Measuring the university–industry–government collaboration in Indian research output

*Rupika, Ashraf Uddin and Vivek Kumar Singh**

Universities, industry and government organizations all play an important role in growth and development of knowledge-based economies in the modern era. These institutions also play a significant role in knowledge creation and its deployment to the benefit of society at large. In this article, we measure and characterize the university–industry–government (UIG) relationship in the research and innovation landscape of India. Research output data for 10 years (2005–14) obtained from Web of Science have been analysed to measure collaboration among different actors of the UIG collaboration network. We have also measured the collaboration variations across different disciplines and identified significant UIG institutional networks. The article presents useful output and analysis, and an informative account of the UIG collaboration network at present.

Keywords: Collaboration, research output, scientometrics, triple helix, university–industry–government.

THE role of university, industry and government has been changing during the recent decades. A university in addition to doing teaching and research also performs functions overlapping with industrial firms, for example call for tenders/bids or running community services or specialized manufacturing units (such as in the case of textiles). Industry no longer performs only output services, processes and products for the people, but also conducts research of its own. Similarly, a government not only frames rules, oversees their implementation and collects taxes and funds, but may also set up administrative bodies to do research or entities that participate in different roles in industrial firms.

The triple helix concept that represents the necessary dynamics among university, industry and government has been elaborated on this basis in the literature. Leydesdorff introduced mutual information as an indicator of the triple helix of relations among university, industry and government, based on the notion of entropy borrowed from Shannon's mathematical theory of information. He interpreted it as a measure of the synergy or information flow between innovation actors. If research activities exploit existing knowledge and produce new ones, the circulation of knowledge between innovation actors ensures its transformation into innovations. Innovation has an important place in industrial development, economic growth and wealth production. It is, therefore, necessary to measure collaboration for research and innovation, just as it is to measure research and innovation.

This article measures and characterizes the university–industry–government (UIG) relationship in the research and innovation landscape of India. Research output data for 10 years (2005–14) have been analysed to measure overall output strength of the UIG collaboration network, UIG instance variations across disciplines, and identify significant UIG institutional networks. The results provide a useful and informative account of the UIG collaborative relationship in India during the recent decade. To the best of our knowledge there has been no previous work of this type in Indian setting.

Background and related work

The notion of triple helix was proposed by Etzkowitz and Leydsdroff in mid-1990s to study the UIG collaboration at local and regional level[1,2]. Three different kinds of triple helix structures were studied, with particular reference to organizations[3]. In triple helix I, the state or government overshadows the university and industry structures. State governs over the university and industry relations. This type of weak structure was found in some Latin American countries. The triple helix II illustrates the working of university, industry and government as isolated bands and separated by strong borders. This structure represents immensely limited relation among the three actors. The most researched triple helix structure, however, is triple helix III, where all three rings overlap each other. Each ring takes part in the role of the others.

Etzkowitz and Leydesdorff employed the model to study the knowledge-based economies. Their result was

Rupika and Ashraf Uddin are in the Department of Computer Science, South Asian University, New Delhi 110 021, India; Vivek Kumar Singh is in the Department of Computer Science, Banaras Hindu University, Varanasi 221 005, India.
*For correspondence. (e-mail: vivekks12@gmail.com)

later used in other works[4] as an operational study for regional development and for incubators in the context of universities in Brazil[5], Sweden[6] and Ethopia[7]. Etzkowitz[8] used triple helix model to study cross-border and co-evolution of technological and institutional alterations in the context of Europe, and compared them with the US scenario. Kivinen and Varelius[9] illustrated the use of triple helix in studying funding pattern in the field of biotechnology in Finland. Etzkowitz *et al.*[3] presented views about the social and economic development changes due to emergence of entrepreneurial universities as connected to UIG. Ye *et al.*[10] studied the consequences of globalization in relation to the triple helix and showed that developed countries have higher UIG collaborations as compared to developing countries. They focused on comparing China and USA. Several other research works on UIG network analysis for different countries/regions have been reported in the literature in the recent past. Realizing the importance and usefulness of the UIG network analysis, Leydesdorff *et al.*[11] introduced an open-source software routine for measuring synergy of relations in triple helix and quadruple helix indicators.

In the Indian context, one of the first works on the theme is by Prathap[12], who evaluated productivity levels of different sectors (UG, UI, IG and UIG) in India and also compared them with data from different countries worldwide. To the best of our knowledge, no follow-up work on this theme has been reported in the literature. A recent useful work on UIG network in the South Asian setting is comparison of ICT infrastructure in South Asian nations[13]. There are some other research works in the Indian and South Asian contexts that are somewhat related to the theme, as they exploited the triple helix model for different purposes. Hossain *et al.*[14] mapped the dynamics of R&D in the context of Bangladesh for the period 1996–2006 using the triple helix model. Datta and Saad[15] studied the outsourcing phenomenon in services domain in India using the triple helix model and social network. Pandey[16] provided an insightful view for communication of innovations in agricultural field in the Indian context at the local level. In the present study, we perform a detailed UIG network characterization of Indian research output during 10 years (2005–2014) and present a comprehensive analytical account.

Data collection and annotations

We have collected data for the research output from India indexed in the *Web of Science* (*WoS*) during the period 2005–2014. The following query has been used to collect the data: CU=India AND LANGUAGE: (English) Indexes= SCI-EXPANDED, SSCI, A&HCI Timespan=2005–2014. A total of 459,164 publication records were obtained. Each of these records contains 60 fields, including metadata like publication type (PT), author names (AU), author address (C1) and WOS subject category (WC). The collected data incorporate records of the type of article, book review, review, meeting abstract, proceedings paper, note, editorial material, letter, etc. and provide a comprehensive coverage of research produce.

The obtained data were computationally processed to remove any duplicate entries and extract unique institution names. We identified a total of 66,968 different institution occurrences mentioned in the data. The first major task was to categorize the names of institutions into three different categories/sectors: university (U), industry (I) and government (G). Three annotators were recruited to tag each institution name into one of the seven tags, namely UA – university from India, IA – industry from India, GA – government organization from India, UO – university outside India, IO – industry outside India, GO – government organization not from India and NA – those institutions which cannot be tagged into any of the six categories. Some institution name entries are incomplete or make no sense, for example. 'A', 'B', 'A Coll', 'Lomb', 'Med Sc', etc. We observed that many of the institutions with smaller number of publications are a result of irregular format in affiliation information or misspelt-data. We removed some of these incorrect entries by deleting from the data all institution names with frequency of occurrence less than 5. Further, the list of the institution names has different entries for the same institution, such as for 'Sam Higginbottom Institute of Agriculture, Technology and Science' (earlier known as 'Allahabad Agricultural Institute'), the names found were: Sam Higginbottom Inst for agriculture, Technology and Science, Allahabad agriculture Inst deemed University, Allahabad Agricultural Inst, Allahabad Agricultural Deemed Univ, and Allahabad Agricultural University. All these names were accordingly tagged into the same category. Finally, a total of 13,312 distinct institution names were tagged manually into the seven categories. Tagging the institution names into different categories/sectors has been a tedious and time-consuming task. Manual tagging of each institution type required a web search to know about the nature of the institution. It took annotators more than 15 days of effort to complete the tagging process.

Identifying UIG network collaboration output

We have identified and measured the UIG collaborations in the research output data for India for the period 2005–2014. Since our main focus is with respect to the Indian settings, we have only taken into account the institution names tagged as 'UA', 'IA' and 'GA', representing Indian university, industry and government institutions respectively. We have identified records that constitute instances of any of the different collaboration sectors: university–industry (UI), university–government (UG), industry–government (IG) and UIG. Figure 1 provides an

overview for the overall UIG collaborations. It is seen that only 135,188 (approx. 29%) of the total 459,164 research papers involve some kind of collaboration. Further, UG is the most prominent collaboration sector. Table 1 lists year-wise publication records and corresponding UI, UG, IG and UIG instances with percentage share. Figure 2 shows the year-wise number of U, I, G, UI, UG, IG and UIG collaborations and presents year-wise growth trend. We observe that in terms of output size, U stands at the top followed by G. These results are different from those observed in 2004 by Prathap, where G is seen as the sector producing most output. This appears to be a natural and welcome change in the Indian settings owing to strong focus on promoting research culture in the universities. In terms of collaboration sector, UG produces more output followed by UI. Thus, UG is the strongest collaboration link followed by UI and IG, which is similar to the status in 2004 (ref. 12). The year-wise output curves for IG and UIG almost overlap each other on the lower side of the plot.

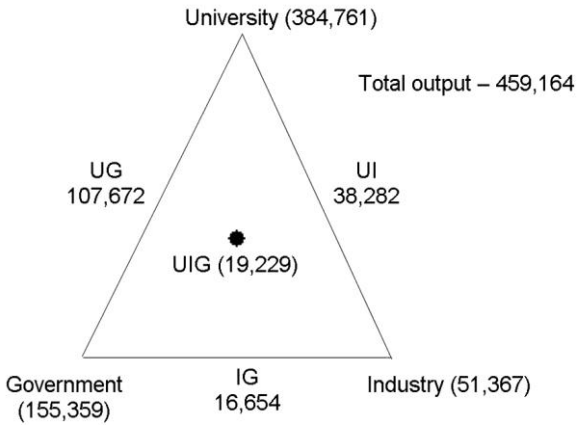

Figure 1. The university–industry–government (UIG) collaboration overview for Indian research output (2005–14).

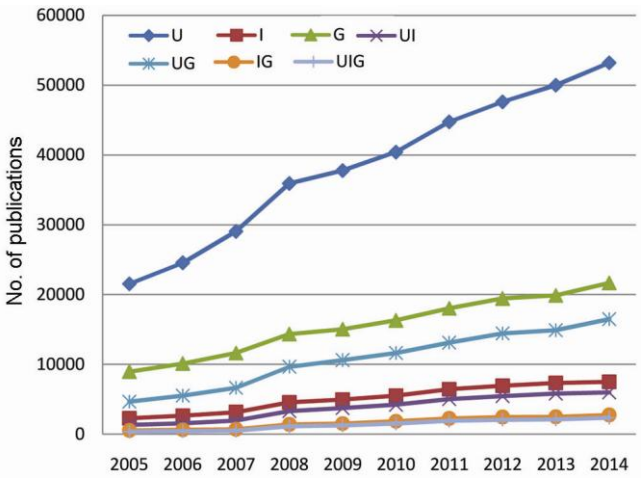

Figure 2. Year-wise research output of UI, UG, IG and UIG collaborations.

Measuring the triple helix indicators

The UIG network can be properly measured using mutual information represented into triple helix at each point in time in terms of probability distribution. Shannon's formula[17] for the theory of distribution is helpful here. Leydesdorff[18] observed that the frequency distribution holds uncertainty, which can be exhibited in terms of bits of information and can be computed using Shannon's formula as follows

$$H(X) = -\sum_{x \in X} p(x) \log p(x). \quad (1)$$

The joint entropy for two discrete variables X and Y is computed as

$$H(X, Y) = -\sum_{x \in X} \sum_{y \in X} p(x, y) \log p(x, y). \quad (2)$$

Here $H(X)$ represents uncertainty. In the context of UIG, entropy values are always positive. The mutual information between two dimensions of the probability distribution is equal to the transmission (T) of the uncertainty and can be defined as follows

$$T(X, Y) = H(X) + H(Y) - H(X, Y). \quad (3)$$

Similarly, for three variables

$$T(X, Y, Z) = H(X) + H(Y) + H(Z) - H(X, Y) - H(Y, Z) - H(Z, X) + H(X, Y, Z). \quad (4)$$

Higher value of T indicates stronger relations, expressed as megabits of information transmission (that mathematically models the expected uncertainty in distribution).

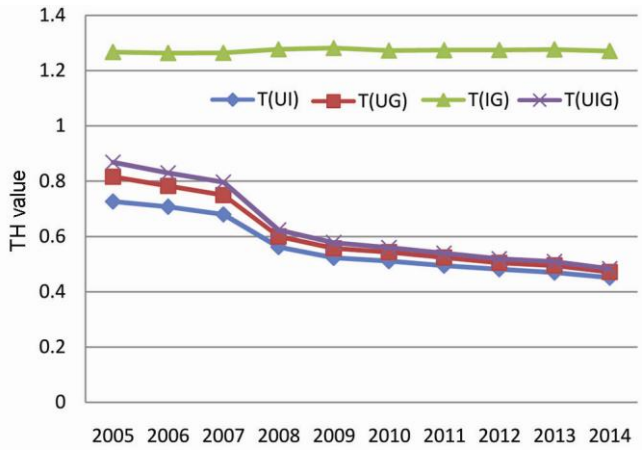

Figure 3. Triple helix values for UI, UG, IG and UIG collaborations.

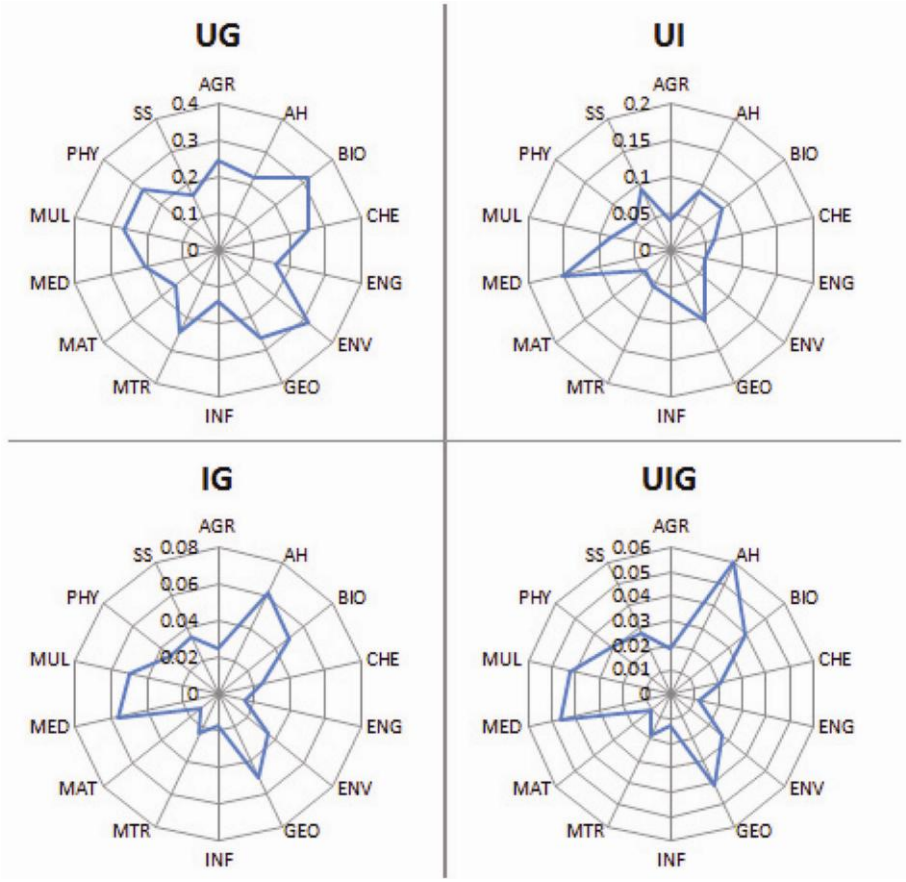

Figure 4. Subject category-wise output variations of different collaboration sectors.

Table 1. Year-wise Publication Records and UI, UG, IG and UIG collaborations

Year	Total records	UI	% Share	UG	% Share	IG	% Share	UIG	% Share
2005	28,167	1321	4.69	4657	16.53	543	1.93	364	1.29
2006	31,645	1549	4.89	5509	17.41	635	2.01	444	1.4
2007	36,906	1934	5.24	6641	17.99	714	1.93	502	1.36
2008	43,425	3309	7.62	9662	22.25	1414	3.26	1125	2.59
2009	44,779	3718	8.3	10,619	23.71	1523	3.4	1263	2.82
2010	48,078	4208	8.75	11,633	24.2	1851	3.85	1525	3.17
2011	52,291	5019	9.6	13,109	25.07	2256	4.31	1926	3.68
2012	55,227	5454	9.88	14,444	26.15	2458	4.45	2078	3.76
2013	57,701	5800	10.05	14,909	25.84	2499	4.33	2122	3.68
2014	60,945	5970	9.8	16,489	27.06	2761	4.53	2361	3.87
Total	459,164	38,282	8.34	107,672	23.45	16,654	3.63	13,710	2.99

UI, University–industry; UG, university–government; IG, industry–government; UIG, University–industry–government.

The *T* values for two dimensions are positive by definition, whereas the values for three dimensions can be positive, negative or zero depending on the distributions[18]. A negative *T* value indicates synergy in the relationship among the UIG and a positive *T* value indicates no synergy.

The bilateral and trilateral alliances for UIG are reckoned by the triple helix indicator for each year and for each subject category. We have computed the triple helix values for UI, UG, IG and UIG. Figure 3 plots the year-wise triple helix values depicting bilateral and trilateral alliances. It can be observed that IG collaboration is low, but persistent throughout the decade, while UI, UG and UIG collaborations have undergone changes. UI is at the bottom of the curves. We can clearly observe a relationship between the difference in publication numbers and triple helix values. For example, when the smallest difference in the number of outputs is seen for IG in Figure 2, the largest triple helix value is seen for IG in Figure 3.

Table 2. Top participant institutions in UI, UG, IG and UIG collaborations

UI	UG	IG	UIG
Indian Institutes of Technology	Indian Institutes of Technology	Indian Institutes of Technology	Indian Institutes of Technology
All India Institute of Medical Sciences	Bhabha Atomic Research Centre	Tata Institute of Fundamental Research	Tata Institute of Fundamental Research
Indian Institute of Science	Council of Scientific and Industrial Research	Bhabha Atomic Research Centre	Bhabha Atomic Research Centre
Guru Nanak Dev University	Tata Institute of Fundamental Research	Council of Scientific and Industrial Research	Council of Scientific and Industrial Research
Bharathidasan University	Indian Institute of Chemical Technology	Panjab University	Panjab University
Tata Institute of Fundamental Research	Indian Institute of Science	University of Delhi	University of Delhi
University of Delhi	University of Delhi	All India Institute of Medical Sciences	All India Institute of Medical Sciences
Panjab University	Indian Association for the Cultivation of Science	Indian Institute of Science	Indian Institute of Science
Bhabha Atomic Research Centre	Banaras Hindu University	UGC–DAE Consortium for Scientific Research	Guru Nanak Dev University
Tata Memorial Hospital	Osmania University	Guru Nanak Dev University	Bharathidasan University

Table 3. Top institution pairs in UI, UG and IG collaborations

Institution 1 (University)	Institution 2 (Industry)	No. of collaborations
Top 5 UI pairs		
Jawaharlal Nehru Technological University	Dr Reddy's Laboratories Ltd	192
Postgraduate Institute of Ophthalmology	Aravind Eye Hospital	99
Maulana Azad Medical College	Lok Nayak Jai Prakash Narayan Hospital	86
Indian Institutes of Technology	Tata Steel Limited	81
Indian Institutes of Technology	Reliance Industries Limited	40
Top 5 UG pairs		
Indian Institutes of Technology	Bhabha Atomic Research Centre	662
Panjab University	Tata Institute of Fundamental Research	610
Indian Institutes of Technology	Tata Institute of Fundamental Research	465
University of Delhi	Tata Institute of Fundamental Research	354
University of Delhi	National Physical Laboratory	281
Top 5 IG pairs		
Tata Memorial Hospital	Advanced Centre for Treatment, Research and Education in Cancer	33
Tata Memorial Hospital	Bhabha Atomic Research Centre	24
National Conservation Foundation	National Centre for Biological Sciences	17
Tata Chemicals Limited	CSIR – National Chemical Laboratory	15
Westbank Hospital	Indian Institute of Chemical Biology	14

Discipline-wise variations

We have also identified discipline-wise variations in the UIG network from India. For this, we have used 14 major subject categories derived from more than 250 provided by the *WoS*. The supplementary data (see online) provide details about these subject categories and mapping from the *WoS* categories. Our aim is to identify which discipline/subject of research involves more collaboration. Also, what kind of collaboration is more prevalent in different disciplines? We have computed the proportion of UI, UG, IG and UIG collaboration outputs for each of the 14 subject categories. Figure 4 shows the discipline-wise variations in collaborated research outputs in each of the collaboration sectors: UG, UI, IG and UIG. We observe that UG collaboration is higher in biological sciences (BIO) and environmental science (ENV). In case of UI collaboration, medical sciences (MED) and geology (GEO) report higher output. The IG collaboration is prevalent in arts and humanities (AH), medical sciences, geology and biological sciences. In UIG collaboration, arts and humanities, medical sciences, multidisciplinary, geology and biological sciences are prominent.

Identifying prominent institution-level UIG networks

UIG network collaborations are the result of participation of different institutions. We have identified the top 10 institutions from university, industry and government sectors which are part of UI, IG, UG and UIG collaborations. For this, we computed the total number of research papers in which pairs of institutions from different sectors participate. For example, in the IG sector all research papers are analysed and those pairs of institutions are extracted where one or more authors are from industry and rest of the authors are from government organization. Table 2 presents the top participant institutions for UI, IG, UG and UIG collaborations. We observe that the Indian Institute of Technology (IIT) is the top participant in all categories of collaboration. In Table 3, we show the most productive five institution pairs for each of the different collaboration categories. The IITs and the Tata Institute of Fundamental Research are the two prominent players with many collaboration output pairs.

Conclusion

We have successfully measured and characterized the patterns of UIG network collaborations in Indian research output during 2005–2014. First, we measured collaborative research output strengths between UI, UG, IG and UIG. Secondly, different collaboration models in UIG collaboration system were analysed, and UI, UG, IG and UIG collaborations characterized in terms of the triple helix model values and percentage output of each collaboration. Top participating institutions and top pairs in each sector have also been identified. Finally, discipline-wise variations in collaboration have been characterized to produce useful inferences. In summary, this article presents an informative and useful account of UIG collaboration network in Indian research output, which could be used for various descriptive and prescriptive purposes.

1. Etzkowitz, H. and Leydesdorff, L., The Triple helix–-university–industry–government relations: a laboratory for knowledge based economic development. *EASST Rev.*, 1995, **14**(1), 14–19.
2. Leydesdorff, L. and Etzkowitz, H., Emergence of a triple helix of university–industry–government relations. *Sci. Pub. Policy*, 1996, **23**(5), 279–286.
3. Etzkowitz, H., Dzisah, J., Ranga, M. and Zhou, C., The triple helix model of innovation: university–industry–government interaction. *Asia-Pac. Tech. Monitor*, 2007, **24**(1), 14–23.
4. Leydesdorff, L., The triple helix of university–industry–government relations. RCLIS Eprints, 2012; available at http://www.leydesdorff.net/th12/th12.pdf
5. Almeida, M., The evolution of the incubator movement in Brazil. *Int. J. Technol. Global.*, 2005, **1**(2), 258–277.
6. Jacob, M., Utilization of social science knowledge in science policy: systems of innovation, triple helix and VINNOVA. *Soc. Sci. Inf.*, 2006, **45**(3), 431–462.
7. Saad, M., Zawdie, G. and Malairaja, C., The triple helix strategy for universities in developing countries: the experiences in Malaysia and Algeria. *Sci. Public Policy*, 2008, **35**(6), 431–443.
8. Etzkowitz, H., The triple helix of university–industry–government: implications for policy and evaluation. Swedish Institute for Studies in Education and Research, Stockholm, Sweden, 2002.
9. Kivinen, O. and Varelius, J., The emerging field of biotechnology – the case of Finland. *Sci. Technol. Hum. Values*, 2003, **28**(1), 141–161.
10. Ye, F. Y., Yu, S. S. and Leydesdorff, L., The triple helix of university–industry–government relations at the country level and its dynamic evolution under the pressures of globalization. *J. Am. Soc. Inf. Sci. Technol.*, 2013, **64**(11), 2317–2325.
11. Leydesdorff, L., Park, H. W. and Lengyel, B., A routine for measuring synergy in university–industry–government relations: mutual information as a triple–helix and quadruple–helix indicator. *Scientometrics*, 2014, **99**(1), 27–35.
12. Prathap, G., Indian science slows down – V: The slack in the university sector. *Curr. Sci.*, 2004, **87**(6), 732–734.
13. Sawyer, B. and Khan, G. F., Mapping ICT knowledge infrastructure in South Asia. *Scientometrics*, 2014, **99**(1), 117–137.
14. Hossain, M. D., Moon, J., Kang, H. G., Lee, S. C. and Choe, Y. C., Mapping the dynamics of knowledge base of innovations of R&D in Bangladesh: triple helix perspective. *Scientometrics*, 2011, **90**(1), 57–83.
15. Datta, S. and Saad, M., Social capital and university–industry–government networks in offshore outsourcing – the case of India. *Technol. Anal. Strat. Manage.*, 2008, **20**(6), 741–754.
16. Pandey, P., Triple helix for communication of innovations: case study of *Bt* cotton in India. *Asian Biotechnol. Develop. Rev.*, 2013, **15**(1), 21–42.
17. Shannon, C., A mathematical theory of distribution. *Bell Syst. Tech. J.*, 1948, **27**(4), 623–656.
18. Leydesdorff, L., The mutual information of university–industry–government relations: an indicator of the triple helix dynamics. *Scientometrics*, 2003, **58**(2), 445–467.

ACKNOWLEDGEMENTS. This research work is supported by UGC-India Research Grant Number 41-642/2012(SR), and DST, India and CONACYT, Mexico Joint Research Grant Number Int/Mexico/P-13/2012.

Received 5 September 2015; revised accepted 11 January 2016

doi: 10.18520/cs/v110/i10/1904-1909