

Social network analysis to influence career development

Eman Abu Khousa¹ · Yacine Atif² 

Received: 28 August 2016 / Accepted: 9 January 2017 / Published online: 18 February 2017
© The Author(s) 2017. This article is published with open access at Springerlink.com

Abstract Social network analysis techniques have shown a potential for influencing graduates to meet industry needs. In this paper, we propose a social-web driven solutions to bridge formal education and industry needs. The proposed career development framework utilizes social network analytics, influence diffusion algorithms and persuasive technology models along three phases: (1) career readiness to measure and visualize the general cognitive dispositions required for a successful career in the 21st Century, (2) career prediction to persuade future graduates into a desired career path by clustering learners whose career prospects are deemed similar, into a community of practice; and (3) career development to drive growth within a social network structure where social network analytics and persuasive techniques are applied to incite the adoption of desired career behaviors. The process starts by discovering behavioral features to create a cognitive profile and diagnose individual deficiencies. Then, we develop a fuzzy clustering algorithm that predicts similar patterns with controlled constraint-violations to construct a social structure for collaborative cognitive attainment. This social framework facilitates the deployment of novel influence diffusion approaches, whereby we propose a reciprocal-weighted similarity function and a triadic closure approach. In doing so, we investigate contemporary social network analytics to maximize influence diffusion across a synthesized social network. The outcome of this social computing

approach leads to a persuasive model that supports behavioral changes and developments. The performance results obtained from both analytical and experimental evaluations validate our data-driven strategy for persuasive behavioral change.

Keywords Social web · Community of practice · Big data · Learning analytics · Computational science · Fuzzy logic

1 Introduction

Higher education has traditionally been notorious for the inefficient use of data to improve the quality and the value of graduates in meeting market needs. The amount of data available to higher education sector surpasses their utilization opportunities, in the wake of big data field emergence. Managing, processing and analyzing these large volumes of data prompted the establishment of learning analytics (LA) field to improve learning processes (Ferguson 2012); and to bridge the gap between education and industry. Higher education graduates are expected to develop a set of qualities or dispositions necessary for building up their competency in embracing change, and making informed career decisions.

Developing such a career-readiness capacity requires a sustained and progressive growth of professional habits. We propose a model that nurtures these dispositions, alongside the formal academic path. Community of practice (CoP) is an alternative informal way to achieve this aim (Jakovljevic et al. 2013). CoP fosters a new form of apprenticeship as students observe and emulate mentors, while engaging in a “learning to be” cycle to master the skill of a field. The practices and norms from established practitioners in

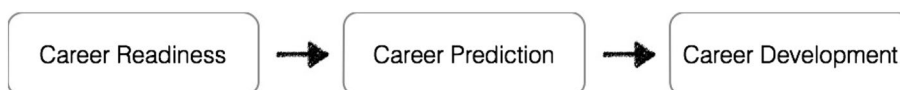
✉ Yacine Atif
yacine.atif@his.se

Eman Abu Khousa
Emanak@uaeu.ac.ae

¹ UAE University, Al Ain, UAE

² University of Skövde, Skövde, Sweden

Fig. 1 Career readiness framework



the field through process fosters the acquisition of apprenticeship experiences. In this research work, we introduce a framework to improve career readiness and enhance the career's success prospects of learners in higher education institutions. Our framework bridges the gap between higher education and industry through an online social structure made up of interconnected CoPs. This structure extends the perspective of educational institutions and develop a joint effort with the industry to leverage education and workforce development. The proposed approach also provides indicators and means for institutions to intervene in order to positively affect career readiness. It also incorporates a persuasive model to incite guide learners adopt career paths that are currently demanded by the industry. Our framework incorporates three major modules as illustrated in Fig. 1: (1) career readiness, to assert professional dispositions, (2) career prediction to identify a domain of employment, and (3) career development to persuade learners sustaining the identified career path.

After the initial formation of CoPs, we build a classification model, which considers each CoP as a class and each learner belonging to a CoP as a member of the corresponding class. The learners' profiles and other attributes, along with the CoP membership information are then used as training data for a classifier. The classification model is highlighted in grey in Fig. 2 as it is outside the scope of this paper. Career development phase aims at constructing CoP-network with a dense community structure to connect learners who belong to the same CoP based on their social similarity. Learners in such networks can benefit from the power of collaborative knowledge towards achieving common career goals and objectives. This module uses social network analytics (SNA) to enhance the properties of ties and how learners develop and maintain these relationships to support their career development. SNA metrics are used to evaluate social influence propagation in the CoP-network; and to identify most influential learners based on their structural position in the network. Those learners are then employed as "persuasive agents" to drive behavioral changes and persuade learners to adopt career paths demanded by the industry.

As a case study consider the global transformational trends induced by data sciences shift. Employers want to influence this shift within education to meet foreseeable workforce needs across derivative disciplines. In Middle East market particularly, computer science and information technology (IT) graduation and post-graduation qualifications, are highly emphasized when looking for

suitable candidates (Bayt 2015), even-though recruiters want to point-out specific data science areas (including data modeller, data analyst and business intelligence developer), and related derivatives (Indeed 2015; Sealey-Morris 2014). It is interesting to observe how the title "analyst" or the skill-term "analytics" appear almost in every job post. One more interesting observation is the emergence of new IT job titles: social medial director and chief data officer (Sealey-Morris 2014). Using this type of input from industry market, we analyze typical higher-educational programs to map our CoP-based apprenticeship model. In our College of IT at UAE University, learners are expected to complete formal education requirements while preparing for a career-in-demand using our proposed apprenticeship model. A Bachelor degree in Computer Science for example includes a range of courses such as programming, operating systems, database systems etc., which are generic qualifications. Our model complements this scenario by immersing learners into areas of professional focus that shape market demands for next years (upon students graduation), at an early stage of their study journey, as illustrated in Fig. 3.

At the first phase, and as an initial stream of generated data sources, College learners fill-out a career profile to capture initial information about current competencies, qualifications, skills; and also list-down potential career

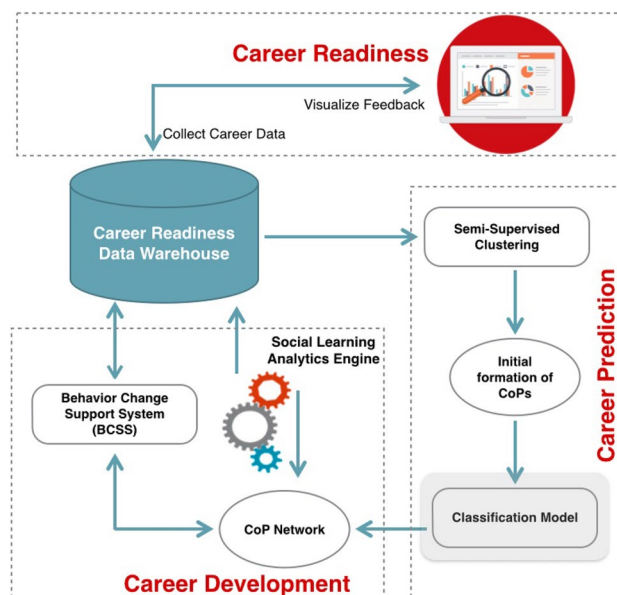
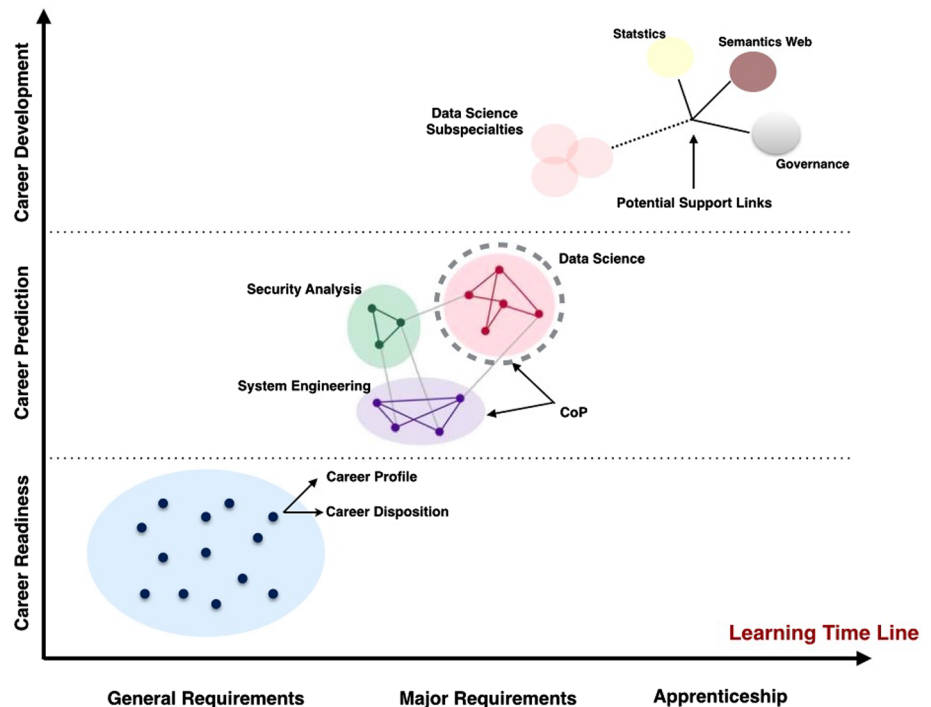


Fig. 2 A high level implementation model of proposed solutions

Fig. 3 The CoP-apprenticeship model: IT scenario



interests. Learners also complete the career readiness instrument in order to evaluate their career dispositions and identify the gaps towards prescribed career aspirations. The provided instrument results in a storehouse view of career dispositions relevant to the general IT practitioner through an integrated portal which captures self-stated learning experiences and converts them into analytical diagnosis that root-out deficiencies and prescribe improvement recommendations.

In the subsequent step, the career prediction module allocates CoP memberships to connect learners who share similar career interests. These interests are analyzed against the market needs in order to map CoP construct to a certain career path (or overlapped paths) that leads to highly-demanded careers (i.e. data science). Each CoP matches a particular professional career and is linked to competencies of that profession. These competencies are extracted from job descriptions and analyzed against career profiles to evaluate the competency gaps and identify priorities for learning development. A similar motivation was suggested in Gaeta et al. (2016) where training activities inside CoP are integrated within the educational-career path. These activities are designed to extend learning objectives beyond formal education and more into employment-market along a personalized learning context to achieve those objectives. Each CoP is assigned academic mentors and/or industry professionals to drive the community towards attaining career-oriented learning objectives. The mentor and the initial members create the structure and operational processes of their community.

This includes sharing experiences and learning resources to sustain the career development within the community in a collaborative effort. To support learning relationships, we reinforce learners' ties via social similarities while participating in CoP-based social network. The members' constant interactions within and outside CoP create a dynamic knowledge container and a repertoire of shared practices and experiences. Our dynamic CoP life-cycle evolves to more focused subdomains by forming a sub-CoP of learners who are "influenced" to steer their initial major into sub disciplines (i.e. data modeler, data analyst, business intelligent analysts, data chief officer), forming new CoPs as illustrated further in Fig. 3. As the community thrives, learners develop domain practices, that lead to further recognize and reach-out other potential members (outside their CoP) who may be driven to migrate to their CoPs, for e.g. statistics, semantic web, data governance.

The remaining sections of this paper are organized as follows. First, in Sect. 2, we provide some background and discuss some related works in the context of career disposition. Then, in Sect. 3, we discuss learning analytics for career dispositions, and reveal in Sect. 4 our predictive analytics approach to CoP formation. Section 5 extends this structure into a broader social network where we integrate our influence propagation approach in the context of our persuasive learning framework. Section 6 shows performance evaluation indicators and results of the approaches discussed in the paper and Sect. 7 concludes the paper with a summary of contributions and suggested future works.

2 Background and related work

2.1 Personality traits and career success

Several research studies relate personality attributes to aspects of career success (Boudreau et al. 2001; Sutin et al. 2009). In particular, the literature on career success identifies five categories of personal traits (termed Big Five personality factors) that influence the general mental ability for career success (McCrae and Costa 1991; Judge et al. 1999; McCrae and John 1998; Seibert and Kraimer 2001; Zimmerman 2008). The contributions of the established Big Five dimensions of personality to career success have been examined by increasing number of studies in the field of human development psychology. For example, conscientiousness was found to be the most consistent predictor of career advancement (Howard and Bray 1990); and that it is related positively to job performance (Salgado 1997). Extraversion was also found to positively affect job and life satisfaction (McCrae and Costa 1991), as well as career promotions (Barrick et al. 1993; Salgado 1997). On the other hand, a negative relationship was correlated between high neuroticism and negative job performance (Salgado 1997), and hence subsequent career dissatisfaction (Furnham and Zacherl 1986).

The need for certain personality traits of career success has been addressed in our research to develop standard criteria that emerge from proactive behaviors in order to instill positive habits or dispositions. In doing so, we noticed a substantial gap between career success and lifelong learning literatures, although 21st century successful professionals build their career path following a sustained lifelong learning aptitude known as learning power (Deakin Crick 2007). These dispositional traits are most likely dominated by the earliest experiences acquired in the family, school and college (Bourdieu 1977; Brunello and Schlotter 2011). Earlier beliefs that personality traits that shape individuals' lifetime are a matter of fate do not apply (Roberts and Mroczek 2008; McAdams and Olson 2010). Therefore, the development of traits that support career success in higher education institutions—or even earlier in schools—appears as prominent to reshape students' prospective traits to manage their future professional growth successfully.

2.2 Learning in social communities of practice (CoPs)

Community of practice (CoPs) is defined as “computer-mediated discussion-form focused on *problems of practice that enable individuals to exchange advises and ideas with others based on common interests*” (Wasko and Faraj 2005). It is an online CoP, although we refer to our computerized method as a CoP in general. CoPs play a central role in knowledge management (KM) strategies and

collaborative learning, and are perceived as an effective mechanism for knowledge creation and exchange. Knowledge in its turn is central to formal education and professional practice. KM literature differentiates between knowledge sharing and knowledge exchange, whereby knowledge sharing can occur in a one-way broadcast form, whereas knowledge exchange occurs at dyadic level and indicates a reciprocal relationship (Pan et al. 2015), hence the social ties within and outside CoPs. Knowledge is also often divided into two distinct entities: explicit knowledge (knowing that) which refers to possession of information and facts; and tacit knowledge (knowing how) which refers to procedural and application form of knowledge (Curran et al. 2009). CoPs are particularly contributing to the creation and exchange of tacit knowledge, whereas formal education tends to focus more on explicit knowledge.

The theory of CoPs lies at the intersection of knowledge transfer and learning processes; and it has become more widespread in higher education due to the benefits derived from collaborative generation of knowledge and cooperative learning activities within and outside the classroom (Jakovljevic et al. 2013; Jawitz 2007; Tight 2004; Mandl et al. 1996). Johnson noted “*the learning evolved form these communities is collaborative, in which the collaborative knowledge of the community is greater than any individual knowledge*” (Johnson 2001). It is also argued that individual learning is enhanced through engagement with others enabling the extension of the individuals' capability to a new and higher level (Vygotsky 1980). Learning in CoPs is recognized as “situated learning” that is defined as knowing how to be in practice rather than knowing about the practice (Lave and Wenger 1991; Wenger 1999; Brown et al. 1989). This involves the process of formation of the individual identity as becoming a member of the community and participate in knowledge development. Supported with the sense of connectedness, knowledge development within CoPs can be continuous, and fluid in a cyclic pattern (Koliba and Gajda 2009). Learning in CoPs also occurs within the context of the cyclical process of DDAE: dialogue, decision-making, action and evaluation (Mindich 2011). Examining DDEA dynamics in CoPs draw a clear link between the group's capacity to evaluate its current practices and learn through this evaluation by talking about it (dialogue), make decisions based on this discussion, and subsequently implement this decision into action (Koliba and Gajda 2009; Gajda and Koliba 2008). Learners participating in CoPs engage and interact to gain knowledge and skills from community members, partly with those who are positioned as mentors or field experts. This view reiterates the apprenticeship model to complement learning in higher education (Warhurst 2003). The demands to extend abstract bodies of knowledge taught in formal higher-education with CoP-based learning is dictated by the need to equip

learners with skills that make them immediately ready to embrace real-world problems upon graduation in order to reduce the current education-market disparities (Jakovljevic et al. 2013; Tynjälä et al. 2003).

Contemporary social networks (SNs) can be employed to build online CoPs within higher education context (Gunawardena et al. 2009; Zhang et al. 2010). Recent research indicates a substantial value of SNs in strengthening student-to-student interactions, enhancing student social engagements, and building campus communities toward improving student learning (Davis III et al. 2012). Facebook, one of the most powerful SNs, enhances the connectedness as well as social learning in higher-education settings (Baran 2010; Qureshi et al. 2015; Selwyn 2009); and information-sharing for knowledge development and innovation (O'Brien and Glowatz 2013). However, seldom has research tapped into the emergence and cultivation of a social structure that emphasizes learners and the network in which this learner navigates in support of his or her professional career practices (Gray et al. 2010).

3 Learning analytics for career dispositions attainment

3.1 Career dispositions

Big Five model identifies five categories of personal traits that influence career success; while lifelong learning dispositions make up the individuals capacity for developing lifelong learning attributes. Career dispositions emerge as the joint set of attitudes and generic skills that engenders professional behaviors; and influences the ability to adapt and respond to changing work situations and environments. They describe the natural tendencies, mind state and preparations of each individual towards a professional practice. We model career dispositions as a 6-dimensional construct captured from the Big Five and Learning Power to comprise (See Fig. 4): (1) *Openness to challenge (OC)* “the degree to which an individual open to new ideas and experiences”; (2) *Critical Thinking (CT)* “the degree to which an individual is investigative, attentive reader/listener, inquisitive, analytical and an evidence-based decision-maker”; (3) *Resilience (R)* “the degree to which an individual is conscientious, determined, assertive and achievement-oriented”; (4) *Learning Relationships (LR)* “the degree to which an individual is cooperative, expressive, agreeable and social oriented”; (5) *Responsibility for Learning (RL)* “the degree to which an individual is dependable, autonomous, motivated, organized and punctual”; and (6) *Creativity (C)* “the degree, to which an individual is intellectual, imaginative, adventuresome, curious and original”.

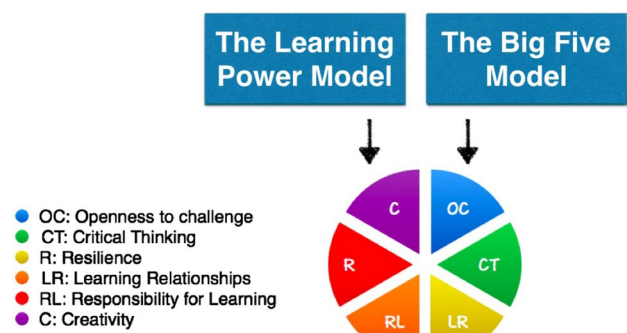


Fig. 4 Career dispositions

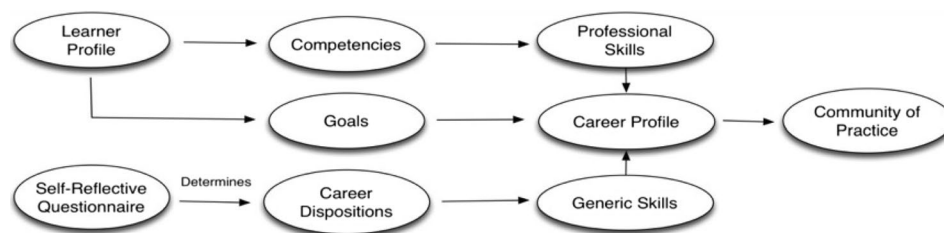
In order to measure career disposition data values, we developed and validated a self evaluation report that is conceptually underpinned by constructs from career success and lifelong learning dispositions literature following our blended approach to career dispositions (Khousa and Atif 2014). Based on this scale, data is routinely captured through the proposed instrument to elicit quantitative reflections of career disposition via multidimensional infrastructure as we will present next.

3.2 Multidimensional career profile

We extract a structure to determine a career readiness construct, labelled career profile (Fig. 5). This construct is designed as a standard mean to collect and access information about learners while they are moving towards a predestined career path. Career profile augments an existing IEEE learner information package (LIP) standard¹ to capture learning data as well as career indicators. Our proposed construct of career profile is structured into three main categories aimed at predicting and assisting learners with their career development throughout their formal education. We use LIP-defined interests, competency, and goal, categories to specify career interests, domain-related qualifications, and long term career objectives of individual learners. We differentiate two types of interests: career interests and social interests. We also introduce a new category labeled professional as a slot for career dispositions ratings and other generic attributes pertaining to career readiness. The multidimensional data attributes reflecting the professional aptitude, career prospects and dispositions of a learner are used to detect a CoP, where members share knowledge, experience and passion for a predicted practice to build capabilities and maintain momentum.

¹ <http://www.imslobal.org/profiles/index.html>.

Fig. 5 Career profile structure



The multidimensionality conceptual view allows better understanding and analysis of data in terms of the subjects (facts) and the different and range of views where a subject can be analyzed from (dimensions). Each dimension is associated with hierarchal levels which contain consolidated data or descriptors; while a fact contains measures (also known as variables or metrics). One fact and several dimensions to analyze it define the data cube or a simple aggregation function (Franconi and Kamble 2004).

4 A predictive analytics approach to community of practice (CoP)

We propose a semi-supervised clustering algorithm as a predictive analytic method to assign learners into common virtual CoPs according to their career interests. The proposed method uses a fuzzy-logic objective function to address issues pertaining to overlapping domains of career interests.

4.1 Semi-supervised clustering analysis

Clustering methods that utilize any side information are said to be operating in a semi-supervised mode (Chapelle et al. 2006). One of the most common methods to specify the side information are in the forms of: (a) pairwise constraints where set of must-link and cannot-link specifies weather point pair connected by the constraint belong or not belong to the same cluster (Davidson and Ravi 2005); and (b) seeding, where some labeled data is used along with large amount of unlabeled data for better clustering (Gu and Lu 2012). Must-link denoted by $c = (x, y)$ and cannot-link denoted by $c \neq (x, y)$, meaning that two instance x and y must be in the same cluster or cannot be in the same cluster, respectively.

4.2 Proposed fuzzy pairwise-constraints K-means (FPKM)

We propose a clustering-based career prediction model that analyzes data from Career Profile in order to predict a hypothetical career practice and bring learners with similar career patterns together into a common cluster. This

process leads to a social structure made up of CoPs that are identified to specifically respond to imminent industrial needs. To solve the cold-start problem of CoP construction, we use the career readiness data warehouse as a source for initializing groups (or clusters) of learners and denote each such cluster as a CoP. In order to conduct this initial grouping process, we apply a semi-supervised clustering technique that brings a seed set of learners into an initial set of CoPs. The seed set consists of learners who achieved high scores in career disposition values that are above a given parameter threshold. There is typically at least one seed member in each cluster (CoP) for which his/her career profile matches the definition suggested by the career ontology that yielded the CoP. The rationale of privileging highly ranked learners in their career dispositions to create dedicated CoPs is driven by the prospects to sustain CoPs. From this initial stage, we infer the use of career disposition values only to provide seed set of new CoPs (including the initial ones).

One challenging problem occurs when and whether a violation of the link constraint should be penalized. In traditional semi-supervised clustering algorithms, a violation of the link constraint is always penalized. Now, as we allow the instances to be associated with multiple labels, a constraint can be violated legitimately. Thus, we re-designed the penalty function of constraints K-means algorithms to allow fuzzy labeling and to estimate if a constraint violation could be legitimate or not. Accordingly, we develop the fuzzy pairwise-constraints K-means (FCKM) algorithm that is presented and evaluated on simulated data in our previous publications (AbuKhousa and Atif 2016; Khousa et al. 2015). The main objective of the FCKM algorithm is to assign learners of overlapped interests to multiple clusters or CoPs. In this paper, we stimulate data from a real world scenario to apply our clustering method in addition to the typically benchmarked data sets.

5 Social learning analytics for persuasive learning

Social learning analytics (SLA) is a distinctive subset of LA, which highlights the social perspective of learning. SLA draws on the significant educational research work evidencing that new skills and ideas are developed and

passed on through interactions and collaboration; and that learning cannot be understood without reference to context. As a group of learners engaged in a joint activity, their success is related a combination of individual knowledge and skills, environment, use of tools and ability to work together (Wells and Claxton 2008). In this paper, we mainly focus on social network analysis (SNA) to analyze the social influence power in CoP-network to achieve persuasive learning. we evaluate social structural regularities that influence individuals' behaviors and actions (Otte and Rousseau 2002).

Social influence is defined as a “*change in an individual's thoughts, feelings, attitudes, or behaviors that results from interaction with another individual or a group*” (Rashotte 2007). Users influence each others to form different networking structures or crowds which directly, explicitly or implicitly, promote information diffusion. Given a SN represented by a graph $G = (V, E)$ with the set of vertices V representing individuals and the set of edges E representing links among individuals, if any node $v_1 \in V$, replicates the action of another node $v_2 \in V$, we may assume that v_2 has influence on v_1 . When the influential node performs an action, nodes connected to it will replicate the action starting information propagation. The influence of a node on other connected nodes can be due to some external factors like trust; the popularity of the information or action or to the prestige or celebrity of the influential node. The relational structure in the network (i.e. role and structural position of individuals) also plays an important part in determining how influential an individual is, in this network (Friedkin 2006; Hu et al. 2015). Therefore, understanding the structure of SNs provides important further insights into how individuals influence each other's behavior.

A very important task for the maximization of diffusion in SNs is to identify the influential nodes or the adopters who can exploit social network effects. Researchers in social influence target to identify a small set of influential individuals (referred to as a seed set) who can ideally maximize the influence across the entire network in minimal time (Fu et al. 2014). The problem of influence maximization can be expressed as follows: “*given a network with influence estimates, how to select an initial set of k users such that they eventually influence the largest number of users in the social network*” (Goyal et al. 2010). Given a SN, a positive integer k , and information diffusion model, the goal is to find such a target set A_k^* of k nodes that maximizes the expected number of adopters of the information if A_k^* initially adopts it. The expected number of nodes influenced by a target set is referred to as its influence degree, and this combinatorial optimization problem is called the influence maximization problem of size k (Kimura et al. 2010).

On similar research direction, persuasive technology emerges as a recent proposal to alter the mindset, attitudes and behaviors of individuals through technologies which create opportunities for persuasive interactions such as those enabled by Web and SNs (Atkinson 2006). The main process used in the psychology of public persuasion involves: (1) collecting facts and opinions of the public and the object of interest; (2) applying diagnostic procedures and statistics to interpret this collected data, and (3) applying various techniques of persuasion to guide the target group towards adopting desired ideas or behaviors (Bernays 1928). In the persuasion paradigm, influence is presented as a detailed argumentation delivered to individual recipients which impact is limited to minimal social interactions. Social influence—in contrast—includes simple information (e.g. “Likes” in FaceBook) about the source's behavior in the network, and delivered across wider social interactions (Wood 2000). A key construct for research in the filed of persuasive technology in related to public persuasion is the *Behavioral Change Support Systems (BCSSs)* defined as an “*information system designed to form, alter or reinforce attitudes, behaviors or an act of complying without using deception, coercion or inducements*” (Oinas-Kukkonen 2010). By definition, BCSS may utilize either computer-human persuasion or computer-mediated persuasion to achieve three outcomes: (1) reinforcement of current attitudes or behaviors; (2) changing an individual's response to an issue; and (3) shaping attitudes and behaviors by formulation of a pattern for a situation when one does not exist beforehand. Computer-human persuasion utilizes some patterns of interaction similar to social communication, whereas computer-mediated persuasion means that individuals are persuading others through computers (e.g. discussion forums, instant messages, blogs, or SNs) (Oinas-Kukkonen and Harjumaa 2009).

In our work, we first construct a CoP-network of a dense community structures; then we apply SNA to select the most influential individuals in this network to be hired as “persuasive agents” to support desired career adoption in higher education.

5.1 Reciprocal-weighted euclidean to construct CoP-network

We propose a reciprocal-weighted euclidean similarity function (*RWD*), which is inspired by the hybrid weighted euclidean function and existing research works on matching user profiles of a network (Carbonell et al. 2014; Raad et al. 2010), to consider self-assigned weights and position orders to profile attributes, when determining similarity measure across profiles. Let L and U be two learner profiles represented by n -dimensional attribute vector, $L = (l_1, l_2, \dots, l_n)$ and $U = (u_1, u_2, \dots, u_n)$ depicting n

pre-established measurements made associated with the learner from n attributes, respectively A_1, A_2, \dots, A_n which represent common social interests. RWD calculates similarity or (the match score used later) of L and U as follows:

$$RWD(L, U) = \sqrt{\sum_i^n (rw_i l_i - rw_i u_i)^2}$$

Where $rw(l) = (w_1, w_2, \dots, w_n)$ is the weighting vector assigned to learner vector L ; and $rw(u) = (w_1, w_2, \dots, w_n)$ is the weighting vector assigned to learner vector U . Using hierarchies as suggested in Carbonell et al. (2014) is the simplest and most convenient method to allow learners to assign order and weight to their attributes by including it in their Career Profile then measure the attributes or interests of a given learner using RWD . In such representation, interests may be associated with an interest level, such as a numeric or graded rating that represents the strength or weakness of the learner's interest in a particular topic. Our proposed reciprocal-weighted interest accumulation networking (RWIAN) algorithm (see Algorithm 1) constructs a CoP-network that employs this profile structure to generate vector attributes of learners within a common CoP, resulting in an overall CoP-network. This process calculates the social similarity using RWD function to generate the similarity matrix (SM) in order to establish (weighted) links between learners with highest similarity.

similar attributes are more often transitive than weak ties. Similarity and transitivity together lead to the formation of cliques (fully connected clusters), which enhance the network structure and community detection as internal edges of community likely to form cliques while inter-community edges unlikely to form cliques. Thus, we propose a triadic-closure approach based on social interests and career interests as presented in Algorithm 2 to enhance the constructed (RWIAN) social networks among learners. In our research work scope, we aim at introducing a systematic approach to build dense CoP-based SNs in which learners can benefit from the power of collaborative knowledge towards achieving career common goals and objectives. In our proposed approach, we filter the initially constructed CoP-network by removing edges of first similarity score below a given threshold ($S \geq t$), and use a second similarity measure as a decisive metric to form the triadic tie if it holds, and only if l and v nodes have strong common interests that are different than the initial interests used to construct the network. For example, if we use Career Interest as the first similarity measure ($S1 = Career - Interests$) to construct the (RWIAN) network: $G_w = (V, E)$, where V is a set of learners and $E \subset V \times V$ is a set of relationships connecting learners of highest $S1$. Three users (l, u, v), where links between l and u ($e_{l,u} = 1$) and between u and v ($e_{u,v} = 1$) exist, form a candidate open triad $OT(l, u, v)$. We

Algorithm 1 Reciprocal-Weighted Interest Accumulation Networking Algorithm (RWIAN)

Input: A set of learners: $L = (l_1, l_2, \dots, l_n)$; t : Similarity threshold

Output: SM (similarity matrix); G_w : weighted directed graph where: nodes are users L & the edges weights are calculated based on n -attributes RWD similarity; Q : Modularity value

Begin

//initialize the similarity matrix(MS) & weighted graph matrix (G_w) to zeros

$SM = 0$

$G_w = 0$

//For every l in L ; find the highest similar learners from L/l by calculating RWD match score (S)

For every $l, u \in L$

 Calculate $S(l, u) = RWD(l, u)$

$SM(l, u) = S(l, u)$

 If $SM(l, u) \geq t$

 //Establish a link between learner l and u ; and assign a weight (w) to the link using S

$G_w(l, u) = s(l, u)$

 end for

Apply CNM algorithm on G_w and generate Q

End

5.2 Triadic closure to enhance community structure in CoP-network

Transitivity of a relation means that when there is a tie between l and u in a network, and also between u and v nodes within the same network, then there is also a tie between l and v . Strong ties that connect individuals with

then use Social Interests as the second similarity measure ($S2 = Social - Interests$), to study whether the $OT(l, u, v)$ will (or not) become a closed triad $CT(l, u, v)$ to establish a link between l and v ($e_{l,v} = 1$). The strength of the newly established link $e_{l,v}$ is then a result of a combination of $S1$ and $S2$.

Algorithm 2 Triadic Closure for Enhanced Community Detection (TC_CNM)

```

Input: Graph  $G$  constructed using  $S1 = Career - Interests$ ; ; Similarity Matrix:  $SM$  of  $S2 = Social - Interests$  ;
Similarity Threshold:  $t$ 
Output: Graph  $G'$  with enhanced community structure
Begin:
For each node  $n$  in  $G$ 
   $neighbors = getNeighbors(n, distance = 2)$ 
  // getNeighbors: gets all the neighbors of a node at given distance, which in this example is 2
  For each  $neighborNode$  in  $neighbors$ 
    if  $edgeExists(n, neighborNode) = False \ \& \ SM[n, neighborNode] > t$ 
       $createEdge(G, n, neighborNode)$ 
  // edgeExists: returns a "True", if there is an edge between given nodes otherwise it returns "False"
  // createEdge: creates a new edge between given nodes and adds it to Graph  $G$ 
  End For
End For
Apply CNM algorithm on  $G'$  and generate  $Q$ 
End
    
```

5.3 Persuasive approach for behavioral adoption

In this section, we describe the underlying structure of our proposed persuasive approach for behavioral change support system for career-adoption (BCSS-CA). We propose to start the persuasion process by identifying the most influential individuals in the CoP-network to be hired as persuasive agents. Persuasive agents be will treated by different persuasive strategies according to their “persuasive profile” to get them to adopt the desired career path. Pervasive agents are then expected to leverage their powerful structure position in the CoP-network to influence other learners to adopt same career paths as required by the persuasion source (i.e. human resource planning authority , recruitment agency).

Our proposed BCSS-CA model aims to stimulate persuasive agents to adopt certain careers based on the recommendations of local human resource authorities that oversee the market needs. The model is a part of career development interactive process which continues to monitor, analyze, and process career-related information; and periodically provides feedback to learners to keep them up-to-data about career opportunities and the special skills set required in the areas of career interests. The designated authorities target first to hire this group of “persuasive agents” to help in influencing other learners to adopt the desired career path. The performance of persuasive agents to get more learners adopting the desired career is monitored by the persuaders in order to provide frequent feedback. This is in alignment to the original process of developing the tacit knowledge and professional skills required for this career. How to stimulate key users to be persuasive agents is discussed next.

5.4 Persuasive agents

Identifying persuasive agents involves creating the “personalized persuasive profile” to develop an understanding of

individual behavior in response to persuasion; and so use suitable means to persuade each agent into adopting the desired career path. They will then naturally influence others in their circle to follow. Typically, the persuasive profile will include measures of “persuasive susceptibility” ; i.e. collections of expected effects of different influence strategies or principles for a specific individual such as authority, liking, and reciprocity. Having the persuasive profile of each agent contributes also in deciding the right motivator, right trigger and right time to persuade him or her. Persuasive agents can then be categorized into different behavioral modalities for the source to design tailored persuasive techniques or messages.

The Fogg behavioral model (FBM) (Fogg 2009) states that for a target behavior to happen, an individual must have sufficient motivation, sufficient ability, and an effective trigger. All three factors must be present at the same time for the behavior to occur. Most specifically, as individual has increased motivation and increased ability, the more likely it is that he or she will perform the target

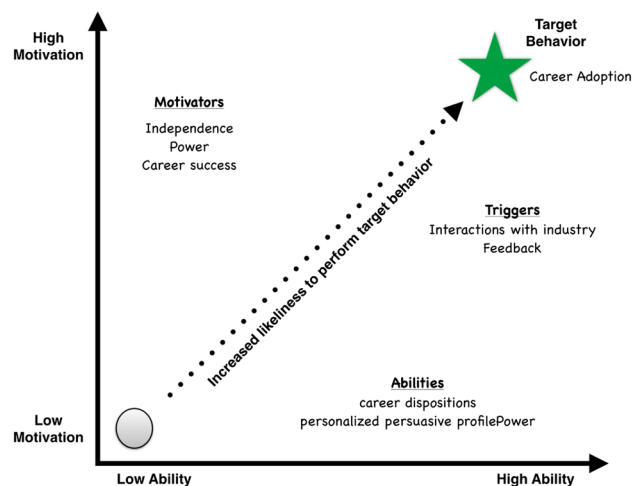


Fig. 6 FBM with persuasive agents

behavior. The FBM model provides us with the insights to use the learners of highest career dispositions as the primary “persuasive agents” as they accordingly will have the required levels of motivation and ability to adopt the target career. Then, based on their structural position in the CoP-network, they will be able to influence other individuals—who should also have some non-zero-level of both motivation and ability—to follow them. As illustrated in Fig. 6, we apply FBM in our design approach by identifying motivators; abilities and triggers as follows:

- *Motivators* Persuasive agents are selected so that they already have the strongest desire to pursue a successful career path. Their structural position in the network fulfills their need to feel they are unique (independence) and to have influence on others (power).
- *Abilities* We differentiate two set of abilities for our persuasive agents: (a) career dispositions; (b) personalized persuasive profile, and (c) powerful network structure. We need the persuasive agents to be able to respond and adopt to the change in career needs as soon as it is delivered from the persuaders. On the other hand, personalized persuasive profile combines the insights of personal traits and persuasion principles to personalize the persuasive intervention in a way that the messages, timing, interfaces, the persuasive strategies and other factors of the BCSS are tailored to one specific individual or agent.
- *Triggers* We identify three types of triggers: (a) interaction with industry, (b) feedback, and (c) recommendations and suggestions. Interactions with industry either via CoP-networks or through industrial insights presented in an analytics dashboard serve as triggers to encourage individuals to adopt future careers when they

most in demands. Another feasible approach to prompting behavioral changes in adopting career paths is to monitor career adoption development; and provide real-time feedback to the individuals. We may enhance the strength of feedback sing instantly processed and visualized results driven from the CoPs monitoring mechanisms. As for recommendations and suggestions, they may come as direct messages from human resources authorities or recruitment agencies.

5.5 SNA and influence diffusion

In our proposed method to identify persuasive agents, we employ social mining techniques in order to select initial seed nodes effectively so as to maximize influence diffusion in CoP-network. For comparative analysis, we use different SNA metrics ($sna - d$) for seeds selection. Influence diffusion models require as input, initial seed nodes (k) which are considered to be influenced (have performed action a already) at the start of the experiment. Algorithm 3 defines the implemented algorithm to mine influence in CoP-network using various selection methods to select k . In Algorithm 3, G is a graph representing the CoP social network, consisting of V nodes and E edges, where $sna - d$ is the structure position of a node with respect to the five measures (degree, betweenness, closeness, coreness, eccentricity) (Scott 2012). First, we initiate seed set S and its influence I_S . The while loop stores a set of nodes with highest $sna - d$ in S ; this set is used to start influence spread process. Next, we execute the influence propagation processes using diffusion models. The maximum influence from S , i.e. I_S is returned.

Algorithm 3 Influence Mining using Various SNA-based Seed Selection Methods

Input
 $G : (V, E, sna - d)$, k : the number of target influential nodes

Output
 Seed set (S) and its influence (I_S)

Method
 $S \leftarrow NULL$
 $I_S \leftarrow NULL$
 while $|S| < k$
 $S \leftarrow S \cup \{v | (sna - d)_v = \max(sna - d); v \in V - S\}$
 end while
 while $V \neq NULL$ do
 $A \leftarrow NULL$
 $V \leftarrow V - S$
 for all $v \in V$ do
 apply Influence diffusion models
 end for
 $S \leftarrow S \cup A$
 $I_S = I_S + |A|$
 end while
 return S
 return I_S

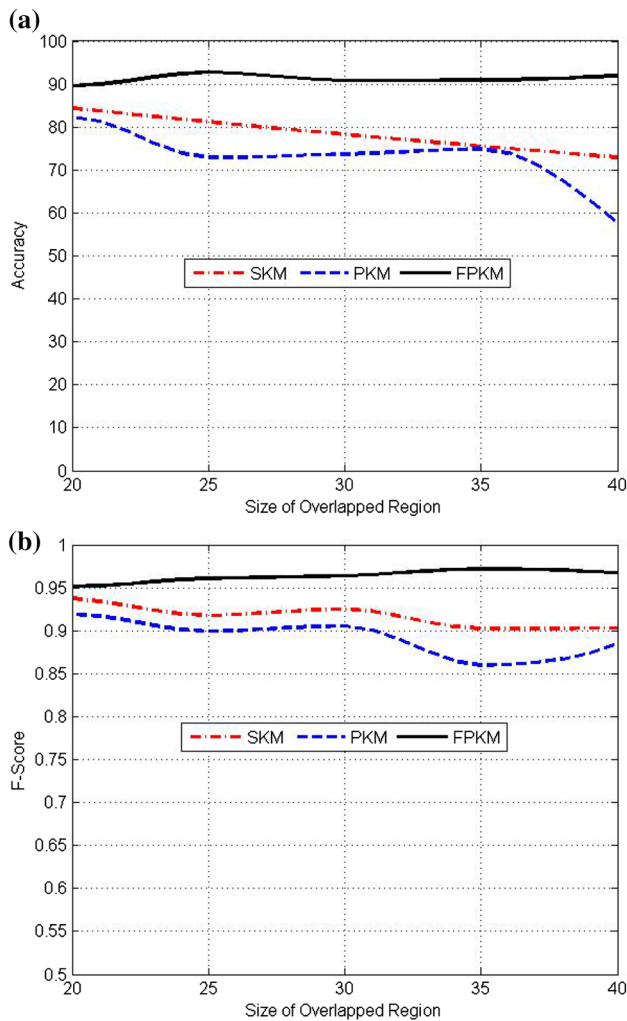


Fig. 7 The performance of three clustering methods on IRIS of different overlapped region size (Number of seeds = 5)

6 Performance evaluation

6.1 Fuzzy semi-supervised clustering to predict CoP

We used MATLAB environment to design two experiments utilizing: (1) IRIS benchmarked data set that is overlapped by nature (Pastizzo et al. 2002); and (2) multidimensional data set simulated from career-related real world scenario. Each experiment consists of several rounds to test two main varying parameters: (1) number of seeds; and (2) degree of overlap between clusters. We compare the results of our proposed methods along two *K*-means candidate methods: (1) Seeded *K*-means (SKM) (Basu et al. 2002); and (2) CVQE-based pairwise-constraints *K*-means algorithm (Davidson and Ravi 2005) (denoted as PKM) that allows constraints violation with certain penalties. To evaluate the performance of the

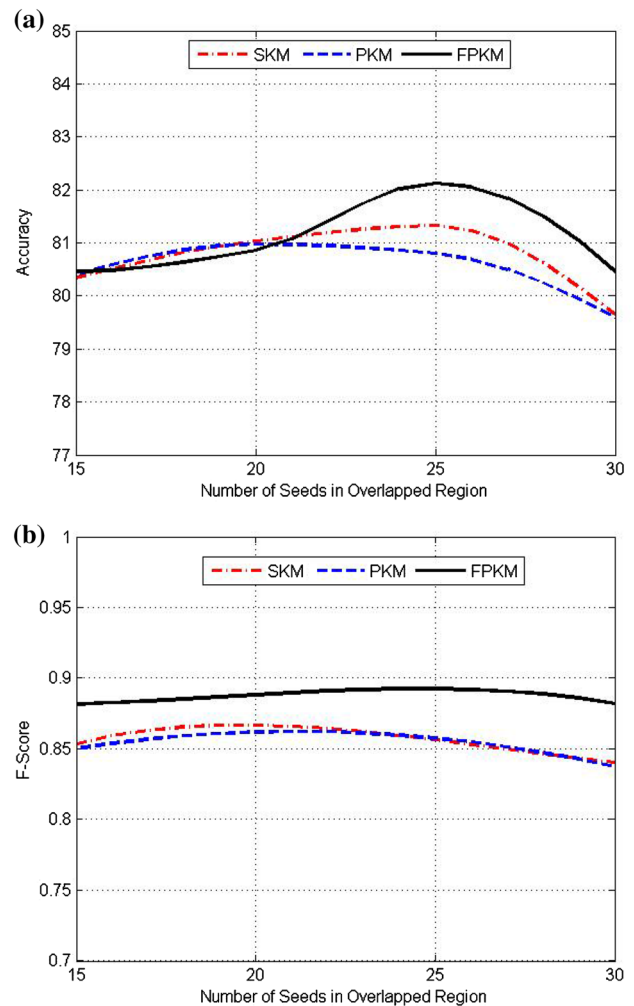


Fig. 8 The performance of three clustering methods on real world scenario of different number of seeds in overlapped region

clustering algorithms, we employ two major external metrics that utilize a priori knowledge about the classification information: (1) total accuracy of clustering to measure the extent to which each cluster contains the correct objects from the corresponding ground-truth category; and (2) total F- Score to measure the effectiveness of clusters retrieval.

6.1.1 Benchmarked data

Figure 7a, b show the total accuracy and total F-Score of all methods when the size of the overlapped region increases. We can see that when the size of the overlapped region increases, in general the performance of the fuzzy algorithm increases or remain stable while the performance of the two baseline methods decrease. This indicates the effects of seeds for accurately recovering the

Table 1 Metrics obtained using absolute similarity threshold filtering

Threshold Δ	Nodes	Edges	Clusters using CNM	Modularity
0.85	100	387	6	0.777
0.80	100	627	4	0.697
0.75	100	853	4	0.627
0.70	100	1149	3	0.553
0.65	100	1492	3	0.487
0.60	100	1800	3	0.419

overlapped region. Without seeds, the clustering becomes harder when the overlapped region becomes larger, and thus the performance of the two baseline methods drops. On the contrary, with the help of seeds, the performance of our fuzzy algorithm does not drop even when the size of the overlapped region increases. Thus, seeds are critical for the clustering of the instances in the overlapped region and it is crucial to allow the seeds or constraints from the overlapped region to be associated with multiple cluster labels. This is indeed the advantage of our method over other non-fuzzy algorithms, or other membership score based fuzzy algorithms which are not able to specify the membership scores for the seeds or constraints.

6.1.2 Data simulated from real world scenario

The performance of the three algorithms for different number of seeds in overlapped region is presented in Fig. 8b showing that the total F-score of our fuzzy method is better than those of the baseline methods. This is because both precision and recall of the fuzzy algorithm increase as the number of seeds in the overlapped region increases. However, when the number of seeds is too large in the overlapped region, the performance of the fuzzy algorithm drops. Also we observe that the total F-score in general increases when the number of seeds increases, as we get more information on the overlapped region from the seeds. The performance drops when the number of seeds is too large. This might be because with too many seeds, we obtain more false positives in the overlapped region. We observed similar patterns for the total accuracy as well (see Fig. 8a).

6.2 Triadic closure to build CoP-social network

We generated a dataset (L) of a group of learners: $L = \{l_i\}_{i=1}^N$ while N is the size of L . Each learner l_i is described by a vector of n -attributes: $l_i = \{x_i\}_{i=1}^n$ that represents his interests and is used to match with other

Table 2 Metrics obtained using k-nearest neighbors

k	Nodes	Edges	Clusters using CNM	CNM_ Modularity (Q)
4	100	250	8	0.819
5	100	304	7	0.772
6	100	354	7	0.786
7	100	408	6	0.754
8	100	461	6	0.737
9	100	515	6	0.708
10	100	678	6	0.698

likeminded learners. The value of each attribute x_i is a value within a given range: $x_i \in [\min(x_i), \max(x_i)]$ and that is generated following a normal distribution. Each attribute is assigned a weight w_i that is a value within a given range: $w_i \in [\min(w_i), \max(w_i)]$ and that is also generated following a normal distribution.

We use the modularity Q as the basic evaluation metric to show the performance of our method. Modularity is a property of a network that specifies the degree of division of that network into coherent communities.

In the network setup experiment, we generate a dataset of 100 learners' vectors consisting of 15 attributes ($n = 15$) to represent hobbies/social interests and sub-interests in areas such as sports, travel, and entertainment. These vectors are used to generate a static network (Size $N = 100$) of similar profiles utilizing the 100×100 similarity matrix (SM) such that l and u are similar and thus linked in the network, only when $SM(l, u) = [S_{lu}] \geq \text{similaritythreshold}(\Delta)$, where S_{ij} indicates the similarity weight between l_i and l_j .

In the network enhancement step, we created a sparse matrix of a specified size ($n \times n$) by combining together n learners vectors. Then, we set two similarity measures ($S1, S2$). $S1$ is used to setup the initial network (G) through a set of attributes (e.g. career interests). $S2$ is further used to add edges to G using another set of attributes (for example: social interests) where these edges are enforced form triads in (G) to produce the enhanced network (G').

We employed CNM community analysis algorithm (Clauset et al. 2004) to compare community structures in the generated network. T-CNM represents community structure using our triadic closure enhancement method. Given (SM), we generated the frequency distribution of the similarity weights ($S1$); and then built several networks by using two different methods. The goodness of the generated network is evaluated using CNM, and T-CNM clustering and Modularity (Q) to assess the obtained communities.

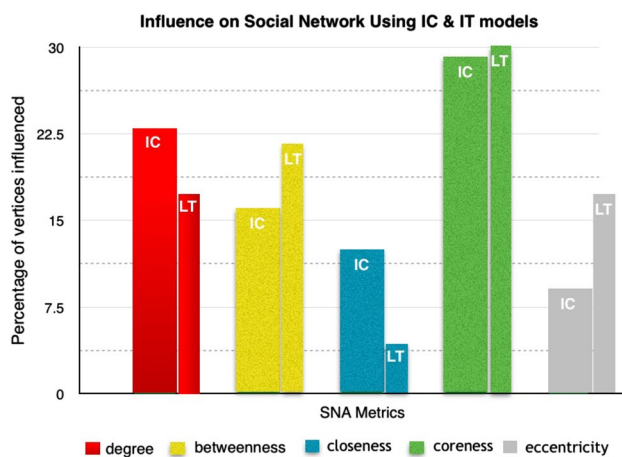


Fig. 9 Percentage of influenced vertices ($k = 5\%$)

6.2.1 Results and discussion

The edges in the original similarity matrix (generated based on career interests i.e. $S1$) are filtered using a threshold Δ to remove weak edges. Different values of Δ were used and the modularity metrics of the obtained networks are provided in Table 1. As indicated in the obtained results, Q value decreases when we include more links of poor similarity weight ($0.60 \leq S1 \leq 0.85$).

The edges in the original similarity matrix are further filtered using the best K-NN strategy (Wu et al. 2008). For every node, its nearest k neighbors are kept in the network, whereas the other edges are removed. Different values of k were used and the metrics for the obtained networks are provided in Table 2. As expected, the more neighbors (k) we include, the more chances to add weak links, resulting in more inter-communities edges than inner-communities edges and so decreasing Q value. On the other hand, a small number of (k) connects nodes which are highly similar in increasingly dense communities and, thus improving Q .

6.3 Social network analytics to identify persuasive agents

Using generated network G' , we applied five SNA methods: degree, closeness, betweenness, closeness, coreness and eccentricity for selection of nodes as initial seed for influence diffusion models. Two widely-used fundamental diffusion models are the independent cascade (IC) model (Kempe et al. 2003) and the linear threshold (LT) model (Watts 2002). In both these models, at a given timestamp, each node is either active (has already adopted the information) or inactive, and each node's tendency to become active increases monotonically as more of its neighbors

become active. Nodes can switch from being inactive to being active, but cannot switch from being active to being inactive. Given an initial set A of active nodes, both model assume that the nodes in A first become active and all the other nodes remain inactive at time-step 0. The diffusion process of active nodes unfolds in discrete time-steps $t \geq 0$. As time unfolds, more and more of neighbors of an inactive node u become active, eventually making u become active, and u 's decision may in turn trigger further decisions by nodes to which u is connected. Thus, the spread of the information through the network G' is represented as the spread of active nodes on G' .

We set the seed nodes or the number of target influential nodes $k = 5\%$; of total nodes. We fix the number of steps (iterations) to be 50. The number of iterations can be considered as an indicator of the time required to influence the entire network as they represent how many iterations it takes to try to influence every node in a network. The ideal seed selection would be with maximum influence and minimum iterations for a network. We measured one primary parameter for the comparative analysis that is the percentage of influenced vertices.

Figure 9 shows the results of influence diffusion using the five SNA selection methods to select initial seed nodes (k) for IC and LT models. The first observation from the graphs for the influence spread is the effects of different seed selection strategies on influence results; and that high coreness nodes have the highest impact in influencing other nodes under both models (29 and 55% respectively); revealing a more central position in the network structure. On the other hand and among centrality measures, closeness—under both models—achieved the least impact to influence other nodes. This is could be attributed to the fact that closeness centrality measures implicitly assume that the underlying social network is strongly connected (Ghosh and Lerman 2010).

Degree and betweenness centrality measures perform well under both model; with degree performs better that betweenness under IC model; and the other way around under LT model. This is can be justified by the technical nature of each model, as that node thresholds are random in LT model, while in the IC model it is the propagations through links that are random. Since degree centrality is a local measure, the small value of node thresholds in LT results in less tendency of nodes to adopt the new behavior when their neighbors do. On the other hand, betweenness measure depends on the number of shortest paths or links that pass through a given node; while IC model operates according to the random propagation probability through these links. Finally, the strategy of using eccentricity measure to select seeds performs well in LT model which was higher by 43% than IC model.

7 Concluding remarks

We developed a scale to measure the lifelong learning capacity and professional skills referred to as “Career Dispositions” for each learner as indicators of preparation level towards practicing in knowledge-based market. We implemented a diagnostic LA tool to analyze and visualize information on collected data; and deliver it to users through a portal structure. Further, we proposed a fuzzy pairwise-constraints K-means (FPKM) algorithm as a predictive LA method to cluster learners of similar career patterns into interconnected (CoPs) that are driven by current industrial needs. We then constructed a CoP-network to support social learning by linking learners who share social interests beside their common career aspirations. We also devised a triadic closure approach to enhance community structures in CoP-network in order to optimize the learning performance of groups of learners. Finally, we incorporated different SNA metrics and concepts of social influence into a persuasive model to instill and track the diffusion of target behavior in a CoP-based network.

Future work involves developing further LA methods to capture, analyze and aggregate learning data that is inherently not interoperable such as learners administrative data, academic performance data, and classroom and online data. While integrating data sets from a variety of unconnected systems can be extremely difficult, it offers better extensive insights that automatically lead to improved capabilities of LA models. Further, we aim to develop LA management model to explore the effectiveness of CoP on improving career preparation outcomes in relation to the set of skills required by the industry for each designated career.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- AbuKhousa E, Atif Y (2016) Virtual social spaces for practice and experience sharing. In: State-of-the-art and future directions of smart learning. Springer, Berlin, Germany, pp 409–414
- Atkinson BM (2006) Captology: a critical review. In: Persuasive technology. Springer, Berlin, Germany, pp 171–182
- Baran B (2010) Facebook as a formal instructional environment. *Br J Educ Technol* 41(6):E146–E149
- Barrick MR, Mount MK, Strauss JP (1993) Conscientiousness and performance of sales representatives: test of the mediating effects of goal setting. *J Appl Psychol* 78(5):715

- Basu S, Banerjee A, Mooney R (2002) Semi-supervised clustering by seeding. In: In Proceedings of 19th international conference on machine learning (ICML-2002). Citeseer
- Bayt (2015) The bayt.com middle east job index survey. Technical report, Bayt
- Bernays EL (1928) Manipulating public opinion: the why and the how. *Am J Sociol* 33(6):958–971
- Boudreau JW, Boswell WR, Judge TA (2001) Effects of personality on executive career success in the united states and europe. *J Vocat Behav* 58(1):53–81
- Bourdieu P (1977) Outline of a theory of practice, Cambridge University Press, Cambridge, UK
- Brown JS, Collins A, Duguid P (1989) Situated cognition and the culture of learning. *Educ Res* 18(1):32–42
- Brunello G, Schlotter M (2011) Non-cognitive skills and personality traits: labour market relevance and their development in education & training systems. IZA Discussion paper. <https://ideas.repec.org/s/iza/izadps.html>
- Carbonell JG, Elchik ME, Simmons J, Assaad A, Pawlowski RJ et al (2014) Matching users of a network based on profile data. *US Patent App.* 14/173,260
- Chapelle O, Schölkopf B, Zien A et al (2006) Semi-supervised learning. MIT Press, Cambridge
- Clauset A, Newman ME, Moore C (2004) Finding community structure in very large networks. *Phys Rev E* 70(6):066111
- Curran JA, Murphy AL, Abidi SSR, Sinclair D, McGrath PJ (2009) Bridging the gap: knowledge seeking and sharing in a virtual community of emergency practice. *Eval Health Prof* 32(3):314–327
- Davidson I, Ravi S (2005) Clustering with constraints: feasibility issues and the k-means algorithm. In: *SDM*, vol 5. SIAM, California, USA, pp 201–211
- Davis CH III, Deil-Amen R, Rios-Aguilar C, Gonzalez Canche MS (2012) Social media in higher education: a literature review and research directions. Report printed by the University of Arizona and Claremont Graduate University
- Deakin Crick R (2007) Learning how to learn: the dynamic assessment of learning power. *Curric J* 18(2):135–153
- Ferguson R (2012) Learning analytics: drivers, developments and challenges. *Int J Technol Enhanc Learn* 4(5–6):304–317
- Fogg BJ (2009) A behavior model for persuasive design. In: Proceedings of the 4th international conference on persuasive technology. ACM, California, USA, p 40
- Franconi E, Kamble A (2004) A data warehouse conceptual data model. In: *IEEE*, p 435
- Friedkin NE (2006) A structural theory of social influence, vol 13. Cambridge University Press, Cambridge
- Fu Y-H, Huang C-Y, Sun C-T (2014) Identifying super-spreader nodes in complex networks. *Mathematical Problems in Engineering* 2015(2015). doi: [10.1155/2015/675713](https://doi.org/10.1155/2015/675713)
- Furnham A, Zacherl M (1986) Personality and job satisfaction. *Pers Individ Differ* 7(4):453–459
- Gaeta M, Marzano A, Miranda S, Sandkuhl K (2016) The competence management to improve the learning engagement. *J Ambient Intell Humaniz Comput*, pp 1–13
- Gajda R, Koliba CJ (2008) Evaluating and improving the quality of teacher collaboration a field-tested framework for secondary school leaders. *NASSP Bull* 92(2):133–153
- Ghosh R, Lerman K (2010) Predicting influential users in online social networks. In: Proceedings of KDD workshop on social network analysis
- Goyal A, Bonchi F, Lakshmanan LV (2010) Learning influence probabilities in social networks. In: Proceedings of the third ACM international conference on Web search and data mining. ACM, NY, USA, pp 241–250

- Gray K, Annabell L, Kennedy G (2010) Medical students' use of facebook to support learning: Insights from four case studies. *Med Teach* 32(12):971–976
- Gu L, Lu X (2012) Semi-supervised subtractive clustering by seeding. In: *Fuzzy systems and knowledge discovery (FSKD)*, 2012 9th International Conference. IEEE, pp 738–741
- Gunawardena CN, Hermans MB, Sanchez D, Richmond C, Bohley M, Tuttle R (2009) A theoretical framework for building online communities of practice with social networking tools. *Educ Media Int* 46(1):3–16
- Howard A, Bray DW (1990) Predictions of managerial success over long periods of time: lessons from the management progress study. In: *Leadership Library of America*, xvii, pp 113–130
- Hu C, Xu W, Shi P (2015) Information diffusion in online social networks: models, methods and applications. In: *Web-age information management*. Springer, Berlin, Germany, pp 65–76
- Indeed (2015) Top 10 in-demand jobs for 2015. Technical report, Indeed
- Jakovljevic M, Buckley S, Bushney M (2013) Forming Communities of practice in higher education: a theoretical perspective. ToKnow Press, Bangkok, Celje, Lublin
- Jawitz J (2007) New academics negotiating communities of practice: learning to swim with the big fish. *Teach High Educ* 12(2):185–197
- Johnson CM (2001) A survey of current research on online communities of practice. *Internet High Educ* 4(1):45–60
- Judge TA, Higgins CA, Thoresen CJ, Barrick MR (1999) The big five personality traits, general mental ability, and career success across the life span. *Pers Psychol* 52(3):621–652
- Kempe D, Kleinberg J, Tardos É (2003) Maximizing the spread of influence through a social network. In: *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, D.C., USA, pp 137–146
- Khousa EA, Atif Y (2014) A learning analytics approach to career readiness development in higher education. In: *Advances in web-based learning—ICWL 2014*. Springer, Berlin, Germany, p 133–141
- Khousa EA, Atif Y, Masud MM (2015) A social learning analytics approach to cognitive apprenticeship. *Smart Learn Environ* 2(1):1–23
- Kimura M, Saito K, Nakano R, Motoda H (2010) Extracting influential nodes on a social network for information diffusion. *Data Min Knowl Discov* 20(1):70–97
- Koliba C, Gajda R (2009) “Communities of practice” as an analytical construct: implications for theory and practice. *Intl J Public Adm* 32(2):97–135
- Lave J, Wenger E (1991) *Situated learning: legitimate peripheral participation*. Cambridge University Press, Cambridge, UK
- Mandl H, Gruber H, Renkl A (1996) 14 communities of practice toward expertise: social foundation of university instruction. In: *Interactive minds: life-span perspectives on the social foundation of cognition*, p 394
- McAdams DP, Olson BD (2010) Personality development: continuity and change over the life course. *Annu Rev Psychol* 61:517–542
- McCrae RR, Costa PT (1991) Adding liebe und arbeit: the full five-factor model and well-being. *Pers Soc Psychol Bull* 17(2):227–232
- McCrae RR, John OP (1998) An introduction to the five-factor model and its applications. *Personality* 60:295
- Mindich D (2011) *Building a learning community: examining the factors associated with the implementation of professional learning communities*. PhD thesis, Stanford University
- O'Brien O, Glowatz M (2013) Utilising a social networking site as a learning tool in an academic environment: advancing practice from information-sharing to collaboration and innovation (ici). *J Teach Learn High Educ* 5(3):1371–1372
- Oinas-Kukkonen H (2010) Behavior change support systems: a research model and agenda. In: *Persuasive technology*. Springer, Berlin, Germany, pp 4–14
- Oinas-Kukkonen H, Harjuma M (2009) Persuasive systems design: key issues, process model, and system features. *Commun Assoc Inf Syst* 24(1):28
- Otte E, Rousseau R (2002) Social network analysis: a powerful strategy, also for the information sciences. *J Inf Sci* 28(6):441–453
- Pan Y, Xu YC, Wang X, Zhang C, Ling H, Lin J (2015) Integrating social networking support for dyadic knowledge exchange: a study in a virtual community of practice. *Inf Manag* 52(1):61–70
- Pastizzo MJ, Erbacher RF, Feldman LB (2002) Multidimensional data visualization. *Behav Res Methods Instrum Comput* 34(2):158–162
- Qureshi IA, Raza H, Whitty M (2015) Facebook as e-learning tool for higher education institutes. *Knowl Manag E-Learn* 6(4):440–448
- Raad E, Chbeir R, Dipanda A (2010) User profile matching in social networks. In: *Network-based information systems (NBIS)*, 2010 13th international conference. IEEE, pp 297–304
- Rashotte L (2007) Social influence. *Blackwell Encycl Soc Psychol* 9:562–563
- Roberts BW, Mroczek D (2008) Personality trait change in adulthood. *Curr Dir Psychol Sci* 17(1):31–35
- Salgado JF (1997) The five factor model of personality and job performance in the european community. *J Appl Psychol* 82(1):30
- Scott J (2012) *Social network analysis: a handbook*. Sage, London, UK
- Sealey-Morris C (2014) Top 20 in-demand information technology careers. Technical report, Degree Query
- Seibert SE, Kraimer ML (2001) The five-factor model of personality and career success. *J Vocat Behav* 58(1):1–21
- Selwyn N (2009) Faceworking: exploring students' education-related use of facebook. *Learn Media Technol* 34(2):157–174
- Sutin AR, Costa PT, Miech R, Eaton WW (2009) Personality and career success: concurrent and longitudinal relations. *Eur J Pers* 23(2):71–84
- Tight M (2004) Research into higher education: an a-theoretical community of practice? *High Educ Res Dev* 23(4):395–411
- Tynjälä P, Välimaa J, Sarja A (2003) Pedagogical perspectives on the relationships between higher education and working life. *High Educ* 46(2):147–166
- Vygotsky LS (1980) *Mind in society: the development of higher psychological processes*. Harvard University Press, Cambridge, Massachusetts, USA
- Warhurst R (2003) Learning to lecture: situating the knowing and learning of higher education teaching. In: *British educational research association annual conference*, Edinburgh
- Wasko MM, Faraj S (2005) Why should I share? examining social capital and knowledge contribution in electronic networks of practice. *MIS Q* 29(1):35–57
- Watts DJ (2002) A simple model of global cascades on random networks. *Proc Natl Acad Sci* 99(9):5766–5771
- Wells G, Claxton G (2008) *Learning for life in the 21st century: sociocultural perspectives on the future of education*. Wiley, New Jersey, USA
- Wenger E (1999) *Communities of practice: learning, meaning, and identity*. Cambridge University Press, Cambridge, UK
- Wood W (2000) Attitude change: persuasion and social influence. *Annu Rev Psychol* 51(1):539–570
- Wu X, Kumar V, Quinlan JR, Ghosh J, Yang Q, Motoda H, McLachlan GJ, Ng A, Liu B, Philip SY et al (2008) Top 10 algorithms in data mining. *Knowl Inf Syst* 14(1):1–37

- Zhang S, Flammer C, Yang X (2010) Uses, challenges, and potential of social media in higher education. In: Cutting-edge social media approaches to business education: teaching with LinkedIn, Facebook, Twitter, Second Life, and Blogs, p 217
- Zimmerman RD (2008) Understanding the impact of personality traits on individuals' turnover decisions: a meta-analytic path model. *Pers Psychol* 61(2):309–348