Visualizing the structure of Iranian chemistry research in SciSearch using author co-citation technique*

Farideh Osareh
Katherine W. McCain

To visualize the structure of Iranian chemistry scientific publications in SciSearch, 43 Iranian and International chemists were identified. They were the highly cited scientists in 7682 Iranian chemistry publications (defined as an article with at least one Iranian author address) indexed in *Science Citation Index* during the period from 1990 to 2006. Co-citation data for these authors from the entire SciSearch file (Dialog, File 34) was collected for the period under study. A PFNet analysis visualized the structure of Iranian chemistry research in SciSearch using author co-citation technique. The results showed that, Iranian and International authors tended to appear in separate sub networks. Geographic and institutional influences, apparently relating in part to institutional affiliation and in part to restricted research topics, appear to underlie the primary structural features of Iranian chemistry in this source and time period.

1. Introduction

Over the past decade, the rapid growth of scientific research in Iran as reflected in counts of publications and citations indexed by the Institute for Scientific Information and in collaborative relationships between Iranian and non-Iranian researchers was studied by Osareh and Wilson [1,2,3,4]. They reviewed the patterns of Iranian indexed scientific publications from 1972 (when there was 1 paper with an Iranian author address in the SciSearch database) until 2002 (when the annual file contained almost 2200 articles). The publication pattern included a sharp decline in publication during and after the Iraq-Iran war (1979-1989) and a dramatic resurgence from 1990 to 2002. The percentage of Iran's publication vis-à-vis the rest of the world also increased during this period. The rise in productivity and visibility may be attributed to a combination of factors that include- improvements in the economic, educational, and political situation; remarkable increase in the number of internationally educated Iranian students, as well as in collaboration between Iranian and international researchers resulting in the increased output in international journals. The inclusion of a number of English language Iranian journals titled *Iranian Journal of Chemistry & Chemical*

Farideh Osareh
School of Education & Psycology
Dept. of Library & Info. Sc.
Shahid Chamran Uinversity,
Ahwaz-Iran,
Osareh.f@gmail.com

Katherine W. McCain
College of Information Science & Technology
Drexel University,
Philadelphia, USA
Kat.McCain@ischool.drexel.edu

* This is a report of an expanded article which has been accepted by JASIST.

Engineering—International English Edition;, Iranian Journal of Science & Technology Transactions A & B; Iranian Polymer Journal; Journal of the Iranian Chemical Society...), into SciSearch and the Web of Science was also identified as a contributing factor.

One persistent finding was the pre-eminence of chemistry as a highly visible research discipline. Chemistry journals figured prominently in Osareh & Wilson's analyses [2, 3, 4] of Iranian scientific productivity and citation visibility, found among the journals most frequently citing the top cited Iranian scientists and also being cited by Iranian scientists. A high proportion of the top cited Iranian scientists in each of the periods studied were found to be chemists.

In addition to the work by Osareh & Wilson, a few other studies examined the current state of research trends in Iranian science in general and chemistry in particular that confirm the rising trend in productivity and the pre-eminence of chemistry in terms of ISI-indexed publication counts in 2000 [5] and the periods 1983-2003 [6] and 1992-2003 [7]. In the period 1994-2003 (approximately), more PhD students in Iran graduated in chemistry than in other basic science disciplines [7]. Yalpani et al. [6] report that almost all of the Iranian publication activity in chemistry has been in organic and analytical chemistry, in contrast to the broader international arena where Biochemistry, Applied and Macromolecular Chemistry and Physical and Analytical Chemistry are emphasized. They also note that Iranian publications in chemistry tend to be published in journals with lower Impact Factors (IF) than the world average (this includes the Iranian journals indexed by ISI); Mehrdad et al.[7] report a similar pattern for IF values calculated at the individual article level. Yalpani and Heydari [8] attribute this high level of publication in low IF journals to a combination of strong incentives for chemistry faculty and doctoral students to publish and poor editorial/ reviewing control at the journal level. Finally, Harirchi et al.[9] surveyed Iranian-international collaboration in Biology, Chemistry, and Physics but do not report survey results for the three disciplines separately.

In the present paper, we extend the earlier analyses of Iranian science from calculations of quantitative science indicators to a more specific, subject-based exploration of Iranian chemistry research. We use Pathfinder Network Analysis—to visualize the intellectual structure of Iranian chemistry research published in ISI-indexed journals over the time period 1990-2006. We are interested in discovering which topic areas within chemistry are of greatest interest to Iranian researchers and which Iranian and international authors are considered key to these topics. A companion study in progress looks even more specifically at the Citation Images and Citation Image Makers [10] of several prominent Iranian chemists who are located in different topic areas.

It should be noted that studies of Iranian scientific research using SciSearch will naturally yield an incomplete picture. Iran's scientific language is Farsi. Thus, since ISI databases focus on English language publications, Iranian research productivity in chemistry would seen as even higher if Iranian scientific publications in Farsi were included along with the few English-language Iranian journals currently indexed in SciSearch. Participation by Iranian scientists in conferences—regional, national and international, — is also not included in SciSearch.

2. Research Methodology

Any Author Citation Analysis (ACA) study begins with a focused list of authors likely to be highly cited in some topic area. In the present case highly cited authors in articles representing Iranian chemistry research over the study period (1990-2006) were extracted using Dialog file 34 (SciSearch) specifying Iran as at least one author's address (s gl=Iran), subject matter of the source journal as Chemistry (s sc= chemistry) and the publication year span (s py=1990:2006). Dialog's RANK command was used to create a list of authors most highly cited in the resulting set of 7682 source articles, of which the top 50 were selected for initial analysis.

2.1 Data gathering

In the process of data gathering, while the initial cited author set was identified using an "Iran-enriched" data set, the co-citation counts for the initial top 50 authors were retrieved from the entire SciSearch database (1990-2006) using standard Dialog commands:

S CA=Shamsipur M and CA=Sharghi H and PY=1990:2006.

All searches were done within the accession number range 00000001-16349341 in spring of 2007. The co-citation counts in a matrix were assembled with identically ordered authors' names on the rows and columns, mean co-citation rate for each author was calculated, and mean of 5 or more co-citations over the 16 year time period was retained. This reduced the number of authors in the sample to 43, that are shown in Table 1, along with an indication of whether the author is "Iranian" (shown in bold phase) or "international" (i.e. all authors who are not clearly associated with an Iranian research institution shown in italics). The table also includes the total number of articles with at least one Iranian byline in which the author is cited. At this stage, there is likely to be some undercounting of "Iranian" citing papers as noted by Wilson and Osareh [4 , p.28] "some collaborations in advanced countries involving scientists from developing countries do not list the home country of the visiting collaborator, particularly in cases where he or she is living for an extended period of time in the advanced country."

Table 1: Highly cited authors in chemistry with an Iranian address

NAME	CITATION COUNT	NATIONAL AFFILIATION
Shamsipur, Mojtaba	**513**	**Iranian**
Firouzabadi, Habib	**512**	**Iranian**
Varma, Rajender S.	*305*	*International*
Iranpour, Nasser	**292**	**Iranian**
Heravi, Majiid M.	**276**	**Iranian**
Izatt, Reed M.	*270*	*International*
Yavari, Issa	**263**	**Iranian**
Olah, Geoerge A.	*263*	*International*
Frisch, Michael J.	*247*	*International*
Bakker, Eric	*245*	*International*
Greene, Theodora W.	*232*	*International*
Ganjali, Mohammad R.	**230**	**Iranian**
Zolfigol, Mohammad A.	**217**	**Iranian**
Umezawa, Yoshio	*212*	*International*
Corey, Elias J.	*205*	*International*
Safavi, Afsaneh	*202*	***Iranian***
Vogel, Arthur I.	*202*	*International*
Hajipoor, Abdol R.	**195**	**Iranian**
Sheldrick, George M.	*195*	*International*
Kamata, Satsuo	*192*	*International*
Buehlmann (Buhlmann), Phillipe	*170*	*International*
Caddick, Stephen	*170*	*International*
Ensafi, Ali A.	**165**	**Iranian**
Ammann, Daniel	*152*	*International*
McKillop, Alexander	*144*	*International*
Yadav, J.S.	*144*	*International*
Tanaka, K.	*138*	*International*
Amini, Mohammad K.	**134**	**Iranian**
Ranu, Brindaban C.	*134*	*International*
Sharghi, Hashem	**131**	**Iranian**
Bard, Allen J.	*128*	*International*
Miller, Jane C.	*126*	*International*
Fakhari, Ali R.	**125**	**Iranian**
Mohammadpoor-Baltork, Iraj	**122**	**Iranian**
Loupy, Andre	*122*	*International*
Tamami, Bahman	**120**	**Iranian**
Katritzky, Alan R.	*119*	*International*
Pedersen, Charles J.	*118*	*International*
Rouhollahi, Ahmad	**115**	**Iranian**
Bandgar, B.P.	*115*	*International*
Shirini, Farhad	**112**	**Iranian**
Dewar, Michael J.S.	*112*	*International*
Becke, Axel D.	*111*	*International*

2.2 Data Analysis

White [11] examined the raw author co-citation matrix to visualize the subject structure of Iran-relevant Chemistry. Schvaneveldt [12] gave Pathfinder Network Analysis (, 1990) to construct a PFNet based on the strongest inter-author raw co-citation counts.

3 Results

Iranian chemical research represented in SciSearch

Fig. 1: Illustrates the growth of the explicitly Iran-authored chemistry literature between 1990 and 2006.

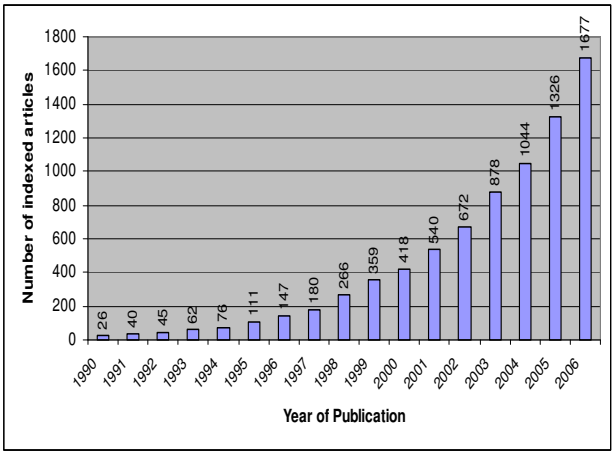

Fig. 1: Growth of Iranian Chemistry articles in SciSearch

The publication count has grown at an approximate rate of 26% per year, with a standard error of 0.004. This should be seen as illustrative and not precise, since the estimate has not been corrected for serial correlation.

It was from this set of papers that the author set in Table 1 was derived. The remainder of the results is not based on this article set but on the matrix of author co-citation counts derived from the entire Dialog File 34 database. It includes citing papers with and without Iranian bylines (and those in which an Iranian co-author does not give his or her home institutional affiliation).

PFNet Analysis

Figures 2 and 3 give two different views of the PFNet linking cited authors by their strongest co-citation counts.

4. Discussion

The discussion with regard to the present research may dwell upon the follwing two points:-

4.1 Iranian chemistry research, 1990-2006: Growth & structure

Figure 1 shows that, over the 16 year time period, the exponential growth in ISI-indexed Iranian chemistry publications documented by Osareh, Wilson and others has continued and that Iranian chemistry research appears to have focused primarily on topics in organic chemistry and, to a lesser extent, analytical chemistry. Iranian chemical research, thus, mirrors the basic trend reported by Schummer [13] for chemistry in general, although his study focused on the pace of discovery of new substances while Iranian chemistry seems to deal more with different methods of synthesis (e.g. solvent free methods, "deoximation" reactions) during the period under study. The findings are also congruent with the report of Yalpani et al. [6] that the young Iranian chemists' studies are published primarily in organic chemistry (50%) and secondarily in analytical chemistry (35%) while organic and analytical chemistry respectively claimed only 9% and 21%, of the world's chemical journal output indexed

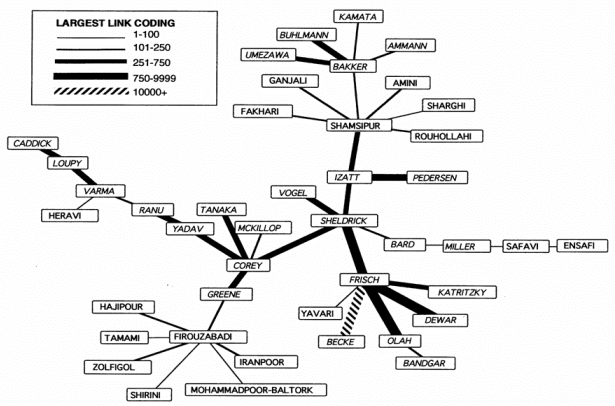

Figure 2 shows the network based on authors' names and strength of co-citation linkage. "Highest linkage" counts ranged from a low of 51 (Shamsipur & Sharghi) to a high of 19632 (Frisch & Becke). The Iranian authors appear primarily as coherent subnetworks with a central figure linking the subnetwork to International authors—Firouzabadi connects one subset to Greene, Shamsipur forms a bridge of Iranian authors between Bakker and Izatt, and Safavi links Ensafi to Bard and thus to Sheldrick and the main network. Heravi and Yavari are single Iranian authors connected to Varma and Frisch, respectively.

Fig. 2: PF Net of Co-Citation Counts for 43 Authors Highly Cited in Iranian Chemistry Articles 1990-2006

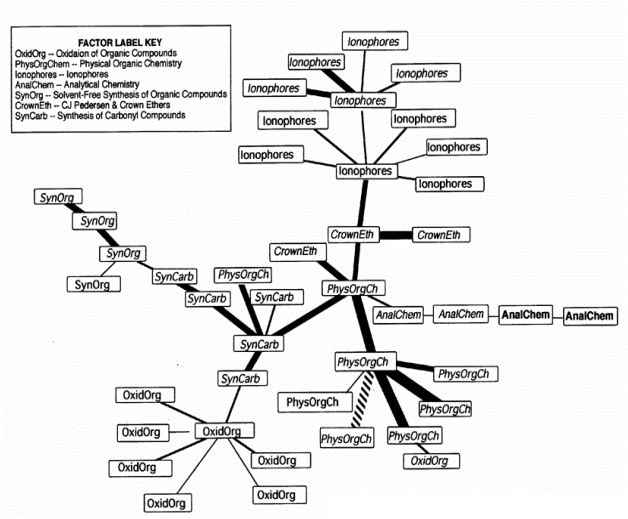

Figure 3: In Figure 3 we have substituted the label of the factor on which each author had the highest loading for that author's name. The authors' primary factor loadings appear as coherent subject groupings—chains or stars—in the PFNet. The subject areas can be split into two subnetworks connected by *Synthesis of Carbonyl Compounds* and *Physical Organic Chemistry*; *Synthesis of Carbonyl Compounds, Solvent-Free Synthesis of Organic Compounds* and *Oxidation of Organic Compounds* on one side and *Physical Organic Chemistry*, *Analytical Chemistry, C.J. Pedersen & Crown Ethers,* and *Ionophores* on the other.

Fig. 3: PF Net of Co-Citation Counts for 43 Authors Highly Cited in Iranian Chemistry Articles with Factor Labels 1990-2006

in ISI. Yalpani et al. [6] attribute this to the specialization of senior Iranian chemists with whom the young chemists get trained, and also to the nature of the existing research facilities and research infrastructure in Iran. The indeginous specialization in chemistry research was also reported by Nagpaul and Pant [14] who observed a similar pattern of high and low emphasis in chemical research in the eleven countries under the purview. For instance , India showed high activity profiles in Inorganic Chemistry, Organic Chemistry, and Polymer Chemistry, Poland in Analytical Chemistry and Chemical Engineering, and Russia in Applied Chemistry, Inorganic Chemistry and Physical Chemistry. Nagpaul and Pant attribute the different patterns of specialization to the national science policies of the various countries.

Looking at the citation rankings in Table 1 (authors highly cited in ISI-indexed Iranian chemistry publications) and the author-labeled PFNet in Figure 2, we can detect only a slight relationship between the level of citedness and position in the network. However, there does appear to be a noticeable geographical/ institutional influence. With the exception of the two most visible Iranian chemists, Firouzabadi and Shamsipur, a high citedness ranking in Table 1 doesn't appear to be reflected in either the number of links an author has in the PFNet or the author's placement. Of the top 10 authors in Table 1, four (Shamsipur, Firouzabadi, Frisch, and Bakker) are linked to 5 or more other authors while Iranpoor, Heravi, and Yavari are peripheral, single-linked authors. Corey and Sheldrick, the other two authors with 5 outlinks, rank 15th and 19th, respectively. We can observe a network "spine" of substantial link counts primarily among International authors—from Caddick to Sheldrick and then branching to (1) Frisch and (2) Izatt. The Iranian authors are primarily located in peripheral subnetworks. Being a Nobel Prize winner does not necessarily place an author in a central position—Corey has a high outlink count but Olah's count is 2 and Pedersen's is 1 (although the strength of connection of these authors to some of their neighbors is high).

4.2 "Quantity over Quality" and the bibliometric structure of Iranian Chemistry

In 2004, Kovacs (2004) published a commentary/letter in *Chemistry and Biodiversity* expressing strong concern about the increasing quantity and lack of quality in Western hemisphere synthetic organic chemistry. In response, Yalpani and Heydari (2005) supported his contention, extended it to non-Western organic chemistry research, and provided bibliometric data for chemistry research in selected third world countries (TWC) over the past three decades. Their discussion focuses primarily on Iran, noting that publication productivity is strongly tied to monetary and other incentives (for faculty) and approval of the PhD (for doctoral students). They report that a content analysis of the publications of faculty and students show (1) a high degree of repetition of

content with minor changes in terminology; (2) repetition of reaction studies with a focus on "deoximation;" (3) a focus by a small number of researchers on *ab initio* calculations (modeling chemical reactions); and (4) a focus by a small number of researchers on the general subject of "oxidation." Their conclusion is that this productive, though highly skewed, pattern of publication, coupled with a decrease in the level of scrutiny by editors and reviewers of low-impact factor international journals, has resulted in a large number of relatively minor, uncited publications useful for vitae and productivity reports but less useful for the advancement of synthetic organic chemistry.

5. Conclusion

Since 1990, Iranian chemistry research, as represented in the ISI-indexed journals, has grown at a rate of roughly 26%. The topic areas are primarily in organic chemistry, and secondarily in analytical chemistry; other major topic areas such as biochemistry, applied chemistry, and chemical engineering are not visible. Author Co-citation Analysis reveals a noticeable influence of cited author national and institutional affiliation in the patterns observed, with most Iranian authors and International authors linked in their own PFNet subnetworks.

Our results illustrate the intellectual structure of Iranian chemistry research but only hint at the forces that underlies this structure. We cannot address Yalpani and Heydari's assertions as to the contributing causes of the publication and citation patterns, but our findings do confirm the patterns' existence in terms of the remarkable publication rate (which can be achieved with many small publications in journals with high acceptance rates), and the use that citing authors make of the oeuvres of chemists highly cited in Iranian chemistry publications. A citation content and context analysis (Small, 1978; McCain & Salvucci, 2006; McCain & Turner, 1989) is beyond the scope of this study but would be a natural next step in assessing how the cited authors' work—both Iranian and International—is actually deployed in the citing authors' writing. Studies of selected Iranian and International authors' Citation Images and Citation Image Makers (White, [10]) would contribute to this as well.

Acknowledgment

Dr. Osareh was supported by Shahid Chamran University, Ahwaz-Iran, as a visiting scholar at the College of Information Science & Technology, Drexel University during the course of this research. We thank Dr. Anthony Addison, Drexel University and Dr. Firouzabadi, Shiraz University, for discussions that helped us understand the chemistry behind the citation patterns. Dr. Roger A. McCain, Drexel University provided the publication growth rate. All errors of fact and interpretation are, however, ours and not theirs.

REFERENCES

1. Osareh, F. and Wilson, C. S. Third world countries (TWC) research publications by disciplines: a country-by-country citation analysis. *Scientometrics,* 39, 1997, 253-266.

2. Osareh, F. and Wilson, C. S. A comparison of Iranian scientific publications in the Science Citation Index: 1985-1989 and 1990-1994. *Scientometrics,* 48, 2000, 427-442.

3. Osareh, F. and Wilson, C. S. Collaboration in Iranian scientific publications. *Libri,* 52, 2002,88-98.

4. Wilson, C. S. and Osareh, F. Science and research in Iran: a scientometric study. *Interdisciplinary Science Reviews*, 28, 2003, 26-37.

5. Moin, M., Mahmoudi, M. and Rezaei, N. Scientific output of Iran at the threshold of the 21st century. *Scientometrics,* 62, 2005, 239-248.

6. Yalpani, M., Heydari, A. and Nehrdad, M. Application of scientometric methods to chemical research in Iran: Reflections on Iran's current science policy. *Scientometrics,* 63, 2005, 531-547.

7. Mehrdad, M., Heydari, A., Sarbolouki, M. and Etemad, S. Basic science in the Islamic Republic of Iran. Scientometrics, 61, 2004,79-88.

8. Yalpani, M. and Heydari, A. Quantity over Quality: *A voice from the third world. Chemistry & Biodiversit*y, 2, 2005, 730-737.

9. Harichi, G., Melin, G. and Etemad, S. An exploratory study of the feature of Iranian co-authorships in biology, chemistry and physics. *Scientometrics,* 72, 2007,11-24.

10. White, H.D. Author-centered bibliometrics through CAMEOS: Characterizations automatically made and edited online. *Scientometrics,* 51, 2001, 607-637.

11. White, H. D. Pathfinder networks and author co-citation analysis: A remapping of paradigmatic information scientists. *Journal of the American Society for Information Science and Technology,* 54, 2003, 423-434.

12. Schvaneveldt, R. W. (Ed). Pathfinder associative networks: Studies in knowledge organization. Norwood, NJ: Ablex, 1990.

13. Schummer, J. Scientometric studies on chemistry I: the exponential growth of chemical substances, 1800-1995. *Scientometrics,* 39,, 1997, 107-123.

14. Nagpaul, P. S. and Pant, N. Cross-national assessment of specialization patterns in chemistry. *Scientometrics,* 27, 1993, 215-235.

15. Kovac, P. Quantity over quality? *An open letter to the community of chemists. Chemistry and Biodiversity,* 1, 2004 , 606-608.

16. McCain, K. W. and Salvucci, L. J. How influential is Brooks' Law? A longitudinal citation context analysis of Frederick Brooks' The Mythical Man-Month. *Journal of Information Science,* 32, 2006, 277-295.

17. McCain, K. W. and Turner, K. Citation content analysis and aging patterns of journal articles in molecular genetics. *Scientometrics,* 17, 1989, 127-163.