Mapping the development in Information Retrieval specialty:
A bibliometric analysis via journals

Ying Ding, Gobinda G. Chowdhury and Schubert Foo
Division of Information Studies, School of Applied Science
Nanyang Technological University, Nanyang Avenue,
Singapore 639798
P143387632@ntu.edu.sg; Asggchowdhury@ntu.edu.sg; assfoo@ntu.edu.sg

Abstract

Analysis of inter-journal citations permits objective evaluation of the network of journals in and around a discipline. The objective of this study is to map the journal's role in the scholarly communication process of IR (Information Retrieval) field and their changes over time by using journal cocitation analysis during the period of 1987-1997. Data are retrieved from Social Science Citation Index (SSCI) and are analyzed (i) covering all the journals in various disciplines which were highly cited by IR researchers and (ii) Library and Information Science journals which were highly cited by IR researchers. The results of this study show that (i) the scholarly communication positions of the journals which are from different areas play an important role in the IR field, (ii) the field of IR is a mature field because the journals used for research communication remained quite stable during the study period. This research also identifies those LIS journals which are highly used for research communication by the IR researchers.

Introduction

A scientific specialty can be defined as a network of people working on the same set of research questions, using the same methods and referring to the same scientific literature (Price 1965). Researchers within a specialty communicate more with one another than with researchers in other communities, and they are expected to refer to one another's work significantly more frequently than to the work of outsiders.

Bibliometric analysis is one of several lines of research which has been used to investigate the nature of changes or development in scholarly activities and associated changes in the intellectual, social, or cognitive structure of scientific specialties (Narin 1976). In the early 1970s, Small (1973) introduced the concept of cocitation analysis as potentially productive method of examining the cognitive/intellectual structure of scientific specialties. From then on, communication patterns of scientific specialties have been studied in terms of co-citations among authors, articles, journal or keywords (Small & Griffith 1974; McCain 1984 1991; Braam, Moed and van Raan 1991). Journals play a vital role in the spread of information within or between specialties. And these journal-to-journal networks form another relation in the fabric of science (Doreian 1988). In a number of studies, it has been shown that journal-to-journal citations can be used as an operational indicator for the disciplinary organization of the sciences (e.g. Cason and Lubotsky 1936; Doreian 1988; Tijssen & van Raan 1990; Everett and Pecotich 1991; McCain 1991; Borgman & Rice 1992; Leydesdorff & Cozzens 1993).

Information Retrieval (IR) is a very important sub-field of library and information science. During 1980s and 1990s, with the rapid development of personal computers and networking through email and the Internet into "information highway", the number and variety of information sources available in different media have increased dramatically. Efficient use of the information resources and the means of choosing the 'best' information are becoming increasingly significant. So, IR is a significant research area, which has profound influence on the world.

Current study is one part of a research project that aims to look into the process of information transfer in the field of information retrieval (IR) through bibliometric analysis of literature. In the first phase of this work, contributions of significant authors in the field of IR, over a period of 11 years 1987--1997, have been analyzed (Ding 1998). Some results of this phase have been published (Ding 1998a) while others in the pipeline (Ding 1998b and 1998c). This paper analyses the role of journals in the process of research communications in IR by using journal cocitation analysis.

The objective of the current research is to map the journal's role in the scholarly communication process of IR field and their changes over time by using journal cocitation analysis during the period of 1987-1997. This study has been conducted from two angles: one from the various journals which were highly cited by IR researchers and the other from the library and information science journals which were highly cited by IR researchers. So, we can not only trace the scholarly communication positions of the journals in IR field which are from different areas and play an important role in IR field, but also show the status of Library and Information Science journals in IR field which most IR researchers are interested in.

Methods
Data gathering

The main bulk of IR research has been carried out by researchers from the disciplines of library and information science, computer science and other smaller related disciplines. Here, we limited our research scope of IR in Library and Information Science (LIS) because of voluminous amount of data that is available for this period of study. So the relevant IR papers were selected from SSCI which includes all the LIS journals (Social Science Citation Index) from 1987 to 1997. A total of 1,466 IR-related papers were selected from 367 journals with 44,836 citations. Two databases (**Source Database** and **Citation Database**) were set up based on these IR papers.

Results

Different journals display different status in specific specialty. In current research, we will discuss the journal's role in scholarly communication process of IR field from two points of view. First, let us look at the role of various journals from different areas, which were highly cited by IR researchers.

Various journals which were highly cited by IR researchers
<u>Journal sample</u>

Fifty highly cited various journals (including some highly cited regular annual publication and proceedings, henceforth called VAR journals) from different specialties were selected as the research sample (Appendix 1). Then the journal cocitation matrix was built up based on the frequencies of these journals' cocitations. The data in the matrix is each journal's profile of cocitation with every other journal on the list. The raw journal cocitation matrix was converted into Pearson correlation matrix indicating the similarity or dissimilarity of each journal pair. This correlation matrix was analyzed using an agglomerative clustering routine (SPSS hierarchical cluster with Ward's method) and multidimensional scaling (MDS) (SPSS ALSCAL) (Norusis 1997). Each cluster was named based on the nature and research scope of these journals in the cluster.

<u>Journal cocitaiton maps</u>

In order to exhibit the VAR journals' scholarly communication positions during the period of 1987-1997 and their changes over this period, the MDS maps for the whole 11-year period and for two different periods (1987-1991, 1992-1997) were generated by SPSS with good fit (Figures 1, 2 and 3). The groups with solid lines are based on the clustering techniques in SPSS. Journals with high correlation coefficient were located together.
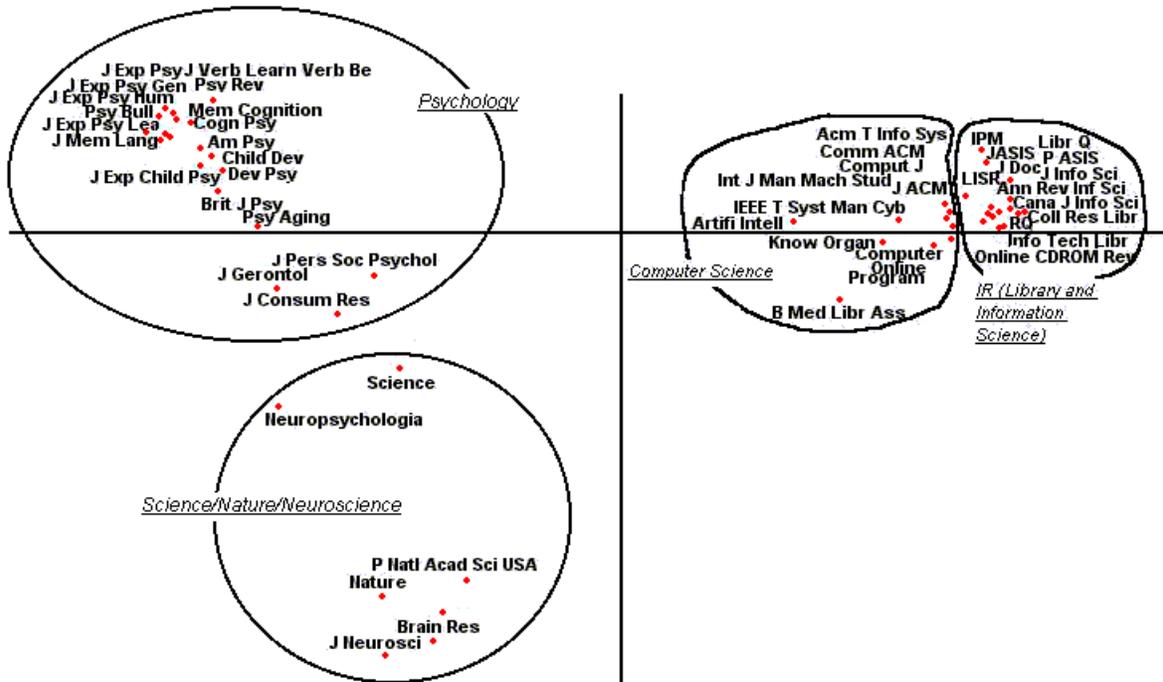
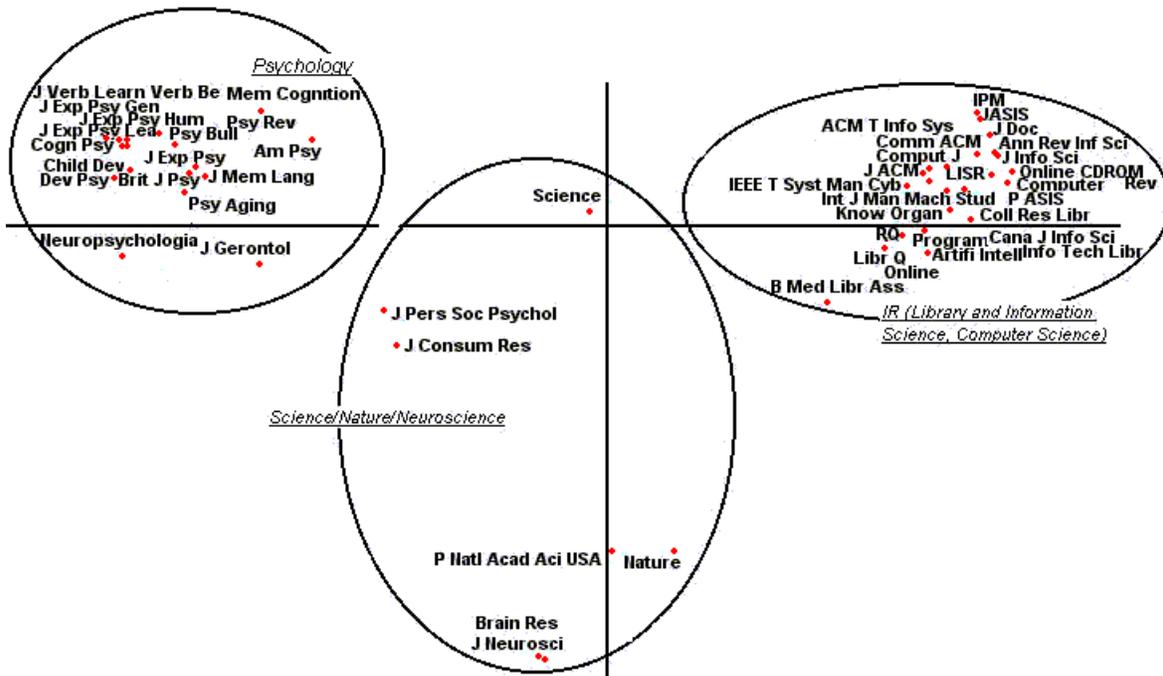Figure 1. MDS map of VAR journals highly cited by IR research (1987-1997)

Figure 2. MDS map of VAR journals highly cited by IR research (1987-1991)

Psychology

J Verb Learn Verb Be
Mem Cognition
Psy Rev
Cogn Psy
J Exp Psy Lea
J Exp Psy Gen
J Mem Lang
J Exp Psy
Psy Bull
J Exp Psy Hum
J Exp Child Psy
Brit J Psy
Dev Psy
Am Psy
Child Dev
J Gerontol
Psy Aging

Science/Nature/Neuroscience

Science
J Pers Soc Psychol
Neuropsychologia
J Consum Res
Nature
Brain Res
P Natl Acad Sci USA
J Neurosci

Computer Science

Comm ACM
ACM T Info Sys
Int J Man Mach Stud
Artifi Intell
Comput J
IEEE T Syst Man Cyb
J ACM
Know Organ
Computer
B Med Libr Ass
Online

IR (Library and Information Science)

IPM
JASIS
J Doc
Ann Rev Inf Sci
Cana J Info Sci
Coll Res Libr
Libr Q
LISR P ASIS
RQ
J Info Sci
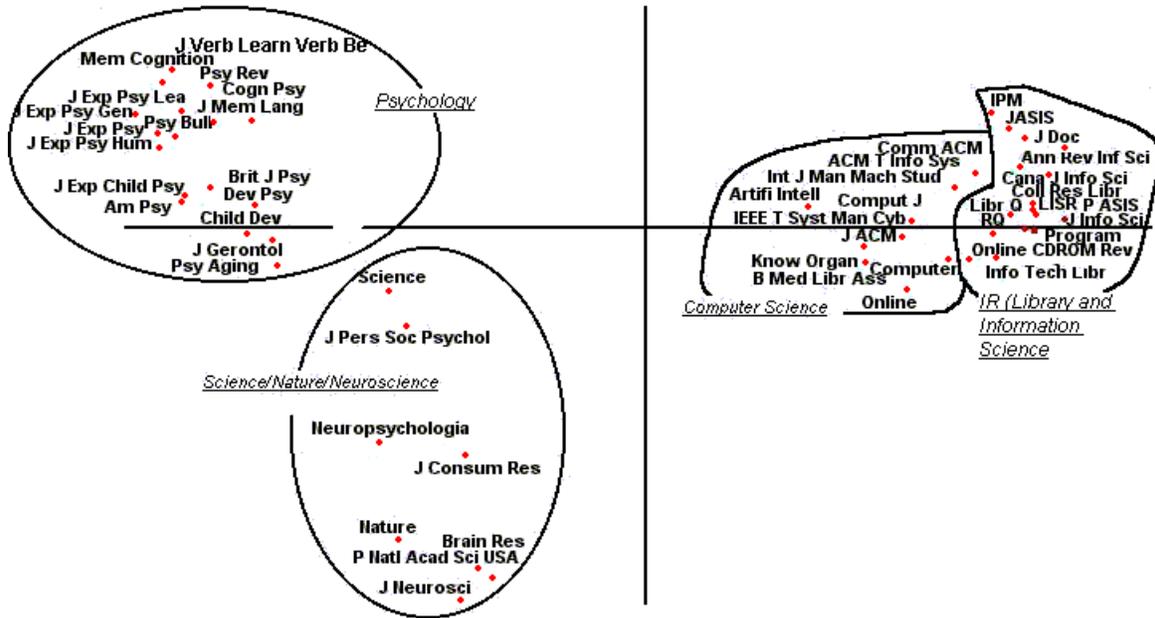Program
Online CDROM Rev
Info Tech Libr

Figure 3. MDS map of VAR journals highly cited by IR research (1992-1997)

During the whole 11-year period, we can see four groups appeared in Figure 1. Psychology, Computer science and Science/nature/neuroscience are the three areas, which have close relationship with IR. Computer science group is located very close to the IR group, which indicate that more and more computer science techniques are applied into IR field, and more and more computer science researchers are becoming interested in IR field. The position of computer science group also indicates that the effect of computer science to IR is much deeper than that of psychology and neuroscience.

The two periods' MDS maps are similar as the whole 11-year period MDS map, especially in 1992-1997. During the first period, computer science and IR field are emerged together into one group and journals in Science/nature/neuroscience group are from science, nature and neuroscience. Most of the journals' positions did not change much during the three periods. Furthermore, the IR group itself seems to be rather stable both in terms of its composition, and in terms of its relation to relevant environments. All these demonstrated the stability of IR field during the period of 1987-1997.

Based on the cited VAR journals (which includes highly cited or less highly cited VAR journals) by IR researcher, we can find that journals from Library and Information Science, psychology and computer science areas are most prominently present during each period. This suggests that IR is strongly related to Computer Science and Psychology, whose techniques are increasingly used as methods in IR specialty. Besides that, we can also find journals from theory-oriented disciplines (cognitive science, neuroscience and so on) or journals from application-oriented disciplines (medical science, chemical science and so on). The interesting point is that areas representing by VAR journals are very broad, and even include management science, consumer science, gerontology, medical science and so on. This indicates that IR techniques are more and more used in other specialties, at the same time, more and more techniques from numerous specialties are bestowed in IR field. So, we can come into conclusion that IR specialty is a multi-disciplinary specialty.

Library and Information Science journals which were highly cited by IR researchers

In this part, we discuss the position of Library and Information Science (LIS) journals in the scholarly communication process of IR field.

<u>Journal sample</u>

Fifty highly cited library and information science journals (including some highly cited regular annual publications and proceedings, henceforth called LIS journal) were selected as the research sample (Appendix 2). Similar data processing techniques, mentioned in the first part (*Various journals which were highly cited by IR researchers*), were used to generate journal cocitation maps in three different periods (1987-1997, 1987-1991, 1992-1997).

<u>Journal cocitation maps</u>

We generated journal cocitation maps of LIS journals in three periods (1987-1997, 1987-1991, and 1992-1997) in order to mark LIS journals' role in IR field and their changes over time (Figures 4, 5 and 6).
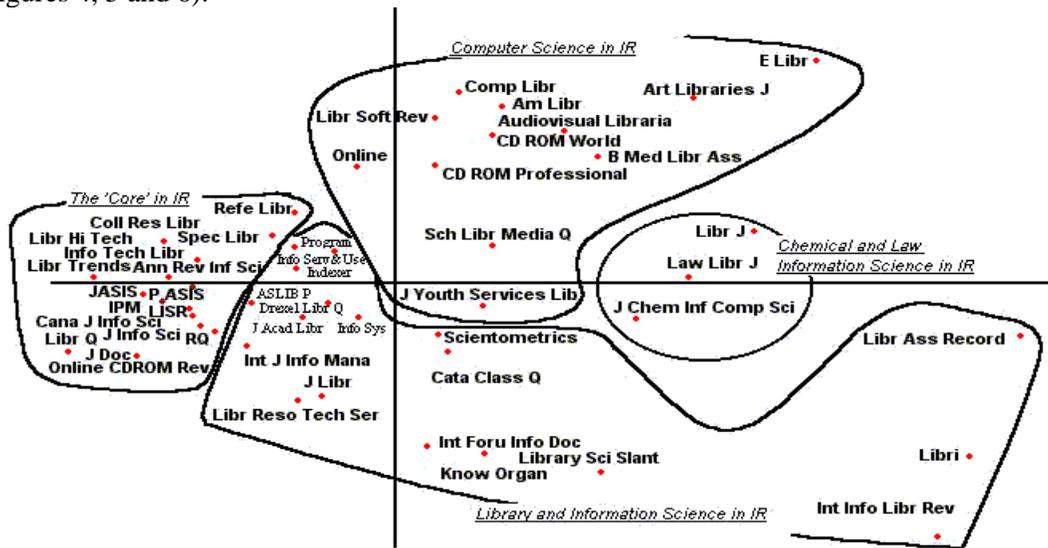

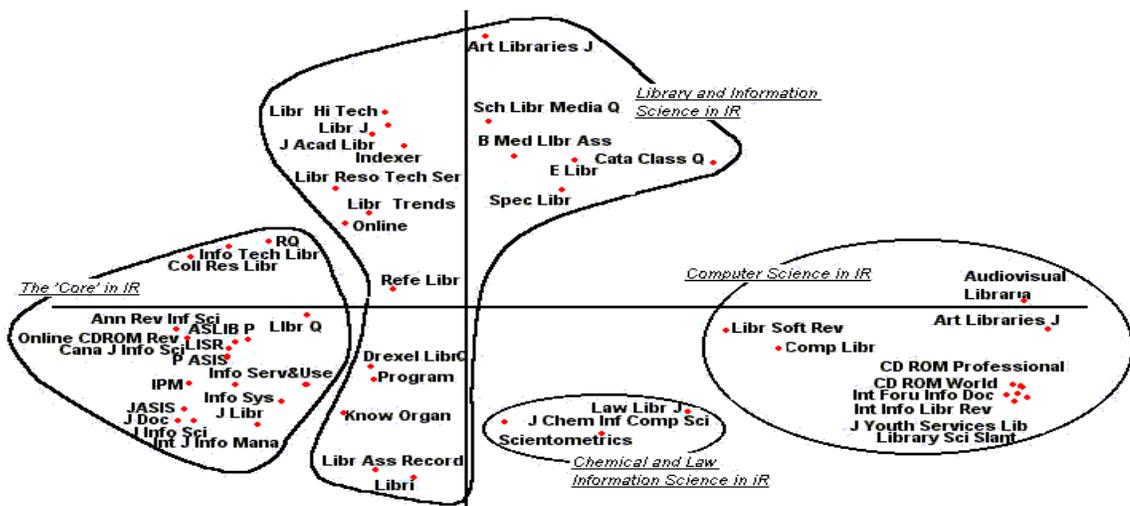Figure 4. MDS map of LIS journals highly cited by IR research (1987-1997)

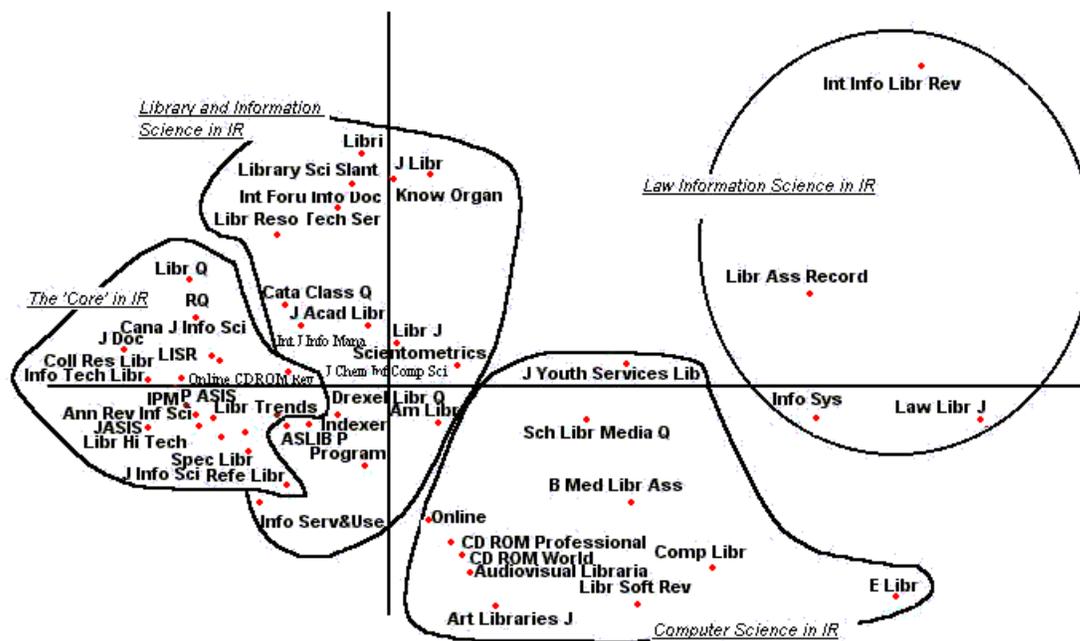Figure 5. MDS map of LIS journals highly cited by IR research (1987-1991)

Figure 6. MDS map of LIS journals highly cited by IR research (1992-1997)

Figure 4 shows the information about the LIS journals' role in the scholarly communication process of IR field during the whole 11-year period. Journals with high cocitation linkage are located together and those representing weak cocitation linkage are placed further apart by the mapping program. We can see four groups in this map. There are two clusters of journals representing IR and LIS research. The core in IR is represented by the journals which are most important in LIS field. The remaining two clusters represent research areas on IR application on chemical, law or computer information science.

Here, we find that 'computer science in IR' group is located very close to 'Library and Information Science in IR' and "The 'core' in IR" groups. This result is coincident with the maps in the first part (Figure 1, 2, and 3) which also show that computer science group is very near the IR group. Some journals in 'Library and Information Science in IR' and "The 'core' in IR" groups also publish papers on computer science or application computer science techniques into IR field. It means that there are some overlapping areas between computer science and LIS. In other words, LIS and computer science disciplines have close relationship between each other, and IR field is one of the most active research areas in both computer science and LIS.

Another group in the map is the journals focusing on chemical science and Law science. This indicates that LIS theory and techniques, especially IR theory and techniques are successfully employed to manage the chemical and law information.

Comparing Figure 5 and 6 with Figure 4, we can easily find that the maps are similar. The same four clusters were repeated in the maps of 1987-1991 and 1992-1997 except that 'Chemical and law information science in IR' is changed into 'Law information science in IR' because *Journal of Chemical Information and Computer Sciences* joins 'Library and Information Science in IR' group. Although some journals' positions in the two period's maps are different, the structure of the clusters does not change. Each of four clusters has its own fixed journal set during different periods, for example: *Information Processing & Management*, *Journal of the American Society for Information Science* and *Journal of Documentation* in "The 'Core' in IR" group; *Program*, *Drexel Library Quarterly* and *Cataloging & Classification Quarterly* in 'Library and Information Science in IR' group; *CD ROM Professional*, *CD ROM World*, *Library*

Software Review and *Computers in Libraries* in 'Computer Science in IR'; *Law Library Journal* and *Journal of Chemical Information and Computer Sciences* in 'Law Information Science in IR' or "Chemical and Law Information Science in IR' group and so on. So we can say that LIS journals' role in scholarly communication process of IR field does not alter much.

Obviously, the journal composition of the IR factor does not change much over times. The citing patterns of journals in the environment of IR journals are also stable. From Table 2, we can find that *Journal of the American Society for Information Science*, *Information Processing & Management*, *Journal of Documentation*, *Annual Review of Information Science and Technology*, *Journal of Information Science, Proceedings of the ASIS Annual Meeting* and so on are the 'core' of core journals in IR field.

The specialty development and changes can be indicated by gradual changes in the journal-set representing the specialty in various periods. If there is no stable journal-set representing a specialty, the specialty can be considered as in a 'preparadigmatic' or critical phase. If the journal-set remains relatively stable over time, this suggests that the specialty is more mature and established (van den Besselaar & Leydesdorff 1996). In each period, IR factor does not consist of heterogeneous journals and maintain the relative coherence. From this point of view, IR can be considered as a maturing specialty where a network of people work on the same set of research questions using the same methods and referring to the same scientific literature. That is to say that the theories and methods of IR are more focused and accepted by a specific community of researchers.

Discussion

In this research, we used journal cocitation analysis to map the VAR journals and LIS journals' role in scholarly communication process of IR field during 1987-1997 and their changes over time. Journal cocitation mapping produced clear, coherent, and reasonable results. It bears an overall resemblance with the author cocitation map of IR (Ding 1998 and1998a).

Based on the journal cocitation network of VAR journals, we found that the citing and cited environment of VAR journal remained stable. Computer science, psychology and science/nature/neuroscience play an important role in IR field. The VAR journal cocitation mapping does not change over time. Furthermore, the IR group itself is rather stable both in terms of its composition, and in terms of its relation to relevant environments. The stability of IR field during the period of 1987-1997 was demonstrated by current research.

Specialties related to IR field are very broad, not only including theory-oriented factors (psychology, cognitive science, science/nature and so on) and application-oriented factors (computer science, online IR, management science, consumer science and so on), but also including management science, consumer science, gerontology, medical science and so on. So, we concluded that IR specialty is a multi-disciplinary specialty.

Based on the journal cocitation analysis of LIS journals, we found that the citing and cited LIS journal environment also remained stable and did not change over time. The journal cocitation mapping of LIS journals indicated that IR has close relationship with computer science, and LIS theory and techniques, especially IR theory and techniques are successfully employed to manage the chemical and law information. The 'core' of core journals in IR field are *Journal of the American Society for Information Science*, *Information Processing & Management*, *Journal of Documentation*, *Annual Review of Information Science and Technology*, *Journal of Information Science, Proceedings of the ASIS Annual Meeting*. This research also indicates that IR can be considered as a maturing specialty during the period of 1987-1997 because of the stable IR journal-set in studying period.

References

Borgman, C. L. & Rice, R. E. (1992). The convergence of information science and communication: A bibliometric analysis. *Journal of the American Society for Information Science* 43: 397-411.

Braam, R. R., Moed, H. F. & van Raan, A. F. J. (1991). Mapping of science by combined co-citation and word analysis. II: Dynamical aspects. *Journal of the American Society for Information Science* 42(4): 252-266.

Cason, H. & Lubotsky, M. (1936). The influence and dependence of psychological journals on each other. *Psychological Bulletin* 33: 95-103.

Ding, Y. (1998). *Bibliometric analysis of information transfer in the field of information retrieval*. First Year Report. School of Applied Science, Nanyang Technological University.

Ding, Y. (1998a). Visualisation of intellectual structure in Information Retrieval: Author cocitation analysis. *International Forum on Information and Documentation* 23(1): 25-36.

Ding, Y. (1998b). Scholarly communication and bibliometrics: I. Scholarly communication model--Literature Review. *International Forum on Information and Documentation* 23(2): 20-29.

Ding, Y. (1998c). Scholarly communication and bibliometrics: II. Scholarly communication process--Literature Review. *International Forum on Information and Documentation*, accepted.

Doreian, P. (1988). Testing structural-equivalence hypotheses in a network of geographical journals. *Journal of the American Society for Information Science* 39(2): 79-85.

Everett, J. E. & Pecotich, A. (1991). A combined loglinear/MDS model for mapping journals by citation analysis. *Journal of the American Society for Information Science* 42(6): 405-413.

Leydesdorff, L. & Cozzens, S. E. (1993). The delineation of specialties in terms of journals using the dynamic journal set of the Science Citation Index. *Scientometrics* 26: 133-154.

McCain, K. W. (1984). Longitudinal author cocitation mapping: the changing structure of macroeconomics. *Journal of the American Society for Information Science* 35(6): 351-359.

McCain, K. W. (1991). Mapping economics through the journal literature: An experiment in journal cocitation analysis. *Journal of the American Society for Information Science* 42(4): 290-296.

Narin, F. (1976). *Evaluative bibliometrics*. Cherry Hill, NJ: Computer Horizons, Inc.

Norusis, M. J. (1997). *SPSS 7.5 guide to data analysis*. Upper Saddle River, N.J.: Prentice Hall.

Price, D. de Solla (1965). Networks of scientific papers. *Science* 149: 510-515.

Small, H. (1973). Cocitation in the scientific literature: a new measure of the relationship between two documents, *Journal of the American Society for Information Science* 24: 265-269.

Small, H. & Griffith, B. C. (1974). The structure of scientific literatures. I. Identifying and graphing specialties. *Science Studies* 4: 147-159.

Tijssen, R. J. W. & van Raan, A. F. J. (1990). Net citation balances: A measure of influence between scientific journals. *Journal of the American Society for Information Science* 41(4): 298-304.

Van den Besselaar, P. & Leydesdorff, L. (1996). Mapping change in scientific specialties: A scientometric reconstruction of the development of artificial intelligence. *Journal of the American Society for Information Science* 47(6): 415-436.

Appendix 1

50 highly cited VAR journals by IR researchers

JOURNAL NAME	CITATION FREQUENCY	ACRONYM
Journal of the American Society for Information Science	2069	JASIS
Information Processing & Management	1735	IPM
Psychological Review	1044	PSY REV
Journal of Documentation	771	J DOC
Journal of Experimental Psychology: Learning, Memory, and Cognition	750	J EXP PSY LEA
Communications of the Association for Computing Machinery	544	COMM ACM
Journal of Verbal Learning and Verbal Behavior	517	J VERB LEARN VERB BE
Memory & Cognition	508	MEM COGNITION
Journal of Personality and Social Psychology	404	J PERS SOC PSYCHOL
Annual Review of Information Science and Technology	312	ANN REV INF SCI
Science	305	SCIENCE
International Journal of Man-machine Studies	304	INT J MAN MACH STUD
Journal of Experimental Psychology: General	304	J EXP PSY GEN
Journal of Experimental Psychology: Human Learning and Memory	277	J EXP PSY HUM
Cognitive Psychology	266	COGN PSY
Journal of Information Science	244	J INFO SCI
Proceedings of the ASIS annual meeting	244	P ASIS
Journal of Experimental Psychology	232	J EXP PSY
Journal of Memory and Language	231	J MEM LANG
Psychological Bulletin	205	PSY BULL
Library Quarterly	170	LIBR Q
Online and CD-ROM Review	170	ONLINE CDROM REV
Nature	168	NATURE
Neuropsychologia	150	NEUROPSYCHOLOGIA
Bulletin of the medical library association	134	B MED LIBR ASS
ACM Transactions on information systems	132	ACM T INFO SYS
Developmental Psychology	130	DEV PSY
Journal of Gerontology	121	J GERONTOL
Library Information Science Research	116	LISR
Journal of the Association for Computing Machinery	115	J ACM
Online	114	ONLINE
American psychologist	112	AM PSY
Journal of Neuroscience	112	J NEUROSCI
Program: Automated Library and Information Systems	110	PROGRAM
Proceedings of National Academy of Sciences (USA)	109	P NATL ACAD SCI USA
Journal OF Experimental Child Psychology	108	J EXP CHILD PSY
RQ	108	RQ
Brain Research	104	BRAIN RES
British Journal of Psychology	104	BRIT J PSY
Knowledge Organization	103	KNOW ORGAN
Psychology and Aging	103	PSY AGING
Artificial Intelligence	101	ARTIF INTELL
College & Research Library	100	COLL RES LIB
Canadian Journal of Information and Library Science	99	CANA J INFO SCI
IEEE Transactions on system, man, and cybernetics	98	IEEE T SYST MAN CYB
Computer	96	COMPUTER
Computer Journal	94	COMPUT J
Information Technology and Libraries	94	INFORM TECHNOL LIBR
Child Development	90	CHILD DEV
Journal of Consumer Research	89	J CONSUM RES

Appendix 2
50 highly cited LIS journals by IR researchers

JOURNAL NAME	CITATION FREQUENCY	ACRONYM
Journal of the American Society for Information Science	2069	JASIS
Information Processing & Management	1735	IPM
Journal of Documentation	771	J DOC
Annual Review of Information Science and Technology	312	ANN REV INF SCI
Journal of Information Science	244	J INFO SCI
Proceedings of the ASIS annual meeting	244	P ASIS
Library Quarterly	170	LIBR Q
Online CDROM Review	170	ONLINE CDROM REV
Bulletin of the Medical Library Association	134	B MED LIBR ASS
Library Information Science Research	116	LISR
Online	114	ONLINE
Program: Automated Library and Information Systems	110	PROGRAM
RQ	108	RQ
Knowledge Organization	103	KNOW ORGAN
College & Research Library	100	COLL RES LIB
Canadian Journal of Information and Library Science	99	CANA J INFO SCI
Information Technology and Libraries	94	INFO TECH LIBR
Library Resources & Technical Services	87	LIBR RESO TECH SER
Cataloging & Classification Quarterly	87	CATA CLASS Q
Library Trends	85	LIBR TRENDS
ASLIB Proceedings	69	ASLIB P
Information Systems	69	INFO SYS
International Journal of Information Management	65	INT J INFO MANA
Journal of Chemical Information and Computer Sciences	64	J CHEM INF COMP SCI
Electronic Library	62	E LIBR
Library Hi Tech	53	LIBR HI TECH
Journal of Academic Librarianship	50	J ACAD LIBR
Libri	34	LIBRI
Scientometrics	33	SCIENTOMETRICS
Library Journal	32	LIBR J
Information Services & Use	32	INFO SERV & USE
School Library Media Quarterly	31	SCH LIBR MEDIA Q
Special Libraries	27	SPEC LIBR
Drexel Library Quarterly	26	DREXEL LIBR Q
Journal of Librarianship and Information Science	25	J LIBR
Law Library Journal	24	LAW LIBR J
Reference Librarian	24	REFE LIBR
Emedia Professional	24	CDROM PROFESSIONAL
Computers in Libraries	20	COMP LIBR
International Information & Library Review	19	INT INFO LIBR REV
International Forum on Information and Documentation	19	INT FORU INFO DOC
Library Software Review	19	LIBR SOFT REV
Library Association Record	18	LIBR ASS RECORD
Indexer	17	INDEXER
CD-ROM World	16	CD ROM WORLD
American Libraries	15	AM LIBR
Audiovisual Librarian	11	AUDIOVISUAL LIBRARI
Art Libraries Journal	10	ART LIBRARIES J
Library Science with a Slant to Documentation and Information Studies	10	LIBRARY SCI SLANT
Journal of Youth Services in Libraries	10	J YOUTH SERVICES LIB