A Multi-criteria Decision Support System for Ph.D. Supervisor Selection: A Hybrid Approach

Mir Anamul Hasan and Daniel G. Schwartz
Department of Computer Science, Florida State University, Tallahassee, FL, USA
hasan@cs.fsu.edu, schwartz@cs.fsu.edu

Abstract

Selection of a suitable Ph.D. supervisor is a very important step in a student's career. This paper presents a multi-criteria decision support system to assist students in making this choice. The system employs a hybrid method that first utilizes a fuzzy analytic hierarchy process to extract the relative importance of the identified criteria and sub-criteria to consider when selecting a supervisor. Then, it applies an information retrieval-based similarity algorithm (TF/IDF or Okapi BM25) to retrieve relevant candidate supervisor profiles based on the student's research interest. The selected profiles are then re-ranked based on other relevant factors chosen by the user, such as publication record, research grant record, and collaboration record. The ranking method evaluates the potential supervisors objectively based on various metrics that are defined in terms of detailed domain-specific knowledge, making part of the decision making automatic. In contrast with other existing works, this system does not require the professor's involvement and no subjective measures are employed.

1. Introduction

Selection of a Ph.D. supervisor is an important step that a student must take at an early stage in their career. Guidance of the supervisor is a major determinant of quality in a doctoral dissertation [1] and thereby plays a very important role in the student's future success. When deciding whether a particular professor is an appropriate person to serve as supervisor, the student should judge the candidate based on a set of criteria that are important in the supervisor selection process. But identifying the important criteria might be challenging for prospective students due to inexperience. Manderson [2] suggests that students should assess their own needs and the capacities and limitations of potential supervisors, when selecting such a supervisor. Phillips et al. [3] suggests to look for positive answers to at least some of the following questions: “Have they published research papers recently? Do they hold research grants or contracts? Are they invited to speak at conferences in home and abroad?”. Students might also be interested in knowing about the quality of journals and conferences where the professors normally publish, their collaborators, their current and previous students’ records, etc. Accordingly, students might well take advantage of a decision support system that identifies the important criteria and guides them in evaluating professors with respect to those criteria.

There has been significant research in areas such as research paper recommender systems, collaborator recommendations, expert search, people search, academic search, etc. These address parts of the Ph.D. supervisor selection problem, but research on supervisor selection covering different types of domain-specific knowledge is scant. Several existing works apply methods available in multi-criteria decision analysis, such as the Analytical Hierarchy Process (AHP) [4], the Analytical Network Process (ANP) [1], and COMplex PRoportional ASsessment of alternatives to Grey relations (COPRAS-G) [5] to solve the supervisor selection problem by structuring it as a multiple criteria decision-making problem. All these works assume that students are mature enough and know enough about the potential supervisors in order to perform objective pairwise comparisons of the candidate supervisors on each identified criterion, but this might not be the case all of the time. For example, many of the criteria considered in these works are subjective in nature, such as commitment and involvement [4], relationship with other faculty members [1], or behaving like a “boss” [5]. This makes it impossible to perform pairwise comparisons if the student has not previously interacted with a particular potential supervisor. Some existing works utilize a combination of both subjective and objective measures of different criteria but also have missed important aspects, such as the professor’s collaboration network, and do not utilize important details about other potentially relevant criteria, such as citations of papers, recent publications, and research grant details (grant amount, duration, role, etc.) [6] [7].
This paper presents a student-centric multi-criteria decision support system for Ph.D. supervisor selection. A set of important criteria to consider is identified and various metrics are defined to evaluate professors objectively with respect to those criteria. The decision support system first retrieves supervisor profiles based on the match between the student’s and professors’ research areas/interests and then re-ranks candidate supervisor profiles based on several other criteria of interest, selected by the user. The system implements two retrieval algorithms, TF/IDF and Okapi BM25, and lets the user choose the retrieval model to apply when recommending potential supervisors. Students can utilize the system to find a list of potential supervisors based on their research interests and other criteria/sub-criteria of interest, concerning a professor’s publications, research grants, and collaborators. Then to learn about their personality and availability before making the final decision, the student can contact the professors, inquire with their current and previous students, and meet and take courses with them if that’s a possibility.

The paper is structured as follows. Section 2 discusses the related works, section 3 details the underlying method, and section 4 presents details on the developed prototype decision support system. After a discussion of the evaluation methods and results in section 5, section 6 concludes the paper.

2. Related Works

Works closely related to Ph.D. supervisor Recommendation: One of the major concerns when selecting a Ph.D. supervisor is to find professors who work in the area of student’s research interests. There has been significant work in the field of research paper recommender systems, where relevant papers are recommended based on some form of inputs. Some of the existing works in research paper recommender systems have utilized user-provided keywords, text snippets, parts of a research paper of interest (such as, title, abstract, bibliography, etc.), or the entire paper, as input to generate recommendations [8]. Some have employed papers that the users had authored [9], tagged [10], browsed [11], or downloaded [12].

Significant research has been done in the field of collaborator recommendation for scholars. Existing works utilize the user’s research interest [13], publications and co-author network [13] [14] [15], academic homepages [14], temporal evolution of research interest, comparative seniority status [16], and so on, to find potential collaborators. Target users of these studies are normally professors in academia or researchers in enterprises, and not Ph.D. students who are looking for supervisors with whom to work. Moreover, most of these studies do not evaluate how influential a professor’s existing collaboration network might be, which also could be of interest to the students.

There have also been works in the field of expert/people search, and academic search to find experts in a particular topic [17] [18]. These are not student centric and cover only parts of the supervisor selection problem.

Works in Ph.D. Supervisor Recommendation: Several existing works have structured the selection problem of Ph.D. supervisor as a multiple criteria decision making (MCDM) problem and applied different methods used in multi-criteria decision analysis to solve it. Ray [4] demonstrated the use of AHP in the selection of doctoral dissertation supervisors. Momeni et al. [1] used ANP in Ph.D. supervisor selection. ANP allows interdependencies among the decision attributes, whereas AHP assumes selection criteria are independent. Datta et al. [5] used another method employed in multi-criteria decision analysis, called COPRAS-G, to select a suitable supervisor. All these works have followed a similar research methodology, where doctoral students were first interviewed to collect a list of factors to consider before selecting a supervisor, and then they were interviewed again to ascertain the relative weights of those factors through pairwise comparison. Then pairwise comparison of the alternatives, i.e., the potential supervisors, is performed with respect to each of the criteria, and finally the alternatives are ranked using a synthesis process.

Zhang et al. [6] presented a Research Analytics Framework for Education (RAF-E), this being a student-centric method for finding and recommending supervisors for new postgraduate students, considering different metrics from 3 dimensions: relevance, connectivity, and quality. Zhang et al. [7] proposed a personality-matching aided approach for supervisor recommendation based on their previous work [6], which integrates objective measurements (relevance, connectivity, and quality) and subjective personality matching, to get a list of supervisors to recommend.

Alarfaj et al. [19] proposed an information-retrieval based supervisor recommendation method, which returns ranked results based on frequency of candidate supervisor name and proximity of user query and supervisor name in pages returned for a user query by an underlying search engine.

The aforementioned works [1] [4] [5] have used purely subjective measures of different criteria, and [6] [7] [19] have employed objective measures but have missed important details and did not consider some of the important criteria as discussed in the foregoing Introduction. Moreover, these works either create supervisor profiles by interviewing professors or require that professors create their own profiles.
3. Proposed Method

The multi-criteria decision support system being proposed here helps students in the selection process of a Ph.D. supervisor by guiding them in identifying and selecting important criteria/sub-criteria to consider and recommending potential supervisors based on that selection. An overview of the proposed method is given in Figure 1.

First, important decision variables, i.e. criteria/sub-criteria to consider, are identified when selecting a Ph.D. supervisor. Then relative weights of those decision variables are calculated through pairwise comparison applying fuzzy AHP by conducting a survey among graduate students. Then documents, i.e. supervisor profiles, are created, collecting data from various relevant sources with respect to those identified criteria/sub-criteria, and then those supervisor profiles are analyzed and indexed in Elasticsearch [20], a document-oriented NoSQL database. To get recommendations, users need to complete a search profile, where they can enter their research interests and custom select the decision variables they think are important in the selection process. The decision support system first retrieves relevant supervisor profiles from the indexed documents based on the user's research interests given in the search profile, utilizing an IR based similarity algorithm (TF/IDF or Okapi BM25) and then re-ranks those candidate supervisor profiles based on the selected criteria/sub-criteria of interest in the search profile, and suggests them to the user. We defined various metrics to objectively measure the identified decision variables and employed the extracted weights of the decision variables in different phases of the final rank computation process.

3.1. Identifying Decision Variables

We identified important factors to consider when selecting a Ph.D. supervisor on the basis of the intensive review of the available prior research in the relevant fields discussed in section 2, together with an analysis the complexities and challenges encountered during the Ph.D. supervisor selection process, and then developed a hierarchical structure for the Ph.D. supervisor selection problem. We identified four main criteria to consider, namely, research area relevance, publication record, research grant record, and collaboration record. These criteria are then further broken down into various sub-criteria, which are presented in Figure 2. In the hierarchy, the overall objective/goal is placed at level 1, criteria at level 2, sub-criteria at level 3, and the decision alternatives at level 4. Additional details regarding the identified criteria and sub-criteria are discussed in section 3.3.

3.2. Determining Weights of Different Criteria and Sub-criteria

AHP is a widely used tool for solving complex multiple criteria decision-making problem involving subjective judgment. Introduced by Saaty [21], this has previously been used in one of the related works concerning Ph.D. supervisor selection [4]. In AHP, weights are calculated via pairwise comparisons of both the criteria and alternatives on a relative importance scale of 1 to 9. As the conventional AHP does not include vagueness for subjective judgements, many studies have included different techniques into AHP to accommodate vagueness, such as fuzzy set theory [22] [23], probability theory [24] [25], and numeric interval
In this study, we used Fuzzy AHP to determine the relative weights of the identified criteria and sub-criteria, since this approach is adequate to explicitly capture the importance assessment for imprecise human judgments. This technique has not previously been used in our problem domain, i.e., selection of a Ph.D. supervisor.

Fuzzy AHP provides a systematic approach to solve multi-criteria decision problem by using the concepts of fuzzy set theory (developed by Zadeh) and hierarchical structure analysis. Many Fuzzy AHP methods have been proposed by various authors. For this study, we used Ayhan’s implementation of Buckley’s method to determine the relative importance of the identified criteria/sub-criteria. This introduces triangular fuzzy numbers into the conventional AHP in order to enhance the degree of judgment of the decision maker.

A triangular fuzzy number is a special fuzzy set \( \tilde{F} \) in a universe of discourse \( U \) and can be defined as \( \tilde{F}=(l,m,u) \), where \( l \) and \( u \) stand for lower and upper value of \( F \) and \( m \) is the mid-value of \( F \). The symbol ‘~’ on a letter is used to indicate that the letter represents a fuzzy set. The membership function \( \mu_{\tilde{F}}(x) \), which associates a real number in the interval \([0,1]\) with each element \( x \) in \( X \), to represent the grade of membership of \( x \) in \( \tilde{F} \) is defined as [30]:

\[
\mu_{\tilde{F}}(x) = \begin{cases} 
\frac{x - l}{m - l}, & l \leq x \leq m \\
\frac{u - x}{u - m}, & m \leq x \leq u \\
0, & \text{otherwise}
\end{cases}
\]

The corresponding linguistic terms and triangular fuzzy number representation of Saaty’s 1 to 9 relative importance scale is depicted in Table 1.

We conducted a survey among computer science graduate students to get the preferences of one criteria/sub-criteria over the other through pairwise comparison. First, the relative weights of each criteria (research area, publication record, research grant record, and collaboration record) are determined. The steps of the procedure are as follows.

The pairwise comparison matrix, \( \tilde{P} \) is computed as:

\[
\tilde{P} = \begin{bmatrix}
\tilde{s}_{11} & \tilde{s}_{12} & \cdots & \tilde{s}_{1n} \\
\tilde{s}_{21} & \tilde{s}_{22} & \cdots & \tilde{s}_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
\tilde{s}_{n1} & \tilde{s}_{n2} & \cdots & \tilde{s}_{nn}
\end{bmatrix}
\]

where \( \tilde{s}_{ij} \) is the averaged preferences of i-th criterion over j-th criterion, defined as \( \tilde{s}_{ij} = \frac{\sum_{k=1}^{K} \tilde{s}_{ik} \tilde{s}_{jk}}{K} \), where \( \tilde{s}_{ik} \) represents the k-th survey taker’s preference of the i-th
criterion over the j-th criterion, and K is the total number of valid survey responses. After that, the geometric mean of the fuzzy comparison values of each criterion is computed as: \( \tilde{r}_i = (\prod_{j=1}^{n} \tilde{s}_{ij})^{1/n} \). Then, the fuzzy weight of each criterion \( \tilde{w}_i \) is computed as \( \tilde{w}_i = \tilde{r}_i \otimes (\tilde{r}_1 \oplus \tilde{r}_2 \oplus \ldots \oplus \tilde{r}_n)^{-1} \), where \( \otimes \) and \( \oplus \) are the fuzzy addition and multiplication operators. Then, the fuzzy weights (\( \tilde{w}_i \)) are de-fuzzified to \( M_i \), where \( M_i \) is the non-fuzzy weight of each criterion and defined as \( M_i = \frac{\tilde{w}_i + Mw_1 + 1}{3} \). Finally, the non-fuzzy weight \( M_i \) is normalized to \( N_i \) to get the final weights of each criterion. \( N_i \) is defined as: \( N_i = \frac{M_i}{\sum_{i=1}^{n} M_i} \). We follow the same procedure to extract the relative importance of each sub-criteria under each of those criteria. Details about the survey are discussed in section 4.

### 3.3. Evaluating the Alternatives

In conventional AHP and Fuzzy AHP, the relative importance of alternatives is calculated through pairwise comparison of the alternatives with respect to the identified criteria and sub-criteria and then alternatives are ranked based on a synthesis process. In the proposed hybrid method, we define metrics (see the following) for the identified criteria/sub-criteria, which makes it possible to evaluate the alternatives objectively and automatically. This differs from typical AHP and Fuzzy AHP applications where the alternatives are evaluated and compared manually.

The proposed two-phase decision support system provides the user with a search interface as depicted in Figure 3. Here the user can enter text data regarding their research interests, specific research interest, title and abstract of a paper of interest, and select the criteria/sub-criteria of their interest. In the first phase, the proposed method retrieves matching candidate supervisor profiles based on relevance between the student’s and supervisors’ research areas/interests. Then in the second phase, the candidate supervisor profiles are re-ranked based on the selected criteria/sub-criteria of interest and presented to the user.

#### 3.3.1. Research Area Relevance (C\(_1\)).

The research area relevance is evaluated based on the following four sub-criteria: relevance of the professor’s broad research interests to those of the student (C\(_{11}\)), relevance of the professor’s specific research interests/topics to those of the student (C\(_{12}\)), relevance of the professor’s and their previous students’ publication and dissertation record to the student’s research interests (C\(_{13}\)), and relevance of the professor’s previously taught courses to the student’s research interests (C\(_{14}\)).

![Figure 3. User Search Interface](image-url)
where \( tf(t \text{ in } d) \) is the term frequency of term \( t \) in document \( d \) and computed as \( tf(t \text{ in } d)= \sqrt{\text{frequency} \cdot \text{idf}(t)} \), \( \text{idf}(t) \) is the inverse document frequency of term \( t \) and computed as \( \text{idf}(t)=1+ \log \frac{\text{numDocs}}{\text{docFreq}(t)+1} \), where \( \text{numDocs} \) is the number of all documents in the collection and \( \text{docFreq}(t) \) is the number of documents containing term \( t \), and \( \text{norm}(d) \) is the normalization factor of a matching document \( d \), which causes higher weights for short documents, computed as \( \text{norm}(d)=\frac{1}{\sqrt{\text{numTerms}}} \).

**Okapi BM25.** Okapi BM25 is a similarity algorithm to score matching documents according to their relevance to a search query and is developed based on the probabilistic retrieval model [33]. In the Okapi BM25 similarity algorithm, the relevance score of a document \( d \) for query \( q \), which consists of terms \( t \) is defined as [32]:

\[
\text{score}(d, q) = \sum \left( \frac{\text{idf}(t) \cdot \frac{tf(t \text{ in } d)(k+1)}{tf(t \text{ in } d) + k(1-b + b \frac{\text{avgl}d}{\text{avgdl}})}}{\text{norm}(d)} \right)
\]

where \( tf(t \text{ in } d) \) is the number of occurrences of term \( t \) in document \( d \), \( \text{avgl}d \) is the average document length over all documents in the collection, \( \text{idf}(t) \) for term \( t \) is computed as \( \text{idf}(t)=\log\left(1+\frac{\text{numDocs}}{\text{docFreq}(t)+0.5}\right) \), where \( \text{numDocs} \) is the number of documents in the collection and \( \text{docFreq}(t) \) is the number of documents containing term \( t \), and \( k \) and \( b \) are the tuning parameters. In our experiment, we used BM25 with standard values for \( k \) (1.2) and \( b \) (0.75) [32].

We utilized the Elasticsearch’s [20] implementations of the TF/IDF and Okpai BM25 similarity algorithms to retrieve relevant documents. We wrote a multi-field search query following the Elasticsearch query DSL [20], where matches in broader research interests (i.e. \( C_{11} \)) is boosted with \( \omega_{11} \), matches in specific research interests (i.e. \( C_{12} \)) is boosted with \( \omega_{12} \), matches in publication record (i.e. \( C_{13} \)) is boosted with \( \omega_{13} \), and matches in taught courses (i.e. \( C_{14} \)) is boosted with \( \omega_{14} \). Here, \( \omega_{11}, \omega_{12}, \omega_{13} \) and \( \omega_{14} \) are the relative weights of the sub-criteria. The weights of the identified criteria/sub-criteria are calculated following the steps discussed in section 3.2 and actual weights used in our system are given in section 4.

**3.3.2. Publication Record (C2).** Academic performance of professors is often measured in terms of number of publications and the quality of journals/conferences where they were published [34][35]. The citation count of a paper can give a rough idea of the paper’s popularity [8].

The factors (sub-criteria) that can affect the publication record criterion are as follows.

**Overall Publication Quality (C21).** The overall publication quality is defined as:

\[
C_{21} = \frac{\sum \text{triples} \cdot \text{weight}}{\sum |\text{triples}| \cdot \text{weight}}
\]

where, \( \text{total no. of papers} , \text{triples} = \text{no. of citations of the i-th paper} , \text{weight} = \text{rank of the journal or conference in which the i-th paper is published according to the CORE ranking database} \) (www.core.edu.au). The CORE ranking database provides rankings of conferences and journals in the computing disciplines.

**Consistency in Publishing (C22).** Another important aspect to consider when evaluating a professor’s publication record is to check how consistent they are in publishing throughout their publishing career [3]. The consistency in publishing measure is defined as:

\[
C_{22} = 1 - \frac{\text{years without publication}}{\text{current year} - \text{year of first published paper}}
\]

**Recent Publication Record (C23).** The recent publication record of a professor is a good indicator of whether a professor is active in research or not, as well as the direction and quality of their current research [3]. The recent publication record measure is defined as:

\[
C_{23} = n^p + \frac{\text{rank of the journal or conference}}{n^p + \frac{n^a + n^b + n^c}{n^p}}
\]

where, \( n^p \) is the number of papers published in last five years, \( n^a \) is the number of papers published in type A journals or conferences in the last five years, \( n^c \) is the number of papers published in type B journals or conferences in the last five years. The types/ ranks of journals and conferences are extracted from the CORE ranking database [www.core.edu.au].

**Publication Record of Professor’s Graduated Students (C24).** The publication record of graduated students of a professor might be of interest to the students, as some might want to be employed in academia/research organizations, where quality and number of publications matter. This sub-criterion is defined as: \( C_{24} = M(G) \), i.e. the median of \( G \), where \( G \) is a set of numbers representing the number of publications of graduated students. The median is used as we assume the sample data size will be small.

Finally, overall publication quality is evaluated as:

\[
C_2 = \omega_{21} C_{21} + \omega_{22} C_{22} + \omega_{23} C_{23} + \omega_{24} C_{24}
\]

where \( \omega_{21}, \omega_{22}, \omega_{23} \) and \( \omega_{24} \) are the relative weights of the sub-criteria. A metric above and hereafter with the symbol ‘’ is assumed to be normalized by scaling it into the range [0,1] based on the corresponding values of the candidate supervisor profiles.
3.3.3. Research Grant Record (C₃). Typically, a Ph.D. student is supported through graduate assistantship, be it teaching or research, which is generally viewed as a means of enhancing the professional development of the student, in addition to providing financial support [36]. These positions are time demanding. Teaching assistants are normally assigned menial types of duties, such as checking assignments and grading tests and quizzes, with occasional greater opportunities for professional development through teaching a course by assuming full responsibility. But those have little or nothing to do with student’s success/progress on their Ph.D. dissertation research. On the other hand, research assistants get the opportunity to be involved in the design and conduct of exciting funded research projects, which helps them develop valuable research skills needed for their graduate study and future career and, in most of the cases, those works become part of their dissertation. A study by Wong et al. [37] found that receipt of a teaching assistantship is less likely to be associated with graduate success than receipt of a research assistantship.

So, the professor’s grant record might be of interest to the students, as research assistants are usually supported through the grant money of professor’s active research grants, funded by different funding agencies. Moreover, research by Bozeman et al. [38] found that professors with more grants and contracts of each type (government and industry) have a greater propensity for industrial involvement than those who have fewer such contracts. A professor’s connection with industry people might also be of interest to some students for future opportunities, like, internships or full-time jobs after graduation.

To evaluate a professor’s research grant record, the following three sub-criteria are identified. 

Research Grant Quality (C₃₁). Research grant quality is evaluated in terms of grant duration, grant amount, and the role played by the professor (PI, Co-PI, etc.) and is defined as:

\[ C_{31} = \frac{\sum_i N_i^g d_i^g r_i^g}{\sum_i N_i^g d_i^g} \]

where, \( N^g \) = total no. of grants, \( d_i^g \) = duration of i-th grant, \( a_i^g \) = amount of i-th grant, and \( r_i^g \) = role in the i-th grant, which works as a boosting factor. We set \( r_i^g = 1 \), if the role is Co-PI and \( r_i^g = 2 \), if the role is PI.

Consistency in Getting Grants (C₃₂). Consistency in getting grants is also taken into consideration when evaluating the research grant record and is defined as:

\[ C_{32} = 1 - \frac{Y_{wog}}{Y_c - Y_{1/2}} \]

where, \( Y_{wog} \) = no. of years without a grant, \( Y_c \) = current year, and \( Y_{1/2} \) = year of the first grant received.

Recent Grant Record (C₃₃). Recent grant record is taken into considerations, as professors who have active grants and contracts, are more likely to be productive [38] and support students as research assistants. This is defined as:

\[ C_{33} = n_y + n_c^g \]

where, \( n_y \) = no. of research grants in last 5 years, and \( n_c^g \) = no. of current grants.

Finally, the overall research grant record is evaluated as:

\[ C_3 = \omega_{31} C_{31} + \omega_{32} C_{32} + \omega_{33} C_{33} \]

where \( \omega_{31}, \omega_{32}, \) and \( \omega_{33} \) are the relative weights of the sub-criteria.

3.3.4. Collaboration Record (C₄). Another aspect to consider when selecting a Ph.D. supervisor is the professor’s collaboration record. Collaboration tends to have positive effects on research productivity [39]. Analyzing 592 scientists’ publications and collaborative activities, Price et al. [40] found that “The most prolific author is also by far the most collaborating, and three of the four next most prolific are also among the next most frequently collaborating”. Working with a professor who has a strong collaboration network might give students the opportunity to be involved in exciting collaborative projects, thereby providing students with the opportunity to interact with and learn from the professor’s collaborators, as well as create new connections in academia/industry. Three sub-criteria are identified that affect the collaboration record criterion. 

Influential Co-authors (C₄₁). To assess the list of co-authors of a professor, we considered the reputation of the co-authors in terms of citations [41] and number of times they have co-authored [41], and give more importance to recent co-authorship. C₄₁ is defined as:

\[ C_{41} = \frac{\sum_{i=1}^{N'} n_{ca}^i b_{ca}^p \log (cc_{ca}^i)}{\sum_{i=1}^{N'} n_{ca}^i b_{ca}^p} \]

where \( N' \) = total no. of co-authors, \( n_{ca}^i \) = no. of co-authored papers with co-author i, \( cc_{ca}^i \) = total number of citations of co-author i, \( b = \) boost factor, which boosts recent involvement (b will be no. of co-authored paper in last five years).

Recent Collaboration Record in Research Papers (C₄₂). This is defined as: \( C_{42} = \) no. of co-authors in last 5 years. Record as Co-PI/Co-I in Research Grants (C₄₃). This is defined as: \( C_{43} = \) no. of research grants as co-PI/co-I in last 5 years.

Finally, the overall collaboration record is evaluated as:

\[ C_4 = \omega_{41} C_{41} + \omega_{42} C_{42} + \omega_{43} C_{43} \]

where \( \omega_{41}, \omega_{42}, \) and \( \omega_{43} \) are the relative weights of the sub-criteria.

Once we get the corresponding scores for research area relevance (C₁), publication record (C₂), research grant record (C₃) and collaboration record (C₄) for the candidate supervisor profiles, the final recommendation score for the potential supervisors is computed as:
Supervisor profiles. The supervisor profiles are very rich, providing useful detailed information about each professor with graphs and charts, when relevant, and can thereby help the user make a more informed decision. More details are reported in a demonstration paper [47].

In the system implementation, we used relative weights of identified criteria and sub-criteria when computing the recommendation score. We conducted a survey to extract those weights, where participants (computer science graduate students in the host department as well as other US universities) were first asked to do pairwise comparison of the identified criteria and then do pair-wise comparisons of the sub-criteria under each criterion. We collected 28 valid survey responses, where 25 survey respondents were Ph.D. students, 3 were M.S. students, and 20 of the 25 Ph.D. students had already selected their supervisor. The weights were extracted following the steps described in section 3.2. The corresponding values are \( \omega_1 = 0.54, \omega_2 = 0.27, \omega_3 = 0.12, \omega_4 = 0.07, \omega_{11} = 0.57, \omega_{12} = 0.25, \omega_{13} = 0.13, \omega_{14} = 0.05, \omega_{21} = 0.53, \omega_{22} = 0.27, \omega_{23} = 0.14, \omega_{24} = 0.07, \omega_{31} = 0.65, \omega_{32} = 0.24, \omega_{33} = 0.12 \) and \( \omega_{41} = 0.65, \omega_{42} = 0.26, \omega_{43} = 0.08 \).

**5. Evaluation**

To evaluate the quality of the recommendations generated by the proposed method, we asked 20 Ph.D. students in the host department to rate the recommendations on a scale of four: not relevant (0), somewhat relevant (1), relevant (2) and very relevant (3) in three different settings (baseline, custom, all) for each retrieval algorithm (TF-IDF, Okapi BM25). The three settings are defined as follows.

**Baseline:** In the baseline method, recommendations are made based purely on research area relevance.

**Custom:** In the custom method, the user can custom select their criteria/sub-criteria of interest, based on which the recommendations will be generated.

**All:** In the all method, all criteria/sub-criteria will be considered when computing the recommendation score.

We used Average Rate (AR) [6] and Normalized Discounted Cumulative Gain (NDCG) [6] as the
evaluation metrics. Table 2 reports AR values of the user ratings for both retrieval models in the three different settings. It can be easily observed that for both retrieval models, the Custom setting of criteria/sub-criteria performs better, and between the TF/IDF and BM25 algorithms, in most cases, the BM25 based algorithm performs better. Table 3 reports the NDCG values, where also Custom settings of the criteria/sub-criteria-based configuration performs better than the Baseline and All settings for both the TF/IDF and BM25 based retrieval models. So, analysis of the results reveals that letting the user custom select the criteria/sub-criteria of interest provides more satisfactory recommendations compared to the Baseline and All criteria/sub-criteria selection settings.

### Table 2. Performance Comparison: AR

<table>
<thead>
<tr>
<th>Evaluation Metric</th>
<th>AR@1</th>
<th>AR@2</th>
<th>AR@3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieval Model</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Criteria Setting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>2.11</td>
<td>2.14</td>
<td>2.07</td>
</tr>
<tr>
<td>Custom</td>
<td>2.26</td>
<td>2.30</td>
<td>2.15</td>
</tr>
<tr>
<td>All</td>
<td>2.17</td>
<td>2.25</td>
<td>2.18</td>
</tr>
</tbody>
</table>

### Table 3. Performance Comparison: NDCG

<table>
<thead>
<tr>
<th>Evaluation Metric</th>
<th>NDCG@1</th>
<th>NDCG@2</th>
<th>NDCG@3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieval Model</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Criteria Setting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>0.81</td>
<td>0.79</td>
<td>0.78</td>
</tr>
<tr>
<td>Custom</td>
<td>0.85</td>
<td>0.82</td>
<td>0.83</td>
</tr>
<tr>
<td>All</td>
<td>0.83</td>
<td>0.80</td>
<td>0.75</td>
</tr>
</tbody>
</table>

### 6. Conclusion

In this paper, we proposed a hybrid method for Ph.D. supervisor selection, which uses detailed domain specific knowledge, keeping the student’s needs in mind. The proposed method retrieves potential supervisors based on the custom selection of criteria/sub-criteria of interest of a user (i.e., a student). This differs from previous works, which retrieve potential supervisors based on all the factors of a curated list of factors, and do not give importance to the fact that not all students might be interested in all the factors. Our evaluation of the proposed method shows that allowing users to select the criteria/sub-criteria of interest provides more satisfaction in the recommendations than recommendations generated purely based on research area relevance and recommendations generated considering all criteria/sub-criteria. Evaluations also show that the Okapi BM25 based recommendations perform better than the TF/IDF based recommendation. Several of the previous methods are not easily scalable, as they require that students perform pairwise comparisons of the professors with respect to the considered criteria/sub-criteria [1] [4] [5]. Our method is easily scalable to larger datasets, however, as it evaluates and ranks the professors automatically based on the defined metrics.

### 7. References


