Investigating the structure of regional innovation system research through keyword co-occurrence and social network analysis

PEI-CHUN LEE
Science and Technology Policy Research and Information Center, National Applied Research Laboratories, Taipei; Graduate Institute of Technology and Innovation Management, National Chengchi University, Taipei, Taiwan

HSIN-NING SU
Science and Technology Policy Research and Information Center, National Applied Research Laboratories, Taipei, Taiwan

ABSTRACT

Research on regional innovation systems (RIS) has evolved into a widely used analytical framework generating the empirical foundation for innovation policy making. The purpose of this research is to shed light on network-based author keyword analysis by integrating social network analysis and bibliometric analysis on the development of RIS research. A total of 432 papers belonging to 36 countries, 276 research institutes, and comprising 1165 keywords, are retrieved from the Web of Science databases for network construction and analysis. The obtained network in this study is capable of providing visual and quantitative insights into the publication trends or knowledge evolution of RIS. Network actors chosen in this study include country, research institute, first author, and keywords. These constitute four types of networks defined in this study: three research focus parallelship (RFP) networks (RFP-country network, RFP-institute network, RFP-author network) and one keyword co-occurrence (KCO) network.

Keywords: Regional innovation system; network theory; knowledge map; centrality; keyword co-occurrence network; research focus parallelship network

1. INTRODUCTION

The concepts of regional innovation system (RIS) have been developed into an important framework for evaluating regional performance in the knowledge-based economy from the early 1990s (Cooke, 1992; Cooke, 2001; Cooke & Morgan, 1994). The important elements and mechanisms of a innovation process have been investigated in view of regional-scaled development. Since the early 1990s, the concept of RIS has drawn much attention from policy makers. and it emerged at a time when policy focused toward systemic promotion of localized learning processes in order to establish the competitive advantage of regions (Asheim & Gertler, 2004). RIS approach has received considerable attention as a promising analytical framework for advancing the understanding of the innovation process in the regional economy (Asheim, Coenen, & Svensson-Henning, 2003; Cooke, Boekholt, & Tödtling, 2000; Leydesdorff, 1998).

A lot of attempts have been made to explore ways of mapping knowledge evolution. Author keyword (keywords specified by author), based analysis as a type of co-word analysis has started to play an important role in understanding the dynamics of knowledge development (Hori et al., 2004; Law & Whittaker, 1992; Edquist, 1997). Author keyword analysis is also used to supplement other analytical methods. For example, morphology analysis is a conventional method of forecasting future technology and identifying technology opportunities. Yoon and Park (2004) argued that morphology analysis is subject to limitations because there is no scientific or systematic way of establishing the morphology of technology. Therefore, keyword-based morphology analysis, which is supported by systematic procedures and quantitative data is thereby proposed as a method for conducting the morphology of technology.

Social network analysis based on keywords has been explored as well. Motter et al. (2002) constructed a conceptual network from the entries in a thesaurus dictionary and consider two words connected if they express similar concepts (Motter et al., 2002). He argued that language networks exhibit the small-world property as a result of natural optimization. Hence, these findings are important not only for linguistics, but also for cognitive science. Author keywords, by presenting the most important core concept of the articles' subject, could provide the information about which research trends are of most concern to researchers. The bibliometric method concerning author keywords analysis was developed recently, which uses the author keywords to analyze which trends of research are infrequent (Chiu & Ho, 2007). The technique of author keywords analysis might be a potential method for monitoring development trends or for the evolution of science, as well as for projecting future research directions.

2. Research question

The scope of a research field has to be constantly evaluated and redefined due to societal and environmental changes over time. In order to examine the fundamental building blocks of research fields and explore directions toward future research, researchers need to review the literature on a regular basis, and if necessary, modify the scope of research fields in order to obtain state-of-the-art insights.

What are the boundaries and contexts of RIS research? The objective of this study is derived from this research question and aims to analyze the academic literature of RIS research with bibliometric and network analysis to achieve the following purposes: 1) to present an overview of RIS research; and 2) to find the research contexts of RIS. To fulfil the aforementioned objectives, network visualization and analysis software is used to obtain a comprehensive overview of a large amount of literature. The results of this paper visually provide several networks as knowledge maps which define the scopes of RIS research, as well as network properties calculations for quantitatively mapping RIS research.

3. Theoretical background

Regional innovation system

The origin of the RIS concept rests in two main bodies of theory and research. The first body of literature is systems of innovation. Built on evolutionary theories of economic and technological change, the system of innovation literature conceptualizes innovation as an evolutionary and social process (2004).

Characteristic of a system approach to innovation is the acknowledgement that innovations are carried out through a network of various actors underpinned by an institutional framework. This dynamic and complex interaction constitutes what is commonly labelled systems of innovation, i.e. systems understood as interaction networks (Edquist, 1997; Kaufmann & Tödtling, 2002). There are increasing variations on this approach developed over time, either taking territories or specific technological clusters as their point of departure.

The second body of literature is regional science which deals both with the role of proximity and the focus on explaining the socio-institutional environment where innovation emerges. From a regional point of view, innovation is localized and locally embedded.

Approaches to RIS research are based on a territorial concept, and demonstrate that the innovation process in a region or cluster is rooted in the structure of the economy and cultural heritage, which includes strong elements of path dependency (Carlsson & Jacobsson, 1997). The context of RIS is therefore mainly characterized by actors located in a specific region rather than a specific sector and by interactions between actors.

According to Isaksen and Hauge's (2002) definition, RIS is: 'a concentration of interdependent firms within the same or adjacent industrial sectors in a small geographic area'. A RIS can, in principle, stretch across several sectors in the regional economy and is more lenient in terms of necessary conditions.

As long as there are firms and knowledge organizations that interact systematically, a RIS can be identified (Isaksen & Hauge, 2002). In brief, a RIS is characterized by co-operation in innovation activity between firms and knowledge creating and diffusing organizations, such as universities, R&D institutes, training organizations, technology transfer agencies, and so forth; and by the innovation supportive culture that enables both firms and systems to evolve over time.

Knowledge paradigm and evolution

When considering biological evolution, evolution is change in the inherited traits of a population of organisms from one generation to the next. When considering the change of human knowledge, knowledge does not exist statically, but emerges only within a context through interaction and evolution. Change also serves as a seed for the production of new knowledge. The scientific knowledge is dynamic; it is in constant evolution, like the knowledge itself. Kuhn (1970) introduced the idea of structure in science in the 1960s. His paradigms were snapshots of the structure of science at specific points in time. He argued that scientific literature is one approach for understanding both scientific evolution and the coming of a new paradigm (Kuhn, Dewey & Neurath, 1970).

Keyword network

The method of co-word analysis is a well-known relational bibliometric method. It originally was applied to make target-oriented retrieves, and later it was used to evaluate and present research outputs (Callon, Courtial & Laville, 1991). Today co-word analysis is often found in connection with information visualization. It allows the relational analysis of documents based on terms and term-groups.

However in this study, we apply the concept of co-word investigation to analyzing co-occurrence of keywords specified by authors, the author keywords. The co-occurrence of author keywords constitutes the network linkage or network ties established in this study. In this case, author keywords are used instead of words retrieved by text-mining to avoid the drawbacks of text mining, i.e. 1) extra effort of expert opinion is required to verify keywords retrieved by text-mining, and 2) text mining results are hard to be reproduced elsewhere due to the use of different coefficients or parameters in equations or computer software. In contrast, the network constructed by author keywords in this study does not require expert opinion and can be easily reproduced without any ambiguity regarding whether network actors have to be linked or not.

4. Methodology

This study integrates social network analysis and bibliometric keyword analysis to draw a picture for the development of RIS research/knowledge, which also can be called a 'regional innovation system knowledge map' where each country, research institute, or researcher that contributed to RIS literature can be positioned and analyzed.

The research process of this study is shown in Figure 1: 1) literature retrieval and filtering; 2) keyword revision and basis statistical analysis; 3) keyword network visualization; and 4) network properties calculation.

Networking of author keyword is based on sufficient relations among keywords. A relation is presented as a 'network tie' in a network. This study provides two methods of network tie generation. These two methods are defined below:

1. Relation between two different papers occurred because these two papers share at least one same keyword. A network generated by this method is defined as RFP-network (research focus parallelship network).
2. Relations among plural keywords occurred because these keywords are listed in the same paper by an author. A network generated by this method is defined as KCO-network (keyword co-occurrence network)

Literature retrieval and filtering

For this research RIS is the research target, trying to map the knowledge evolution of RIS. However, alternative terminologies other than RIS are possibly used in literature; e.g., 'regional system of innovation', 'industrial cluster', or just 'regional innovation'. To reach maximum and precise coverage, the term 'regional innovation' is used for paper retrieval from the literature database. Science Citation Index (SCI) and Social Sciences Citation Index (SSCI) literature databases of the Web of Science are used for paper retrieval, since SCI and SSCI are both generally regarded as important indicators for quality papers. Papers with the term 'regional innovation' in their title or abstract are retrieved from the Web of Science database. Retrieved papers are then carefully reviewed, and those not related to the field of RIS are precluded. Finally, a total of 432 papers are obtained from the Web of Science. The paper retrieval time is Jan. 01, 2009.

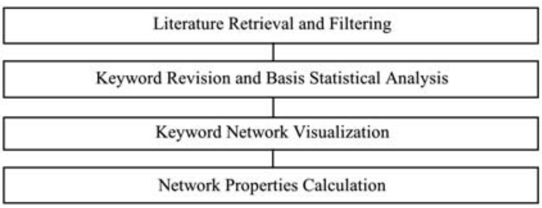

FIGURE 1. RESEARCH PROCESSES

The 432 papers are from 73 journals, and the top 10 journals that publish more than 10 papers are listed in Table 1.

Keyword revision and basis statistical analysis

Due to the fact that different words can be used for describing the same meaning, it is necessary to standardize words – for example: 1) words are standardized to their singular forms; 2) technique, technologies, technology are standardized to technology; and 3) regional systems of innovation, RIS, and industrial cluster are standardized to 'regional innovation system'. A total of 1165 keywords are obtained after standardization. The top 30 high occurrence keywords are listed in Table 2.

TABLE 1: STATISTICS OF PAPERS FROM TOP 10 JOURNALS

Rank	Top 10 journals	No. of papers	%
1	Regional Studies	109	25.23
2	Research Policy	56	12.96
3	International Journal of Technology Management	39	9.02
4	Technovation	26	6.01
5	Tijdschrift voor Economische en Sociale Geografie	16	3.7
6	Journal of Economic Geography	16	3.7
7	European Urban and Regional Studies	15	3.47
8	Economic Geography	11	2.54
9	Papers in Regional Science	11	2.54
10	Journal of Evolutionary Economics	10	2.31

TABLE 2: TOP 30 HIGH OCCURRENCE KEYWORDS

Rank	Keyword (total 1165)	Occurrence
1	Innovation	102
2	Regional development	32
3	Regional innovation system	32
4	Cluster	30
5	Regions	27
6	R&D	26
7	Network	23
8	Biotechnology	19
9	Entrepreneurship	17
10	Innovation system	17
11	Agglomeration	16
12	Knowledge spillovers	16
13	Spillover	16
14	Knowledge	15
15	Regional growth	15
16	Innovation policy	14
17	Patents	13
18	Technology transfer	12
19	SME	11
20	Economic geography	10
21	Industrial district	10
22	Competitiveness	9
23	Regional policy	9
24	China	8
25	Economic development	8
26	Embeddedness	8
27	Regional innovation	8
28	Social capital	8
29	Globalization	7
30	Learning	7

Definition and visualization of keyword network

The method of establishing networks in this study is based on author keywords and keyword co-occurrence, which is further divided into two types of networks, RFP network and KCO network. These are defined below:

1. **RFP network** (Research focus parallelship network): The relation between two different papers occurred because these two papers share at least one keyword. For example, a paper is used as a network actor (network node) and any of two actors sharing one same keyword will be linked. This is based on an assumption made in this study that keywords represent the core research of a paper and sharing the same keyword implies these two papers partially overlap. The two papers are thus regarded as a pair of parallel papers and the constructed network is defined as RFP-network (research focus parallelship network). However, the network node is not necessarily the paper, it can also be selected from different actors, e.g. first author, research institute, country, by which papers are published. Hence, there are three types of RFP networks in this study, noted below:
   - RFP-country network: Research focus parallelship network with country as the network actor
   - RFP-institute network: Research focus parallelship network with research institute as the network actor
   - RFP-author network: Research focus parallelship network with first author as the network actor

   In this study, RFP-country, RFP-institute, and RFP-author networks are investigated in order to understand parallelship of knowledge evolution of RIS at macro, meso and micro levels, respectively.

2. **KCO network** (Keyword co-occurrence network): Relations of author keywords are formed because author keywords specified by authors are listed in the same paper. Author keywords listed in the same papers are linked together because they are all terms that can be used to represent the core of a research paper and stronger relations to each other can be expected. Keywords in the same paper share equal importance for the paper. Author keywords with higher network centrality are those closer to the core of knowledge of RIS.
   - KCO network: Keyword co-occurrence network

In this study, KCO network is investigated in order to understand co-occurrence of keywords in RIS papers at micro level.

Network properties calculation

Computer software is used to visualize RFP network and KCO network and then network properties are subsequently calculated. In social network theory, centrality is used to estimate the influence of actors. Centrality as an indicator can be used to understand to what degree an actor is able to obtain or control resources. Brass and Burkhardt (1992) indicated that network centrality is one source of influence from the viewpoint of organizational behaviour. A person with higher centrality in an organization is always the one with higher influence (Brass & Burkhardt, 1992). Freeman (1979) suggested three methods of centrality measurement for a network: 1) degree centrality, 2) between centrality, and 3) closeness centrality (Freeman, 1979). Network properties are calculated by the above three methods in this study in order to understand the power of influence of first author, research institute, and country. A social network can be either a directed network or an undirected network, but networks constructed in this research are undirected networks because there no in-and-out concept exists behind any linked keywords, e.g. no causal relation, position difference, resource exchange, flows, or diffusion.

1. *Degree centrality*

Network nodes (actors) which are directly linked to a specific node are in the neighborhood of that specific node. The number of neighbors is defined as the nodal degree, or degree of connection. Granovetter (1973) suggested that nodal degree is proportional to probability of obtaining resource (Granovetter, 1973). Nodal degree represents the degree a node (actor) participates the network. This is a basic concept for measuring centrality. Degree centrality is defined as the number of direct linkages between actor i and other actor.

$$d(i) = \sum_j m_{ji}$$

$m_{ij} = 1$ if actor i and actor j are linked

2. *Between centrality*

The concept of betweenness is a measure of how often an actor is located on the shortest path (geodesic) between any other two actors in the network. Those actors located on the shortest path between other actors are playing roles of intermediary that help any two actors without direct contact reach each other indirectly. Actors with higher between centrality are those located at the core of the network.

$$b(i) = \sum_{j,k\neq 1} \frac{g_{jik}}{g_{jk}}$$

g_{jk} : the shortest path between actor j and actor k

g_{jik} : the shortest path between actor j and actor k that contains actor i

3. *Closeness centrality*

The closeness centrality of an actor is defined by the inverse of the average length of the shortest paths to/from all the other actors in the network. Higher closeness centrality indicates higher influence on other actors.

$$c(i) = \sum_{j=1}^{N} \frac{1}{d_{ji}}$$

d_{ji} : the shortest path between actor j and actor i

5. Results and discussion

Paper sample analysis

Figure 2 shows the number of retrieved RIS papers published annually and a gradual increase can be observed.

Among all the retrieved 432 papers, USA is the country with the largest number of papers (76), then UK (61), Germany (51), Netherland (35), Canada (20), Spain (19), Wales (18), Italy (15), etc. (Table 3). A total of 36 countries have publication involved in this research.

Research institutes which publish papers total of 276. Table 4 lists research institutes that pub-

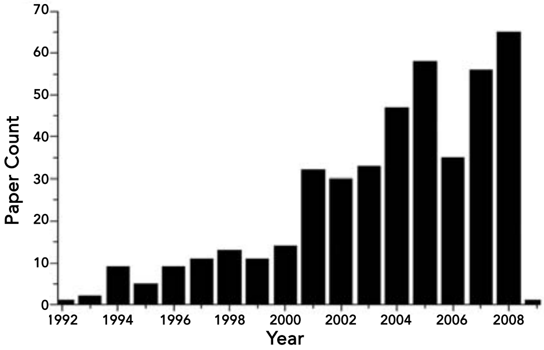

FIGURE 2. NO. OF PAPERS PUBLISHED EACH YEAR

lish more than three papers. The research institute with the largest number of paper is Cardiff University.

Keyword network analysis

Keywords of the retrieved 432 papers are used as the basis for network construction to obtain three RFP networks by the use of different network actors, e.g. country, research institute, first author

TABLE 3: PAPER PUBLISHED BY DIFFERENT COUNTRIES WITH MORE THAN THREE PAPERS

Rank	Countries (total 36)	No. of papers
1	USA	76
2	England	61
3	Germany	51
4	Netherlands	35
5	Canada	20
6	Spain	19
7	Wales	18
8	Italy	15
9	Sweden	15
10	France	14
11	Austria	12
12	China	11
13	Finland	10
14	Korea	10
15	Denmark	8
16	Australia	6
17	Belgium	6
18	Portugal	5
19	New Zealand	4

TABLE 4: TOP RESEARCH INSTITUTE WITH HIGH PAPER QUANTITY WITH MORE THAN THREE PAPERS FROM A TOTAL OF 276 RETRIEVED PAPERS

Rank	Research institute	No. of papers
1	Cardiff University	16
2	University Cambridge	8
3	University Utrecht	8
4	University Toronto	7
5	Fraunhofer Institute Syst & Innovat Research	6
6	Coventry University	5
7	University Reading	5
8	Eindhoven University Technology	5
9	Unavailable	4
10	Queens University Belfast	4
11	Max Planck Institute Economics	4
12	University Amsterdam	4
13	Georgetown University	4
14	SUNY Buffalo	4
15	University California Los Angeles	4
16	Vienna University Econ & Business Admin	3
17	Queens University	3
18	Oxford Brookes University	3
19	University Manchester	3
20	University Newcastle Upon Tyne	3
21	University Sussex	3
22	University Warwick	3
23	Technical University Bergakad Freiberg	3
24	University Jena	3
25	Technion Israel Institute Technology	3
26	Seoul National University	3
27	Erasmus University	3
28	Chalmers University Technology	3
29	Lund University	3
30	George Washington University	3
31	Georgia Institute Technology	3
32	Harvard University	3
33	Indiana University	3
34	Stanford University	3
35	University Michigan	3

(individual paper), and KCO network by the use of keyword as network actor.

1) RFP-country network: Papers are classified by country, and any pair of countries with the

same keyword(s) is/are linked together. A total of 36 network actors and 409 network ties are obtained. As shown in Fig. 3, American and European countries are the major countries that contribute the most to this field; Asia-Pacific countries are China, Korea and Australia.

2) RFP-institute network: Papers are classified by research institute, and any pair of research institutes with the same keyword(s) are linked together. A total of 276 network actors and 5692 network ties are obtained and shown in Fig. 4.
3) RFP-author network: Any two first authors with the same keyword are linked together. A total of 432 network actors and 9031 network ties are obtained and shown in Fig. 5.
4) KCO network: Each keyword is treated as a network actor; keywords listed within the same papers are linked together. A total of 1165 network actors and 3543 network ties are obtained (Fig. 6). Several small groups of separated sub-networks are on the right side of Fig. 6. This is because the keywords for those sub-networks are different from those in the major network; so, keywords of the same paper in these sub-networks are mutually linked without connection to the major network and shown as isolated sub-networks.

Network properties calculation

Network properties are analyzed on these obtained four networks to obtain degree centrality, between centrality, and closeness centrality of the network actors. Therefore, a knowledge evolution map, in terms of country, research institute and first author, can be obtained, positioning each actor in RIS research field can be quantitatively and visually identified.

For RFP-country network (Fig. 3), countries with top 10 network properties are listed in table 5. Germany has the highest centrality and then Korea, Netherlands, England, Finland, and Spain. The number of papers that each country

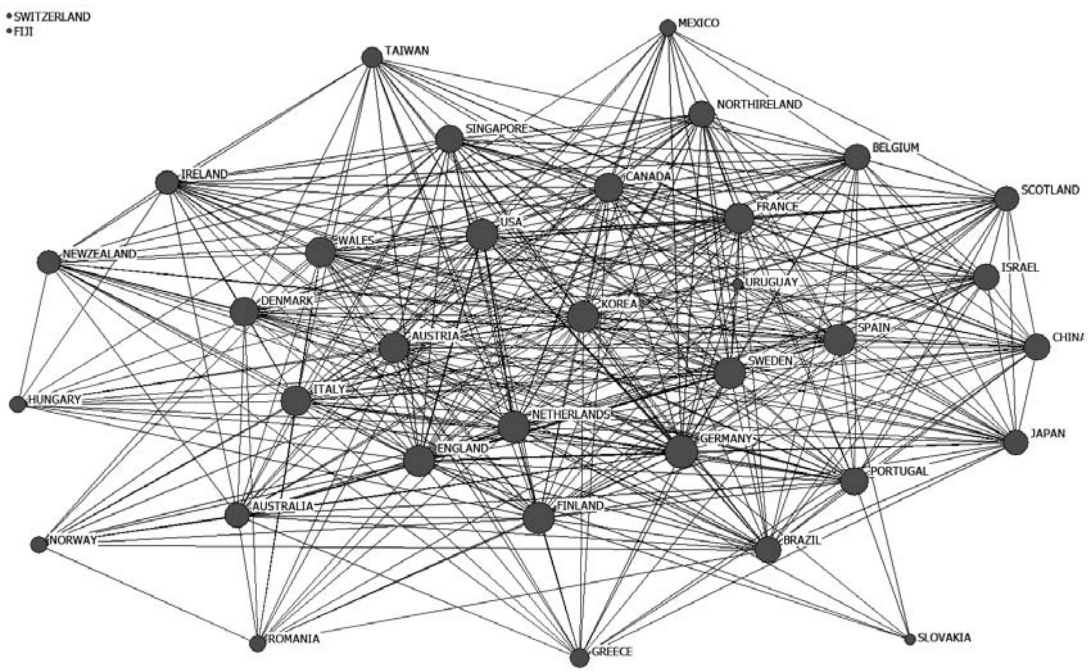

FIGURE 3. RFP-COUNTRY NETWORK: RESEARCH FOCUS PARALLELSHIP NETWORK WITH COUNTRY AS NETWORK ACTOR.

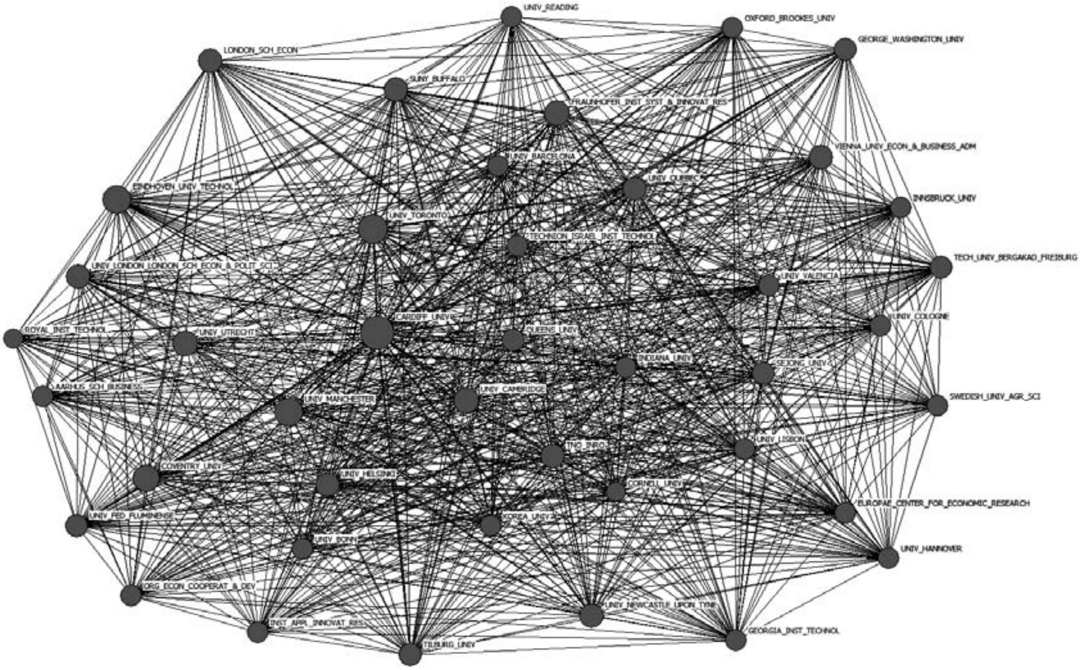

Figure 4. RFP-institute network: research focus parallelship network with research institute as network actor (only shows actors with degree centrality larger than 90).

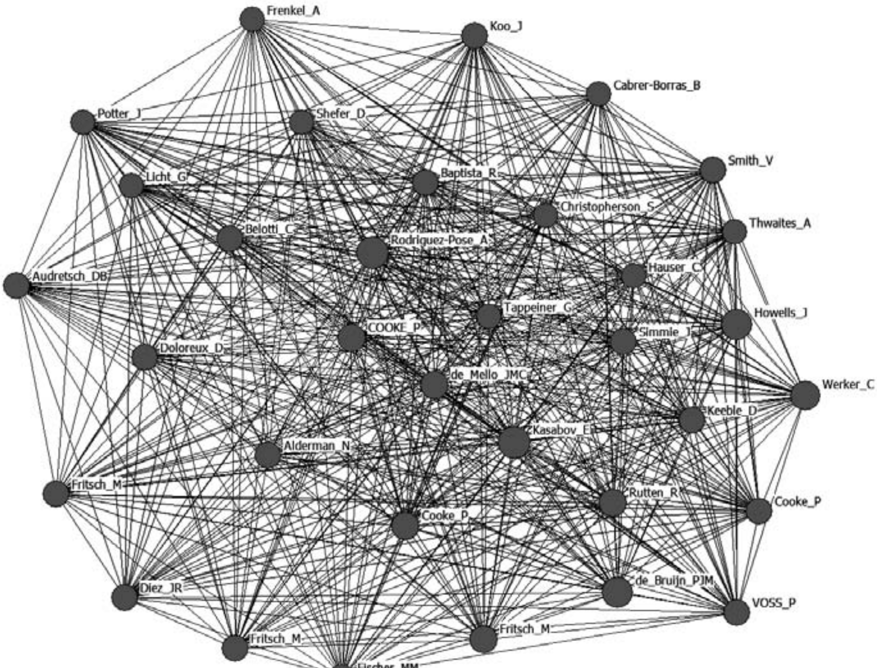

Figure 5: RFP-author network: research focus parallelship network with first author as network actor (only shows actors with degree centrality larger than 120).

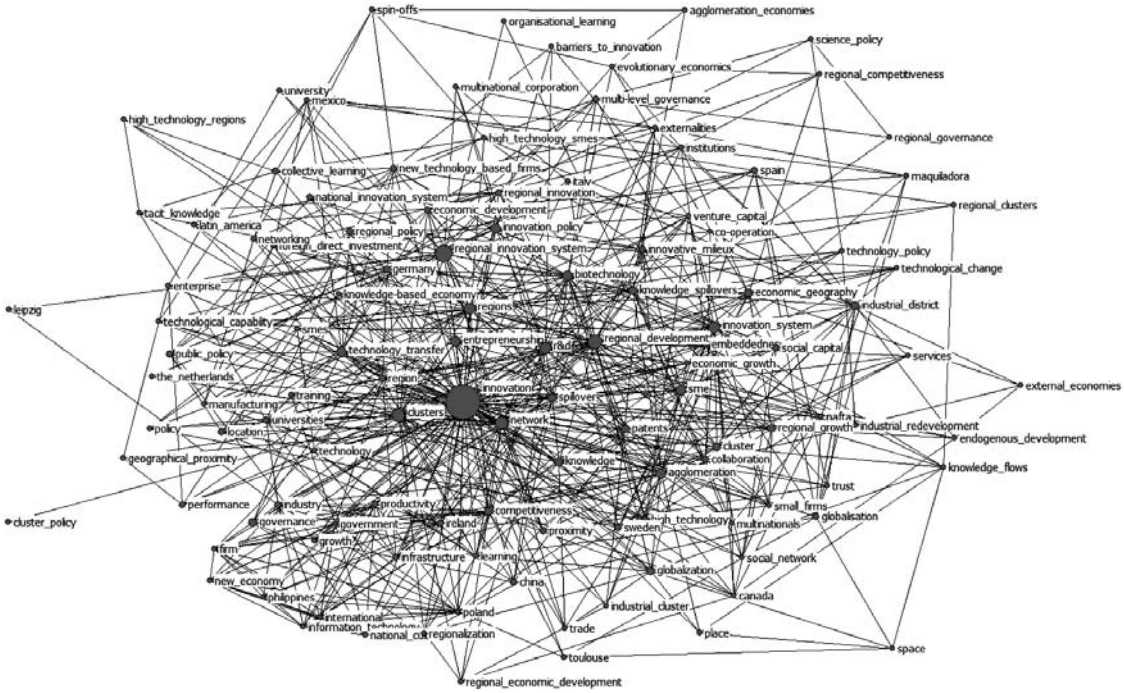

Figure 6: KCO network- keyword co-occurrence network with keyword as network actor (only shows actors with degree centrality larger than 10).

contributes to this field is different, but it is easily anticipated that countries with more papers tend to have more linkages to other countries by their larger number of papers. Hence, countries with more papers have higher centrality, and are thus positioned at the core of the network. Countries with more papers shown in Fig. 3 are consistent to countries with higher centrality calculated in Table 5. However, the number of papers for US is ranked no. 1, but the centralities are ranked no. 8 while England ranks no. 2 in papers and no. 4 in centralities. European countries have higher centralities than the US. The ranking of paper and centrality measurement for Korea is outstanding and it is also the only Asian country ranked within the top 10.

Table 5: Top 10 countries with highest network properties

Rank	Degree centrality	Between centrality	Closeness centrality
1	Germany (51)	Germany (51)	Germany (51)
2	Korea (10)	Spain (19)	Spain (19)
3	Netherlands (35)	Netherlands (35)	Netherlands (35)
4	England (61)	England (61)	England (61)
5	Finland (10)	Korea (10)	Korea (10)
6	Spain (19)	Sweden (15)	Sweden (15)
7	Sweden (15)	Finland (10)	Finland (10)
8	USA (76)	USA (76)	USA (76)
9	Italy (15)	France (14)	France (14)
10	Austria (12)	Austria (12)	Austria (12)

Note: the number within bracket indicates the number of publications of the country.

For RFP-institute network (Fig. 4), institutes with top 10 network properties are listed in Table 6. The highest centralities are Cardiff University, University of Toronto, University of Manchester, University of Utrecht, etc. Similar to the previous country observation, research institutes with more papers are also supposed to have higher centralities. Research institutes with more papers shown in Table 4 are consistent to those with higher centralities calculated in Table 6.

For RFP-author network (Fig. 5), first authors with top 10 centralities are shown in Table 7, Rodriguez-Pose, Kasabov, Howells, de Bruijn, Werker, Rutten, etc. The author with the top degree centrality is Andrés Rodríguez-Pose from the Department of Geography and Environment, London School of Economics (LSE). His are: no. 1 degree centrality, no. 1 between centrality, and no. 2 closeness centrality. This researcher and his affiliation (the degree centrality of the Department at LSE is ranked no. 9) and country (three centrality measurements for England are all ranked no. 3) all perform pretty well.

For keyword co-occurrence network with keyword as network actor, Fig. 6, keywords with the top 20 centralities are listed in Table 8. Due to the research target-RIS set in this study, the implications of keywords with higher network centralities are expected to be lexically similar to the term 'cluster'. This is why RIS, regional development, cluster, network, regions, etc. are found to be top keywords in table 8, and the other keywords in Table 8 all indicate implications of technology or innovation. A total of 21 keywords in Table 8 can be categorized into four groups: 1) region names; 2) issues; 3) policies; and 4) the technology field. Only China is in the first group, region name, implying that China is the region with the most attention from RIS research. The second group is issues, e.g. knowledge spillovers, entrepreneurship, regional growth, patents, Small and medium enterprises (SME), globalization, etc. Those are all keywords directly related to regional development activities and match the research target of this study. This also implies that these issues have a relatively important association with RIS research. The third group is policy related such as innovation policy or regional policy. The fourth group is related technology field such as

Table 6: Top 10 research institutes with highest network properties

Rank	Degree centrality	Between centrality	Closeness centrality
1	Cardiff University (16)	Cardiff University (16)	Cardiff University (16)
2	University Toronto (7)	University Utrecht (8)	University Toronto (8)
3	University Manchester (3)	University Manchester (3)	Eindhoven University Technology (5)
4	Eindhoven University Technology (5)	George Washington University (3)	University Manchester (3)
5	Coventry University (5)	Fraunhofer Institute Syst & Innovat Research (6)	Coventry University (5)
6	University Cambridge (8)	University Toronto (7)	University Utrecht (8)
7	Fraunhofer Institute Syst & Innovat Research (6)	Coventry University (5)	University Cambridge (8)
8	University Utrecht (8)	University Cambridge (8)	Fraunhofer Institute Syst & Innovat Research (6)
9	London School Economics (1)	Eindhoven University Technology (5)	University Helsinki (2)
10	University London London School Economics & Political Science (2)	University Newcastle Upon Tyne (3)	London School Economics (1)

Note: the number within bracket indicates the number of publications of the research institute.

TABLE 7: TOP 10 AUTHORS WITH HIGHEST NETWORK PROPERTIES

Rank	Degree centrality	Between centrality	Closeness centrality
1	Rodriguez-Pose A	Rodriguez-Pose A	Kasabov E
2	Kasabov E	Howells J	Rodriguez-Pose A
3	De Bruijn PJM	De Bruijn PJM	Howells J
4	Howells J	Kasabov E	De Bruijn PJM
5	Werker C	Werker C	Werker C
6	Rutten R	Audretsch DB	Rutten R
7	Cooke P	Chung SY	De Mello JMC
8	Cooke P	Carayannis EG	Cooke P
9	Fritsch M	Licht G	Fritsch M
10	De Mello JMC	Belotti C	Keeble D

biotechnology, implying biotechnology development receives higher concern in RIS research.

6. DISCUSSION AND CONCLUSION

RIS research has been undertaken since the early 1990's, and this study provides an analytical method to visually map and quantitatively analyze the context, e.g. publication trend, knowledge evolution, and research focus, of global RIS research. A total of four types of networks: 1) RFP-country network; 2) RFP-institute network; 3) RFP-author network; and 4) KCO-network are proposed in this study with country, institute, author, and keyword as network actors, in order to allow an understanding of different scale RIS research contexts. The research contexts of RIS research at different levels provide full-spectrum implications to sectors of government, industry,

TABLE 8: TOP 20 KEYWORDS WITH HIGHEST NETWORK PROPERTIES

Rank	Degree centrality	Between centrality	Closeness centrality
1	Innovation	Innovation	Innovation
2	Regional innovation system	Regional development	Regional development
3	Regional development	Regional innovation system	Biotechnology
4	Clusters	Clusters	Clusters
5	Network	Regions	Network
6	R&D	R&D	Regional innovation system
7	Regions	Biotechnology	Regions
8	Innovation system	Network	Innovation policy
9	Biotechnology	Innovation system	R&D
10	SME	Regional growth	Knowledge spillovers
11	Agglomeration	SME	Spillover
12	Entrepreneurship	Entrepreneurship	SME
13	Technology transfer	Innovation policy	Entrepreneurship
14	Innovation policy	Agglomeration	Agglomeration
15	Industrial district	Globalization	Knowledge
16	Knowledge spillovers	Knowledge spillovers	Innovative milieu
17	Spillover	Technology transfer	Competitiveness
18	Regional growth	Industrial district	Social capital
19	Competitiveness	Regional policy	Regional growth
20	Globalization	Economic geography	Globalization

Table 9: Classification for top 20 keywords with highest network properties

Classification	Core keyword (top 20 keywords with highest network properties)
Region name	China
Issue	Knowledge spillovers, R&D, cluster, network, entrepreneurship, regional growth, patents, technology transfer, SME, economic geography, industrial district, competitiveness, economic development, embeddedness, social capital, globalization, innovative milieu
Policy	Innovation policy, regional policy
Technology field	Biotechnology

academy, and research institutes. This research avoids the use of complicated text-mining techniques which may lead to ambiguity due to the selection of the coefficients or parameters used in equations or computer software. Further, expert opinion for distinguishing keywords and stop words are not necessary because keywords specified by authors in each paper are used for identifying network linkage/tie, and thus the controversy over keyword identification is avoided. Furthermore, the method is a quick and straightforward approach for investigating evolution of plural actors that can be linked as a network.

However, there are two important issues to be noted as limitations in relation to paper sampling in this study:

1) This study maps knowledge evolution for RIS research by the use of paper retrieved from the Web of Science database. But research results might be a bit biased because RIS articles retrieved in this study are mainly English. It is obvious not fair even though the database covers most of the important journals in the world. Also, the term RIS used as the research target might be expressed differently in other papers which are not considered in this study.
2) Due to the fact that the method proposed by this study only considers author keywords, it might lead to unavoidable limitations. There are usually only 3–6 keywords specified by author in a paper, some papers do not even have such author keywords, and the 3–6 author keywords might not be sufficient to cover the core concepts of a paper even thought they are likely to be the most important keywords carefully selected by author. To avoid this problem, we expect the limitation can be minimized if a relatively large number of sample papers can be used for constructing networks when investigating macroscopic or country and industry-level analysis where detailed insight into a single paper is not quite necessary.

Therefore, this research is experimental in character and the total 432 papers retrieved in this study may or may not be sufficient to reach the scale of macroscopic analysis. However, the main purpose of this study is to propose and demonstrate a method in which investigating full-spectrum implications of RIS knowledge evolution is possible. Obtaining knowledge evolution of RIS research by different methods and comparing the results obtained in this study is desireable to not only correlate research results from different approaches but also to evaluate whether 432 papers are enough for future macro-scaled investigations.

The four types of networks investigated in this study can be future advanced by the using computers if detailed insight on network structure is necessary. For example, symbols of network actors or thickness/length of network ties can be a function of network attributes to obtain more information from the same network structure (Cambrosio, Keating, & Mogoutov, 2004). The research structure can be converted to a two dimensional contour map by the use of Van Eck and Waltman's algorithm (Van Eck & Waltman, 2007; Lee, Su, & Wu, 2010).

The network linkage obtained in this study is based on author keywords because author keywords are carefully selected by authors to represent the core concepts of research papers.

However, in addition to author keywords, there is still an abundance of rich and complex information that can be extracted from research papers. It will be useful to implement content analysis which is an important advance in network analytics (Smith & Humphreys, 2006; Criscuolo, Salter, & Sheehan, 2007; Tseng, Lin & Lin, 2007), to extract more information from research papers. But the issue is that different content analysis algorithms lead to different results and make objective evaluation difficult.

To advance our investigation, what can be done in the future are:

1) The similarity between linked network actors can be calculated, so the obtained similarity can be used as a function of network tie, e.g. thickness of tie and length of tie.
2) Network properties such as degree centrality, between centrality, or closeness centrality can be used as network actors' properties such as actors' nodal size or colour, e.g. network actor nodal size can be proportional to its Degree Centrality, to allow more informative visualization.
3) Measuring and predicting the future of various branches of knowledge is worth discussion. From the obtained evolution context, the scenario for RIS research in the future can possibly be projected.
4) Implement content analysis by text-mining technique to retrieve more information from a large number of research papers.

References

Asheim, B., Coenen, L., & Svensson-Henning, M. (2003). *Nordic SMEs and regional innovation systems*. Final Report, Nordic Industrial Fund, Oslo.

Asheim, B., & Gertler, M. (2004). Understanding regional innovation systems. In: J. Fagerberg, D. Mowery and R. Nelson eds *Handbook of Innovation*. Oxford University Press, Oxford.

Brass, D. J., & Burkhardt, M. E. (1992). Centrality and power in organizations, In: N. Nohria & R. G. Eccles (Eds.), *Networks and organizations: Structure, form and action*, pp. 191-215. Harvard Business School Press, Boston MA.

Callon, M., Courtial, J. P., & Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemsitry. *Scientometrics*, *22*(1), 155-205.

Cambrosio, A., Keating, P., & Mogoutov, A. (2004). Mapping collaborative work and innovation in biomedicine: A computer-assisted analysis of antibody reagent workshops. *Social Studies of Science*, *34*(3), 325.

Carlsson, B., & Jacobsson, S. (1997). Diversity creation and technological systems: a technology policy perspective. In: Edquist, C. (Ed.), *Systems of innovation: technologies, institutions and organizations*, pp. 266-294. Pinter, London.

Chiu, W. T., & Ho, Y. S. (2007). Bibliometric analysis of tsunami research. *Scientometrics*, *73*(1), 3-17.

Cooke, P. (1992). Regional innovation systems: competitive regulation in the new Europe. *Geoforum*, *23*, 365-365.

Cooke, P. (2001). Regional innovation systems, clusters, and the knowledge economy. *Industrial and Corporate Change*, *10*(4), 945-974.

Cooke, P., Boekholt, P., & Tödtling, F. (eds) (2000). *The governance of innovation in Europe - regional perspectives and global competitiveness.* Pinter, London.

Cooke, P., & Morgan, K. (1994). The regional innovation system in Baden-Württemberg. *International Journal of Technology Management*, *9*(3), 394-429.

Criscuolo, P., Salter, A., & Sheehan, T. (2007). Making knowledge visible: Using expert yellow pages to map capabilities in professional services firms. *Research Policy*, *36*(10), 1603–1619.

Edquist, C. (1997). *Systems of innovation: technologies, institutions and organizations*. Routledge, London.

Edquist, C. (2004). Systems of innovation–a critical review of the state of the art. In: J. Fagerberg, D. Mowery and R. Nelson eds *Handbook of Innovation.* pp. 181–208. Oxford University Press, Oxford.

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

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1360-1380.

Hori, K., Nakakoji, K., Yamamoto, Y., & Ostwald, J. (2004). Organic perspectives of knowledge management: Knowledge evolution through a cycle of knowledge liquidization and crystallization. *Journal of Universal Computer Science*, *10*(3), 252-261.

Isaksen, A., & Hauge, E. (2002). *Regional clusters in Europe–Observatory of European SMEs report No. 3*. European Communities, Luxembourg.

Kaufmann, A., & Tödtling, F. (2002). How effective is innovation support for SMEs? An analysis of the region of Upper Austria. *Technovation*, *22*(3), 147-159.

Kuhn, T. S., Dewey, J., & Neurath, O. (1970). *The structure of scientific revolutions*. University of Chicago Press, Chicago.

Law, J., & Whittaker, J. (1992). Mapping acidification research: A test of the co-word method. *Scientometrics*, *23*(3), 417-461.

Lee, P. C., Su, H. N., & Wu, F. S. (2010). Quantitative mapping of patented technology: The case of electrical conducting polymer nanocomposite. *Technological Forecasting & Social Change*, 77 (3), p.466-478

Leydesdorff, L. (1998). Reply about using co-words. *Journal of the American Society for Information Science*, *49*(1), 98-99.

Motter, A. E., de Moura, A. P. S., Lai, Y. C., & Dasgupta, P. (2002). Topology of the conceptual network of language. *Science Phys Rev E*, *65*, 065102.

Smith, A. E., & Humphreys, M. S. (2006). Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping. *Behavior Research Methods*, *38*(2), 262.

Tseng, Y. H., Lin, C. J., & Lin, Y. I. (2007). Text mining techniques for patent analysis. *Information Processing and Management*, *43*(5), 1216–1247.

Van Eck, N. J., & Waltman, L. (2007). Bibliometric mapping of the computational intelligence field. *International Journal of Uncertainty Fuzziness and Knowledge Based Systems*, *15*(5), 625-645.

Yoon, B. C., & Park, Y. T. (2004). Morphology analysis approach for technology forecasting. In *2004 IEEE International Engineering Management Conference, 2004. Proceedings* (Vol. 2).

Received 19 May 2009 *Accepted 29 March 2010*