CrossMark

Influence of network-based structural and power diversity on research performance

Alireza Abbasi[1] · Mahdi Jalili[2] · Abolghasem Sadeghi-Niaraki[3,4]

Abstract

With an increase in studies of actors' collaborative activities on their performance very few have examined the role of the structural influence of actors on their cohorts' performance. In this study, we argue that the collaborative process involves social influence and social capital embedded within relationships and network structures amongst direct partner. We examine whether such influence is directly associated with an actor's productivity and performance. In particular, we aim to understand how structural position of an actor's cohorts and their structural and power diversity in a network influence the productivity and performance. Network-based centrality measures are used to evaluate the structural position. To investigate how partners of an actor in a collaboration network can influence their productivity and/or performance, we evaluate their structural diversity in the network by measuring the variance of their centrality measures. We evaluate if the variance of centrality measures of partners of a scholar influence their productivity or performance.

Keywords Network science · Co-authorship network · Big data · Social influence · Research performance

Introduction

Network science and theory of complex networks have found many applications in various disciplines ranging from engineering to computer science, biology, medicine and social sciences (Barabási 2013). This is mainly due to advances in data collection and analysis tools and infrastructures. A major application of network science is in computational social sciences and social networks analysis and mining. Large-scale social network data analytics has various applications such as recommendation systems (Javari and Jalili 2015), link prediction (Lü and Zhou 2011; Jalili et al. 2017), information cascade (Jalili and Perc

✉ Abolghasem Sadeghi-Niaraki
dras.niaraki@gmail.com

[1] School of Engineering and IT, UNSW Australia, Canberra, Australia

[2] School of Engineering, RMIT University, Melbourne, Australia

[3] Faculty of Geomatics Engineering, K.N. Toosi University of Technology, Tehran, Iran

[4] Department of Computer Science and Engineering, Sejong University, Seoul, Republic of Korea

2017), node ranking (Shahriari and Jalili 2014) and citation analysis (Saarela et al. 2016). Analysis of social influence has become a major trend in computational social systems (Peng et al. 2017). It has potential applications in modern information and service-oriented industries. One can use nodes' internal characteristics and their social relations in the network to measure their social influence (Zhang et al. 2017).

An interesting application of network analytics tools is to analyse big data of scholars' activities and to measure their research performance. Individuals and team performance is measured through a set of metrics relating to their activities and contextual performance. Likewise, scientists and scholars' performance is evaluated based on measures such as their research productivity (in terms of number of publications), teaching evaluations, and attained grants income (Jiang 2008; Abramo and D'Angelo 2016; Waltman 2016). Appropriately evaluating scientists' research performance and identifying the underlying factors affecting it have been discussed in the literature. Perhaps the easiest approach would be counting scientists' publication activities (i.e. the number of publications), a proxy for their performance. On the other hand, citations received by publications have been considered as a measure of quality of the work and consequently the summation (or average) of citations a scholar receive has been proposed as an alternative performance metrics for scientists. More recently, Hirsch (2005) argued the drawbacks of previous metrics and proposed h-index as a metric which combines both quantity (publication count) and quality (citation count) of scholars' publication as a fair metric for measuring their research performance (Hirsch 2005). Due to human's limited capacity to gain and use knowledge and also accessibility to restricted resources, collaborating with others is required for knowledge creation and co-creation (Demsetz 1991) which is often expressed in a form of publication. That is why many scientific outputs are a collective effort of often geographical diverse scholars which has been shown to have a positive effect on the impact of the output in terms of citation (Abbasi and Jaafari 2013).

Although most of the bibliometric studied focus on measuring scholar's performance based on their publication activity and citation count received, Abbasi et al. (2010) highlighted the importance of scholars' collaboration activity and its effect of their research performance. Furthermore, numerous studies (Yan and Ding 2009; Takeda et al. 2010; Zhuge and Zhang 2010) have exposed how the position of scholars' in their collaboration networks (e.g. co-authorship), measured often through the network centrality metrics such as degree or betweenness, can effect on their performance. In addition, recently new hybrid centrality measures have been proposed (Abbasi 2013; Abbasi and Hossain 2013) for an actor in a network which have been used to measure actors' cohorts' influence and shown to have positive association with their performance.

This research, as another effort to demonstrate the prominence of scholars' position in their collaboration network, investigates the association between scholar's productivity and performance (measured by publication count and h-index) with not only the scholars' structural position, but also their partners' (direct contacts) structural position and also their structural diversity and power diversity in their collaboration network. We use actors' co-authors (partners) centrality measures variance to check their influence considering different structural position metrics in the network (i.e., degree, closeness and betweenness centrality measures). In particular, we investigate the difference in influence by connecting to co-authors with similar or divers roles (positions) in a network. We extract two real datasets for this purpose and analyse their network properties.

The remainder of the paper is organized as follows. Section 2 reviews social influence and social capital theories and previous studies on using those theories and social network analysis in investigating collaboration networks. Section 3 discusses our data and the

metrics proposed to use for measuring scientists' social influence and capital and their diversity. The correlation between the proposed measures and scholar's research performance is presented and discussed in Sect. 4. The paper concludes with a brief review of our finding and approach in Sect. 5.

Backgrounds

Theories of social influence and social capital

The theory of social influence expresses that people behavior is deliberately or involuntarily influenced by others (Strang 2000). Social influence process explains why individuals' behavior (actions and activities) adapts to the ones they interact with (Friedkin 1998; Crandall et al. 2008). Thus, social influence can be defined as "the change in behavior that one person causes in another, intentionally or unintentionally, as a result of the way the changed person perceives themselves in relationship to the influencer, other people and society in general" (Imran et al. 2013). Friedkin (1998) proposed a theoretical framework for 'social influence' as a network process. He treated 'social influence' as "proportional to the strength of direct and short indirect ties linking actors". As Strang (2000) has noted, Friedkin's theoretical approach enhanced the analysis of cohesion within network analysis.

This process can be seen in academic collaboration as well; some scholars have the ability to persuade their colleagues to follow their approaches or ideas when working collaboratively on a research project (Takeda et al. 2010). Expanding previous research on the role of scientists' position in their collaboration networks on their performance (Abbasi et al. 2011), here we try to recognize that the scholars' social influence diversity, which can be exposed through the variance of their centrality measures, on their co-authors' structural position in their co-authorship network and their performance.

Network theories and social network analysis

Social network analysis (SNA) is to map and measure relationships between actors (nodes) of (social) network. SNA provides both a visual and a mathematical analysis of human-influenced relationships. The social environment can be expressed as patterns or regularities in relationships among interacting units (Wasserman and Faust 1994). Each social network can be represented as a graph made of actors (e.g. individuals, organizations, information) that are tied by one or more specific types of relations, such as financial exchange, friends, trade, and Web links. A link between any two actors exists, if there is a relationship between them. If the actors represent people, a link means that those two people know each other in some way.

Not all nodes of a network have the same role in its structure and function; some nodes are central (vital) nodes. Bavelas (1950) pioneered investigations of formal properties of centrality as a relationship between structural centrality and influence in group process. Conducting an experiment with small group of students to solve puzzles, he found that groups with centralized structures (e.g., an actor collaborating with all or most of other members of the group which can be visualized as a 'star' or 'Y' network) were more successful in conducting simple tasks. On the other hand, where tasks were far more complex, a hybrid structure (i.e. combination of centralized and decentralized structures) was more appropriate. To quantify the importance of an actor in a social network, various centrality measures have been proposed over the years (Scott 1991). Freeman (1979)

defined centrality in terms of node degree centrality, betweenness centrality, and closeness, each having important implications on outcomes and processes.

The concept of 'centrality' originated in sociometric literature of the 'star' (Scott 1991). A central point was one with many direct connections to other points. The simplest and easiest way of measuring point centrality is by the degree of the various points in the graph. The degree of an actor in a network is simply the number of other actors connected directly to it. Inevitably, a central actor is not necessarily placed in the center of the network (Scott 1991). Freeman (1979, 1980) was the first who suggested closeness centrality as a global measure in terms of the distance among various actors in the network, although Sabidussi (1966) had been used 'sum distance' (i.e., the sum of the shortest path between any particular pair of actors in a network) to all other actors in the network as the same concept. Freeman (1979) also proposed another betweenness centrality that measures the number of times a particular actor lies 'between' the various other actors's shortest paths in the network. Betweenness centrality is defined more precisely as "the number of shortest paths (between all pairs of points) that pass through a given point" (Borgatti 1995).

While these measures are widely used to investigate the role and importance of networks, in summary, each is useful with significant socio-technical implications: Degree centrality indicates as an indicator of an actor's activeness and/or popularity. Closeness centrality reflects the accessibility of an actor to all other actors, and betweenness centrality can be regarded as an indicator of actors' ability in controlling the flow of resources within the network and reflects the gate-keeping or brokerage role and behavior of actors. Freeman (1979) found that centrality has a significant structural factor influencing leadership, satisfaction, and efficiency. Further reviews on centrality measures are covered by Abbasi et al. (2011).

Data, measures and methods

Datasets

For the analysis at an exploratory level, two real co-authorship networks are used. The first dataset, originally extracted and used in Abbasi et al. (2014), shapes the first co-authorship network of researchers in the field of "information science" by extracting the publication records having "information science" in their title, keywords or abstracts, published during the time period of 2001 and 2010 from the Scopus bibliometric database. After cleansing the data such as merging universities and departments with variant names (e.g. abbreviations and/or misspellings) in the original extractions, the resulting database contained 4579 articles published in 1392 journals and conference proceedings (indexed by Scopus) reflecting the contributions of 10,254 authors. We refer to this dataset as IS henceforth.

The second dataset, created and used in Abbasi and Jaafari (2013), extracts the publication records of top 10 journals of the "Information Science & Library Science" which are nominated based on their 2009 impact factor as enumerated in Web of Science *Journal Citations Report*. The meta-data of these shortlisted journal publications between 2000 and 2009 (inclusive) were mined, from Scopus by limiting the results to only 'articles' and 'reviews'. The records with no information for author(s) and their affiliation were disregarded. The dataset includes 4837 publications written by 10,820 authors. We will refer to this dataset as ISLS henceforth.

Centrality measures

Demographic data of the collaboration networks extracted from these datasets are shown in Table 1. As it is seen, both networks have almost low average degree and high clustering coefficient, indicating high levels of local connectedness in these collaboration networks. The average of closedness centrality for both networks is also close but there is a big difference between the betweenness centrality measures in the two networks: the average betweenness centrality of ISLS is 45.20 while it is 1.33 for IS. That reflects the topology of the two network is different. Figures 1 and 2 show distribution of actors' degree, clustering coefficient, betweenness and centrality measures for IS and ISLS networks, respectively. Let us briefly define these centrality measures. Let us denote adjacency matrix of an undirected and unweighted network by $A = [a_{ij}]$, where $a_{ij} = a_{ji} = 1$, if there is a link between actors (which are authors in these datasets) i and j, and $a_{ij} = a_{ji} = 0$, otherwise. Note that if two actors are coauthors in a paper, a link is assigned between them. Furthermore, we have $a_{ii} = 0$ indicating that there are no self-loops in the networks. Node degree is the simplest centrality measure and is defined as

$$C_{\mathrm{D}}(i) = \sum_j a_{ij}, \tag{1}$$

where $C_{\mathrm{D}}(i)$ is degree of node i. Indeed, degree is obtained by counting the number of connection links to a node. Clustering coefficient $CC(k)$ of node k measures its local connectedness and is defined as

$$CC(k) = \frac{\sum_{i,j} a_{ij} a_{ik} a_{jk}}{k_k(k_k - 1)} \tag{2}$$

Clustering coefficient indeed shows the strength of local connections in the network. It is obtained by counting all triples and dividing it by all possible structures. Many real networks including social networks have been shown to have rather high levels of clustering coefficient. Here we also consider two other centrality measures that are based on path lengths: betweenness and closeness centrality measures. Let us denote the edge between actors i and j by e_{ij}. Actor-betweenness centrality $C_{\mathrm{B}}(i)$ is a centrality measure of actor i in a graph, which shows the number of shortest paths making use of actor i (except those between the ith actors with the other nodes), and is compute as

$$C_{\mathrm{B}}(i) = \sum_{j \neq i \neq k} \frac{\Gamma_{jk}(i)}{\Gamma_{jk}}, \tag{3}$$

where Γ_{jk} is the number of shortest paths between actors j and k in the graph, and $\Gamma_{jk}(e_{ij})$ is the number of these shortest paths passing through actor i. Closeness of an actor indicates

Table 1 Demographic information of the datasets used in this research

Network	N	$<C_{\mathrm{D}}>$	$<CC>$	$<C_{\mathrm{B}}>$	$<C_{\mathrm{Cl}}>$
IS	10,254	4.69	0.76	1.33	5.24
ISLS	10,820	3.44	0.67	45.20	7.82

N network size, C_D degree, C_B betweenness centrality, C_{Cl} closeness centrality, CC clustering coefficient, $<.>$ means average
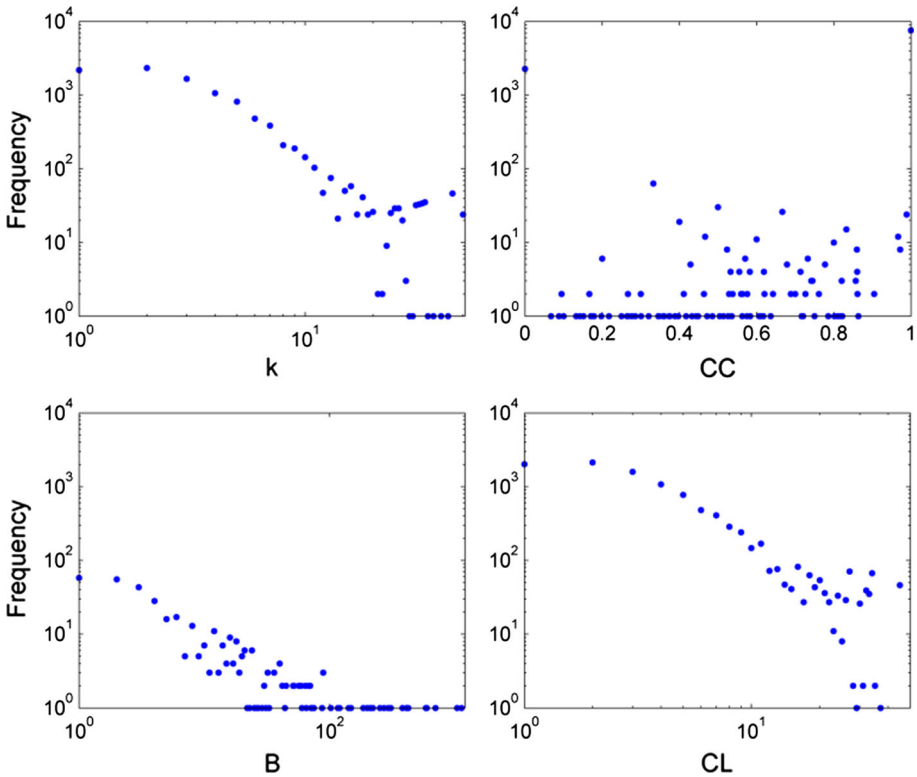
Fig. 1 Distribution of degree, clustering coefficient, betweenness and closeness centrality in IS. *k* degree, *CC* clustering coefficient, *B* betweenness centrality, *CL* closeness centrality

the inverse of its farness from all other nodes (i.e., closeness to the network center), and is calculated as

$$C_{xl}(i) = \sum_j (L_{ij})^{-1} \tag{4}$$

where L_{ij} is the length of the shortest path between actors i and j.

As Figs. 1 and 2 show both networks have scale-free degree distribution with heavy tails. In other words, while a small portion of actors have large degrees, many have rather small degree values. Many real-world networks have been shown to have power-law degree distribution, including biological, technological, ecological and social networks. The actors have distributed clustering coefficients from medium to high values. Betweenness and closeness centrality values can also be approximated as a power-law distribution with exponential start and heavy tails. UCINET is used for measuring network measures.

Performance measures

To evaluate scientists' performance, previous studies recommended publication count and citation count as indicators of scholars' productivity and performance accordingly. Hirsch
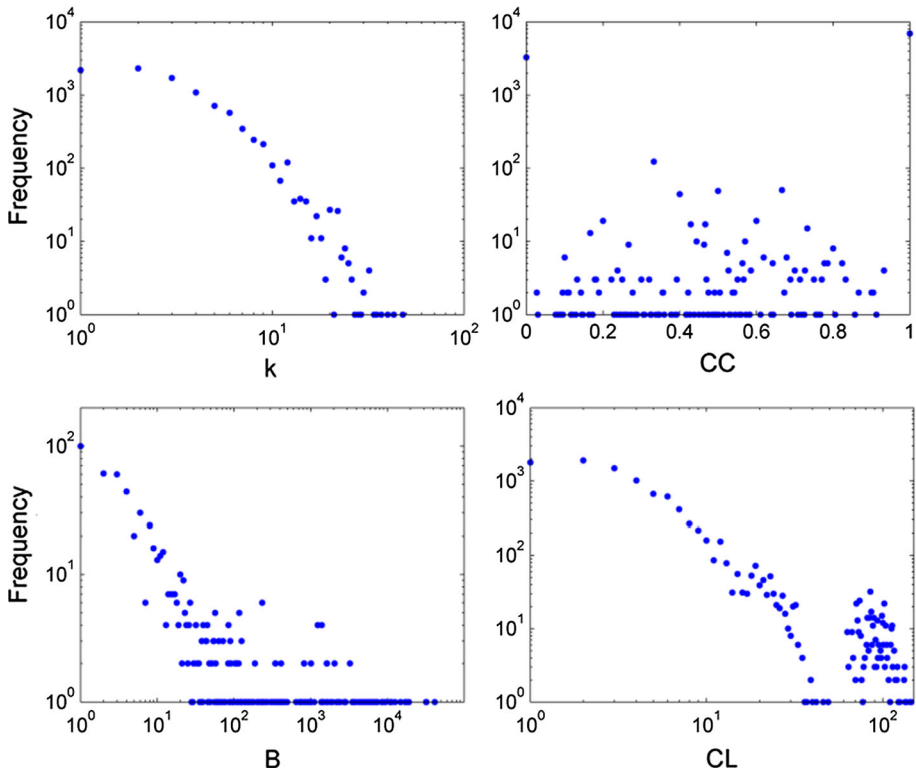
Fig. 2 Distribution of degree, clustering coefficient, betweenness and closeness centrality in ISLS. Other designations are as Fig. 1

(2005) proposed the *h*-index as a metric to combine in a simple way scholars' research quantity (i.e. publications' count) and quality (i.e. citations' count). A researcher with an index of *h* has *h* published papers which have been cited at least *h* times (Hirsch 2005). Although there is still discussion on the trustworthiness of the *h*-index and its variants (Haque and Ginsparg 2009) for measuring research performance, it is still the most commonly used metric amongst researchers. Thus, *h*-index is considered in this research as a citation-based proxy for research performance of scholars. Also, we consider the number of publications as a proxy for productivity of scholars. Table 2 shows some descriptive information (i.e., the minimum, average, and maximum values) of the productivity (the number of publications) and performance (i.e. *h*-index) measures for both datasets.

Table 2 Descriptive information of performance measures for the datasets

Network	N	Publication Count			*h*-index		
		Min	Avg.	Max	Min	Avg.	Max
IS	10,254	1	1.10	9	0	0.84	9
ISLS	10,820	1	1.18	36	0	1.08	12

Measuring structural and power diversity

Structural diversity

To quantify to what extent a scholar is influenced by her co-authors (research partners), this research proposes new metrics which consider the classic centrality measures (i.e. degree, closeness and betweenness) of the co-authors of actors as a proxy for their structural positions. To define scholar's social influence metrics, it is considered that centrality measures of each actor in their co-authorship network are known as an attribute of the actor in the network. Then, we define social influence of a scholar as the variance of her direct co-authors' centrality measures. A similar approach has been applied in Abbasi (2013) and Abbasi and Hossain (2013) by proposing new hybrid centrality measures for an actor considering her partners' centrality measures into account, but here we investigate the difference as a means of representing influence by connecting to partners with similar or diverse structural positions in the network.

We consider three types of structural measures for the actors: *i*) their centrality values measure by their degree, clustering coefficient, betweenness or closeness centrality measures, *ii*) their partners' mean centrality values obtained by taking average over the partners' centrality measures, and *iii*) diversity of their partners' centrality obtained by calculating the standard deviation of partners' centrality measures. Thus, for instance, the diversity of social influence measure based on the degree centrality of partners of actor *i*, $PVC_D(i)$—short for partner' variance of degree centrality, can be defined as the variance of degree centralities of *i*' partners. To calculate the variance, we first measure the partners' degree centrality mean and then as per the following equations calculate the variance. The three measures can be shown as below where N_i is the number of node *i*'s partners (i.e., the number of direct neighbors of node *i*). PMC_D, PMC_C and PMC_B show the mean of degree centrality, closeness centrality and betweenness centrality of *i*' partners, respectively.

$$PVC_D(i) = \frac{1}{N_i - 1} \sum_j a_{ij} (PMC_D - C_D(j))^2, \tag{5}$$

$$PVC_C(i) = \frac{1}{N_i - 1} \sum_j a_{ij} (PMC_C - C_C(j))^2, \tag{6}$$

$$PVC_B(i) = \frac{1}{N_i - 1} \sum_j a_{ij} (PMC_B - C_B(j))^2, \tag{7}$$

where a_{ij} shows the weight of the link between actors *i* and *j* in weighted networks and will be 1 in binary networks.

Power diversity

On the other hand, to measure the available power diversity to actors in the network we use Power-Tie-Diversity Index (PTDI) defined in Abbasi et al. (2014) as a social capital measure. Diversity and power has been discussed in the social capital literature [for detail discussion and references please refer to Abbasi et al. (2014)] as factors of valued resources. Therefore, social capital is defined as the frequency and diversity of partners (directly connected actors) considering their power (e.g. high performance). Therefore,

having the power (here measured by h-index) of actors in a social network, social capital of an actor in a network can be measured considering the added value of her direct contacts in addition to their quantity and frequency of contacts with each partner (e.g. tie strength).

In order to consider all the variables (i.e. Power, diversity and tie strength), Power Diversity Index (PDI) of an actor has been proposed in Abbasi et al. (2014) as the largest number such that an actor's top-h partners have each at least a power value (e.g. h-index) of h. In other words, PDI is the 'h-index of partner's h-indices'. In another effort, actors' tie strengths to their partners have been taken into consideration in PTDI. To measure PTDI, first partners' "power-strength" is defined as the h-index of each partner multiplied by the strength of the tie between that partner and the actor. Thus, the actor's PTDI is the largest number such that her top-h partners have each at least the power-strength of h.

As discussed in Abbasi et al. (2014), it is expected that actors with high PTDI value, reflecting having frequent connection to more powerful actors, have also more power. "Accordingly, this reflects individuals' power and influence on transmitting and controlling information as well as the popularity of an individual based on popularity of its direct contacts" (Abbasi et al. 2014). We obtain correlations of PTDI with the actors' research productivity and performance measures.

Results and discussions

Table 3 shows Spearman rank correlation test between the scholars' productivity/performance (as measured by their publication count and h-index) and their network-based centrality (degree, clustering coefficient, betweenness and closeness centralities), and structural and power diversity measures for both IS and ISLS datasets. The results reveal a similar behaviour for actors' productivity and performance in terms of the association with all four groups of variables. This shows actors' network-based measures correlation and actors' output is independent of the choice of variable as a proxy for their scholarly activity either quantity (e.g. publication count) or quality (e.g., citation-based measures such as h-index).

The positive significant correlation between actors' productivity/performance and the first group of parameters which reflect actors' structural position in the collaboration network show that actors with good structural position in their collaboration (i.e., having more partners, being close on average to others and bridging other scholar's connection in the network) have higher productivity and performance. This findings support previous research on the effects of actors collaborative activities and the role of their position in the collaboration network on their performance (Abbasi et al. 2011).

While actors' centrality measures show a significant positive correlation with their productivity and performance, the correlation for clustering coefficient is negative. This means that the scholars with stronger interconnection between their partners (i.e. having high clustering coefficient) have lower research performance. Access to strongly interconnected partners often provides redundant knowledge and skills to the actor and this might not be paid off by the resources (e.g. time) she needs to use to build and maintain these relationships and can affect her performance. This can explain the negative correlation between actors' clustering coefficient and their performance.

The positive significant association results for the second group of network-based parameters (i.e., partners' structural position), obtained by averaging the centrality of partners, show that actors collaborating with partners whom have good structural position (i.e., connecting to many actors, being close to others and being in a bridging position among otherwise unconnected actors) have higher productivity and performance measure.

Table 3 Spearman rank correlation test between scholars' productivity/performance (as measured by their publication count and h-index) and their network-based centrality measures, structural and power diversity

Network-based measures	IS dataset		ISLS dataset	
	Productivity (pub. count)	Performance (h-index)	Productivity (pub. count)	Performance (h-index)
Actors' structural position				
C_D	**0.185***	**0.176***	**0.214***	**0.196***
C_{CI}	**0.145***	**0.170***	**0.179***	**0.182***
C_B	**0.143***	0.076	**0.322***	**0.273***
CC	**− 0.243***	− 0.031	**− 0.235***	**− 0.154***
Partners' structural position				
PMC_D	**0.176***	**0.173***	**0.190***	**0.179***
PMC_C	**0.181***	**0.178***	**0.200***	**0.193***
PMC_B	**0.198***	**0.149***	**0.231***	**0.210***
Partners' structural diversity				
PVC_D	0.041	0.096	0.026	0.073
PVC_C	0.062	0.103	0.046	0.094
PVC_B	0.123	0.035	**0.151***	**0.145***
Partners' power diversity				
PTDI	**0.194***	**0.573***	**0.246***	**0.423***

*Correlation is significant at the 0.001 level (2-tailed)

This shows an actor's connectedness to well-positioned partners may have an impact on her performance.

The performance/productivity measures showed significant positive correlation only with partners' structural diversity based on betweenness centrality in ISLS dataset. This indicates if an actor collaborates with partners with the diverse range of structural positions, she will have higher productivity and performance, although with less correlation magnitude than the other measures. This finding is a bit surprising as one might expect collaboration with only a group of high-profile partners (such as partners who are connected to many others or being close to all others on average) would be more useful than connecting to diverse range of actors whom for example some of them have a high number of partners and some very few. However, it is important to note that the significant correlation between productivity/performance and PVCB in ISLS could be influenced by a higher dispersion of CB in the ISLS data set.

Finally, the positive significant correlation result between actors' power diversity metric measured by PTDI and their productivity and performance reveals that actors' social capital (having connection to high performance partners) is an important factor that helps them to achieve high performance and productivity as well. PTDI has much higher correlation values as compared to other structural metrics, indicating its more significant role in the scholars' research productivity and performance.

Conclusions

Social network analysis and mining has become one of the core topics in computer science and computational social sciences. In this paper we studied the influence of network-based structural and diversity measures on scholar's research productivity and performance.

Developing new metrics to assess the structural position and diversity of collaborators of a scholar and also examining their association with scholars' performance and productivity have been the main novel contributions of this study. The study was conducted on two real datasets of research publications.

Three types of structural measures (scholar's centrality, mean and variance of their partner's centrality measures) and a power diversity measure were correlated with scholar's research productivity (as measured by the publication count) and performance (as measured by h-index). The results showed significant positive correlation for all centrality measures expect the clustering coefficient for which we found negative correlations. Partner's structural position related measures (i.e. PMC_D, PMC_C, PMC_B) also showed a positive significant correlation, reflecting that structural position of partners of a scholar can affect on her performance and productivity. On the other hand, collaborators' structural diversity measures have not been found significantly correlated with the performance and productivity measures except for PVC_B in one of the dataset which could be as a result of the higher dispersion of betweenness centrality measures in that data set. The outcome of this research are significant in applying tools available in graph theory and network science for describing underlying principles in scholars' research productivity and research.

References

Abbasi, A. (2013). h-Type hybrid centrality measures for weighted networks. *Scientometrics, 96*(2), 633–640.

Abbasi, A., Altmann, J., & Hossain, L. (2011). Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics, 5*(4), 594–607.

Abbasi, A., Altmann, J., & Hwang, J. (2010). Evaluating scholars based on their academic collaboration activities: Two indices, the RC-index and the CC-index, for quantifying collaboration activities of researchers and scientific communities. *Scientometrics, 83*(1), 1–13.

Abbasi, A., & Hossain, L. (2013). Hybrid centrality measures for binary and weighted networks. In R. Menezes, A. Evsukoff, & M. C. González (Eds.), *Complex networks* (Vol. 424, pp. 1–7). Berlin: Springer.

Abbasi, A., & Jaafari, A. (2013). Research impact and scholars' geographical diversity. *Journal of Informetrics, 7*(3), 683–692.

Abbasi, A., Wigand, R., & Hossain, L. (2014). Measuring social capital and its influence on individual performance. *Library & Information Science Research, 36*(1), 66–73.

Abramo, G., & D'Angelo, C. A. (2016). A farewell to the MNCS and like size-independent indicators. *Journal of Informetrics, 10*(2), 646–651.

Barabási, A.-L. (2013). Network science. *Philosophical Transactions of the Royal Society A, 371,* 20120375.

Bavelas, A. (1950). Communication patterns in task-oriented groups. *Journal of the Acoustical Society of America, 22,* 725–730.

Borgatti, S. (1995). Centrality and AIDS. *Connections, 18*(1), 112–114.

Crandall, D., Cosley, D., Huttenlocher, D., Kleinberg, J., & Suri, S. (2008). *Feedback effects between similarity and social influence in online communities*. New York: ACM.

Demsetz, H. (1991). The theory of the firm revisited. In O. E. Williamson, S. G. Winter, & R. H. Coase (Eds.), *The nature of the firm: Origins, evolution, and development* (pp. 159–179). New York: Oxford University Press.

Freeman, L. C. (1979). Centrality in social networks conceptual clarification. *Social Networks, 1*(3), 215–239.

Freeman, L. C. (1980). The gatekeeper, pair-dependency and structural centrality. *Quality & Quantity, 14*(4), 585–592.

Friedkin, N. E. (1998). *A structural theory of social influence*. Cambridge: Cambridge University Press.

Haque, A., & Ginsparg, P. (2009). Positional effects on citation and readership in arXiv. *Journal of the American Society for Information Science and Technology, 60*(11), 2203–2218.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the USA, 102*(46), 16569–16572.

Imran, M., Elbassuoni, S. M., Castillo, C., Diaz, F., & Meier, P. (2013). *Extracting information nuggets from disaster-related messages in social media*. Baden-Baden: ISCRAM.

Jalili, M., Orouskhani, Y., Asgari, M., Alipourfard, N., & Perc, M. (2017). Link prediction in multilayer online social networks. *Royal Society Open Science, 4*(2), 160863.

Jalili, M., & Perc, M. (2017). Information cascades in complex networks. *Journal of Complex Networks, 5*(5), 665–693.

Javari, A., & Jalili, M. (2015). A probabilistic model to resolve diversity-accuracy challenge of recommendation systems. *Knowledge and Information Systems, 44,* 609–627.

Jiang, Y. (2008). Locating active actors in the scientific collaboration communities based on interaction topology analyses. *Scientometrics, 74*(3), 471–482.

Lü, L., & Zhou, T. (2011). Link prediction in complex networks: A survey. *Physica A, 390,* 1150–1170.

Peng, S., Wang, G., & Xie, D. (2017). Social influence analysis in social networking big data: Opportunities and challenges. *IEEE Network, 31*(1), 11–17.

Saarela, M., Kärkkäinen, T., Lahtonen, T., & Rossi, T. (2016). Expert-based versus citation-based ranking of scholarly and scientific publication channels. *Journal of Informetrics, 10*(3), 693–718.

Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika, 31*(4), 581–603.

Scott, J. (1991). *Social network analysis: A handbook*. Beverley Hills: Sage.

Shahriari, M., & Jalili, M. (2014). Ranking nodes in signed social networks. *Social Network Analysis and Mining, 4*(1), 1–12.

Strang, D. (2000). Review: A structural theory of social influence. *JSTOR, 45,* 162–164.

Takeda, H., Truex III, D., & Cuellar, M. (2010). Evaluating scholarly influence through social network analysis: The next step in evaluating scholarly influence. In *AMCIS 2010 proceedings* (Vol. 573).

Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics, 10*(2), 365–391.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge, NY: Cambridge University Press.

Yan, E., & Ding, Y. (2009). Applying centrality measures to impact analysis: A coauthorship network analysis. *Journal of the American Society for Information Science and Technology, 60*(10), 2107–2118.

Zhang, C., Lu, T., Chen, S., & Zhang, C. (2017). Integrating ego, homophily, and structural factors to measure user influence in online community. *IEEE Transactions on Professional Communication, 60*(3), 292–305.

Zhuge, H., & Zhang, J. (2010). Topological centrality and its e-science applications. *Journal of the American Society for Information Science and Technology, 61*(9), 1824–1841.