ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645
Indian Journal of Science and Technology, *Vol 8(S1), 333-339, January 2015* 10.17485/ijst/2015/v8iS1/59311

Comparing Analysis Study of Centrality Indices using Paper Information on Secondary Battery

Young-il Kwon, Dae-hyun Jeong*, Yeongho Moon and Jae-young Yoo

Korea Institute of Science and Technology Information;
ylkwn@kisti.re.kr, yhmoon@kisti.re.kr, yoojy@kisti.re.kr, gregori79@kisti.re.kr

Abstract

Background/Objectives: Secondary battery is expanding large secondary battery for various applications. In this study, joint research trends were analysed using network analysis in order to investigate the R&D of secondary batteries. **Methods/Statistical Analysis:** Degree centrality and betweenness centrality among the network analysis methods, the complex degree centrality was used to analyse the joint research network. Furthermore, the relationships among the degree centrality, betweenness centrality, and complex degree centrality were analysed. The analysis results show that 78 out of 91 countries carried out joint research, with the exclusion of 13 countries that did not participate in international joint research. **Results:** The joint research between China and USA institutions was most active, and China actively conducted joint research with Asian countries. A cluster analysis of the countries participating in joint research found that there were seven clusters in total, and the USA, Germany, France, and the U.K. played central role in each cluster. A correlation analysis of the centrality indices analysis results by country showed strong positive (+) correlation among the three indices. Furthermore, a regression analysis showed that the greater the complex degree centrality was, the greater the degree centrality and betweenness centrality became, and the increasing rate of the betweenness centrality was very high. **Conclusion/Application:** This study is meaningful that the correlations between complex degree centrality and other centrality indices were analysed in network analysis and investigated joint research status.

Keywords: Centrality Indices, Network Analysis, PFNet, Regression, Secondary Battery

1. Introduction

The technical development of secondary batteries is moving from small secondary batteries for notebook computers and smart phones toward the performance improvement and expansion of the storage capacity of medium and large secondary batteries that are used in electric cars and energy storage devices. For medium and large lithium ion batteries in particular, new businesses using secondary batteries for electric cars to energy storage devices are being developed. In other words, the paradigm of secondary battery business is shifting from mobile devices to medium and large secondary batteries applied to electric cars, the storage of power generated from renewable energy, and emergency power storage devices[1]. In this study, international joint research status in the secondary battery field was analysed through the analysis of international joint research network in secondary batteries.

Since a couple of decades ago, network analysis methods have been used in some scientific areas[2]. Social network consists of a group of actors such as individuals and organisations, and social network analysis methods are used to describe various social phenomena[3]. The social network analysis developed in various analysis methods in various areas. There are largely three types of social network analysis methods. The first is social measurement analysis method that contributed to the

* *Author for correspondence*

technical development of social network research using graph theory methodology; the second is the method analysing the relationships among people; and the third is the method of analysing community structures in tribe and village units. German scholars who moved to the USA in 1930s researched cognitive psychology and social psychology. Influenced by the gestalt theory by Wolfgang Köhler, German scholars researched the social measurements and dynamic structures between groups. They investigated the group structures of organisations and the distribution paths of information and ideas through case studies. During the same period, anthropologists and sociologists who belonged to Harvard University succeeded the theory of Alfred R. Radcliffe-Brown who was a British anthropologist and actively researched the interdependency among components of social structures. The research of these scholars about social communities is being regarded as very important. Their main interest was to reveal the unofficial relationships of a social system and the importance of interpersonal relationships. Meanwhile in the U.K. scholars who succeeded the Radcliffe-Brown's theory around Manchester University studied the frictions and collisions of the members of a social system. They applied these ides to the ethnic societies of Africa and later to the small garden cities in the U.K. They made considerable achievements by applying mathematics to various sociological theories. The social network analysis methodology was dramatically developed by Harrison White from Harvard University in 1960s. Harrison White conducted studies on social structures based on mathematics. He synthesised the research results of previous researchers in North America and his synthesised theory was later developed into a systematic research of social networks by his students. The most actively discussed among the analysis methods using the network theory is the network's degree centrality measurement method. Network centrality analysis is classified into degree centrality, closeness centrality, and betweenness centrality between nodes. Degree centrality is the measurement of how central a node is in a network. With degree centrality, the degree of linkage between a node that forms the network and other nodes directly linked to the node is measured, and the number of linked nodes is the absolute measuring criterion. That is, this technique is used to identify how central in a network each node is linked. Degree centrality can be divided into in-degree centrality and out-degree centrality depending on the direction of the link of the two nodes. When this is formulated, the degree centrality $D_c(p_k)$ of a given node p_k can be calculated as follows as the sum of other nodes that is adjacent to p_k. Where, $(p_i, p_k) = 1$ means that p_i and p_k are linked, while $(p_i, p_k) = 0$ means that they are not linked.

$$D_c(p_k) = \frac{\sum_{i=1}^{n} p_i, p_k}{n-1} \qquad (1)$$

Closeness centrality represents the closeness of a node to another node in a network, regardless of direct or indirect. This centrality is used to analyse the degree of closeness of information, and it is measured by a total of the shortest distances from all nodes to a node. That is, contrary to the degree centrality, it measures the centrality by summing up all the distances not only from the directly linked nodes to a node in a network, but also that of the nodes that are indirectly linked to the node. The node with the smallest sum of all pathway distances has the highest closeness centrality in the network, making it the network's central node. This means the node can get close to others in the shortest path according to the characteristics of the network, and it can monitor the flow of the information. When it is mathematized, the closeness centrality (C_c) can be calculated as follows by taking the inverse of the sum of the shortest distances between nodes. Where, $d(p_i, p_k)$ represents the shortest distance between p_i and p_k .

$$C_c = \frac{1}{\sum_{i=1}^{n} d(p_i, p_k} \qquad (2)$$

Contrary to the closeness centrality analysis, betweenness centrality measures the degree of betweenness of one node and another chosen node that is linked to the first node. The betweenness centrality of a certain node expresses the ratio between the shortest distances from other pairs of nodes to the certain node and the number of the existence of the certain node in the actual shortest distance. That is, betweenness centrality measures the degree of mediating function between other nodes, and it can be expressed in a formula like below. Where, $i<j$ and $b_{ij}(p_k) = d_{ij}(p_k)\, d_{ij}$. The denominator is shortest distance between p_i and p_j, and the numerator is the shortest distance between p_i and p_j including p_k.

$$B_c(p_k) = \frac{2\sum_{i}^{n}\sum_{j}^{n} b_j(p_k)}{n^2 - 3n + 2} \qquad (3)$$

While degree centrality means that the node performs a central function in the network in the organisation, closeness centrality indicates the degree of direct and indirect relationships in an overall network, and betweenness centrality indicates the aspect of control that mediates the flow of information[4].

The first study on network centrality measurement method was attempted by Harold Leavitt et al.[5] who discussed in detail about leadership, the satisfaction of a participating group, and the efficiency of an organisation. Cohn et al.[6] used network centrality to understand political integration in the diversity of Indian social life. He investigated how a multiracial country like India could be managed and as a result, revealed that various phenomena in Indian society were interconnected through networks in various aspects. Pitts[7] examined the importance of centrality in the process of communication for urban development. He restructured the network of Russian transportation in the 12th century and revealed that Moscow which had the highest centrality among Russian cities was more excellent in terms of transportation and communication compared to other cities. Czepiel[8] used the network centrality measurement method to explain the diffusion pattern of technical innovation in the steel industry. He found that in the network among companies, the higher the centrality of a company was, the more active early adopter of the new casting process the company was introduced[5–8].

Centrality parameters like these are to be applied to a binary network, while it cannot be applied well in a weighted network in which the weight is loaded with links. Most data subject to bibliometric analysis such as simultaneous quoting or simultaneous linkage are expressed via the frequency matrix, which can become a weighted network when it is plotted as a network[9].

In this study, we used complex degree centrality by improving the existing degree centrality so that it analyses the centrality of each country in secondary battery areas.

Complex degree centrality is calculated using following formula:

$$CDC_A = \sqrt{"\boldsymbol{D}_A"\times\boldsymbol{R}_A} = \sqrt{2^{H(A)}\times\boldsymbol{R}_A} \quad (4)$$

In this formula, "DC_A" is a term that reflects the number of coauthors of A, and "TR_A" is a term that reflects the number of co-authorships. "DC_A", the "number" of coauthors of researcher (A) is measured by applying entropy as an exponent of $2(2^{H(A)})$. When the number of co-authorships of the researcher (A) and the coauthor (i) is K_i, the entropy H (A) is measured like follows:

$$H(A) = -\sum_{i=1}^{\tilde{z}}\left\{\frac{K_i}{\boldsymbol{R}_A}\times\log_2\frac{K_i}{\boldsymbol{R}_A}\right\} \quad (5)$$

For Complex Degree Centrality (CDC_A), a higher value is deducted when a researcher, a research institute, or a country has a higher number of co-authorships, and performs a joint research with more parties in equal portion[10–13].

2. Methods of Analysis

In this study, we extracted the paper data from the area of secondary battery worldwide using Scopus Database, established a network between countries using the nationality information of the authors, and compared the complex degree centrality to the existing centrality measurement.

To establish a network using the nationalities of the authors, we established the co-occurrence matrix and constructed the structure of the network based on the matrix as shown in Figure 1.

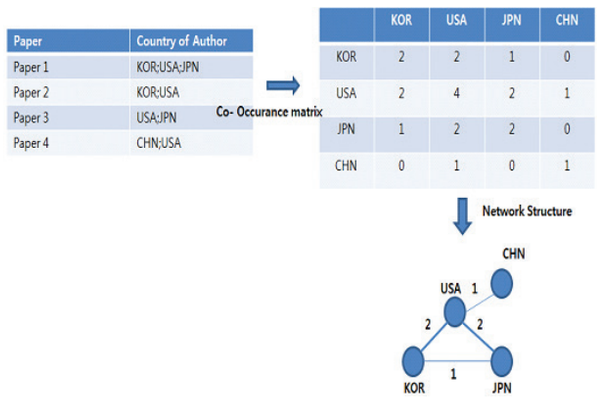

Figure 1. Networking using the countries of the authors (Example).

3. Analysis Results

When the structure of the network between the nationalities of authors in the area of secondary battery was analysed, it was found that of all of the countries that have published at least one paper, a total of 78 countries

are performing joint researches, with the exception of 13 countries that have not participated in any international joint research (Figure 2).

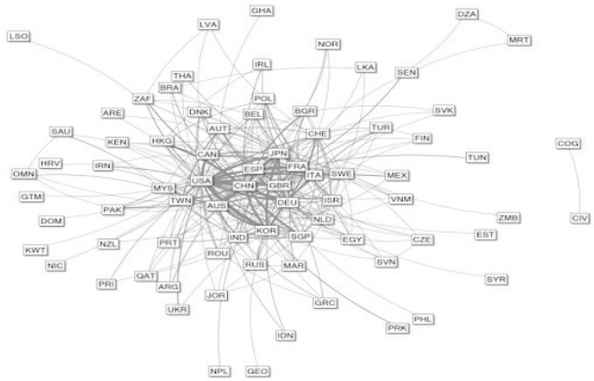

Figure 2. Network structure between the countries of the authors.

Figure 3 enlarged the center of the network structure of countries that performed 10 or more international joint researches. Korea (KOR) was found to actively perform researches with USA, China (CHN), Japan (JPN), Australia (AUS) and Italy (ITA).

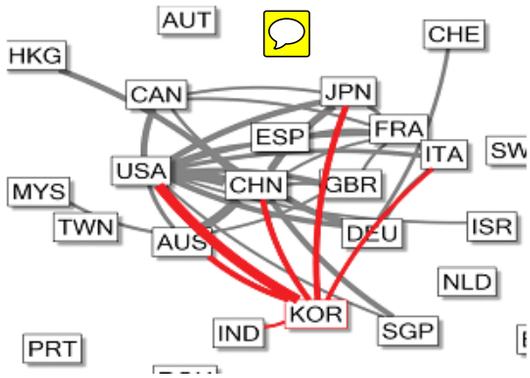

Figure 3. Enlarged center of the network of joint research (with 10 or more joint researches).

The result of analyses of the Complex Degree Centrality (CDC), Degree Centrality (DC), and Betweenness Centrality (BC) of the international joint research network structure is shown in Table 1.

USA showed a high Complex Degree Centrality (CDC), Degree Centrality (DC) and Betweenness Centrality (BC). Japan (JPN) had a degree centrality of 0.44, second to USA, while the complex degree centrality was lower than that of China (CHN), France (FRA), and Germany (DEU). It was analysed that USA was performing the most joint researches and with the most countries, playing a role of mediator in joint researches.

It was discovered that China, France, and Germany were performing research with various countries taking equal parts, and France was found to play a strong mediating role in joint research organisations.

Table 1. Centrality analysis by country in the area of secondary battery

Ranking	Country	Number of papers	Number of countries making joint researches	CDC	DC	BC
1	USA	2,185	47	91.91	0.63	0.32
2	CHN	3,495	30	74.80	0.40	0.07
3	FRA	339	31	54.42	0.41	0.13
4	DEU	424	30	54.12	0.40	0.08
5	JPN	1,316	33	52.34	0.44	0.09
6	KOR	734	25	45.39	0.33	0.07
7	GBR	252	27	43.82	0.36	0.06
8	AUS	310	23	41.66	0.31	0.04
9	ITA	201	21	32.42	0.28	0.02
10	IND	320	16	26.96	0.21	0.04
11	CHE	95	19	26.32	0.25	0.05
12	SWE	65	20	24.26	0.27	0.05
13	SGP	155	15	22.74	0.20	0.03
14	ESP	142	15	22.48	0.20	0.01
15	CAN	188	13	22.10	0.17	0.01
16	ISR	57	15	19.92	0.20	0.00
17	MYS	72	12	16.09	0.16	0.01
18	AUT	56	11	15.98	0.15	0.03
19	RUS	112	13	14.00	0.17	0.03
20	TWN	293	8	13.26	0.11	0.00
21	BGR	45	10	12.71	0.13	0.01
22	DNK	17	10	12.08	0.13	0.01
23	EGY	23	10	11.94	0.13	0.03
24	POL	37	9	11.75	0.12	0.01
25	HKG	56	8	11.14	0.11	0.00
26	BEL	26	8	10.86	0.11	0.00
27	NLD	46	7	10.43	0.09	0.00
28	ZAF	29	7	9.29	0.09	0.03
29	TUR	31	7	8.77	0.09	0.00
30	PAK	14	6	8.54	0.08	0.00

After the joint research frequencies among 42 countries that published 10 or more theses was normalized using the cosine coefficient, the Pathfinder Network (PFNet)[13]

and PNNC clusters were analysed with the WNET program to divide 42 countries into 7 clusters: Cluster C1 around China (CHN); Cluster C2 around the USA (USA) which also includes Japan (JPN) and Korea (KOR); Cluster C3 around Germany (DEU), Cluster C4 around France (FRA); Cluster C5 around U.K. (GBR); Cluster C6 including South Africa (ZAF); and Cluster C7 including Egypt (EGY). Japan conducted the largest number of studies in terms of the number of joint research with China, but when normalized with the cosine coefficient, the intensity of joint research with Korea was stronger than that with China (Figure 4).

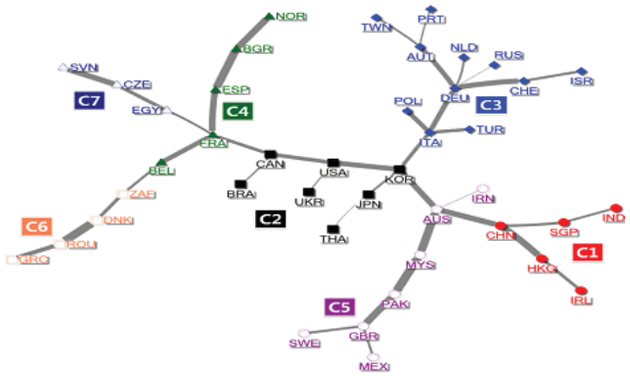

Figure 4. Joint research cluster analysis normalised with cosine coefficient (PFNet, PNNC).

In this study, we analysed the correlation between Complex Degree Centrality (CDC) and Betweenness Centrality (BC). In the result, it was analysed that complex degree centrality, Degree Centrality (DC), and betweenness centrality had high correlation, and it was significant in confidence level of 0.01. That is, it was analysed that a country executing joint researches in the secondary battery area with more countries in equal part are playing mediating roles in the relevant network area (Table 2).

Table 2. Analysis on correlation between centralities

	CDC	DC	BC
CDC	1		
DC	0.97	1	
BC	0.848	0.864	1

* Significant in confidence level 0.01 interval

Based on the correlation analysis performed above, regression analysis was performed on the centrality indices. In the result, the Complex Degree Centrality (CDC) and Degree Centrality (DC) were found to be related by a linear function, (Figure 5). The regression analysis showed that R^2 was 94.1%, the F value of variance analysis was 444, and the significance probability (p) was 0.000. Thus, the regression equation was statistically significant (Table 3).

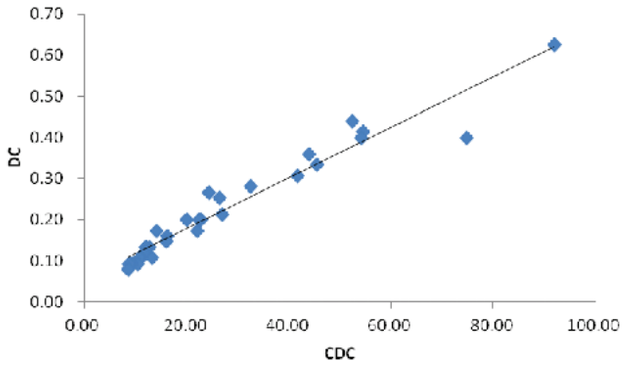

Figure 5. Regression analysis between Complex Degree Centrality (CDC) and Degree Centrality (DC).

Table 3. Regression analysis between CDC and DC

Model	Unstandardised Coefficients		F value	R^2
	B	Standard error	(p)	
CDC	0.006	0.000	444 (0.000)	0.941

Also, the Complex Degree Centrality (CDC) and Betweenness Centrality (BC) were related by an exponential function (Figure 6). The regression analysis showed that R^2 was 72.1%, the F value of variance analysis was 72.22, and the significance probability (p) was 0.000. Thus, the regression equation was statistically significant (Table 4).

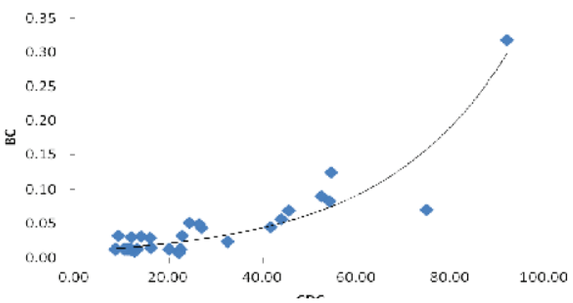

Figure 6. Regression analysis between Complex Degree Centrality (CDC) and Betweenness Centrality (BC).

Table 4. Regression analysis between CDC and BC

Model	Unstandardised Coefficients		F value (p)	R^2
	B	Standard error		
CDC	0.039	0.005	72.22 (0.000)	0.721

The regression analysis between the Degree Centrality (DC) and the Betweenness Centrality (BC) resulted in an exponential function. The regression analysis showed that R^2 was 78.0%, the F value of variance analysis was 99.31, and the significance probability (p) was 0.000. Thus, the regression equation was statistically significant (Figure 7, Table 5).

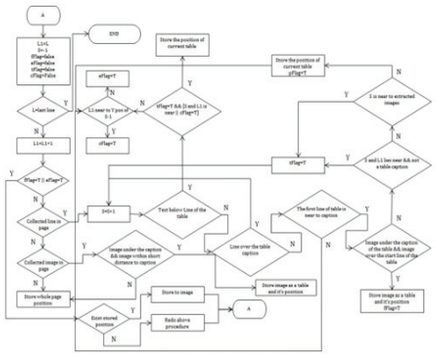

Figure 7. Regression analysis between Degree Centrality (DC) and Betweenness Centrality (BC).

Table 5. Regression analysis between DC and BC

Model	Unstandardised Coefficients		F value (p)	R^2
	B	Standard error		
DC	6.44	0.646	99.31 (0.000)	0.780

4. Conclusion

In this study, we used the information in papers from secondary battery areas to establish a network between countries, and we performed the analysis using various centrality indices. In the result, USA had the highest index of all centrality indices in the secondary battery areas. That is, USA was found to be performing joint research with the highest number of countries and the most diverse countries in the secondary battery area, playing the role of mediator in the joint research network. Korea was also found to be actively performing joint researches with USA, China, Japan, Australia and Italy.

When the complex degree centrality and betweenness centrality of each country were compared, China was found to be ranked relatively higher in complex degree centrality than in degree centrality, while Japan was found to have higher ranking in degree centrality than in complex degree centrality. This means that China is performing researches with relatively more diverse countries.

In the result of correlation analysis between centrality indices, the complex degree centrality, degree centrality, and betweenness centrality had a strongly positive (+) relationship. That is, when a country is performing various researches with various countries, the country plays the role of hub and mediator in the joint research network. In the result of the regression analysis between centrality indices, the complex degree centrality and degree centrality were related linearly, while the complex degree centrality and betweenness centrality were related exponentially.

In the correlation between the degree centrality and the betweenness centrality, the higher the degree centrality of a country was, the greater the betweenness centrality was in the corresponding joint research network. In other words, the country that conducts joint research with a larger number of countries in the joint research network transmits faster the latest technology in secondary batteries to other countries. This study is meaningful in that the correlations between complex degree centrality with other centrality indices were analysed in a joint research network analysis and valuable in that it investigated international joint research status in the secondary battery field which is growing these days.

5. Acknowledgement

This study was supported by the National Research Foundation of Korea Grant founded by the Korean Government (MSIP) (2013M1A2A 3000319).

6. References

1. Hwang KH, Joo IJ, Koh BY. An analysis of future skill demand using patent analysis – with a focus on the car batteries as industry convergence, Korea Research Institute for Vocational Education & Training, 2011.
2. Strogatz SH. Exploring complex networks. Nature. 2001; 410: 268–76.

3. Borgatti SP. Dynamic social network modeling and analysis, Workshop Summary and Papers. 2003. p. 241–52.
4. Sohn D-H. Social Network Analysis. Kyeongmunsa. 2009.
5. Leavitt HJ. Some Effects of Certain Communication Patterns on Group Performance [Unpublished Ph.D. thesis] Cambridge, MA: Massachusetts Institute of Technology; 1949.
6. Cohn BS, Marriott M. Networks and centres of integration in indian civilization. J Social Res. 1958; (1):1–9.
7. Pitts FR. A graph theoretic approach to historical geography. Prof Geogr. 1965; (17):15–20.
8. Czepiel JA. Word of mouth processes in the diffusion of a major technological innovation. J Market Res. 1974; 11(2):172–80.
9. Freeman LC. Centrality in social networks conceptual clarification, Social Networks. 1979; 1:215–39.
10. Lee JY. A study on document citation indicators based on citation network analysis, J Korean Soc Libr and Informat Sci. 2011; 45(2):119–43.
11. Lee JY. A study on the network generation methods for examining the intellectual structure of knowledge domains. J Korean Soc Libr Informat Sci. 2006; 40(2):333–55.
12. Lee JY. Centrality measures for bibliometric network analysis. J Korean Soc Libr and Informat Sci. 2006; 40(3):191–214.
13. Schvaneveldt RW. (Ed.). Pathfinder associative networks: Studies in knowledge organisations. Norwood: Ablex Pub. Corp; 1990.