Job Seeker To Vacancy Matching Using Social Network Analysis

PAVITHRA M[1], Mr. P. ANBUMANI[2]
[2] Assoc.Prof.
[1, 2] Krishnasamy College of Engineering and Technology,
Cuddalore.

Abstract- *Social network analysis is the investigation of social structures by using methods such as graph theory and machine learning. Social networks characterize networked structures in terms of nodes (i.e., individuals) and their relationships to each other as acquaintances, colleagues, collaborators and/or classmates. Through these relationships, one can find their ties with their connections, professions, and the degree of the ties. Networking sites such as LinkedIn® and Researchgate® also contain more information of the knowledge of connections about the skill of an individual. The purpose of this study is to identify methods that measure the skills, expertise and experience of a job seeker and to investigate importance of using social networking data as input to user modeling that determines the strength of skills to be used for recommending matching job vacancies. Result of preliminary experiment using social network data in skill measurement shows consistent improvement in accuracy of matching job seekers tovacancies.*

I. INTRODUCTION

In this increasing globalization of businesses and industries, shift of the economy to knowledge economy, expansion ofInformation Communication Technology (ICT) andemergence of knowledge workers, personalized job vacancy recommendation using knowledge extracted from diverse sources is of paramount importance. To achieve that, a deeper look on the application of Knowledge Management (KM) techniques to capture useful knowledge from large unstructureddatacancontributesignificantlytosolveproblems of job seeker to vacancy matching. KM is the process of identifying, capturing, creating, utilizing, storing, and dissemination of knowledge. and provisions; iii) job seekers do not incorporate complete information in their application.Ontheotherhand,employeesquitorgetfiredafter they assume the position because they under- or overstated their suitability to the jobduring their application [2].This paper is set out to contribute to the current literature by discussing a prototype online job matching and recruitment system that integrates social networking data, as part of jobseeker modeling system for vacancy matching and recommendation.

This paper is organized into eight sections. Section one introduces the scope and the motivation of the study.In Section II, we present the context – the jobseeker-to-vacancy matching system setup. In that section we briefly describe the building blocks of the context-aware job recommendation systems and their interaction with the objective of providing where this study fits into the big picture. In Section III, we discuss social networking and the role it plays in job recommendation. Section IV describes the methods of skill weighting in the process of jobseeker modeling with special emphasison social network analysis. Section V describes how the matching process is performed. In Section VI, we discuss the result of preliminary experiment with and without the use of social network data. Section VII, briefly presents a review of related works and describes the difference between our study and the related works. Finally, in Section VIII we present our conclusion and outline the future outlooks of this study.

II. VACANCY RECOMMENDATION

This study is part of the research on the development of personalized vacancy recommendation that involves i) job vacancy modeling through data from occupational standards [3], ii) job seeker modeling through skill, qualification and experience data from the job seeker’s self-assessment as well as evaluation by connections of individuals; and iii) matching the job vacancy to the job seeker, as depicted in Fig. 1.

Traditional recruitment involves the evaluation of job seeker's training qualifications, personal statement of purposes, and recommendations from supervisors and/or peers. Online recruitment systems seem to lack one or more of the last two key measures – self-evaluation and evaluation by others – of a job seeker. However, online systems could incorporate self-evaluationviaskillsurveyandconnectionevaluationviasocialnet working.evaluationviaskillsurveyandconnectionevaluationvias ocial networking.

Fig. 1 Components of Context-Aware Vacancy Recommendation

This paper mainly connects the role recommendation letter plays in recruitment in real life to online recruitment systems and explores how we can incorporate this information from social networking data into online job matching and recruitment systems.

III. USING SOCIAL NETWORKING FOR E-RECRUITMENT

Recommendation and references (either in the form of letters or otherwise) play a big role in getting more information about (and assessing the background of) a jobseeker from the supervisors'/coworkers' point of view. Social network is an online communication platform for people who are connected in some ways – be it as relatives, colleagues, classmates or simple acquaintances – expedited by the emergence of ICT resultedinfacilitationofdifferentcommunicationplatformsfor people.

These days, social networking sites are increasingly becomingkeysourceofdeterminingthebehaviorofindividualsbasedon their connections, contributions, and reactions. Social networking is playing an important role in the day-to-day life of individuals and organizations. There are a number of social networking platforms being used online with varying focus. Amongthem,Facebook®[4],Twitter®[5],andLinkedIn®[6] are the most widely used platforms for social and professional purposes. Others such as Researchgate® [7] and Academia.edu® [8] are used predominantly for scientific purposes.

Though not integrated in recruitment systems, social networking platforms are being utilized by employers as main tool to search for their prospective employees. These sites are good sources not only of employee profiles, but also their connections together with the witness of their connections about those prospective employees. According to [9], that conducted a survey on the extent to which social networking sites are used for recruitment, 45% of companies use twitter, 80% use LinkedIn®, and 50% use Facebook® to find talent.

This study is focused on extracting useful knowledge that emanates from relationships between users of social networking sites – the knowledge of each other's skills, expertise, experiences and attitude – to use in evaluating the match between job seekers and job vacancies. Social networks are increasingly being used in recruitment and selection processes and hence can be integrated to online recruitment (also known as e-recruitment) systems per se.

ExpansionofsocialnetworkingsitessuchasLinkedIn®[6]andResearchgate®[7]tohavemillionsofusersaroundtheworldis evident for its importance as source of information about job seekers. For example, in LinkedIn®[6], connections of users provide useful information about one user in two ways: i) by endorsing skills of a member in the networking sites with whichtheyareconnectedtoandii)bywritingrecommendation that explains how they evaluate the user they collaborated with. On the other hand, in Researchgate® [7] connections of users provide information about one user by i) approving questions asked by a particular user, ii) approving answers provided by a user, iii) evaluating the publication of a user,and soon.

The importance of social networking as a means to better understand a job seeker is discussed in a number of studies. "Professional networking sites such as LinkedIn® Corp. and Jobster Inc. are making it easier for employers to get in touch with people who have worked with job candidates in the past or know them personally." [10]

IV. JOB SEEKER MODELLING

The technologies used to perform the analysis and matching are i) machine learning – to extract, analyze, and present the result; ii) web mining – to scrap data online; and iii) user interaction – to allow the user to add or remove features that characterizehim/herand/oradjusttheweightsassignedtoeach skillfeature.

A. Determining and Selecting Skills Features

We extract features – that characterize the job seeker – from his/her connections on social networking sites and forums (i.e., from the endorsements and contributions) to build template for job seeker modeling. Endorsements and up votes are used as means to distinguish the strength of a particular skill from others. For example, using endorsement, as shown in Fig. 2, this user is more suited to skill Artificial Intelligence than Algorithms.

B. Measuring the skills of job seeker

A job seeker can acquire different skills at different times and tend to forget some of these skills over time because of lack of practice according to the power law of forgetting [11, 12, and 13].Forthisreason,measuringtheextenttowhichtherelevant skill is readily put into action is important to decide whether a job seeker possessing a particular skill is fit to a vacancy that requires thatskill.

Measuring the skill of a job seeker involves assigning weights to skills of individuals based on: subjective measure, semi- subjective measure, and objective measure as explained below.

'Subjective measure – skill of a user can be assigned weights based on self-assessment by job seekers themselves, i.e., by letting the user provide the extent of his/her skills on a scale.
Semi-subjective measure – in Semi-subjective measure can be done by using i) rated measure using information on recency and duration of experience/training ii) rated measure using endorsement of skills (e.g., recommendation) by networks of the individuals. Quantifying 'endorsements' by networks after collecting the endorsement information.

Objective measure – objective measure is obtained when administering a short quiz for each skill to the job seeker to determine the skill level.

Skill weighting is done using these three measures, namely, score 1, denoted by x_1, measures the subjective evaluation of theskilllevelprovidedbyjobseeker'sself-assessment;score2, denoted by x_2, measures the semi-objective evaluation of the skill level of a job seeker calculated from the assessment of his/hercollaboratorsinsocialnetwork;andscore3,denotedby x_3, measures the objective evaluation of the skill level of a job seeker calculated from the results of quizzes designed for assessment of skills of the job seeker. As shown in Fig. 3, subjective(i.e.,x_1),semi-subjective(i.e.,x_2)andobjective(i.e., x_3) skill measures together with their associated weights are provided as input to the system that in turn calculates the cumulative measures of a particular skill as the sum of the product of the score and the weight associated with it. The result of this study is useful for determination of whether the job seeker's skills are relevant to job vacancy or not and to whatextent.

C. Ranking andWeighting

Ranking and weighting provide means of how the features are weighted. These weights can come from various sources. First, they can come from the user (i.e., by self-evaluation) such that the user will give estimation of weights to the features that are selected according to the skill/ expertise levels. Second, they can also come from recommendations and social networking sites endorsements of skills of the job seeker. Third, they can be determined by testing the user on sample tests that measure the level of knowledge of the user on the skill. Forth, they can bedeterminedbythetrainings,andcertificatesobtained.These sources are then subject to further testing using the recency of the trainings, and certificates because the skill obtained by these trainings or confirmed by these certificates fade away with time [11, 12, and 13]. Thus, the more recent the attended trainings are the higher the weight of the feature. Finally, they can also be determined by the contributions the job seeker made in the form of publications, or online forums. These contributions will have more weight when they have more up-votes or morecitations. Similarly, the relevance of a skill varies with i) its age, i.e., the time span between the time of job matching and the time the skill was acquired and ii) whether the user is actively using the skill on the current job, i.e., the period between the time of job matchingandthelasttimetheskillwasusedoriii)whetherthe user has used the skill actively for longer period and has not used the skill for shorttime.

In the first case, the older the skill, the less its relevance for the job matching, and hence it will be penalized by a lesser weight as shown in the skill measure (cf. Fig. 3) because the skillmighthavefadedawaywithtime.Inthesecondcase,however,theoldertheskillthebetterwillbeitsrelevancetothevacancybecause theskillhasbeenusedforlongerperiodthatitiseasily available for use on thejob.

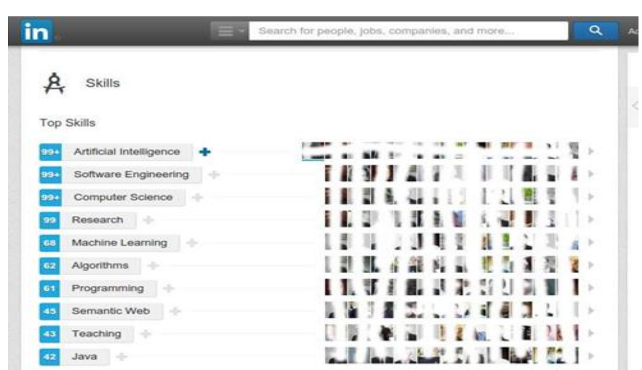

Fig. 2 LinkedIn® Endorsements of Skills [6]

A. Normalization

The values of endorsements are normalized to be within the rangebetween0and1insuchawaythattheskillthatdoesnot have any endorsement will be rated 0 and the skill with the maximum endorsement will have skill rate of 1. These ratings are local to the single job seeker as they will not be used to compare a job seeker against other job seekers. Rather they will be used to measure the degree of match between the job seeker and job vacancy for which the job seeker has a relevant skill.

B. Configuration

The template for modeling the job seeker needs to be highly configurable so that the user can add a new feature, remove a feature that the user thinks does not characterize him/her anymore. Modify the feature weight in response to updates in the skill recency or other weighting criteria.
With these weighted features we can model the user as a job seeker to determine the most-fit job vacancy to recommend according to the cumulative results from the weights of the features.

V. MATCHING OF JOB SEEKER TO VACANCY

Jobseekerrepresentationisaccomplishedusingfeaturev ectors of skills and qualifications extracted from their CVs and social network data. Likewise, vacancies are represented using feature vectors extracted and transformed form ‘required’ and ‘desired’ requirements in vacancies and enriched by data from occupational standards[3].

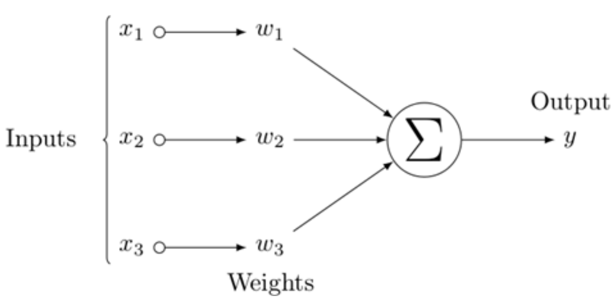

Fig. 3 Aggregating Skill Measures

Matching job seekers with vacancies is done using bi- directional matching algorithm described in [3] taking corresponding feature vectors of vacancies with that of job seekers instead of standard job descriptions. For example qualification section of job seeker information is matched with required qualification of vacancies in order to compute the similarityofcorrespondingfeatures.Similarly,job‘location’in vacancies is matched with location mention of job seeker’s ‘preferences’.

VI. RESULT AND DISCUSSION

To demonstrate the extent of usefulness of social networking data in job vacancy recommendation, we conducted an experiment and presented the preliminary results.

Data of 990 job vacancies was collected from online open access sources using a web crawler [23] (also known as a web spider or web robot) – a program or automated script which browses the World Wide Web in a methodical, automatedmanner.Vacancydatawascollectedviacrawlingfortnig htlyin the period between the first of August 2016 and thirtieth of November 2016. The two-week interval is chosen because of the lifetime of a job vacancy, i.e., the time expired vacancies are removed and new ones appear. Even when the crawling was done fortnightly, there is redundancy of the collected data because some vacancies may not expire and the crawler gets them more than once. For that reason, de-duplication is one of the key preprocessing tasks to remove all redundancies. After preprocessing was done to work only on distinct jobvacancies – thereby reducing the size of the data – 110 vacancies were used to match with job seekers.

Data of individuals from professional network that contains professional details of persons (i.e., attributes) and information produced due to interaction of these individuals on topics (i.e.,relations) is also collected from online open access sources. After cleaning noisy data and eliminating individuals that have no data in either attributes, relations or both, we obtained 65 records of persons with data about themselves and professional connections. From 65 job seekers with relevant skills for the jobs in 110 vacancy advertisements, to test the system, we utilized A/B Testing [25] which measures the difference in performance of the system when a certain parameter is included (as compared to the performance in the absence of the parameter). For example, the result shown in the confusion matrix in Table 1 is when matching job seeker with vacancy in the absence of social networking data. Whereas, the result shown in Table 2 is when matching job seeker with the same job vacancy using social networking data included.

Evaluation is done using accuracy measure as shown in (1) becauseitisanevaluationmetricthatgivesasinglemeasureof quality[24].Theaccuracymeasurewhichistheratioofcorrectmatc hes (i.e., the number of correctly matched job seekers) to total number of job vacancies to evaluates the effectiveness, such that *CM* is the number of individuals that are actually correct matching to the job under consideration, *CNM* is the numberofindividualsthathaverelevantskillsbutnotmatched, *NCM* is the number of individuals thatdo

TABLE 1. PRELIMINARY RESULT (WITHOUT SOCIALNETWORK)

Predicted → **Actual ↓**	Matched	Not Matched
Correct	7	9
Incorrect	4	45

TABLE 2. PRELIMINARY RESULT (WITH SOCIALNETWORK)

Predicted → **Actual ↓**	Matched	Not Matched
Correct	10	6
Incorrect	3	46

not have relevant skills but matched, and *NCNM* is the number of individuals that do not have relevant skills and are not matched. Then,

Accuracy = CM+NCNM /--(1)

CM+CNM+NCM+NCNM+ð

where, ð stands for the number of distinct vacancies less the sum of *CM, CNM, NCM, NCNM*, i.e., the number of vacancies that did not show up in the confusion matrices (cf. Table 1 and Table 2) which are excluded during clustering phase as described in the bidirectional matching algorithm [3].

The experiment clearly shows promising result in achieving accuracy improvement that the number of incorrect matches are reduced (from 4 to 3) while the number of correct matches are increased (from 7 to 10), i.e., the percentage accuracy due to addition of social network data increased from 47.27% to 50.91% as shown in Table 1 and Table 2 in matching job seekers to job vacancies.

Thus, inclusion of social network data in skill measurement shows consistent improvement in accuracy, i.e., 65 out of 110 job vacancies have positive improvement in matching to job seekers, as shown in Fig. 4, improving overall matching accuracy by 59.1%.

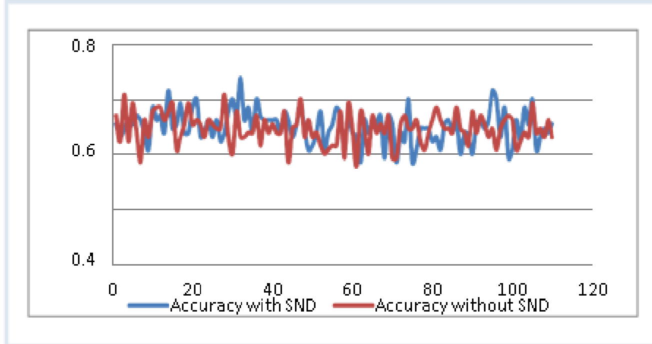

Fig. 4. Matching accuracy with and without Social Network Data (SND)

VII. RELATED WORKS

A number of studies have been carried out on job recommendation systems [14] in general and on application of social network analysis on recruitment in particular. This work is an extension of [3] and [15] which explores the study on matching job vacancies with job seekers using data from occupational standards, and job seeker's self-assessment respectively.

Studiesontheusageofsocialnetworkstofacilitaterecruit ment in the knowledge society in Bulgaria confirmed the role socialnetworksplayinsupportingrecruiterswithdemonstratedres ult of a survey of social media usage[16].

Reference [10] summarized how the online contacts can be used by recruiting managers to get more information about potential candidates for a job.

Reference [17] researched on using directed graph for endorsement deductions and ranking. They developed an algorithm that adds new weighted arcs to the digraph of endorsements based on the relation of endorsements.

Reference [18] explored theoretical foundations of social network analysis and practical application of it with particular emphasis to the information sciences. Reference [19] conducted study on job preference assessment.

Reference [20] studies the role of individual's social network one's job search behavior, the extent to which the strength of connections affect job outcomes from online social networks and compares the results with that of traditional job searching methods.

In[21],socialnetworkisusedtorecruitsubjectsfortesting and counseling through utilization of social diffusion, network stability, choice and training of network members. Only the relation (as opposed to attributes) aspect of social network is utilized in [21]. Moreover, [21] did not use social networking datatorankthesubjects,ratheronlytoreachthem.

Reference [22] investigated the benefits (and challenges thereof) of e-recruitment and the role of social network for this purpose. Reference [22] identifies companies that utilize social media in the process of recruitment to select and rank job applicants. In addition, [22] explored the legal issues that are related to social network in recruitment and selection. While discussing the benefits of e-recruitment and the role of social networking in recruitment and selection of job applicant, [22] did not discuss the specific benefits of social network data in the integrated e-recruitment systems, i.e., the attribute and relation data of job seekers is not integrated into the e- recruitment system, rather recruiters used social networking sites to get more information about the applicants.

According to [23] we can obtain 'attribute' and 'relational' data from social networks. Attributes refer to the data about attitudes, opinions, and other qualities that characterize the participants in the relationship while relation refers to the contacts, connections and ties to show structural topology and attachments of individuals in the network.

Although a number of literatures have discussed clear evidences about the importance of social networking data to get more insight about a job seeker, employers have not made extensive use of it, apart from mere exploration of jobseeker profile on social networking sites. This is partly because little was done on integration of social networking data in online job matching and recruitment systems.

In this research we used both types of social networking data described in [23], i.e., attributes and relations, to model characteristics of a job seeker and the strength of connection he/she has in the network with other peers, respectively.

VIII. CONCLUSION AND OUTLOOK

The paper discusses the general framework of online recruitment system and provides the role of social networking data as one component of the system to improve the accuracy of skill measurement in jobseeker modeling. It also explored the importance of social networking data in the jobseeker-to-vacancy matching in online recruitment systems and presents methods of measuring skills. Furthermore, it explored the role of relationships between users and their knowledge of the skill oftheirconnectionsindeterminingtheskilllevelstomodeljob seekers which will then be used as input to context-aware job recommendation system showing promising results of preliminaryexperiment.

Usecaseinwhichoursystemcanbeappliedincludepromotion(internal mobility), taking into account employee information stored in the human resource (HR) department such as salary information, service period, supervisor evaluation, etc. Using skillanalysisforstaffpromotionisinterestingusecasetoapplythesystemininternallywithinanorganizationmatching.

There are a number of future works in this study. First, implementation of prototype in its entirety and its evaluation needs to be done to test its effectiveness when used in matching with job vacancies with actual usage by job seeker. Second, it is imperative to also consider qualifications and certifications in the analysis process. Finally, experiment on varied use cases is important to measure its actual usability in the real world.

IX. ACKNOWLEDGEMENT

Theauthorswouldliketoacknowledgethefinancialsupportof the Eduworks Marie Curie Initial Training Network Project (PITN-GA-2013-608311) which is part of the European Commission’s 7th FrameworkProgramme.

REFERENCES

[1] Belloni, M. Brugiavini, A., Meschi, E., and Tijdens, K.G. Measurement Error in Occupational Coding: An Analysis on Share Data,2014.

[2] Branham, L. The 7 Hidden Reasons Employees Leave: How to Recognize the Subtle Signs and Act before it’s too Late. AMACOM Div American Mgmt Assn.2012.

[3] Chala, S., Ansari, F. and Fathi, M. A Framework for Enriching Job Vacancies and Job Descriptions through Bidirectional Matching. In Proceedings of the 12th International Conference on Web Information Systems and Technologies (WEBIST2016) Volume 2, pages 219-226 ISBN: 978-989-758-186-1, SCITEPRESS – Science and Technology Publications, Lda.

[4] Facebook. [Online].http://www.facebook.com [Accessed: 08- Aug-2016].

[5] Twitter. [Online].http://www.twitter.com [Accessed: 08-Aug- 2016].

[6] LinkedIn®. [Online].http://www.linkedin.com [Accessed: 08- Aug-2016].

[7] Researchgate®. [Online]. http://www.researchgate.net [Accessed:08-Aug-2016].

[8] Academia.edu [Online].http://www.academia.edu [Accessed: 08-Aug-2016].

[9] "Sample Individual Skills Profile," Individual Skills Profile Sample. [Online].https://www.itsyourskills.com/sample/individual-skills- profile. [Accessed:08-Aug-2016].

[10] Athavaley, A. Job References You Can't Control. The Wall Street Journal, Volume 27.2007

[11] Anderson, J. R.. Learning and Memory. John Wiley New York, 2000

[12] Kluge, A. The acquisition of knowledge and skills for taskwork and teamwork to control complex technical systems: A cognitive and macroergonomics perspective. Springer.2014

[13] Kahana, M. J. and Adler, M. Note on the Power Law of Forgetting. University of Pennsylvania, unpublished note.2002

[14] Panniello, Umberto and Tuzhilin, Alexander and Gorgoglione, Michele. Comparing context-aware recommender systems in terms of accuracy and diversity. User Modeling and User- Adapted Interaction, Volume 24, Number 1-2, Pages 35-65, Springer.2014.

[15] Chala, S.A, Ansari, F., and Fathi, M. (2016). Towards Implementing Context-aware Dynamic Text Field for Web- based Data Collection, Int. Journal of Human Factors and Ergonomics [inpress]