2016 IEEE International Conference on Computational Science and Engineering, IEEE International Conference on Embedded and Ubiquitous Computing, and International Symposium on Distributed Computing and Applications to Business, Engineering and Science

Levelized Taxonomy Approach for The Job Seeking/Recruitment Problem

Michaël Guedj
UTEC INT, School of IT and New Technologies
UTEC – CCI Seine-et-Marne
Emerainville, 77436 Marne-la-Vallée Cédex 2 (France)
michael.guedj@seineetmarne.cci.fr

***Abstract*—Recently, job recommendation has attracted a lot of research attention; the aim being to get a sorted list of relevant candidates for an applicant (job sekeer or recruiter). To an effective matching, the utilisation of semantic technology has shown good results. Particularly, the use of taxonomy hierarchizing the skill following a relation of inheritance. However request to the user to weighing the skills is a barrier to an usability and an efficiency of such methods on the user point of view. This paper intends to provide a first answer for such a problem.**

I. Introduction

Recommender systems emerged as an independent research area in the mid-1990s, when researchers and practitioners started focusing on recommendation problems that explicitly rely on the notion of ratings as a way to capture user preferences for different items [1] [6]. The Internet-based recruiting platforms become a primary recruitment channel in most companies. Recommendation systems are widely used on the Internet to assist customers in finding the products or services that best fit with their individual preferences. Recommender systems are being broadly accepted in various applications to suggest products, services, and information items to latent customers. In order to improve the e-recruiting functionality, many recommender system approaches have been proposed [22] [2]. Different from traditional recommendation systems which recommend items to users, job recommender systems recommend one type of users (*e.g.*, job applicants) to another type of users (*e.g.*, recruiters). In particular, job recommender system is designed to retrieve a list of job positions to a job applicant based on his/her preferences or to generate a list of job candidates to a recruiter based on the job requirements.

II. Related Work

Several techniques and approaches have been proposed to construct automatic online recruitment systems. The traditional keyword based techniques mainly depend on exact matching between keywords extracted from job posts and candidate resumes; they suffer from low precision because they usually ignore the underlying semantic aspects of the terms that are extracted from both job posts and resumes [3]. Relevance models are usually built from known relevant resumes to a specific job post [13]. A major problem of these approaches is their low precision when tested against large-scale real-world datasets [24]. A number of machine learning algorithms are exploited in the online recruitment domain for data analysis and information extraction [10] [21] [11] [9]. These algorithms include neural networks [21], clustering [9], decision trees [19], and support vector machines [11]. As reported in [10], the main drawback of machine learning approaches is that they produce high error rates as they rely on manually-developed training corpora. To effectively locate and match individuals and positions, within or from outside an organization, it is important to use semantic technology [7] [15]. The authors of [5], [15], [20] [23] propose automatic recruitment systems that employ semantic resources that have been built based on integrated classifications and standards. [18] uses extra probabilistic edges to extend the applications with skills that the applicant possibly possesses. Particularly, [5] and [15] uses [25]'s similarity measure to evaluate the degree of match between job offers and applicants. [8] uses a similar approach but employs different similarity measures related to different scenarios. The semantic approaches have shown interesting results in accomplishing the matching process. However, development of complete and reliable ontologies that capture up-to-date knowledge about specific domains remains tricky [16] [14]. Mixed approaches also have been proposed using the previous ones or/and using technics from other fields [12]; such approaches are however heavily dependent on previously presented ones.

III. Semantic Matching

Semantic matching is a technique which combines annotations using controlled vocabularies with background knowledge about a certain application domain. In [5], the domain specific knowledge is represented in the form of various concept hierarchies and can be used to determine the semantic similarity between concepts [17]. Thus a comparison between job descriptions and applicants profiles is possible based on their semantic similarity and not merely relying on the containment of keywords.

This approach sustained especially by [5] and [8] is based on ideas from [25] and [4].

A. Hierarchy of Skills

The background knowledge of the recruitment domain is represented in a machine understandable format that allows to compare job descriptions and applicant profiles based on

978-1-5090-3593-9/16 $31.00 © 2016 IEEE
DOI 10.1109/CSE-EUC-DCABES.2016.222

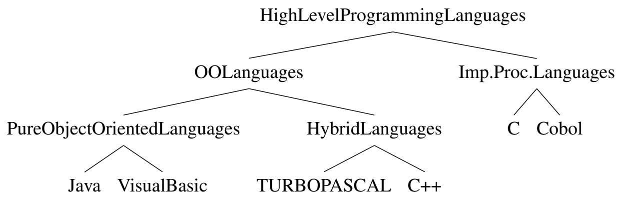

Fig. 1. Skills Hierarchy (taken from [15])

their semantic similarity [17] instead of merely relying on the containment of keywords. Such a format is usually [5] [15] [18] a hierarchy of skills so-called taxonomy, which captures the inheritance relationship. An example of such a taxonomy is given in Figure 1 (and is taken from [15]).

B. The Concept Similarity

The similarity between two concepts (nodes in the taxonomy) c_1 and c_2 is determined by the distance $d_c(c_1, c_2)$ between them, which reflects their respective positions in the concept hierarchy. This concept similarity is defined in [25] (then used by [5] and [8]) as:

$$sim_c(c_1, c_2) = 1 - d_c(c_1, c_2)$$

Every node in a concept hierarchy is assigned a milestone value [25], which is calculated with the formula:

$$milestone(x) = \frac{1}{2k^{level(x)}}.$$

where:

- k is a factor larger than 1 indicating the rate at which the value decreases along the hierarchy;
- $level(x)$ is the depth of the node x in the hierarchy.

Since the distance between two given concepts in a hierarchy represents the path from one concept to the other one over the closest common parent ccp, the distance is calculated as shown below [25]:

- $d_c(c_1, c_2) = d_c(c_1, ccp) + d_c(c_2, ccp)$
- $d_c(c, cpp) = milestone(ccp) - milestone(c)$

This model implies two assumptions:

1) The semantic differences between upper level concepts are bigger than those between lower level concepts (in other words: two general concepts are less similar than two specialized ones);
2) The distance between brothers is greater than the distance between parent and child.

C. The Competence Level Similarity

[5] provides means for specifying required competence levels (cl) in job postings. Hence, the method provided by [5] not only considers taxonomic similarity of concepts but also compares competence levels in order to find the best match. The competence level similarity is determined by the following formula [5]:

$$sim_p(cl_1, cl_2) = \begin{cases} 1 - \alpha(cl_1 - cl_2) & \text{if } cl_1 > cl_2 \\ 1 & \text{otherwise} \end{cases}$$

where $0,25 \leq \alpha$ is a factor indicating the rate at which the value of sim_p decreases with increasing deviation between competence levels.

D. Global Similarity

The approach sustained by [5] gives employers also the opportunity to specify the importance of different job requirements. The concept similarity is then justified by the indicated weight, *i.e.* the similarity between more important skills will have greater influence on the similarity between a job position posting and an applicants profile. Putting all together, the formula for calculating the similarity of a job position posting (c) and a job position seeker (s) is (from [5]):

$$Sim(c, s) = \sum_{i \in I} w(c_i). \max_j \Big[sim_c(c_i, s_j).sim_p(p(c_i), p(s_j)) \Big]$$

where $\sum_{i \in I} w(c_i) = 1$ and $p(c_i)$ (resp. $p(s_j)$) is a competence level in c_i (resp. s_j).

Each required skill from the job position posting (c_i) is compared with each skill in an applicants profile (s_j). This includes the calculation of both concept and competence level similarities. The similarity values of the best matching pairs are multiplied by the corresponding weight and summed up yielding the final similarity.

IV. Proposal

To an effective matching, the utilisation of semantic technology has shown good results. Particularly, the use of taxonomy hierarchizing the skill following a relation of inheritance. However request to the user to weighing the skills is a barrier to an usability and an efficiency of such methods on the user point of view. One intends to provide a first answer for such a problem.

We propose to prioritize the skills inside the taxonomy itself, while ensuring the inheritance relationship. Thus each level of each concept matches a level of competence. The global similarity formula can be thereby written as follow:

$$Sim(c, s) = \sum_{i \in I} level(c_i). \max_j \Big[sim_c(c_i, s_j).sim_p(level(c_i), level(s_j)) \Big]$$

It is also natural to consider $\alpha = \frac{1}{depth\_max+1}$, where $depth\_max$ is the depth of the hierarchy (taxonomy). Such an approach matches the common sense which states that a specialized knowledge/profile worths more than a general one. It follows that it is unnecessary to require applicants to weighting the skills (procedure rather cumbersome). Indeed, any node being associated with a level of competence, the weighting of each requested skill is derived naturally.

Note that choices, or even interpretations, must be made by the designer, on the conception of this levelized hierarchy, which is not necessarily obvious, and therefore requires the use of an expert. But such a preliminary work is able to simplify the tasks of the user, who no longer is obliged to perform a weighting: on the desired requested profile in the case of a recruiter; or on the alleged profile in the case of a job seeker.

Figure 2 gives an example of such a levelized taxonomy. Figure 3 gives the results of the ranking of some profiles, using the similarity measure Sim on the levelized taxonomy from the Figure 2, and where the applicant is some company requesting a candidate with good knowledge in JEE and Big Data expressed by the string "JEE_++__Big_Data_++"); the adding of skills being indicated by: "__".

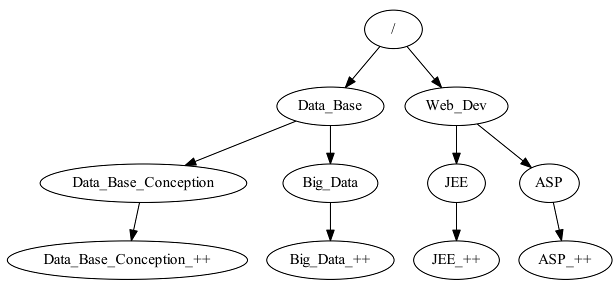

Fig. 2. Example of Levelized Taxonomy

Profile	Similarity (%) to JEE_++__Big_Data_++
JEE_++__Big_Data_++	100.0
JEE__Big_Data_++	87.5
JEE_++__Big_Data	87.5
ASP_++__Big_Data_++	83.3
JEE_++__Data_Base_Conception	79.2
ASP__Big_Data_++	79.2
JEE__Big_Data	75.0
JEE_++__Data_Base	75.0
ASP_++__Big_Data	70.8
JEE__Data_Base_Conception	66.7
ASP__Big_Data	66.7
JEE__Data_Base	62.5
ASP_++__Data_Base_Conception	62.5
ASP__Data_Base_Conception	58.3
ASP_++__Data_Base	58.3
ASP__Data_Base	54.2

Fig. 3. Similarity with the profile JEE_++__Big_Data_++

Note that the example of the Figure 1 does not match the inheritance relationship with the competence level; as well as does not consider that the concept-nodes of the hierarchy at a certain depth have the same value of competence. For instance, still considering Figure 1, if we had considered that the hierarchy had been levelized, then one would have deduced that the skill in VisualBasic had been more valuable than the skill in C; because then level(C) $<$ level(VisualBasic). More generally, if x and y are two skills and $p(x)$ (resp. $p(y)$) is a competence level in x (resp. in y)); then note that, the knowledge of $(x, p(x))$ and $(y, p(y))$ is not sufficient to operate an automatic weighting, *i.e.* to automatically deduct an order between $(x, p(x))$ and $(y, p(y))$; think *e.g.* of (Word, medium) and (Java, medium).

V. Conclusion

Our ongoing work takes into account more complex cases. Our proposal presented here has a counterpart: the design of the levelized taxonomy. Indeed, each node at the same level (depth) of the taxonomy has to be considered at the same level of competence. We are currently working on a design methodology for the levelized taxonomies, related to the job seeking/recruitment problem. We investigate also the implications of our approach in the field of machine learning.

Acknowledgment

Our work would not have been possible without the support of the CCI Seine-et-Marne, Mr Fehd Bensaid Director of the UTEC and Mr Frédéric Bourcier Director of the IT and new technology department of the UTEC.

References

[1] Gediminas Adomavicius and Alexander Tuzhilin. Context-aware recommender systems. In *Recommender Systems Handbook*, pages 191–226. 2015.

[2] S. T. Al-Otaibi and M. Ykhlef. A survey of job recommender systems. In *International Journal of the Physical Sciences, vol. 7(29)*, pages 5127–5142, July 2012.

[3] Nicholas J. Belkin and W. Bruce Croft. Information filtering and information retrieval: Two sides of the same coin? *Commun. ACM*, 35(12):29–38, December 1992.

[4] A. Billig and K. Sandkuhl. Match-making based on semantic nets: The xml-based approach of baseweb. In *Proceedings of the workshop on XML-Technologien fr das Semantic Web*, pages 39–51, 2002.

[5] Christian Bizer, Ralf Heese, Malgorzata Mochol, Radoslaw Oldakowski, Robert Tolksdorf, and Rainer Eckstein. The impact of semantic web technologies on job recruitment processes. In *Wirtschaftsinformatik 2005: eEconomy, eGovernment, eSociety, 7. Internationale Tagung Wirtschaftsinformatik 2005, Bamberg, 23.2.2005 - 25.2.2005*, pages 1367–1382, 2005.

[6] Seng W. Loke Chuong C. Vo, Torab Torabi. A survey of context-aware recommendation systems.

[7] Simona Colucci, Tommaso Di Noia, Eugenio Di Sciascio, Francesco M. Donini, Marina Mongiello, and Marco Mottola. A formal approach to ontology-based semantic match of skills descriptions. *J. UCS*, 9(12):1437–1454, 2003.

[8] Maryam Fazel-Zarandi and Mark S Fox. Semantic matchmaking for job recruitment: an ontology-based hybrid approach. *Proceedings of the 8th International Semantic Web Conference*, 2009.

[9] Wenxing Hong, Siting Zheng, Huan Wang, and Jianchao Shi. A job recommender system based on user clustering. *Journal of Computers*, 8(8), 2013.

[10] Remy Kessler, Nicolas Bechet, Mathieu Roche, Juan-Manuel Torres-Moreno, and Marc El-Beze. A hybrid approach to managing job offers and candidates. *Inf. Process. Manage.*, 48:1124–1135, 2012.

[11] Remy Kessler, Juan Manuel Torres-Moreno, and Marc El-Beze. *MICAI 2007: Advances in Artificial Intelligence: 6th Mexican International Conference on Artificial Intelligence, Aguascalientes, Mexico, November 4-10, 2007. Proceedings*, chapter E-Gen: Automatic Job Offer Processing System for Human Resources, pages 985–995. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.

[12] A. B. Kmail, M. Maree, M. Belkhatir, and S. M. Alhashmi. An automatic online recruitment system based on exploiting multiple semantic resources and concept-relatedness measures. In *Tools with Artificial Intelligence (ICTAI), 2015 IEEE 27th International Conference on*, pages 620–627, Nov 2015.

[13] Victor Lavrenko and W. Bruce Croft. Relevance based language models. In *Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '01, pages 120–127, New York, NY, USA, 2001. ACM.

[14] M. Maree and M. Belkhatir. A coupled statistical/semantic framework for merging heterogeneous domain-specific ontologies. In *2010 22nd IEEE International Conference on Tools with Artificial Intelligence*, volume 2, pages 159–166, Oct 2010.

[15] Malgorzata Mochol, Holger Wache, and Lyndon J. B. Nixon. Improving the accuracy of job search with semantic techniques. In *Business Information Systems, 10th International Conference, BIS 2007, Poznan, Poland, April 25-27, 2007, Proceedings*, pages 301–313, 2007.

[16] Paslaru E. Mochol, M. and B. Simperl. Practical guidelines for building semantic erecruitment applications. In *International Conference on Knowledge Management*, IKNOW06, Special Track: Advanced Semantic Technologies (AST'06), pages 1–8, Graz, Austria, 2006.

[17] Jonathan Poole and J. A. Campbell. A novel algorithm for matching conceptual and related graphs. In *Proceedings of the Third International Conference on Conceptual Structures: Applications, Implementation and Theory*, ICCS 95, pages 293–307, London, UK, UK, 1995. Springer-Verlag.

[18] Gábor Rácz, Attila Sali, and Klaus-Dieter Schewe. *Foundations of Information and Knowledge Systems: 9th International Symposium, FoIKS 2016, Linz, Austria, March 7-11, 2016. Proceedings*, chapter Semantic Matching Strategies for Job Recruitment: A Comparison of New and Known Approaches, pages 149–168. Springer International Publishing, Cham, 2016.

[19] K. Ramar and N. Sivaram. Article: Applicability of clustering and classification algorithms for recruitment data mining. *International Journal of Computer Applications*, 4(5):23–28, July 2010. Published By Foundation of Computer Science.

[20] Francisco García Sánchez, Rodrigo Martínez-Béjar, Leonardo Contreras, Jesualdo Tomás Fernández-Breis, and Dagoberto Castellanos Nieves. An ontology-based intelligent system for recruitment. *Expert Syst. Appl.*, 31(2):248–263, 2006.

[21] Chung-Kwan Shin, Ui Tak Yun, Huy Kang Kim, and Sang Chan Park. A hybrid approach of neural network and memory-based learning to data mining. *Trans. Neur. Netw.*, 11(3):637–646, May 2000.

[22] Z. Siting, H. Wenxing, Z. Ning, and Y. Fan. Job recommender systems: A survey. In *Computer Science Education (ICCSE), 2012 7th International Conference on*, pages 920–924, July 2012.

[23] F. Trichet, M. Bourse, M. Leclere, and E. Morin. Human resource management and semantic web technologies. In *Information and Communication Technologies: From Theory to Applications, 2004. Proceedings. 2004 International Conference on*, pages 641–642, April 2004.

[24] Xing Yi, James Allan, and W. Bruce Croft. Matching resumes and jobs based on relevance models. In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '07, pages 809–810, New York, NY, USA, 2007. ACM.

[25] Jiwei Zhong, Haiping Zhu, Jianming Li, and Yong Yu. Conceptual graph matching for semantic search. In *Proceedings of the 10th International Conference on Conceptual Structures: Integration and Interfaces*, ICCS '02, pages 92–196, London, UK, UK, 2002. Springer-Verlag.