Baležentis, and Willem KM Brauers. Personnel selection based on computing with words and fuzzy multimoora. *Expert Systems with applications*, 39(9):7961–7967, 2012.

[15] Krisztian Balog, Leif Azzopardi, and Maarten de Rijke. A language modeling framework for expert finding. *Information Processing & Management*, 45(1):1–19, 2009.

[16] Krisztian Balog, Yi Fang, Maarten de Rijke, Pavel Serdyukov, Luo Si, et al. Expertise retrieval. *Foundations and Trends® in Information Retrieval*, 6(2–3):127–256, 2012.

[17] Chumki Basu, Haym Hirsh, William W Cohen, and Craig G Nevill-Manning. Technical paper recommendation: A study in combining multiple information sources. *J. Artif. Intell. Res.(JAIR)*, 14:231–252, 2001.

[18] Adil Baykasoglu, Turkay Dereli, and Sena Das. Project team selection using fuzzy optimization approach. *Cybernetics and Systems: An International Journal*, 38(2):155–185, 2007.

[19] Joeran Beel, Bela Gipp, Stefan Langer, and Corinna Breitinger. paper recommender systems: a literature survey. *International Journal on Digital Libraries*, 17(4):305–338, 2016.

[20] Ismat Beg and Tabasam Rashid. Topsis for hesitant fuzzy linguistic term sets. *International Journal of Intelligent Systems*, 28(12):1162–1171, 2013.

[21] Gleb Beliakov, Tomasa Calvo, and Simon James. Aggregation of preferences in recommender systems. In *Recommender systems handbook*, pages 705–734. Springer, 2011.

[22] David M Blei and J Lafferty. Topic models. text mining: theory and applications, 2009.

[23] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022, 2003.

[24] DM Blei and J Lafferty. Text mining: Classification, clustering, and applications. *chapter Topic Models, Chapman & Hall/CRC*, 2009.

[25] Jesús Bobadilla, Fernando Ortega, Antonio Hernando, and Abraham Gutiérrez. Recommender systems survey. *Knowledge-based systems*, 46: 109–132, 2013.

[26] Jesús Bobadilla, Francisco Serradilla, Antonio Hernando, et al. Collaborative filtering adapted to recommender systems of e-learning. *Knowledge-Based Systems*, 22(4):261–265, 2009.

[27] Joan Borràs, Antonio Moreno, and Aida Valls. Intelligent tourism recommender systems: A survey. *Expert Systems with Applications*, 41(16):7370–7389, 2014.

[28] Pierre Bourdieu. *Distinction: A social critique of the judgement of taste*. Routledge, 2013.

[29] Christopher Brewster, Simon Jupp, Joanne Luciano, David Shotton, Robert D. Stevens, and Ziqi Zhang. Issues in learning an ontology from text. *BMC Bioinformatics*, 10(5):S1, May 2009. ISSN 1471-2105. doi: 10.1186/1471-2105-10-S5-S1. URL https://doi.org/10.1186/1471-2105-10-S5-S1.

[30] Robin Burke. Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, 12(4):331–370, 2002.

[31] Robin Burke. Multisided fairness for recommendation. *arXiv preprint arXiv:1707.00093*, 2017.

[32] Robin Burke and Maryam Ramezani. Matching recommendation technologies and domains. In *Recommender systems handbook*, pages 367–386. Springer, 2011.

[33] Francisco Javier Cabrerizo, Ignacio Javier Pérez, Francisco Chiclana, and Enrique Herrera-Viedma. Group decision making: Consensus approaches based on soft consensus measures. In *Fuzzy Sets, Rough Sets, Multisets and Clustering*, pages 307–321. Springer, 2017.

[34] Paula Caligiuri and David C Thomas. From the editors: How to write a high-quality review, 2013.

[35] Lourdes Canós and Vicente Liern. Soft computing-based aggregation methods for human resource management. *European Journal of Operational Research*, 189(3):669–681, 2008.

[36] Peter Cappelli. Your approach to hiring is all wrong outsourcing and algorithm won't get you the people you need. *HARVARD BUSINESS REVIEW*, 97(3):47–57, 2019.

[37] Christer Carlsson and Robert Fullér. Fuzzy multiple criteria decision making: Recent developments. *Fuzzy sets and systems*, 78(2):139–153, 1996.

[38] Walter Carrer-Neto, María Luisa Hernández-Alcaraz, Rafael Valencia-García, and Francisco García-Sánchez. Social knowledge-based recommender system. application to the movies domain. *Expert Systems with applications*, 39(12):10990–11000, 2012.

[39] M. Casabayó. *Shopping behaviour forecasts: experiments based on a fuzzy learning technique in the Spanish food retailing industry*. PhD thesis, University of Edinburgh, 2005.

[40] Jonathan Chang, Sean Gerrish, Chong Wang, Jordan L Boyd-Graber, and David M Blei. Reading tea leaves: How humans interpret topic models. In *Advances in neural information processing systems*, pages 288–296, 2009.

[41] Shuo Chang and Aditya Pal. Routing questions for collaborative answering in community question answering. In *2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013)*, pages 494–501. IEEE, 2013.

[42] Laurent Charlin and Richard S Zemel. The Toronto paper matching system: an automated paper-reviewer assignment system. In *ICML: Workshop on Peer Reviewing and Publishing Models (PEER)*, volume 10, 2013. URL https://openreview.net/pdf?id=caynafZAnBafx.

[43] Ling-Show Chen and Ching-Hsue Cheng. Selecting is personnel use fuzzy gdss based on metric distance method. *European Journal of Operational Research*, 160(3):803–820, 2005.

[44] Yen-Liang Chen, Li-Chen Cheng, and Ching-Nan Chuang. A group recommendation system with consideration of interactions among group members. *Expert systems with applications*, 34(3):2082–2090, 2008.

[45] Ching-Hsue Cheng, Jia-Wen Wang, and Ming-Chang Wu. Owa-weighted based clustering method for classification problem. *Expert Systems with Applications*, 36(3):4988–4995, 2009.

[46] Philipp Cimiano, Andreas Hotho, and Steffen Staab. Learning concept hierarchies from text corpora using formal concept analysis. *Journal of artificial intelligence research*, 24:305–339, 2005.

[47] Don Conry, Yehuda Koren, and Naren Ramakrishnan. Recommender systems for the conference paper assignment problem. In *Proceedings of the third ACM conference on Recommender systems*, pages 357–360. ACM, 2009.

[48] G Sena Daş and Tolunay Göçken. A fuzzy approach for the reviewer assignment problem. *Computers & Industrial Engineering*, 72:50–57, 2014.

[49] Andre de Korvin, Margaret F Shipley, and Robert Kleyle. Utilizing fuzzy compatibility of skill sets for team selection in multi-phase projects. *Journal of Engineering and Technology Management*, 19(3-4):307–319, 2002.

[50] Christian Desrosiers and George Karypis. A comprehensive survey of neighborhood-based recommendation methods. In *Recommender systems handbook*, pages 107–144. Springer, 2011.

[51] Paul DiMaggio. Classification in art. *American sociological review*, pages 440–455, 1987.

[52] Gideon Dror, Yehuda Koren, Yoelle Maarek, and Idan Szpektor. I want to answer; who has a question?: Yahoo! answers recommender system. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1109–1117. ACM, 2011.

[53] Susan T Dumais and Jakob Nielsen. Automating the assignment of submitted manuscripts to reviewers. In *Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 233–244. ACM, 1992.

[54] Pragya Dwivedi and Kamal K Bharadwaj. e-learning recommender system for a group of learners based on the unified learner profile approach. *Expert Systems*, 32(2):264–276, 2015.

[55] Kate Ehrlich, Ching-Yung Lin, and Vicky Griffiths-Fisher. Searching for experts in the enterprise: combining text and social network analysis. In *Proceedings of the 2007 international ACM conference on Supporting group work*, pages 117–126. ACM, 2007.

[56] Michael S Evans. A computational approach to qualitative analysis in large textual datasets. *PloS one*, 9(2):e87908, 2014.

[57] Evanthia Faliagka, Athanasios Tsakalidis, and Giannis Tzimas. An integrated e-recruitment system for automated personality mining and applicant ranking. *Internet research*, 22(5):551–568, 2012.

[58] Evanthia Faliagka, Lazaros Iliadis, Ioannis Karydis, Maria Rigou, Spyros Sioutas, Athanasios Tsakalidis, and Giannis Tzimas. On-line consistent ranking on e-recruitment: seeking the truth behind a well-formed cv. *Artificial Intelligence Review*, 42(3):515–528, 2014.

[59] Hui Fang and ChengXiang Zhai. Probabilistic models for expert finding. In *European Conference on Information Retrieval*, pages 418–430. Springer, 2007.

[60] Maryam Fazel-Zarandi, Hugh J Devlin, Yun Huang, and Noshir Contractor. Expert recommendation based on social drivers, social network analysis, and semantic data representation. In *Proceedings of the 2nd international workshop on information heterogeneity and fusion in recommender systems*, pages 41–48. ACM, 2011.

[61] Ignacio Fernández-Tobías, Matthias Braunhofer, Mehdi Elahi, Francesco Ricci, and Iván Cantador. Alleviating the new user problem in collaborative filtering by exploiting personality information. *User Modeling and User-Adapted Interaction*, 26(2-3):221–255, 2016.

[62] P Festa. A brief introduction to exact, approximation, and heuristic algorithms for solving hard combinatorial optimization problems. In *2014 16th*

International Conference on Transparent Optical Networks (ICTON), pages 1–20. IEEE, 2014.

[63] J. Figueira, S. Greco, and M. Ehrgott. *Multiple Criteria Decision Analysis: State of the Art Surveys*. Springer Verlag, Boston, Dordrecht, London, 2005.

[64] Yanjie Fu, Bin Liu, Yong Ge, Zijun Yao, and Hui Xiong. User preference learning with multiple information fusion for restaurant recommendation. In *Proceedings of the 2014 SIAM international conference on data mining*, pages 470–478. SIAM, 2014.

[65] Gayatree Ganu, Yogesh Kakodkar, and AméLie Marian. Improving the quality of predictions using textual information in online user reviews. *Information Systems*, 38(1):1–15, 2013.

[66] José Luis García-Lapresta. A general class of simple majority decision rules based on linguistic opinions. *Information sciences*, 176(4):352–365, 2006.

[67] Susan Gauch, Mirco Speretta, Aravind Chandramouli, and Alessandro Micarelli. User profiles for personalized information access. In *The adaptive web*, pages 54–89. Springer, 2007.

[68] Pierdaniele Giaretta and N Guarino. Ontologies and knowledge bases towards a terminological clarification. *Towards very large knowledge bases: knowledge building & knowledge sharing*, 25(32):307–317, 1995.

[69] Judy Goldsmith and Robert H Sloan. The ai conference paper assignment problem. In *Proc. AAAI Workshop on Preference Handling for Artificial Intelligence, Vancouver*, pages 53–57, 2007.

[70] Carlos A Gomez-Uribe and Neil Hunt. The netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems (TMIS)*, 6(4):13, 2016.

[71] Zülal Güngör, Gürkan Serhadlıoğlu, and Saadettin Erhan Kesen. A fuzzy ahp approach to personnel selection problem. *Applied Soft Computing*, 9 (2):641–646, 2009.

[72] Yue Guo, Stuart J Barnes, and Qiong Jia. Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tourism Management*, 59:467–483, 2017.

[73] Anika Gupta and Deepak Garg. Applying data mining techniques in job recommender system for considering candidate job preferences. In *Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on*, pages 1458–1465. IEEE, 2014.

[74] David Hartvigsen, Jerry C Wei, and Richard Czuchlewski. The conference paper-reviewer assignment problem. *Decision Sciences*, 30(3):865–876, 1999.

[75] Joan C Henderson. Food tourism reviewed. *British food journal*, 111(4): 317–326, 2009.

[76] Thorsten Hennig-Thurau, Kevin P Gwinner, Gianfranco Walsh, and Dwayne D Gremler. Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet? *Journal of interactive marketing*, 18(1):38–52, 2004.

[77] Francisco Herrera and Luis Martínez. A model based on linguistic 2-tuples for dealing with multigranular hierarchical linguistic contexts in multi-expert decision-making. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 31(2):227–234, 2001.

[78] Seth Hettich and Michael J Pazzani. Mining for proposal reviewers: lessons learned at the national science foundation. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 862–871. ACM, 2006.

[79] Sunghyup Sean Hyun. Predictors of relationship quality and loyalty in the chain restaurant industry. *Cornell Hospitality Quarterly*, 51(2):251–267, 2010.

[80] Mehmet Kabak, Serhat Burmaoğlu, and Yiğit Kazançoğlu. A fuzzy hybrid mcdm approach for professional selection. *Expert Systems with Applications*, 39(3):3516–3525, 2012.

[81] Yordan Kalmukov. Architecture of a conference management system providing advanced paper assignment features. *arXiv:111.6934*, 2011.

[82] Yordan Kalmukov. Describing papers and reviewers' competences by taxonomy of keywords. *arXiv:1309.6527*, 2013.

[83] Maryam Karimzadehgan and ChengXiang Zhai. Constrained multi-aspect expertise matching for committee review assignment. In *Proceedings of the 18th ACM conference on Information and knowledge management*, pages 1697–1700. ACM, 2009.

[84] Maryam Karimzadehgan, ChengXiang Zhai, and Geneva Belford. Multi-aspect expertise matching for review assignment. In *Proceedings of the 17th ACM conference on Information and knowledge management*, pages 1113–1122. ACM, 2008.

[85] Robert Kass and Tim Finin. Modeling the user in natural language systems. *Computational Linguistics*, 14(3):5–22, 1988.

[86] Alecos Kelemenis and Dimitrios Askounis. A new topsis-based multi-criteria approach to personnel selection. *Expert systems with applications*, 37(7):4999–5008, 2010.

[87] Tsvi Kuflik, Judy Kay, and Aaron Quigley. Preface to the special issue on ubiquitous user modeling and user-adapted interaction. *User Modeling and User-Adapted Interaction*, 25(3):185–187, 2015.

[88] J Richard Landis and Gary G Koch. The measurement of observer agreement for categorical data. *biometrics*, pages 159–174, 1977.

[89] Kurt Lewin. The research center for group dynamics at massachusetts institute of technology. *Sociometry*, 8(2):126–136, 1945.

[90] Baochun Li and Y Thomas Hou. The new automated ieee infocom review assignment system. *IEEE Network*, 30(5):18–24, 2016.

[91] Lei Li and Tao Li. Meet: a generalized framework for reciprocal recommender systems. In *Proceedings of the 21st ACM international conference on Information and knowledge management*, pages 35–44. ACM, 2012.

[92] Ming Li, Lu Liu, and Chuan-Bo Li. An approach to expert recommendation based on fuzzy linguistic method and fuzzy text classification in knowledge management systems. *Expert Systems with Applications*, 38(7):8586–8596, 2011.

[93] Xinlian Li and Toyohide Watanabe. Automatic paper-to-reviewer assignment, based on the matching degree of the reviewers. *Procedia Computer Science*, 22:633–642, 2013.

[94] Shu-Chuan Liao, Kuo-Fong Kao, I-En Liao, Hui-Lin Chen, and Shu-O Huang. Pore: a personal ontology recommender system for digital libraries. *The Electronic Library*, 27(3):496–508, 2009.

[95] Duanduan Liu, Wei Xu, Wei Du, and Fuyin Wang. How to choose appropriate experts for peer review: An intelligent recommendation method in a big data context. *Data Science Journal*, 14, 2015.

[96] Ou Liu, Jun Wang, Jian Ma, and Yonghong Sun. An intelligent decision support approach for reviewer assignment in r&d project selection. *Computers in Industry*, 76:1–10, 2016.

[97] Xiang Liu, Torsten Suel, and Nasir Memon. A robust model for paper reviewer assignment. In *Proceedings of the 8th ACM Conference on Recommender systems*, pages 25–32. ACM, 2014.

[98] Pasquale Lops, Marco De Gemmis, and Giovanni Semeraro. Content-based recommender systems: State of the art and trends. In *Recommender systems handbook*, pages 73–105. Springer, 2011.

[99] Jie Lu, Dianshuang Wu, Mingsong Mao, Wei Wang, and Guangquan Zhang. Recommender system application developments: a survey. *Decision Support Systems*, 74:12–32, 2015.

[100] Brian Lucey. Ten tips from an editor on undertaking academic peer review for journals. *Available at SSRN 2331281*, 2013.

[101] Pasi Luukka and Mikael Collan. Fuzzy scorecards, fhowa, and a new fuzzy similarity based ranking method for selection of human resources. In *Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on*, pages 601–606. IEEE, 2013.

[102] Pasi Luukka and Onesfole Kurama. Similarity classifier with ordered weighted averaging operators. *Expert Systems with Applications*, 40(4): 995–1002, 2013.

[103] Susan A Lynham. The general method of theory-building research in applied disciplines. *Advances in developing human resources*, 4(3):221–241, 2002.

[104] Tariq Mahmood and Francesco Ricci. Improving recommender systems with adaptive conversational strategies. In *Proceedings of the 20th ACM conference on Hypertext and hypermedia*, pages 73–82. ACM, 2009.

[105] Jochen Malinowski, Tobias Keim, Oliver Wendt, and Tim Weitzel. Matching people and jobs: A bilateral recommendation approach. In *System Sciences, 2006. HICSS'06. Proceedings of the 39th Annual Hawaii International Conference on*, volume 6, pages 137c–137c. IEEE, 2006.

[106] Lucas Marin, David Isern, and Antonio Moreno. Dynamic adaptation of numerical attributes in a user profile. *Applied Intelligence*, 39(2):421–437, September 2013. ISSN 0924-669X. doi: 10.1007/s10489-012-0421-5. URL http://dx.doi.org/10.1007/s10489-012-0421-5.

[107] Jorge Martinez-Gil, Bernhard Freudenthaler, and Thomas Natschläger. Recommendation of job offers using random forests and support vector machines. In *Proceedings of the Workshops of the EDBT/ICDT 2018 Joint Conference (EDBT/ICDT 2018)*, 2018.

[108] Sebastia Massanet, Juan Vicente Riera, Joan Torrens, and Enrique Herrera-Viedma. A new linguistic computational model based on discrete fuzzy numbers for computing with words. *Information Sciences*, 258:277–290, 2014.

[109] Mark T Maybury. Discovering distributed expertise. *Regarding the "Intelligence" in Distributed Intelligent SystemsMITRE*, 2007.

[110] Frank McSherry and Ilya Mironov. Differentially private recommender systems: Building privacy into the netflix prize contenders. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 627–636. ACM, 2009.

[111] Juan Julián Merelo-Guervós and Pedro Castillo-Valdivieso. Conference paper assignment using a combined greedy/evolutionary algorithm. In *Parallel Problem Solving from Nature-PPSN VIII*, pages 602–611. Springer, 2004.

[112] Stuart E Middleton, David De Roure, and Nigel R Shadbolt. Ontology-based recommender systems. In *Handbook on ontologies*, pages 779–796. Springer, 2009.

[113] David Mimno and Andrew McCallum. Expertise modeling for matching papers with reviewers. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 500–509. ACM, 2007.

[114] Rosana Montes, Ana M Sánchez, Pedro Villar, and Francisco Herrera. A web tool to support decision making in the housing market using hesitant fuzzy linguistic term sets. *Applied Soft Computing*, 35:949–957, 2015.

[115] Juan Antonio Morente-Molinera, Ignacio J Pérez, M Raquel Ureña, and Enrique Herrera-Viedma. On multi-granular fuzzy linguistic modeling in group decision making problems: a systematic review and future trends. *Knowledge-Based Systems*, 74:49–60, 2015.

[116] Daniel Müllner. Modern hierarchical, agglomerative clustering algorithms. *arXiv preprint arXiv:1109.2378*, 2011.

[117] Cataldo Musto, Marco de Gemmis, Giovanni Semeraro, and Pasquale Lops. A multi-criteria recommender system exploiting aspect-based sentiment analysis of users' reviews. In *Proceedings of the eleventh ACM conference on recommender systems*, pages 321–325. ACM, 2017.

[118] Efendi Nasıbov and Cagin Kandemır-Cavas. Owa-based linkage method in hierarchical clustering: Application on phylogenetic trees. *Expert Systems with Applications*, 38(10):12684–12690, 2011.

[119] Mahmood Neshati, Zohreh Fallahnejad, and Hamid Beigy. On dynamicity of expert finding in community question answering. *Information Processing & Management*, 53(5):1026–1042, 2017.

[120] Tulasi K Paradarami, Nathaniel D Bastian, and Jennifer L Wightman. A hybrid recommender system using artificial neural networks. *Expert Systems with Applications*, 83:300–313, 2017.

[121] Moon-Hee Park, Han-Saem Park, and Sung-Bae Cho. Restaurant recommendation for group of people in mobile environments using probabilistic multi-criteria decision making. In *Asia-Pacific Conference on Computer Human Interaction*, pages 114–122. Springer, 2008.

[122] William Pasmore. Action research in the workplace: The socio-technical perspective. *Handbook of action research*, 2:38–48, 2006.

[123] Andreas Pesenhofer, Rudolf Mayer, and Andreas Rauber. Improving scientific conferences by enhancing conference management systems with information mining capabilities. In *Digital Information Management, 2006 1st International Conference on*, pages 359–366. IEEE, 2006.

[124] Luiz Pizzato, Tomek Rej, Thomas Chung, Kalina Yacef, Irena Koprinska, and Judy Kay. Reciprocal recommenders. *ITWP 2010*, page 53, 2010.

[125] Luiz Pizzato, Tomasz Rej, Joshua Akehurst, Irena Koprinska, Kalina Yacef, and Judy Kay. Recommending people to people: the nature of reciprocal recommenders with a case study in online dating. *User Modeling and User-Adapted Interaction*, 23(5):447–488, 2013.

[126] Jarosław Protasiewicz, Witold Pedrycz, Marek Kozłowski, Sławomir Dadas, Tomasz Stanisławek, Agata Kopacz, and Małgorzata Gałężewska. A recommender system of reviewers and experts in reviewing problems. *Knowledge-Based Systems*, 106:164–178, 2016.

[127] Sohrab Rahimi, Clio Andris, and Xi Liu. Using yelp to find romance in the city: A case of restaurants in four cities. In *Proceedings of the 3rd ACM SIGSPATIAL Workshop on Smart Cities and Urban Analytics*, page 3. ACM, 2017.

[128] Daniel Ramage, Christopher D Manning, and Susan Dumais. Partially labeled topic models for interpretable text mining. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 457–465. ACM, 2011.

[129] Al Mamunur Rashid, George Karypis, and John Riedl. Learning preferences of new users in recommender systems: an information theoretic approach. *Acm Sigkdd Explorations Newsletter*, 10(2):90–100, 2008.

[130] Paul Resnick and Hal R Varian. Recommender systems. *Communications of the ACM*, 40(3):56–58, 1997.

[131] Francesco Ricci, Lior Rokach, and Bracha Shapira. Introduction to recommender systems handbook. In *Recommender systems handbook*, pages 1–35. Springer, 2011.

[132] Philippe Rigaux. An iterative rating method: application to web-based conference management. In *Proceedings of the 2004 ACM symposium on Applied computing*, pages 1682–1687. ACM, 2004.

[133] Lauren A Rivera. Hiring as cultural matching: The case of elite professional service firms. *American sociological review*, 77(6):999–1022, 2012.

[134] Marko A Rodriguez and Johan Bollen. An algorithm to determine peer-reviewers. In *Proceedings of the 17th ACM conference on Information and knowledge management*, pages 319–328. ACM, 2008.

[135] Rosa M Rodríguez, Luis Martínez, and Francisco Herrera. Hesitant fuzzy linguistic term sets. In *Foundations of Intelligent Systems*, pages 287–295. Springer, 2011.

[136] Rosa M Rodriguez, Luis Martinez, and Francisco Herrera. Hesitant fuzzy linguistic term sets for decision making. *IEEE Transactions on Fuzzy Systems*, 20(1):109–119, 2012.

[137] Vasile Rus, Nobal Niraula, and Rajendra Banjade. Similarity measures based on latent dirichlet allocation. In *International Conference on Intelligent Text Processing and Computational Linguistics*, pages 459–470. Springer, 2013.

[138] Sara L Rynes. "getting on board" with amj: Balancing quality and innovation in the review process, 2006.

[139] Germán Sánchez-Hernández, Francisco Chiclana, Núria Agell, and Juan Carlos Aguado. Ranking and selection of unsupervised learning marketing segmentation. *Knowledge-Based Systems*, 44:20–33, 2013.

[140] Carol Saunders. Editor's comments: Looking for diamond cutters. *MIS Quarterly*, pages iii–viii, 2005.

[141] Ho Sun Shon, Sang Hun Han, Kyung Ah Kim, Eun Jong Cha, and Keun Ho Ryu. Proposal reviewer recommendation system based on big data for a national research management institute. *Journal of Information Science*, 43 (2):147–158, 2017.

[142] Nicholas D Sidiropoulos and Efthymios Tsakonas. Signal processing and optimization tools for conference review and session assignment. *IEEE Signal Processing Magazine*, 32(3):141–155, 2015.

[143] Thushari Silva, Zhiling Guo, Jian Ma, Hongbing Jiang, and Huaping Chen. A social network-empowered research analytics framework for project selection. *Decision Support Systems*, 55(4):957–968, 2013.

[144] Brent Smith and Greg Linden. Two decades of recommender systems at amazon. com. *Ieee internet computing*, 21(3):12–18, 2017.

[145] Julia Storberg-Walker. From imagination to application: Making the case for the general method of theory-building research in applied disciplines. *Human Resource Development International*, 9(2):227–259, 2006.

[146] Milan Straka and Jana Straková. Tokenizing, pos tagging, lemmatizing and parsing ud 2.0 with udpipe. In *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pages 88–99, Vancouver, Canada, August 2017. Association for Computational Linguistics. URL http://www.aclweb.org/anthology/K/K17/K17-3009.pdf.

[147] Yong-Hong Sun, Jian Ma, Zhi-Ping Fan, and Jun Wang. A hybrid knowledge and model approach for reviewer assignment. In *System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on*, pages 47–47. IEEE, 2007.

[148] Yong-Hong Sun, Jian Ma, Zhi-Ping Fan, and Jun Wang. A hybrid knowledge and model approach for reviewer assignment. *Expert Systems with Applications*, 34(2):817–824, 2008.

[149] Wenbin Tang, Jie Tang, Tao Lei, Chenhao Tan, Bo Gao, and Tian Li. On optimization of expertise matching with various constraints. *Neurocomputing*, 76(1):71–83, 2012.

[150] Devendra Kumar Tayal, PC Saxena, Ankita Sharma, Garima Khanna, and Shubhangi Gupta. New method for solving reviewer assignment problem using type-2 fuzzy sets and fuzzy functions. *Applied intelligence*, 40(1):54–73, 2014.

[151] Álvaro Tejeda-Lorente, Carlos Porcel, Eduardo Peis, Rosa Sanz, and Enrique Herrera-Viedma. A quality based recommender system to disseminate information in a university digital library. *Information Sciences*, 261: 52–69, 2014.

[152] Yuan Tian, Pavneet Singh Kochhar, Ee-Peng Lim, Feida Zhu, and David Lo. Predicting best answerers for new questions: An approach leveraging topic modeling and collaborative voting. In *International Conference on Social Informatics*, pages 55–68. Springer, 2013.

[153] Vicenç Torra. The weighted owa operator. *International Journal of Intelligent Systems*, 12(2):153–166, 1997.

[154] Vicenç Torra and Yasuo Narukawa. The interpretation of fuzzy integrals and their application to fuzzy systems. *International journal of approximate reasoning*, 41(1):43–58, 2006.

[155] Louise Travé-Massuyès, Francesc Prats, Mónica Sánchez, and Núria Agell. Relative and absolute order-of-magnitude models unified. *Annals of Mathematics and Artificial Intelligence*, 45(3):323–341, 2005.

[156] Anne S Tsui and John R Hollenbeck. Successful authors and effective reviewers: Balancing supply and demand in the organizational sciences. *Organizational Research Methods*, 12(2):259–275, 2009.

[157] Kun Tu, Bruno Ribeiro, David Jensen, Don Towsley, Benyuan Liu, Hua Jiang, and Xiaodong Wang. Online dating recommendations: matching markets and learning preferences. In *Proceedings of the 23rd International Conference on World Wide Web*, pages 787–792. ACM, 2014.

[158] Raquel Ureña, Francisco Chiclana, Juan Antonio Morente-Molinera, and Enrique Herrera-Viedma. Managing incomplete preference relations in decision making: a review and future trends. *Information Sciences*, 302:14–32, 2015.

[159] Andrew H Van de Ven et al. *Engaged scholarship: A guide for organizational and social research*. Oxford University Press on Demand, 2007.

[160] Chong Wang and David M Blei. Collaborative topic modeling for recommending scientific articles. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 448–456. ACM, 2011.

[161] Fan Wang, Ben Chen, and Zhaowei Miao. A survey on reviewer assignment problem. In *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, pages 718–727. Springer, 2008.

[162] G Alan Wang, Jian Jiao, Alan S Abrahams, Weiguo Fan, and Zhongju Zhang. Expertrank: A topic-aware expert finding algorithm for online knowledge communities. *Decision Support Systems*, 54(3):1442–1451, 2013.

[163] Yong Wang, Xiaolei Ma, Yunteng Lao, and Yinhai Wang. A fuzzy-based customer clustering approach with hierarchical structure for logistics network optimization. *Expert Systems with Applications*, 41(2):521–534, 2014.

[164] Max Weber. *From Max Weber: essays in sociology*. Routledge, 2013.

[165] Pinata Winoto and Tiffany Y Tang. The role of user mood in movie recommendations. *Expert Systems with Applications*, 37(8):6086–6092, 2010.

[166] Dianshuang Wu, Jie Lu, and Guangquan Zhang. A fuzzy tree matching-based personalized e-learning recommender system. *IEEE Transactions on Fuzzy Systems*, 23(6):2412–2426, 2015.

[167] Peng Xia, Benyuan Liu, Yizhou Sun, and Cindy Chen. Reciprocal recommendation system for online dating. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*, pages 234–241. ACM, 2015.

[168] Wanhong Xu. Supervising latent topic model for maximum-margin text classification and regression. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 403–414. Springer, 2010.

[169] Yunhong Xu, Jian Ma, Yonghong Sun, Gang Hao, Wei Xu, and Dingtao Zhao. A decision support approach for assigning reviewers to proposals. *Expert Systems with Applications*, 37(10):6948–6956, 2010.

[170] Ronald R Yager. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *Systems, Man and Cybernetics, IEEE Transactions on*, 18(1):183–190, 1988.

[171] Ronald R Yager. Quantifier guided aggregation using owa operators. *International Journal of Intelligent Systems*, 11(1):49–73, 1996.

[172] Ronald R Yager. Nonmonotonic owa operators. *Soft computing*, 3(3):187–196, 1999.

[173] Ronald R Yager. Modeling prioritized multicriteria decision making. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 34(6): 2396–2404, 2004.

[174] Ronald R Yager and Janusz Kacprzyk. *The ordered weighted averaging operators: theory and applications*. Springer Science & Business Media, 2012.

[175] Kun-Woo Yang and Soon-Young Huh. Automatic expert identification using a text categorization technique in knowledge management systems. *Expert Systems with Applications*, 34(2):1445–1455, 2008.

[176] Shuo Yang, Mohammed Korayem, Khalifeh AlJadda, Trey Grainger, and Sriraam Natarajan. Combining content-based and collaborative filtering for job recommendation system: A cost-sensitive statistical relational learning approach. *Knowledge-Based Systems*, 136:37–45, 2017.

[177] Dejian Yu, Wenyu Zhang, and Yejun Xu. Group decision making under hesitant fuzzy environment with application to personnel evaluation. *Knowledge-Based Systems*, 52:1–10, 2013.

[178] Hongtao Yu, Chaoran Liu, and Fuzhi Zhang. Reciprocal recommendation algorithm for the field of recruitment. *JOURNAL OF INFORMATION &COMPUTATIONAL SCIENCE*, 8(16):4061–4068, 2011.

[179] Atila Yüksel and Fisun Yüksel. Measurement of tourist satisfaction with restaurant services: A segment-based approach. *Journal of vacation marketing*, 9(1):52–68, 2003.

[180] Lotfi A Zadeh. A computational approach to fuzzy quantifiers in natural languages. *Computers & Mathematics with applications*, 9(1):149–184, 1983.

[181] Jun Zeng, Feng Li, Haiyang Liu, Junhao Wen, and Sachio Hirokawa. A restaurant recommender system based on user preference and location in mobile environment. In *2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*, pages 55–60. IEEE, 2016.

[182] Chenbin Zhang, Hongyu Zhang, and Jianqiang Wang. Personalized restaurant recommendation method combining group correlations and customer preferences. *Information Sciences*, 454:128–143, 2018.

[183] Yong Zheng. Multi-stakeholder recommendation: Applications and challenges. *arXiv preprint arXiv:1707.08913*, 2017.

[184] Yong Zheng, Tanaya Dave, Neha Mishra, and Harshit Kumar. Fairness in reciprocal recommendations: A speed-dating study. In *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization*, pages 29–34. ACM, 2018.