Visualizing the Structure of Scientific Output of Iranian Scholars in Science Citation Index (SCI) during 2000-2006

İranlı Bilim İnsanlarının Bilimsel Üretim Yapısını Görselleştirme: Science Citation Index (SCI) Verilerine Dayalı Bir Araştırma (2000-2006)

Farideh Osareh

Shahid Chamran University, Faculty of Education and Psychology, Department of Library and Information Science, Ahwaz-Iran. fosareh@yahoo.com

Maryam Keshvari

Shahid Chamran University, Ahwaz-Iran. ma.keshvari@gmail.com

Abstract: *To visualize the structure of Iranian scientific output in Science Citation Index (SCI), accessible via Web of Science (WOS), during 2000-2006, we used scientometric techniques and HistCite software. The number of Iranian documents indexed in SCI during the study period was 24,480. Generally HistCite analyzes citation data on two different levels: based on citations in WOS (Global Citation Scale), and citations in collection of retrieved documents (Local Citation Scale). The results of this study showed that, in the study period a total of 8 clusters have been formed on the two levels (GCS and LCS): Clusters 1 and 2 (with 3 sub-clusters) in GCS and clusters 3, 4, 5, 6, 7, and 8 in LCS. The subject area of whole clusters was chemistry, but different areas of this discipline. The prominent subject area in our study was organic chemistry. The most effective document in this study was an article by Zolfigol with 123 global citations and 71 local citations. The subject category of cluster 1 was analytical chemistry and membrane electrodes. Cluster 2 consists of 3 sub-clusters (sub-cluster 1 hydrocarbons, sub-cluster 2 in the field of oxidation and nitrogen, and sub-cluster 3 catalysts). Cluster 4 was crystal structure, cluster 5 electrochemical analysis, cluster 6 macro cycles, cluster 7 aliphatic and aromatic complexes and the 8th cluster was polymers.*

Keywords*: Scientometrics, historiographical map, scientific output, citation indexes, Iran*

Öz: *Web of Science (WOS) aracılığıyla erişilen Science Citation Index'te (SCI) İran'ın 2000-2006 yılları arasındaki bilimsel üretim yapısını görselleştirmek için bilimetrik teknikler ve HistCite yazılımını kullandık. Araştırma kapsamına giren süre içinde SCI'de dizinlenen İran'a ait belge sayısı 24.480'di. Genelde HistCite atıf verilerini iki farklı düzeyde, WOS'taki atıflara dayalı olarak (Küresel Atıf Ölçeği) ve erişilen belge dermesindeki atıflara dayalı olarak (Yerel Atıf Ölçeği) analiz etmektedir. Bu araştırmanın sonuçları iki düzeyde (KAÖ ve YAÖ) toplam 8 küme oluştuğunu göstermektedir: KAÖ'de Küme 1 ve Küme 2 (3 alt küme ile birlikte), ve YAÖ'de Küme 3, 4, 5, 6, 7 ve 8. Bütün kümelerin konu alanı kimya, ancak kimya disiplininin farklı alanlarıydı. Çalışmamızda öne çıkan konu alanı organik kimya idi. Bu araştırmadaki en etkin belge 123 küresel atıf ve 71 yerel atıfla Zolfigol'un makalesiydi. Küme 1'in konu kategorisi analitik kimya ve membran elektrotlarıydı. Küme 2, üç alt kümeden oluşmaktadır (hidrokarbonlar, oksitlenme ve nitrojen, ve katalistler). Küme 4 kristal yapısı, Küme 5 elektrokimyasal analiz, Küme 6 makro devirler, Küme 7 alifatik ve aromatik kompleksler ve Küme 8 polimerler ile ilgiliydi.*

Anahtar sözcükler*: Bilimetri, tarihyazımsal harita, bilimsel üretim, atıf dizinleri, İran*

Introduction

The issue of scientific production was first introduced at the University of Tehran in 1978 but was not pursued. In 1993-1994 this topic was studied more seriously, and in 1997 the citation role in evaluating scientific collaborations was the main subject area of the "Conference of Methodology and Research Techniques" which was carried out by the research deputy of the University of Tehran (Osareh, 1997). After 1999, scientific production was considered by the Iranian Ministry of Science, Research and Technology (MSRT). According to the law passed by MSRT, each Iranian who publishes an article in a Thomson ISI Journals receives a considerable reward (Moosavi Movahedi, Kiani Bakhtiari, & Khan Chamani, 2003).

The government supports and encourages the attention and tendency of Iranian authors towards internationalized Iranian research output, and this has caused a gradual increase in Iranian scientific output (Osareh & Wilson, 2002).

A simple search in WOS on Iranian scientific output showed that an exponential increase has happened in recent years. For instance, the number of Iranian publications in SCI in 2000 was only 1371, while this rate increased to 5630 in 2005, 13440 in 2008 and 16492 in 2009. In this article, we decided to study Iranian scientific output using scientometric indicators.

Importance and Necessity of Research

One of the important factors for sustainable development in each country is the number of the scientific products indexed in international accredited databases. Based on the importance of scientific products on national and international scales, this paper is trying to visualize scientific output of Iranian scholars in WOS during 2000-2006. To this end, the subject areas of established clusters will be identified, and the effective authors and articles in both GCS and LCS will be recognized. We also try to explore the participation rate of Iranian scholars in creating these products. In other words, in this research we consider the scientific products of Iran in the Thomson ISI. We should draw your attention to the Iranian national language which is Farsi; therefore a lot of Iranian scientific output is in Farsi so does not have a chance to be indexed in Thomson ISI databases. Hence, in this research we have mapped Iranian non-Farsi publications in the Science Citation Index via WOS.

Aims and Purposes

The main purpose of this research is to visualize the structure of the scientific products of Iranian scholars indexed in Thomson ISI accessible via WOS during 2000-2006. According to this map, we will study the history of science and recognize the effective authors, based on citations on both LCS and GCS. To reach the above goals we should answer the following questions.

Research Questions

1. How is the historiographical map of Iranian scientific output in WOS during 2000-2006?
2. What are the most important scientific clusters formed in Iranian scientific output during 2000-2006? And what are their subject areas?
3. Who are the most productive and effective Iranian authors on both citation levels (GCS & LCS) during 2000-2006?
4. What are the most effective articles based on both GCS and LCS citation levels and their publication dates?

Background

Osareh and Wilson (2002) analyzed international collaboration of Iranian scientific publications in SCI during 1995-1999. The results show that Iran's publication output in science and technology increased dramatically in the SCI during 1995-1999. One of the most important and significant factors that caused this rise seems to be the government's research policies in the last few years of their study. In 1996, the Iranian government announced the first national research call for papers and continued it for the following years. The researchers selected topics according to their areas and started working with large research grants. This can lead the researchers towards the research topics needed by the government. Another result of this study shows that Iran's main international collaborators are authors with institutional affiliations in the US or the UK. However, it is obvious that Iran is looking more and more for collaborative partners elsewhere. Collaboration with authors in Canadian and Australian institutions has increased either in absolute numbers, relative percentages or both.

Osareh and McCain (2008) tried to draw the intellectual structure of Iranian chemistry research in Science Citation Index (SCI). The results of this research showed that since 1990, Iranian chemistry research, as represented in the SCI, has grown at a rate of roughly 26% and 7 major clusters, Oxidation of Organic Compounds**,** Physical Organic Chemistry, Ionosphere, Analytical Chemistry, Solvent-Free Synthesis, CJ Pedersen and Crown Ethers, Synthesis of Carbonyl Compounds, were identified. The topic areas were primarily in organic chemistry, and secondarily in analytical chemistry; other major topic areas such as biochemistry, applied chemistry, and chemical engineering were not seen.

Research Methodology and Data Gathering

The research method for this study was the scientometric method. The population of this research comprises 24480 documents produced by Iranian authors, indexed in SCI during 2000-2006.

Data were gathered and analyzed in 3 steps using 3 tools. In the first step, data were extracted by using SCI and via WOS in plain text format. In the second step, data were recognized by ISI.exe software, and in the third step data were registered into an Excel spreadsheet and made ready for analysis.

To draw the historiographical map of Iranian scientific output in SCI, we used HistCite software. This software is a product of ISI. Its input contains plain text files extracted from WOS and its output contains a graphical image of scientific outputs (Garfield, Paris, & Stock, 2006).

Data Analysis

A) Describing data

The analysis of the data revealed that there were totally 24480 documents produced and indexed in SCI by Iranian authors during 2000-2006. The publication year, document type and the language of the documents were analyzed and are displayed in Table 1.

The publication rate of Iranian scientific products increased from 2000 to 2006: 5.85 percent of the total for the period (1417 documents) were published in 2000 and 24.74 percent (5995 documents) in 2006. In other words, Iran increased its annual productivity 4.23 fold from 2000-2006.

The results of the analysis of the type of documents showed that the documents are in 13 different formats. The most frequent format was the article with 21513 (87.88%) titles, followed by meeting abstracts with 2139 (8.74%) titles (Table 1).

As can be seen in Table 1, Iranian scientific products in SCI during 2000-2006, were published in five different languages. English with 24499 (99.88%) documents ranked top, followed by French and German with 18 (0.07 %) titles, and 10 (0.04 %) titles respectively.

Table 1. Iranian scientific products by language, format and publication year

Publication year			Type of documents			Language		
Year	**Frequency**	**Percent**	**Format**	**Frequency**	**Percent**	**Language**	**Frequency**	**Percent**
2000	1,417	5.85	Article	21,513	87.88	English	24,449	99.88
2001	1,775	7.32	Meeting Abstract	2139	8.74	French	18	0.07
2002	2,411	9.95	Literature Criticism	336	1.37	German	10	0.04
2003	3,234	13.35	Review	209	0.85	Russian	2	0.01
2004	4,062	16.77	Editor Review	200	0.83	Italian	1	0
2005	5,335	22.02	Correction	61	0.25	**Total**	**24,480**	**100.0**
2006	5,995	24.74	News	17	0.07			
Total	**24,229**	**100.0**	Biography	2	0.01			
			Database Review	1	0			
			Reprint	1	0			
			Bibliography	1	0			
			Book Review	0	0			
			Software Review	0	0			
			Total	**24,480**	**100**			

B) Historiographical Map

A historiographical map has been drawn based on two separate levels, using HistCite: 1. Global citations scale (GCS). 2. Local citations scale (LCS). For the GCS map, the data sample was based on 300 documents (nodes). For the LCS map, due to the high number of links, and to have a clear graph we, drew the map with only 200 nodes.

Analyzing the Clusters of Iranian Scientific Products in SCI during 2000-2006

Because of the length of the map, we decided to divide it by clusters and identify the clusters one by one. The results of the research showed that on the GCS level there were only 2 clusters; cluster 2 had 3 sub-clusters due to the diversity of the subject areas. In the LCS map we observed 6 clusters. The subject fields of clusters in the GCS map were "membrane electrode" (cluster 1); the 3 sub-clusters of cluster 2 were in "operation on hydrocarbons", "oxidation and nitrogen" and "catalysts". In the LCS map, clusters were in "operation in organic chemistry", "crystal structure", "electrochemical analysis", "macro cycles", "aliphatic & aromatic complexes" and "polymers". Firouzabadi and Heravy each participated in 3 clusters of which two had the same subject areas: "operation in hydrocarbons" and "operations in organic chemistry". The subject category of the 3rd cluster for Firouzabadi was "aliphatic & aromatic complexes", while Heravy participated in "catalysts" as the 3rd cluster. It should be noted that the first authors in all clusters were Iranian.

Scientific Clusters in SCI on GCS Map:

Considering the GCS map, 2 clusters have been observed in Iranian scientific output in SCI during 2000-2006. Clusters 1 and 2 will be defined shortly. It should be noted that due to the subject diversity of cluster 2, it has formed 3 sub-clusters.

Cluster 1

This cluster has been established by the collaboration of 7 Iranian authors (Shamsipour, Ganjali, Mousavi, Shahrokhian, Javanbakht, Mashhadizadeh and Bagheri) during 1999-2005, in the area of "membrane electrode" in analytical chemistry. Figure 1 shows cluster 1, with the top 5 articles based on the number of citations and links which are shown in bold numbers in this Figure.

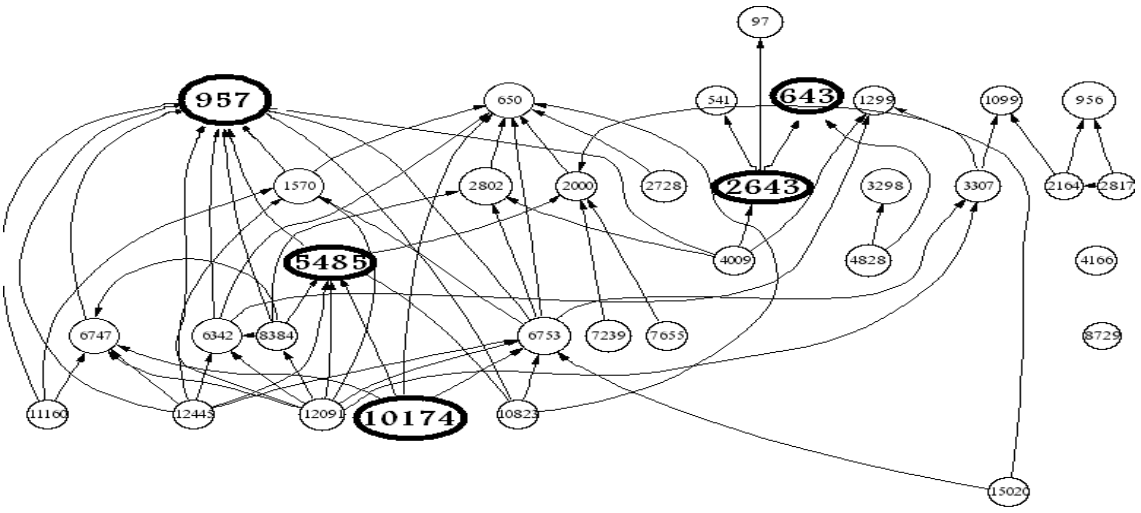

Figure 1. Cluster 1 in membrane electrode area on GCS map

In this cluster, the most effective document (considering the number of citations received and links to it) is by Shamsipour (957) with 102 GCS. This document with 66 LCS is also the most effective document in LCS map.

Cluster 2

As was mentioned and can be seen in Figure 2, cluster 2 is a large and separable cluster. Therefore, it was divided in 3 sub-clusters due to the diversity of subject areas. Each sub-cluster will be analyzed separately.

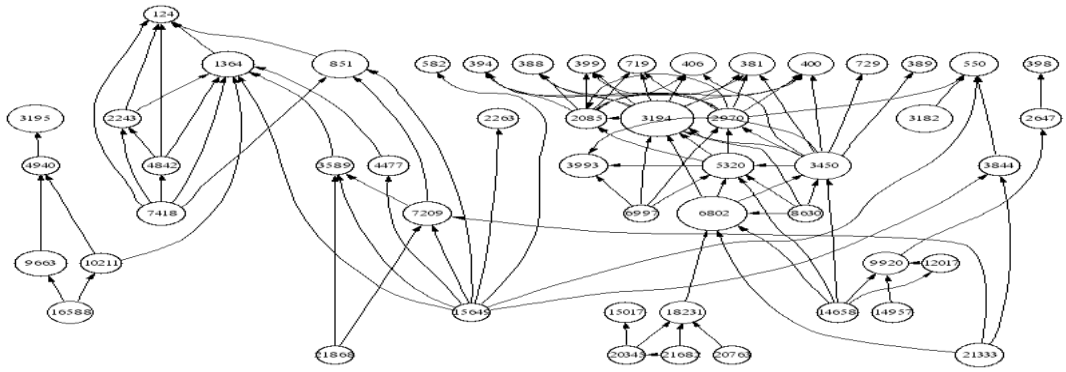

Figure 2. Cluster 2 on GCS map

Sub-cluster 1 from Cluster 2

Figure 3 shows the first sub-cluster of cluster 2. This sub-cluster is established by collaboration of 7 Iranian authors (Kaboudin, Karimi, Azizi, Firouzabadi, Habibi, Saidi, and Heravy) during 1999-2006. The subject area of this cluster is "operation on hydrocarbons". The most effective document (3195) is by Kaboudin with 76 GCS. In Figure 3 the top 5 most effective documents in this sub-cluster are in bold.

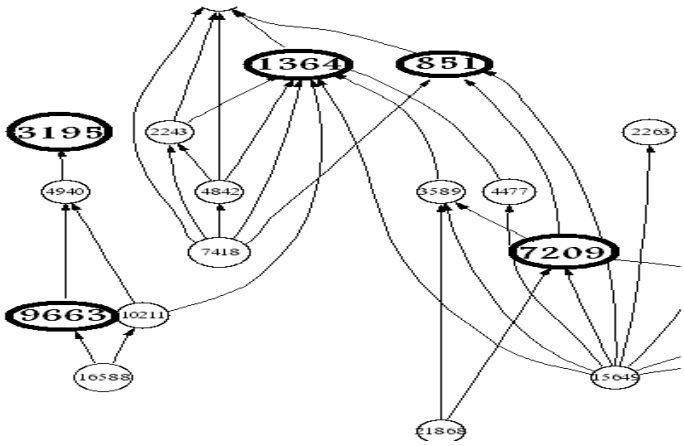

Figure 3. Sub-cluster 1 from cluster 2 in operation on hydrocarbons on GCS

Sub-cluster 2 from Cluster 2

The subject area of this sub-cluster is "oxidation, nitrogen and catalysts" in the area of organic chemistry. The most important document in this sub-cluster is indicated by the number 3194 in Figure 4 and titled "Silica sulfuric acid/NaNO2 as a novel heterogeneous system for production of thionitrites and disulfides under mild conditions" by Zolfigol (2001). It received 123 GCS and 71 LCS. This document received the highest number of citations in both parts: GCS and LCS.

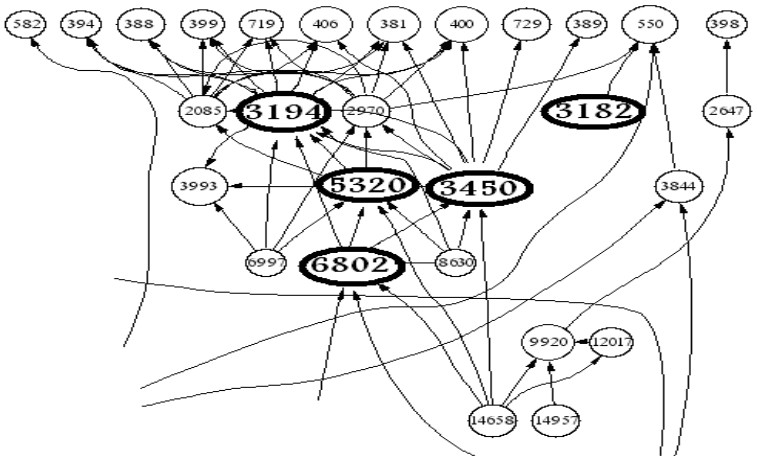

Figure 4. sub-cluster 2 of cluster 2 on GCS map

As can be seen in Figure 4 sub-cluster 2 is a big sub-cluster and has been established by collaboration of 10 authors (Zolfigol, Salehi, Firouzabadi, Shirini, Heravy, Khosropur, Mohamadpour, Balterak, Khodayi, Iranpour, and Sadeghi) by 28 documents during 2000-2006. Zolfigol with 14 documents produced the most articles in this sub-cluster. All of the first authors in these documents are Iranians. The top 5 most important documents from this sub-cluster are bold in Figure 4.

Sub-cluster 3 from Cluster 2

This is a small sub-cluster with only 5 documents and by collaboration of 2 Iranian authors (Heravy and Bamoharram). It was established during 2005-2006 in the "Catalysts" subject area. The most effective document in this cluster is document number 18231 from Heravy with 47 GCS.

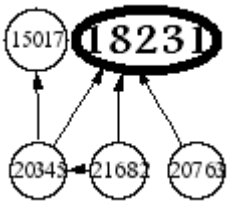

Figure 5. Sub-cluster 3 from cluster 2 on GCS map

Clusters in SCI in Local Citation Scale

According to local scales, there are 6 clusters in scientific products of Iranian authors in SCI during the research period. Here we analyze them.

Cluster 3

This is another big cluster with 29 documents and collaboration of 8 authors (Zolfigol, Shirini, Salehi, Firouzabadi, Sadeqi, Keypour, Mirjalili, and Heravy). This cluster was established during 2000-2005 based on LCS. Zolfigol produced 21 documents out of 29 in this cluster. As was mentioned his document number 3194 received the most number of citations based on LCS and GCS. The subject category of this cluster is "oxidation, nitrogen and catalysts" in Organic Chemistry. First authors in all documents of this cluster are Iranian.

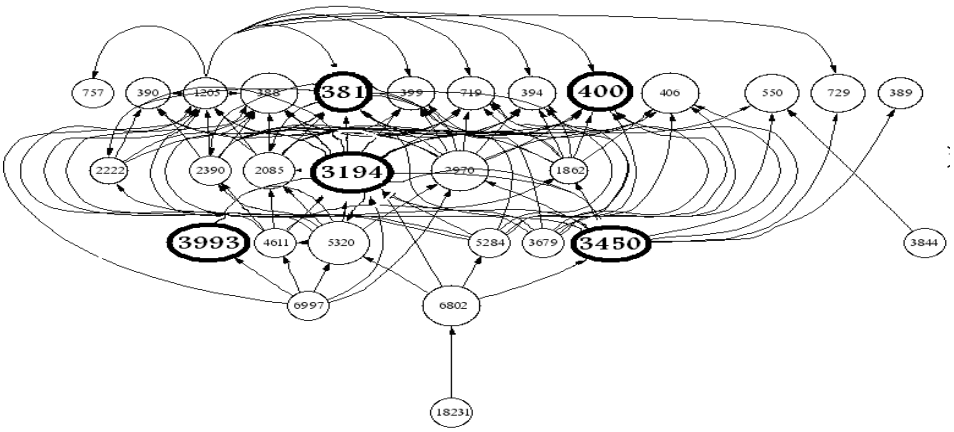

Figure 6. Cluster 3 in oxidation, nitrogen and catalysts on GCS map

Cluster 4

This cluster has 9 documents produced by 2 authors (Moghimi and Ranjbar). It was established during 2001-2005 and its subject area is "crystal structure". The most effective document in this cluster is document 4170, by Moghimi, with 28 LCS.

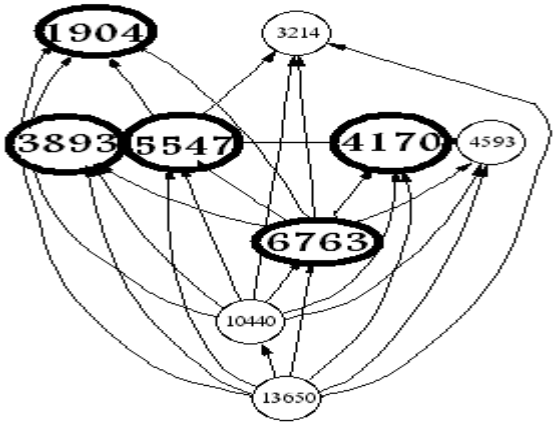

Figure7. Cluster 4 in Crystal structure on LCS map

Cluster 5

The subject area of this cluster is "electrochemical analysis". It has 29 documents with collaboration of 7 Iranian authors (Shamsipour, Shahrokhian, Ganjali, Javanbakht, Musavi, Rahmani and Bagheri). This cluster was established during 1999-2004 based on LCS.

Figure 8. Cluster 5 in Electrochemical analysis based on LCS map

Cluster 6

This cluster has 7 documents by one author (Salavati Niasri). It was in "macro cycles" and has been established during 2003-2005. The most effective documents in this cluster are numbers 10714 and 11621 which each received 21 LCS.

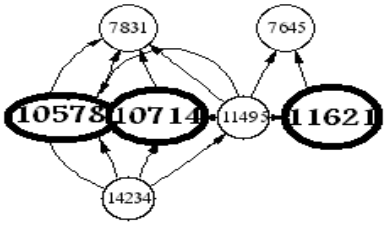

Figure 9. Cluster 6 in Macro cycles on LCS map

Cluster 7

This cluster has 9 documents by 4 authors (Karimi, Firouzabadi, Azizi and Saidi) and was established during 1999-2004 in the area of aliphatic and aromatic complexes. The most effective document in this cluster is document number 1364 with 22 citations in local scale. The information of this cluster is available in Figure 10. All of the first authors in this cluster are from Iran.

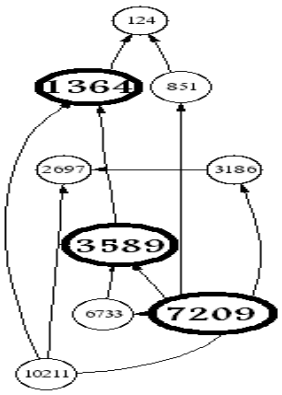

Figure 10. Cluster 7 in aliphatic and aromatic on LCS map

Cluster 8

Cluster number 8 was established by 17 documents by 2 Iranian authors (Hajipour and Malekpour) during 2000-2004, in the subject area "Polymers". The most effective documents in this cluster are documents number 868 with 54, and documents number 558 and 852 each with 52 LCS.

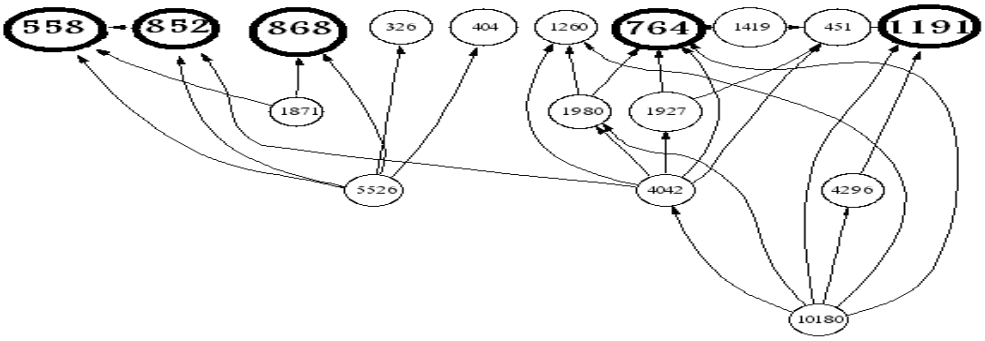

Figure 11. Cluster 8 in polymers on LCS map

Conclusions and Suggestions

This article found that, the total rate of scientific production in the period of this research was increasing and the scientific products of Iran from 2000 to 2006 showed a growth rate of 4.23 percent. The most similar research to this study was done by Osareh and McCain (2008) in which all clusters were about chemistry. These results were proved in our research and shown that generally, Iranian chemists have a tendency to produce scientific documents first in the field of organic chemistry and second in analytical chemistry. The most participant authors in the clusters of this study were Firouzabadi and Heravy who each participated in 3 clusters. Of those one was in the 3rd sub-cluster of cluster 2 (in GCS) and 2 other clusters in LCS. Among all 8 clusters, 1 cluster was formed by only 1 author. Two clusters were formed by publications of 2 authors. Other clusters had more than 2 authors. The results of this study also showed the influential Iranian authors and articles during 2000-2006 in WOS. The subject areas of clusters were recognized as follows:

"membrane electrode", "operation on hydrocarbons", "operation oxidation and nitrogen", "organic chemistry" and "catalysts". In LCS map, clusters were in "operation in organic chemistry", "crystal structure", "electrochemical analysis", "macro cycles", "aliphatic & aromatic complexes" and "polymers".

Using information visualization in different scientific disciplines could be useful for specialists as well as policy makers. The specialists at a glance can see which subject areas in their discipline have been under research by their colleagues, and which areas have been less attendant during a specific time. The results of such studies would let the policy makers allocate the budgets to subject fields with more confidence. The results of such studies also would be helpful for scientists and young specialists who can save their time by reading the works of key authors and influential scientific output in their disciplines.

References

Garfield, E., Paris, S., & Stock, W.G. (2006). HistCited™: A software tool for informatic analysis of citation linkage. *Information Wissenschaft und Praxis 57*, 391-400.

Moosavi Movahedi, A.A., Kiani Bakhtiari, A., & Khan Chamani, J. (2003). Methods of production and dissemination of scientific findings. *Rahyaft*, *31*, 5-19. (in Farsi).

Osareh, F. (1997). Bibliometrics. *Journal of School of Education & Psychology of Shahid Chamran University*, *8*(4), 90-97. (in Farsi)

Osareh, F. & McCain, W. (2008). The structure of Iranian chemistry research, 1990-2006: An author cocitation analysis. *Journal of the American Society for Information Science and Technology*, *59*(13), 2146-2155.

Osareh, F. & Wilson, C.S. (2002). Collaboration in Iranian scientific publications. *Libri*, *52*, 88-98.