Volume 13, Issue 2, Page 30-33, 2024 https://doi.org/10.46810/tdfd.1397626 Research Article

BİNGÖL UNIVERSITY

Türk Doğa ve Fen Dergisi
Turkish Journal of Nature and Science
www.dergipark.gov.tr/tdfd

Expert Recommendation System to Promote Development-Focused Universities

Bayraktar Mehmet Yasar[1*], Kaya Mehmet[2]

[1](Bingöl Üniversitesi, Mühendislik Fakültesi,Bilgisayar Mühendisliği,1200, Bingöl,Türkiye)
[2](Fırat Üniversitesi, Mühendislik Fakültesi,Bilgisayar Mühendisliği,2300, Elaziğ,Türkiye)
Mehmet Yaşar Bayraktar ORCID No: 0000-0003-3182-120X
Mehmet KAYA ORCID No: 0000-0003-2995-8282

**Corresponding author: mehmetyasarbayraktar@hotmail.com*

(Received: 29.11.2023, Accepted: 19.04.2024, Online Publication: 28.06.2024)

Keywords
Finding Experts, Automatic Candidate Selection, Expert Suggestion,

Abstract: With the support given to university education in recent years, newly opened or developing universities are trying to specialize in various fields within the framework of the regional development mission. In order to achieve success, there is always a need for experts -who do successful studies and have a wide perspective- in the field sought. Reaching candidates who are experts in their field is a difficult process since it cannot be expressed with a linear/linear mathematical formula. In particular, the increasing number of recent studies, articles and reports has made the selection of experts on the subject difficult and systems that automatically find experts are needed. Reaching candidates who are experts in their field is a difficult process since it is relative. In particular, the increasing number of recent studies, articles and reports has made the selection of experts on the subject difficult and systems that automatically find experts are needed. With this study, the scientific studies published by the academicians and the co-authors in the studies will be evaluated and expert candidates "suitable for the criteria sought" will be listed. Results will be given on the basis of scoring scientific studies for the appropriate researcher criteria. To evaluate a researcher's score, an object function of the expert will be made by looking at the filtered keywords over the articles and studies written by that person. In this application, using the 'Dergi Park' database, suitable expert candidates were scored for the filtered words.

Kalkınma Odaklı Üniversiteleri Destekleyecek Uzman Öneri Sistemi

Anahtar Kelimeler
Uzman Bulma, Otomatik Aday Seçimi, Uzman Öneri

Öz: Son yıllarda üniversite eğitimine sağlanan desteklerle, yeni açılmış ya da gelişmekte olan üniversiteler, bölgesel kalkınma misyonu çerçevesinde çeşitli alanlarda uzmanlaşmaya çalışmaktadırlar. Bu çerçevede başarıya ulaşmak amacıyla, alanında yetişmiş başarılı ve geniş görüş açısına sahip olan akdemiysen/uzmanlara her zaman ihtiyaç duyulmaktadır. Konusunda uzman adaylara ulaşmak ise doğrusal/lineer bir matematiksel formül ile ifade edilemediğinden dolayı zorlu bir süreçtir. Özellikle, son zamanlarda yapılan çalışma, makale ve raporların gün geçtikçe artması konuyla ilgili uzman seçimini zorlaştırmış ve otomatik uzman bulan sistemlere ihtiyaç duyulmuştur. Yapılacak bu çalışmayla akademisyenlerin yayınlandığı bilimsel çalışmalar ve çalışmalardaki ortak yazarlar değerlendirilerek "aranan kriterlere uygun" uzman adayları listelenecektir. Uygun araştırmacı kriteri için bilimsel çalışmaların puanlanması esasına göre sonuç verilecektir. Bir araştırmacının puanı değerlendirilirken, o kişinin yazdığı makaleler ve çalışmalar üzerinden, filtre edilen anahtar kelimelere bakılarak, uzmana ait objektif bir sıralama yapılacaktır. Bu uygulamada "Dergi Park" sisteminin veri tabanı kullanılarak, filtrelenen kelimelere uygun uzman adayların puanlandırılması sağlanmıştır.

1. INTRODUCTION

The increase in the speed and use of the internet and computers day by day has led to the need to store almost every existing piece of information in offline or online databases. The explosion in knowledge in the database brings to mind the idea of how these resources can be turned into a profit. While the previously used database techniques were based on querying and filtering, data mining techniques aim to discover the hidden and useful information stored in the data [1].

Data mining is the discovery of invisible or unnoticed knowledge from large-scale data. Since the aim of this study is to select experts on the subject, determining the knowledge, skills, and areas of expertise; it is crucial to choose the sources correctly to obtain meaningful information.

In addition, with the augmentation in the use of computers and technological tools, almost every organization and company has started to store their employees' areas of expertise, talents and contact information in databases. Especially in academic communities, information about the person, the schools and universities he or she attended, personal background information, specific fields of study and certificate information are kept in related tables. So, meta-databases could be used as first step to determine expert area of a person.

As a second step, documents, personal blogs, news and websites will provide much more detailed information about the expert. This process can be called as secondary source selection and document collection about the expert.

Experts often cannot express themselves well or they may use very general or specific expressions when describing their areas of expertise. In particular, updating information is dependent on people, updating the process is time-consuming and undesirable, and the information accessed in the database does not reflect the person relatively.

In addition, while data are being collected, they contain some errors called noise caused by people, environment and equipment [2]. In addition, obtaining data from different sources leads to inconsistent or incomplete information. These noisy and incomplete data obtained from these sources can be somewhat removed by using some data preprocessing techniques.

The demand for specialists must be obtain from external sources because of ever-increasing need for information in organizations and companies, and the limited number of experts in these area. For this purpose, the evaluation of expert candidates outside the organization, which is called external resources, is of great importance.

In general, expert represents people who have knowledge about a particular subject, are self-educated, have successful studies and have a broad perspective. Sometimes, they are called people who have special talents, mastery and dexterity in addition to their knowledge. The need for specialists is increasing, especially in rapidly and continuously developing fields such as health, informatics and technology. For this purpose, it is desired to create automatic expert finder systems by making use of many kinds of data sources and expert knowledge.

Previous research suggests expert-finding systems using many techniques such as information extraction, data mining, and probabilistic modeling. P@NOPTIC Expert [3], Expertise Browser(ExB) [4], and Expert Finder [5] are some of the models that is developed as software in this regard.

The process of finding experts includes three important steps such as 'Selecting Experts' Source', 'Creating Expert Model' and 'Matching Searched Criteria with Expertise Topics' [6].

2. LITERATURE WORK

Meta databases serve to keep the skills and expertise of employees or members. Thus, general profile information of experts can be created. Moreover; projects the candidates were worked on, previous experience and background information of the candidates are recorded in the databases and certain information about the candidate is extracted. In a study conducted at NASA, a system called SAGA [7], which helps to find experts, is developed by using various educational data and profile information.

Although the process of finding talent and expertise by using the main databases seems relatively easy, the fact that experts do not like to update their knowledge or there is incomplete or erroneous information causes the solution to become complicated and expensive. In addition, the fact that the expertise information in the system is very general and irregular are the factors that make it difficult to score experts.

2.1. Document Collection of an Expert

Having knowledge of related documents belonging to the candidate is also an important factor in defining the candidate's area of expertise. The information in the related publications, reports, mails, forums and web pages can provide a detailed inference about the expert. Dumais and Nielsen reached 117 conferences and tried to find their areas of expertise by looking at the abstracts of the articles published by the researchers [8]. Moreover, Balog developed a model that finds experts by using various documents such as speech, mail, web page, and report by making use of the local document warehouse [9]. Campbell et al., on the other hand, tried to determine the fields of expertise of the users by making use of the sender and receiver information, as well as the e-mail contents and headers by using inter-mail connections [10]. However, a problem that should not be forgotten is that when determining the subject related to the mail, subject drifts (topic-drift) may be encountered. In addition, in the document collection process, name similarities are another issue that needs to be studied.

3. MATERIAL AND METHOD

In order to find the areas of specialization of a candidate, it is necessary to collect knowledge about the candidate. This type of information can be obtained from major databases (personal information, talent, expertise), online/offline documents (such as publications, forums, emails, reports, and web pages) belonging to the

candidate, or candidate's relational networks (such as social, academic and business networks).
When we are searching for an expert, we generally ask these types of questions 'Who is an expert on this area' or 'Who can make this job/work better?' [6]. With in these types of queries, client is looking for an expert in a specific field. Algorithm structures such as 'Creative Probabilistic Model', 'Lay-Based Model', 'Network-Based Model', and 'Hybrid Model' are widely used in expert finding problems. The expert inference algorithm usually answers questions such as 'Finding the Candidate's Areas of Expertise', 'Exploring Evidence of Expertise', and 'Inferring the Relationship between the Questioned Subject and the Expert Candidate' [11].

There are three inputs/criteria that are important to expert search:

- Expert Candidate (EC)
- Document(D)
- Subject(S), Term in the Query (t,t ∈S)

While the expert candidate represents the people who can be experts in the subject to be sought, the documents characterize the online/offline resources that may belong to that expert. The subject represents the expert search query/criteria. While this criterion is sometimes a few keywords, sometimes it can be in a summary text to represent the expertise to be sought.

Figure 1 shows web interface for finding expert candidates who have worked on 'bee venom'. Additionally, by selecting the search source, it is possible to specify which sources data can be retrieved from.

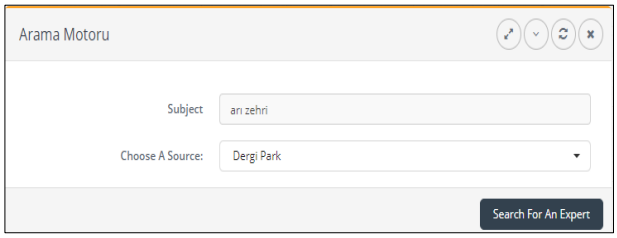

Figure 1: Determination of Expert Keywords

Figure 2 demonstrates the flowchart of the expert candidate extraction algorithm suitable for the filter. After keywords are entered into the application, article information containing these words is pulled from the specified data source. Considering each article and author information, it is placed in the fitness function.

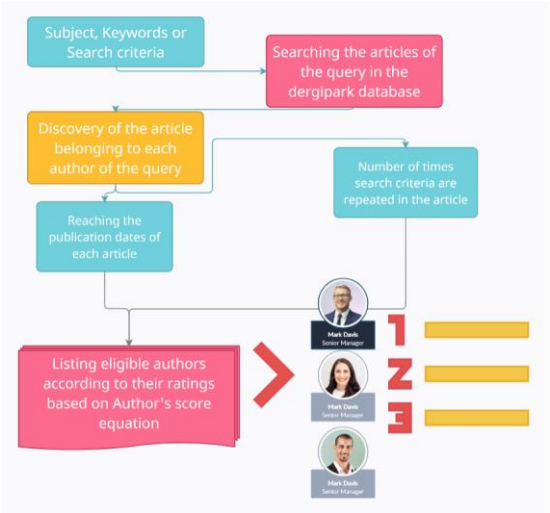

Figure 2: Workflow of sorting expert candidates suitable for the filter

In Equation 1, how the scores of each candidate author will be calculated is expressed mathematically. Thus, the content of the article written by an author is evaluated and scored according to the publication date and the number of co-authors. Here, the 'alpha(α)' coefficient can be made on a user basis by changing the publication date according to its importance. For example, in some fields (agriculture, agriculture, etc.), the time of publication of the article may not be very important, while in the field of technology and computers, the 'alpha(α)' coefficient can be changed to increase the value of time.

$$AuthorScore_i = \sum_{i=0}^{i=AuthorPaperCount} (20 - (2023 - PublishDate) * \propto * \left(\left(1 + \frac{1}{SharedAuthorCount}\right) * KeywordsRepetationCount\right) \quad (1)$$

In Figure 3, the candidates who wrote an article in the "Journal Park" database in the "Bee venom" field are listed. Universities and organizations that need experts in this field can contact these authors and sign successful projects with academicians who are experts in their fields.

ID	Name	Score	TimeBasedScore	RepetitionCount	WeightedScore
0	Ahmet DODOLOĞLU	30	270	80	127,222222222222
0	Erkan TOPAL	7	128	33	114,416666666667
0	Nurullah ÖZDEMİR	10	170	30	113,333333333333
0	Sevgi KOLAYLI	11	137	36	102,666666666667
0	Vural BAŞARAN	10	150	30	100
0	Yusuf ÇELİK	10	140	30	93,3333333333333
0	Ralitsa BALKANSKA	10	110	40	91,6666666666667
0	Borislav KASHAMOV	10	110	40	91,6666666666667
0	Engin KILIÇ	10	190	20	84,4444444444444
0	Meltem UÇAR	4	72	23	84

Showing 1 to 10 of 486 entries Previous 1 2 3 4 5 49 Next

Figure 3: Listing successful authors in the field of 'bee venom'

4. DISCUSSION

When companies and universities search for expert candidates in their field, they generally use keywords and examine articles and studies written on the subject. Examining these studies and finding connections between authors not only took a long time, but also left the process of finding an expert to personal preference. In this study, using the 'Dergi Park' database, connections were revealed between the articles written by expert candidates and the searched keywords. In addition, the up-to-dateness of the expert candidate's work was taken into consideration, taking into account the date of acceptance of the articles and the order of the authors in the article. Thus, a system that evaluates expert knowledge based on time, author and keyword was obtained.

3. RESULTS

Finding experts, which is a difficult and time-consuming process, is automated using the contents of academic articles. Thus, it was transformed into an application that companies, organizations and universities can benefit from.

An expert evaluation system is developed by including the effect of academic articles. In particular, a more dynamic system is proposed to replace the web-sites where experts are registered, such as 'kariyer.net' and 'Linked-in', which allow searching for experts by paying fees and which need constant updating.

In addition, the fact that expert candidates remain general or specific when describing themselves and their expertise, and the need for continuous updating of their CV information makes the process of finding experts difficult. However, with the use of academic network structure and personal reports, documents and articles, it is possible for the expert to evaluate the candidate's work area and abilities in a more up-to-date and objective way.

As a result, a new expert recommendation system is presented by ensuring that the academicians who are experts in their fields that universities and organizations may need are listed in an objective way, preventing the institutions from spending too much time and effort.

REFERENCES

[1] N. T. ,. M. Y. Serkan SAVAŞ, «Veri Madenciliği Ve Türkiye'deki Uygulama,» İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi, 2011.

[2] A. OĞUZLAR, «Veri Ön İşleme,» Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, pp. 67-76, 2003.

[3] N. C. P. d. V. S. Peter Bailey, «Overview of the TREC 2007 Enterprise Track,» TREC, 2007.

[4] A. &. H. J. Mockus, «Expertise browser: a quantitative approach to identifying expertise,» %1 içinde Proceedings of the 24th international conference on software engineering, 2002.

[5] M. Maybury, «Collaborative Virtual Environments for Analysis and Decision Support,» Communications of the ACM, 2001.

[6] W. H. D. W. L. Shuyi Lin, «A survey on expert finding techniques,» Springer Science+Business Media New York, p. 260, 23 11 2017.

[7] Becerra-Fernandez, «Locating expertise at nasa: developing a tool to leverage human capital,» Knowledge Management Review, pp. 34-37, 2001.

[8] J. N. Susan T. Dumais, «Automating the assignment of submitted manuscripts to reviewers,» Proceedings of the 15th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 233-244, 1992.

[9] L. A. d. R. Krisztian Balog, «Formal Models for Expert Finding in Enterprise Corpora,» Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval, 2006.

[10] P. P. M. C. D. Christopher S. Campbell, «Expertise identification using email communication,» Proceedings of the twelfth international conference on Information and knowledge management , 2003.

[11] Y. F. d. R. S. S. Krisztian Balog, Foundations and Trends in Information Retrieval, Hanover, MA 02339,USA: now Publishers Inc., 2012.