InsERT: the Inspirational Expert Recommender Tool

Jennifer Nguyen
GREC Research Group
Industrial Engineering School
Universitat Politècnica de Catalunya
Barcelona, Spain
Email: jennifer.nguyen@upc.edu

Germán Sánchez-Hernández, Núria Agell
GREC Research Group
ESADE Business School
Ramon Llull University
Barcelona, Spain
Email: {german.sanchez, nuria.agell}@esade.edu

Cecilio Angulo
GREC Research Group
Automatic Control Department
Universitat Politècnica de Catalunya
Barcelona, Spain
Email: cecilio.angulo@upc.edu

Abstract—The continued growth in enterprise social networks is fueled by the need to enable productivity and innovation. Reducing the constraints to communication and knowledge sharing of a globally distributed workforce will facilitate the workflow. People finder systems are one of the main solutions in enterprise social networks which are reducing these constraints leading to time and cost savings. Finding expertise efficiently helps organizations to unlock knowledge within the enterprise, solve problems, and identify collaborators. However, the following challenges still exist: validating expertise, determining responsiveness and accessibility, and managing expert profiles. In this paper, we propose the fuzzy OW A technique as a novel approach to ranking expertise candidates. Then, we propose fuzzy OW A as a ranking method. Hence, finally, we present our ERS software tool (InsERT) as part of a larger enterprise social network implementation.

I. INTRODUCTION

For global organizations, finding an expert can be challenging as experts are disperse and vary in level of knowledge of a topic. Their knowledge is difficult to qualify and changes frequently [1]. In addition, experts can be culturally isolated where business units are disconnected from one another [2]. This need supports the continued investment organizations are making in enterprise social networks (ESN). According to a study by IDC (International Data Corporation, the premier global provider of market intelligence), worldwide enterprise social network market revenue is expected to increase from 1.24 billion in 2013 to 3.5 billion by 2018 [3]. As one of the solutions of ESN, expertise recommender systems (ERS) help users to identify informed people. In this paper we explore the application of (ERS) in open innovation marketplaces and in enterprise settings. First, we define dimensions to evaluate expertise candidates. Then, we propose fuzzy OW A as a ranking method. Hence, finally, we present our ERS software tool (InsERT), a fuzzy system software application developed within COLLAGE, a European sponsored enterprise social network implementation.

In open innovation a firm’s R&D crosses not only internal boundaries but disciplines. It is an interactive process of knowledge generation and transfer between internal and external firms. Open innovation marketplaces broker relationships between seekers and solvers of challenges. Seekers have a problem which they need to solve and solvers are a community of people with the right skills to discover innovative ideas to address them. As intermediaries, on-line marketplaces are intended to provide access to a wide range of participants, facilitate connections that might not have been possible in a closed network, and find appropriate solutions, thereby lowering the cost of search and knowledge exchange [4]. Beyond the local geographical area, intermediaries reach audiences which are distant and diverse [5]. Consequently, intermediaries connect people who were previously unknown to one another. Despite the assistance of open innovation marketplaces, the process of matching seekers and solvers remains a challenge. Because the primary function of an intermediary is to identify the “right partner”, we analyze dimensions for an ERS within the context of open innovation.

Organizations with disparate and dissimilar databases have siloed information further hampering the ability to share and reuse knowledge [6]. Complicating the search even more, seekers of expertise do not necessarily have clearly defined requirements and often have difficulty evaluating the quality of an expert’s knowledge [1]. Therefore, organizations like those in COLLAGE, are looking to incorporate expert recommender systems into their knowledge management initiatives. Expert Recommender Systems (ERS), also called Expert Finding Systems (EFS), Expertise Location Systems (ELS) [1], or expertise retrieval connect people to areas of expertise [7]. ERS support users’ search by enabling them to find expertise rapidly and inexpensively. Notable ERSs include MITRE’s Expert Finder, IBM’s SmallBlue, INDURE’s FacFinder, and NASA’s POPS.

There are several limitations of current approaches to expertise finding [7]. First, information retrieval ERSs focus on the strength of the relationship between a person and a topic rather than a person’s expertise. In other words, if the system only captures a person’s name occurring with a particular topic, it may miss that the topic was only referenced rather than at the center of the paper. Second, the ERSs in Balog’s review of expertise retrieval systems focused on content matching, neglecting contextual factors that users and experts take into account before connecting [7]. Searchers assess social and organizational factors to evaluate a person’s responsiveness and suitability in providing information [2]. Therefore, users of the system require personalized views tailored to their information needs. Third, the development of personal profiles remains a challenge as they are often not kept up to date. Furthermore, as highlighted by Ackerman et al., the ability to collect data is influenced by the organizational culture in which an ERS is implemented [8]. Some organizational systems have access to limited information due to privacy concerns. Fourth, it is a challenge to measure a person’s expertise objectively [8].
In this paper, we propose an ERS with four dimensions intended to address the first three challenges. We build upon the four dimensions in current ERSs to evaluate a candidate’s expertise: expertise, qualifications, proximity, and availability. Through initial research with an open innovation intermediary, we identified challenges in finding the “right” partner. We extend the proximity dimension and include a fifth dimension for responsiveness. The dimensions were integrated into our ERS as a set of criteria to evaluate each candidate. Applying fuzzy OWA to the evaluation criteria, we develop a new theoretical model for ERS able to aggregate available information instead of relying on filtering methods.

The rest of the paper is organized as follows. First, the role, benefits, and challenges of ERSs are highlighted. Second, the properties of existing ERSs are explained. Third, selected dimensions are defined for an expertise recommender system. Fourth, we introduce fuzzy OWA to ERSs, develop our theory, and explain our methodology. An illustrative example of the current pilot implementation of the system is provided. We close with an example of a business application, conclusions, and areas for future research.

II. OVERVIEW OF EXPERT RECOMMENDER SYSTEMS

For global organizations, expertise is likely to be distributed across remote locations. The ability to identify and assemble this information for a competitive advantage is critical in a global economy [1]. Employees search for expertise when they need to resolve a problem [7]. However, the search may be time consuming amid the volumes of available information surrounding the problem. For some organizations, the search begins with an employee’s personal network [2]. Yet, these networks are not always large enough to reach people with the right information [7] or expertise [2]. In the case of a new hire, where personal networks have not been established, finding expertise is more difficult [7].

According to [9] there are five situations in which people seek an expert as a source of information: 1) access to information is not in public documentation, 2) specify information to solve a problem, 3) leverage others’ expertise, 4) interpretation of discovered information, and 5) socialization.

ERS systems have four main functions: 1) collecting data, 2) preprocessing and indexing data, 3) modeling and retrieving data, and 4) actionable data [7]. Data may be collected from sources internal and external to the firm including corporate directories, email, and publications. Preprocessing and indexing is required to ensure that the same person is appearing in the evidence sources. A person’s name may appear differently in different documents. In addition, this step addresses the challenge of applying weight to different data sources removing bias associated with the length and size of different sources [10], as well as, self-ratings in personal profiles. Modeling and retrieval ranks the quality of a match between the candidate and the user’s query. Interaction design refers to advanced search features and presentation of a candidate. Enhancements in both areas assist users in deciding whom to contact. Therefore, ERSs often display candidate rankings, documents, publications, projects, social networks, and organizational relationships.

A. Existing Commercial Tools

To date, several ERSs have been developed. In the following, a list of existing expertise recommender systems that come as standard is provided, as well as a description of their main characteristics.

Who Knows [11] finds people with appropriate expertise through latent semantic indexing of their work products. In a nutshell, when a user enters search text, Who Knows returns a list of people whose profiles match this text.

Yenta [12] analyzes people’s email archives, communications, news postings, and other types of documents to create a user profile. When the user enters a query, Yenta searches for individuals with profiles that match the query. One of Yenta’s weaknesses is that it matches people with similar interests without taking into account their expertise.

Expertise Recommender (ER) [13] uses locally meaningful data to recommend sets of potential answerers for queries. ER recommends expertise based on both the best expertise in a particular area and context fit with the seeker. One important characteristic of this system is that it rates experts based on their work products, which are mined by ER, rather than on their areas of interest or on peer ratings.

MITRE’s Expert Finder [14] was created within the MITRE Corporation to identify experts within topic domains. The system ranks the expert by the number of times his/her name is associated with specific terms found in corporate documents, newsletters, communications and so forth. The user enters a key word search and the system returns the top ranked experts [1]. The Expert Locator, implemented in MITRE Corporation, retrieves information from several activity spaces. It applies weights to the documents at the activity space, evidence type, and sub-activity space evidence levels [15].

Expert Finder [16] creates user profiles based upon their work. When a user queries for a skill set, the system searches other users’ profiles for them and returns experts whose skills are slightly more advanced than the user’s.

Expertise Recommender using Web Mining [17] obtains a person’s expertise by dynamically extracting data from semi-structured web documents. It consists of the following five main components: web crawler, expertise extractor, referral chain builder, knowledge base, and web interface.

APOSDE’s People Recommender Service [18] is a service integrated into APOSDE’s platform1. The system automatically detects the user’s current work task and relevant domain concept taking into account both the candidate and the user’s profile.

NASA POPS [19] Leveraging RDF as the exchange and modeling format and Semantic Web, NASA POPS aggregates information from multiple databases. Users can search by organization, project, and/or competency. The system displays a list of people matching these attributes along with the contact

---

1APOSDE (Advanced Process-Oriented Self-Directed Learning Environment) was partially funded under the 6th framework programme (FP6) for R&D of the European Commission within the Information Society Technologies (IST) work program 2004 under contract no. IST-027023. See http://www.aposdle.org.
information and social network connecting the candidate and
the user.

IBM SmallBlue [20], [21] uses outbound emails and IM
chat to analyze the social network for “who knows what
and whom”. It maps search strings to keywords, identifies
candidates matching these keywords and ranks the results by
relevance weighting and social network structure.

INDURE FacFinder [10] collects information on faculty from
university profiles, NSF awards, and faculty homepages. The
system indexes each document in entirety and applies weights
to the proximity of terms. It considers the order of terms, the
data source, and document rank when applying weights.

StrangersRS [22] scores people based on their familiarity and
similarity to one another. The system recommends people with
similar interests but are unfamiliar with each other.

B. Current Dimensions of ERS

The existing commercial tools discussed above focus on
finding the person with the “right level of expertise”. We
analyzed these recommenders for dimensions used to evaluate
candidate expertise. The dimensions can be represented by
four categories: expertise of the candidates, their sub-skills
or qualifications, their proximity to the user and their current
availability. Table I details the characteristics of the most
relevant systems in the field of people recommenders in terms
of the four dimensions.

The **expertise dimension** reflects the topics in which the
candidate has a certain degree of expertise. For instance,
candidate A is knowledgeable in computer science. This informa-
tion can be analyzed both explicitly and implicitly. Explicit
expertise can come from topics which candidates have selected
for themselves or skill ratings which candidates have received
from others. Implicit expertise is the automatic processing of
documents from different sources, such as forums, papers,
presentations, email, and chat in order to find keywords which
will reflect the candidate’s level of expertise in a specific topic.

The **qualification variables** capture specific areas where
the the candidate has applied their expertise in a topic. For
instance, candidate A is experienced with databases. Similar
to the expertise dimension, candidate qualifications can be col-
lected explicitly and implicitly. Analyzing implicit information
on the qualification variable also provides the frequency with
which a candidate communicates about a topic. Some ERSs
use this as an indicator of the level of expertise in a particular
qualification or a sub-skill.

The **proximity information** is used to measure the distance
between the candidate and the user. Some ERSs mentioned
here, reference email and chat exchanges, co-authorship of
publications, common project teams, and organizational struc-
ture to infer a candidate’s social network and his connection
to the user.

Finally, it is proposed to incorporate the **dimension of avail-
ability**, which informs about the current availability of each
candidate. Some ERSs retrieve information from candidates’
calendars to infer availability.

C. Existing Ranking Models

Balog [7], categorized existing models of expertise retrieval
to associate query topics with people into four main groups:
generative probabilistic, discriminative, voting, and graph-
based models. In the generative probabilistic model, candidates
are ranked based on the probability of being an expert on a
topic. However, the accuracy of generative models depends on
the validity of the assumptions. In particular, the independence
assumption between candidates and terms, ignores the relation-
ship between terms and candidates which appear in the same
document. In other words, candidates are assumed to have
equal expertise in all topics in the document. Discriminative
models estimate the binary conditional probability that a query
topic and a candidate are related. These models require fewer
model assumptions allowing the data to “speak for itself”.
They have been preferred over generative models in informa-
tion retrieval applications. Specifically, the application of
support vector machines were shown to outperform language
modeling, the state-of-the-art generative model for IR due to its
ability to learn features [23]. The Voting Model, demonstrated
by MacDonald and Ounis, for expert search aggregates scores
from a single ranking of documents into a single ranking
of candidates [24]. Graph-based models infer associations
between query topics and candidates by analyzing expertise
graphs [7].

III. The Inspirational Expertise Recommender
Tool (InsERT)

Currently, as seen in Section II, recommender systems
focus on finding the person with the “right level of expertise”.
However, ERSs should recommend the “right person” based
on an appropriate mixing and an optimal matching of the
characteristics of the candidates and the preferences of the user.
In this section we introduce the expertise recommender tool
InsERT and emphasize the types and sources of information
used in the pilot application.

A working version of InsERT is currently being tested
and its interface, re-designed. It has been developed in PHP
under MySQL and MongoDB databases. The initial version
of InsERT was framed in the European research project
COLLAGE and it is currently integrated in this environment,
thus delegating some of the tasks to other COLLAGE services,
such as security and external data modules.

The main contribution of this work is the design and develop-
ment of an ERS based on fuzzy OWA to aggregate uncertain
information. In addition, we propose a set of dimensions for
future ERSs.

This paper proposes fuzzy OWA as an alternative model to
ERSs. When the user enters a search query with his preferences
for an expert, a set of specific requirements is generated
and each candidate’s profile is matched to the individual
requirements for a partial score. The partial score is aggregated
across all requirements for a particular candidate as a global
score. This score is then used to rank the final candidate list.

The following subsections reference each step required
for the generation of a recommendation: III-A) the sources
of information from which the candidates’ profiles are built;
modules called Peers. The Peer Profile aggregator based on the development of standalone services and therefore it relies on the information obtained from another service: P2PUM. A mash-up of services and therefore it relies on the information obtained from another service: P2PUM.

### A. Sources of information

InsERT was initially designed as a service integrated within a mash-up of services and therefore it relies on the information obtained from another service: P2PUM. This service is a Social Profile aggregator based on the development of standalone modules called Peers. The strength of this methodology is that the entire profile of a user is not aggregated in a single location, but stored in multiple locations according to the peers involved. It also ensures privacy and safety of the stored information.

P2PUM supports a set of different peers, each related to a single source of information. It currently offers access to the following social networks peers: LinkedIn, Twitter, and Facebook. It also includes a peer for COLLAGE, which provides some information about the interaction between users and COLLAGE tools. Finally, the Explicit peer allows users to state explicitly their profiles. When a user enables a certain peer to create his profile, it retrieves the information from the corresponding source. The user is able to review his personal information and select which information he wants exposed by the peer. This benefit enables the user to protect his privacy.

The seeker of expertise can define the exact skill and level of expertise desired. Figure 1 depicts InsERT’s user interface which allows the seeker to specify her preferences in order to get a proper recommendation. InsERT is currently using P2PUM. Specifically, it uses Explicit peer to instruct the LinkedIn peer to get the topics in which a candidate has received endorsements from his connections. Each endorsed topic in LinkedIn is considered to be a skill in InsERT.

### Table I. Dimensions of existing ERS

<table>
<thead>
<tr>
<th>System</th>
<th>Authors</th>
<th>Variables used for profiling candidates</th>
<th>Aggregation / Filters</th>
<th>User preferences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who Knows</td>
<td>Streeter &amp; Lochbaum, 1998</td>
<td>Semantic indexing</td>
<td>No</td>
<td>N/A</td>
</tr>
<tr>
<td>Yenta</td>
<td>Foner, 1997</td>
<td>Semantic indexing</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ER</td>
<td>McDonald &amp; Ackerman, 2000</td>
<td>Automatic</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>MITRE’s Expert Finder/Expert Locator</td>
<td>Mattos et al., 1998; D’Amore, 2008</td>
<td>Frequency of association</td>
<td>Yes</td>
<td>Organizational &amp; activity space</td>
</tr>
<tr>
<td>Expert Finder</td>
<td>Vavau &amp; Lieberman, 2000</td>
<td>Frequency of use of Java items</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>APOSDLE’s People Recom. Serv.</td>
<td>Lokaiczyk et al., 2007</td>
<td>Knowledge levels (competency)</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>NASA POPS</td>
<td>Grove &amp; Schain, 2008</td>
<td>Project and competency profile, Semantic Web</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>IBM SmallBlue</td>
<td>Lin et al., 2007; Lin et al., 2009</td>
<td>Frequency of association</td>
<td>Yes</td>
<td>Organizational and social distance</td>
</tr>
<tr>
<td>INDURE FacFinder</td>
<td>Fang et al., 2008</td>
<td>Frequency of association and user profile</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>StrangersRS</td>
<td>Guy et al., 2011</td>
<td>Similarity btw. user and candidate profiles</td>
<td>Yes</td>
<td>Organizational and social distance</td>
</tr>
</tbody>
</table>

**FIG 1.** User interface to define expert preferences

The seeker can define her preferences for each dimension. Currently, InsERT supports the four dimensions discussed in Section II-B. First, the seeker can select a set of required skills. Her selection can be made implicitly, as exhibited in Figure 1 or explicitly from a preloaded list, as displayed in Figure 2.

**FIG 2.** User interface to explicitly select required skills

Second, from a qualifications point of view, the seeker can select from a preloaded list the tools with which a
Moreover, there is another partial order relationship in a partial order, since there are pairs of non-comparable labels.

Let 
\[ S_m = \{ B_1, \ldots, B_m \} \]
where the basic labels chosen to describe a variable is dependent on the characteristics of each represented variable and not fixed. Each basic label corresponds to a linguistic term such as “none” < “very low” < “low” < “medium” < “high” < “very high” (m = 6). The complete universal universe of descriptions for the order-of-magnitude space \( OM(m) \), with granularity \( m \), is the set 
\[ S_m = S_m^* \cup \{ [B_i, B_j] | B_i, B_j \in S_m, i < j \} \]
where the labels \([B_i, B_j]\) with \( i < j \) are defined as \([B_i, B_j] = \{ B_i, B_{i+1}, \ldots, B_j \}\) and named non-basic labels. The order considered in the set of basic labels \( S_m^* \) induces a partial order \( \leq \) in \( S_m \) defined as:

\[ [B_i, B_j] \leq [B_r, B_s] \iff (B_i \leq B_r \text{ and } B_j \leq B_s) \]

where \([B_i, B_j] = B_i\).

This relationship is an order relationship in \( S_m \), but of a partial order, since there are pairs of non-comparable labels. Moreover, there is another partial order relationship in \( S_m \), “to be more precise than”. Given two qualitative labels \( X_1 \) and \( X_2 \) in \( S_m \), we say that \( X_1 \) is more precise than \( X_2 \) if, and only if, \( X_1 \subset X_2 \). The least precise label (most abstract description) is \( ? = [B_1, B_m] \) and most precise label is a basic label.

A distance between a basic label and any other label is considered. This distance is inspired by the concept of distance defined over a set of linguistic elements \( S_m \) [25].

In this paper, we model the seeker’s preferences by employing the following labels. Skills, qualifications, and tools are represented by qualitative nominal variables. Proximity and skill levels are expressed by basic labels: \( C_6^* = \{ “disabled” \text{, “low” \text{, “high”} \} \} \) and \( C_6 = \{ “none” \text{, “very low” \text{, “low” \text{, “medium” \text{, “high”} \text{, “very high” \} \}} \} \) respectively.

The system carries out a process for expressing requirements according to the preferences stated by the seeker. Skills levels are defined as \([C_1, C_6]\), where \( C_1 \) is the basic label of the corresponding requirement. Tools and proximity remain as requirements. An extra requirement is added to indicate the current availability of the candidate. Its default setting is the basic label \( C_6 = “very high” \) (maximum availability is required).

We represent the candidates’ profiles as a list of the following items. Skills and tools are expressed as qualitative nominal variables. Skills are associated with their respective levels. Both skills levels and availability are represented by a basic label from \( C_6 \).

C. Coping with User Requirements

Each candidate is assessed according to each of the seeker’s requirements. We compute an index indicating the level at which each candidate fulfills each requirement.

Each skill selected by the seeker is analyzed. The index defined is high when a candidate’s level in a skill is greater than or equal to the required skill level. If not, the distance between the labels of the requirement and the candidate is used. It is important to note that this index also enables the seeker to select “none” in the preferences interface to express his/her preference for candidates without a certain skill in their profiles.

Next, each candidate’s profile is evaluated for the presence of the required tools. Each candidate’s proximity with respect to the seeker’s preferences defines a matrix of distances between the users in terms of distances between their departments or institutions. Finally, a distance between a candidate’s availability and the availability preference (usually “very high”) is employed.

All partial scores are aggregated for each candidate to give him/her a global score. Figure 3 displays an example of a final rank of candidates recommended by InsERT. Candidates are represented by rows and requirements by columns, where the last column is the global score (“Relevance”) of each candidate.

An experiment was carried out employing skills obtained from LinkedIn endorsements of some of the participants of COLLAGE. This data was provided by COLLAGE peer of P2PUM service. A random set of 300 candidates with a mean of five skills per candidate was defined, along with a small set of tools, each candidate’s availability, and the distance between the candidates who were representing ten random institutions.

Figure 3 displays InsERT’s recommendation for three candidates when the following query was made: high expertise in Market Research, Machine Learning and Programming, knowledge in the tool Bing, low proximity to our institution and the implicit requirement of high availability. As shown, candidates #74 and #72 obtain quite similar global scores while fulfilling the requirements in very different ways: candidate #72 receives good partial scores for the required skills, while candidate #74 obtains good scores in the other variables.

D. OWA operator

The OWA operators are a type of weighted mean that enables the weights of each variable to be tuned by their relative importance. Therefore, variable values are ordered before being weighted.

Definition 1: An OWA operator of dimension \( n \) is a mapping \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) such that [26]:

\[ f(x_1, \ldots, x_n) = \sum_{i=1}^{n} w_i x(i), \]

where \( x(i) \) are the same values as \( x(i) \) ordered from the largest to the smallest, and \( w_i \) are a set of weights such that \( w_i \in [0, 1], \).
Fig. 3. Example of the current output interface of InsERT. The color intensity indicates the candidate’s ability to fulfill the user’s requirement. For example, candidate #72 strongly matches the user’s requirement for the skills Machine Learning and Availability, partially matches the skills Market Research and Programming, and does not match Bing and low proximity requirements.

\[ \sum_{i=1}^{n} w_i = 1 \text{ and } w_i = Q\left(\frac{i}{n}\right) - Q\left(\frac{i-1}{n}\right). \quad (3) \]

In this paper we consider the fuzzy linguistic quantifier “most of”, represented by the Regular Increasing Monotone (RIM) function \( Q(r) = r^{\frac{1}{2}} \).

IV. ADAPTING INSERT TO A BUSINESS APPLICATION

Finding expertise within a global organization is challenging, although necessary. Managers look for internal employees for multiple reasons including filling positions, forming project teams, and solving problems. However, when information is siloed and scattered it is difficult to know where the information resides. ERSs can facilitate this search. In this section, we show how InsERT can be adapted to a business application. Specifically, we look at the challenges in open innovation marketplaces, the role of InsERT, and refine the proximity dimension.

Open innovation intermediaries emerged from two significant challenges which prevented enterprises from benefiting from external innovation resources. First, companies needed to identify how to access external knowledge for innovation and second, they needed to detect and isolate who could help them to innovate [27]. Open innovation intermediaries assist companies with transferring their R&D outside of the organization enabling them to remain competitive, agile, and cost effective. However, due to the unsystematic nature of partner identification, realizing transactions presents a managerial challenge [28]. As an intermediary’s main goal is to make the process of bringing in ideas efficient and to reduce the cost of knowledge transactions, the formation of strong ties is relevant. This efficiency approach makes the nature of the ties that are formed between the solver and the seeker very weak [29]. As Hansen discovered in his study, tie weakness and knowledge complexity interact [30]. Weak ties, such as the ties formed by using open innovation intermediaries’ channel, are regarded as good at bringing in ideas but they are also seen as problematic for transferring complex knowledge. Especially relevant is the effect of these weak ties on the transfer of tacit knowledge, which is revealed through its application [31] and is sticky by nature [32].

To understand the challenges of open innovation intermediaries we looked at previous research and conducted preliminary interviews with an intermediary and two of the university technology transfer offices participating on the platform. The challenges identified can be separated into two groups: systematic and non-systematic. Systematic challenges refer to mechanisms inherent in recommendation systems (RS) such as passive search, filtering for appropriate application, personalizing search results, and matching technology or expertise to technology needs. Non-systematic challenges refer to characteristics of the potential partner(s). In general, ERSs focus on finding the person with the “right level of expertise” rather than “the right person”. Although, intermediaries aim to find the “uniquely prepared mind” to solve the problem [33], the main challenge of intermediary networks is that seeker firms need to identify the “right” partner to build strong ties in order to make knowledge transfer efficient. Therefore, we add a fifth dimension called responsiveness and extend the proximity dimension discussed above.

Responsiveness and suitability of a candidate can play roles in the efficient transfer of knowledge. However, as Erlich et al. point out, one of the challenges of expertise recommenders is to evaluate a candidate’s responsiveness and suitability in providing information [2]. IBM’s SmallBlue, addresses this issue by including information about the shortest path from the user to the candidate. In addition, the system provides information about candidates’ interests and activities [2]. We propose to measure the responsiveness of a candidate based on their behavior. In open innovation marketplaces this information can be captured implicitly by analyzing the interaction between the solver and the intermediary platform. It consists of analyzing available information like: quantity of solutions that a solver provides a seeker, quantity of solutions selected for development by the seeker, time it takes a solver to express interest in solving a problem, and the number of problems a
solver attempts to solve.

We propose two measures to provide contextual factors in a user’s decision to contact a candidate: perceived and language distances. These factors assist in determining a candidate’s responsiveness and suitability in providing expertise. In addition, when applied to organizations they can help to cross organizational boundaries. Siloed organizations resemble the disparate solver community in open innovation marketplaces.

As intermediaries connect disparate networks, there is a high likelihood that partners will be geographically dispersed and need to perform in virtual teams. The perceived dimension refers to seven of the eleven factors in Lojeski et al.’s multidimensional construct of perceived distance called Virtual Distance [34]. The ones which apply to open innovation intermediaries include physical distance, temporal distance (difference in time zones), relational distance, cultural distance, social distance, relationship history and technical skill. The authors found that Virtual Distance© had a direct effect on goal clarity, trust, innovation, and organizational citizenship behavior. We map perceived distance to the proximity dimension.

Culture is a system of shared meanings, which West and Graham operationalized as language [35]. Their study concluded that the greater the linguistic distance from the focal language, the greater the difference in managerial values. Therefore, we propose to map language to the proximity dimension.

Thus, the system should allow the user to select the requirements that best fit his/her needs. Based on the user’s requirements, the system should recommend expertise constructed from a ranking of individuals’ similarities and differences.

Expert recommender systems can address two of the main challenges that open innovation platforms are currently facing, matching and building strong ties between firms and experts. If resolved, they will contribute directly to intermediaries’ value creation. It has been proposed that ERSs can help to fill this need by applying specific dimensions. Furthermore, ERSs will transfer the ownership of search from the solver to the seeker providing greater control over when partners will be matched. By accelerating the process of finding the right solver for a challenge, ERSs reduce the costs of partner identification and innovation lead time for the seeker, and increase the community of participants for the intermediary while growing their revenue stream.

V. CONCLUSION AND FUTURE RESEARCH

In this paper, we analyzed existing commercial expertise recommender systems, existing ranking models, and current ERS dimensions. We presented a new model for an expertise recommender system using linguistics and fuzzy aggregation. Our proposal defines a way to assess candidate profiles according to the user’s preferences and determine a partial assessment. Previous systems allowed the user to adjust the weights of activity spaces in order to differentiate their importance in the assessment of the candidate’s profile [15]. The more weight given to a particular activity the more consideration it was given in the final candidate score. However, a user may not know how much weight to give an activity. A user may be better able to define the level of expertise required in relative terms like low, medium, and high. Therefore, we proposed a linguistic approach to describe the user requirements. These linguistic terms are used to match the user requirements with the candidate’s profiles. We then proposed to aggregate the partial assessments into a global index that reflected the overall adequacy of the candidate’s profile with respect to the user’s requirements. The aggregation method proposed is fuzzy OWA.

Our work continues in two directions. We are working on theoretic aspects of the operator and business case applications of ERSs. From a theoretical point of view, we are considering other linguistic quantifiers to guide the computation of weights used by the OWA operator, identifying the properties of quantifiers such as “at least α” or “almost all” to evaluate their suitability for each type of value being aggregated. In addition, we are analyzing different types of fuzzy OWA operators in order to extend the aggregation function considered in this paper to take into account the imprecision in the fusion process.

From the practical point of view, we are working towards testing InsERT in two kinds of real cases. We are analyzing data from the Tilburg University Expert Collection© [36], which includes a large list of topics associated with a set of teachers or researchers. In addition, we are contacting the editorial boards of two upcoming conferences to study the application of InsERT to the paper-reviewer assignment problem.

Finally, as discussed in section IV, ERSs play an important role in management. We are developing a web tool to support an on-line application of InsERT. Its implementation into a existing business case will allow us to validate our model.

ACKNOWLEDGMENT

The research reported in this paper is partially supported by the SENSORIAL Research Project (TIN2010-20966-C02-01), funded by the Spanish Ministry of Science and Information Technology, the PATRICIA Research Project (TIN2012-38416-C03-01) funded by the Spanish Ministry of Economy and Competitiveness, and by the European Commission funded project COLLAGE: Creativity in Learning through Social Computing and Game Mechanics in the Enterprise, (GA318536) 2012-15.

REFERENCES


1http://ilps.science.uva.nl/tu-expert-collection/