Jointly published by Akadémiai Kiadó, Budapest and Springer, Dordrecht

Scientometrics, Vol. 68, No. 2 (2006) 303–324

Patterns of scientific collaboration between Japan and France: Inter-sectoral analysis using Probabilistic Partnership Index (PPI)

YASUHIRO YAMASHITA,[a,c] YOSHIKO OKUBO[b,c]

[a] *Evaluation and Analysis Office, Yamagata University, Yamagata (Japan)*
[b] *Laboratoire Stratégie & Technologie, Ecole Centrale Paris, Paris (France)*
[c] *National Institute of Science and Technology Policy (NISTEP) (Japan)*

In this article we present an indicator – Probabilistic Partnership Index (PPI) – for use in measuring scientific linkages. This indicator is based on the Monte-Carlo simulation which provides a standard model to each network established in collaboration between two countries. Any relationship that occurs within a (whole) network can be projected to a standard model respectively and thus PPI is useful in examining individual networks within complex exchanges. We investigate inter-sectoral cooperation between France and Japan for the period of 1981–2004, by classifying every research unit appearing in the data set by its sector. We examine international collaborative patterns, domestic collaborative patterns and multilateral relationships established within the French-Japanese cooperation. We also compare PPI with the classic collaborative linkage indexes – Jaccard Index, Salton-Ochiai Index and Probabilistic Affinity Index – in order to describe the specificity of the new indicator. Our hope is that PPI will prove to be a useful and complementary tool for the analysis of international collaboration.

Introduction

The use of co-authored articles to study the nature of collaboration and communication patterns in research activities has been practised for several decades. Price and de Beaver were probably the first to use such data to measure research

Received October 25, 2005

Address for correspondence:
YOSHIKO OKUBO
Laboratoire Stratégie & Technologie, Ecole Centrale Paris
362 bis rue de Vaugirard, 75015 Paris, France
E-mail: okuboy@attglobal.net

0138–9130/US $ 20.00

collaboration, describing collaboration provocatively as "a method for squeezing papers out of the rather large population of people who have less than a whole paper in them".[1]

However, the use of co-authorship data to examine international scientific collaborative activities is more recent; some early studies appeared in the late 1970s, but it was not until the 1990s that international co-authorship analyses abounded and the use of this tool diversified. Rosen and de Beaver[2] demonstrated that the rise of teamwork in its modern form is inherent to the professionalization of science. Frame and Carpenter[3] brought to light a correlation between the quantity of international collaboration and the scientific size of a country. This indicator has been taken up by the National Science Foundation for use in its periodical publication, *Science & Engineering Indicators*,[4] and it was re-assessed more recently by Melin.[5]

Some early work on international scientific collaboration was carried out by Tijssen and Moed,[6] the French Laboratoire d'Evaluation et de Prospective Internationales[7] and Schubert and Braun.[8] A co-authorship indicator has been used to analyze international collaborative patterns and the development of internationalization in science and technology.[9–13] Numerous studies have drawn on co-authorship data to observe publication patterns and international networks of one nation at various levels and in diverse disciplines. Examples include studies at the country-level – Mexico,[14] China,[15] Japan[16] and India[17] among many others; at the sector or institutional-level;[18–21] at the authors' level,[22–24] or different levels combined.[25] It was also used to investigate regional collaboration networks such as in Latin America[26] and in the European Community.[27–30] The role of external research collaboration in small science systems was explored by Thorsteinsdottir.[31] All of these studies made use of indicators devised from international co-authorship data.

On the other hand, the effect of "proximity" on the density of international partnerships has been investigated. The effect of geographical proximity on the strength of scientific links between two partners has been demonstrated,[29,32] while Miquel and Okubo[33] and Zitt et al.[13] showed models in which cultural proximity overrides geographical proximity. "Social proximity", meaning the similarity of status in the social stratification of science has been found to be significant at the individual level.[34]

In parallel, methods to measure collaborative linkages were developed. The application of similarity coefficients is one such method. While many similarity coefficients were proposed in various research fields, the Jaccard[35] Index and the Salton-Ochiai[36] Index (also known as a cosine index) have been the most commonly used in scientometric studies. More recently, Zitt et al. introduced a probabilistic indicator to measure the strength of scientific linkages between partners.[13]

While co-authorship is by now well-established as an international indicator, much remains to be done in the use of co-authorship to further our understanding of international cooperation. For example, until now, investigation of international collaboration patterns at the sector level has largely focused on one of two cross-border partners, as a way of investigating the collaborating actors of a particular country. In order to understand the linkage patterns of bilateral cooperation, it is also necessary to observe collaborative linkages *between sectors belonging to different countries*. It is also vital to examine domestic networks created respectively by two countries, as well as multilateral exchanges with 3rd countries, undertaken within bilateral collaborations. In such analysis, it will be useful to re-assess existing indicators that measure the strength of scientific linkages, so that the prominent features of each index may be brought to light.

The objective of this article is to analyze international cooperation by use of an indicator, the Probabilistic Partnership Index (PPI). We will investigate inter-sectoral cooperation between France and Japan for the period of 1981–2004, by classifying every research unit appearing in the data set by its sector. We will examine international collaborative patterns of the two countries and investigate domestic collaborative patterns and multilateral relationships established within the French-Japanese cooperation. We will use PPI indicators to examine cooperation which is composed of multi-level networks. We will compare PPI with the classic collaborative linkage indexes – Jaccard Index, Salton-Ochiai Index and Probabilistic Affinity Index – in order to describe the specificity of the new indicator. Our hope is that PPI will prove to be a useful and complementary tool for the analysis of international collaboration.

Methods

Data

We used Japan-France Collaboration Citation Report (JFCCR) data set compiled by Thomson Scientific, Philadelphia. It contains the number of publication and citation of international collaboration between Japan and France, for the period of 1981 to 2004. The JFCCR is extracted from the Web of Science which covers data from the Science Citation Index (5500 journals) and the Social Science Citation Index (1800 journals). Due to the changes in the counting method at the Thomson Scientific undertaken in 2004, the data of 2003 onward seem particularly affected by this new policy.[*] As the changes and the effects are not clear at this moment, our analysis is based on 5-years or on all-period-combined.

Four types of documents are included in the data set – articles, letters, notes and reviews. There are 9,830 articles co-authored between research units in Japan and France and published during the period of 24 years investigated.

[*] Information obtained through personal communication with Thomson Scientific, October 25, 2004 10:43.

Sectors

Research units in the data set are classified into sectors: (U)niversity, (P)ublic, (C)orporation, (N)on-profit, (H)ospital, (O)thers, Un(k)nown and (I)nternational. The countries of origin of units are classified into: (F)rance, (J)apan and (O)ther (C)ountries. The first letter of each category represents countries and sectors and they are combined as: JU, FN, FH etc., except for Other Countries (OC).

In the ISI data, the institutional addresses are divided into org(anization), dep(ar)t(ment), lab(oratory) and sect(ion). We have classified the research units (institutions) listed in the org column into sectors according to the abbreviations attached to the institutions such as: UNIV, COLL, SA, LTD, HOSP etc. For the research units which could not be classified into sectors by this method, internet search, telephone calls, and consultation with several experts were undertaken one-by-one. 1.3% of the total of institutions (331 French and 81 Japanese units) could not be identified and are gathered in the Unknown category.

Whole counting method

One of the prominent characteristics of the structure of the French R&D institutions is that there are numerous *mixed institutions* or *joint institutions* of various research organizations. In 2002, for example, 90% of all the laboratories or units related to Centre National de la Recherche Scientifique (CNRS), the largest research institution in France, are joint organizations with universities, Grandes Ecoles (French higher education for Pro Master degree) and other public research organizations.[37] The INRA and the INSERM also have the mixed structure, as do many other public research organizations. In the same manner, hospital CHU (Centre Hospitalier Universitaire) is also a mixed structure with universities and public organizations. In the recent years, the State has encouraged research implemented by joined efforts of diverse establishments.[38] The condition of the agglomeration varies according to the joint unit, but in most cases the mixed units are composed of researchers who come from different organizations with conditions of work determined individually by each organization of origin. The mixed units are indicated as UMR, UA, URA, UPRES, EMI, EA etc. before their institution names.

In our study, the French mixed units or joint laboratories are classified into two sectors: public institution and university. The hospital CHU is classified into 2-3 sectors: hospital and university and/or public institution. Some Grandes Ecoles, such as the Ecole Polytechnique, are classified into university and public institution.

We have used the whole counting method and attributed 1 participation to every sector in a research unit. The participation of one unit at the Ecole Polytechnique, for example, is counted as 2 participations, 1 university and 1 public institution.

In contrast, such mixed structure is rare in Japan except for the following institutions: hospitals administrated by universities or some public research institutions, such as the University of Tokyo Hospital and the National Cancer Center Hospital (attributed 1 hospital and 1 university or public); a research consortium the Joint Research Center for Atom Technology (JRCAT, attributed as both public and non-profit) and some research programs of the Japan Science and Technology Corporation (JST, now Japan Science and Technology Organization, attributed as both public and another sector).

This way of counting, i.e. attributing credits to every sector in a single research unit, was selected for the present study in order that the sector-comparison can be undertaken "symmetrically" between Japan and France, in spite of the great difference of the institutional structure between the two countries. We attempt to acknowledge all sector-participants under a common research project and describe collaborative profiles which are the closest to the reality.

There are total of 39,324 sector-linkages identified in 9,830 co-authored articles investigated.

The problems we have encountered in the classification of sectors were first the difficulty of comparing two different R&D systems between Japan and France, the major difference being the existence of the mixed structure in France that hardly exists in Japan. The second difficulty was to compare the evolution of sectors for the period as long as 24 years, for the simple reason that institutions change. The turnover of laboratories in the CNRS is approximately 7-8% per year; the structures and names of institutions change and with it the labels representing structures such as UPR, UMR, UA change. Some public laboratories such as the CNET (France) and the NTT (Japan) are privatised, firms come and go, and it is rare that start-up companies last or remain status quo for some decades. In addition, the changes occur differently in the two countries. During the period examined, the ISI's database set also changed. These factors rendered the comparison of R&D development for a large number of years at the sector level particularly difficult.

Collaboration indexes

Jaccard and Salton-Ochiai Indexes. Salton-Ochiai and Jaccard Indexes are used to measure relative overlap of links. Both indexes can be calculated either by the number of co-authored articles or collaborative links established in co-authorships. Meyer et al.[39] compared them and other six similarity coefficients – Sorensen-Dice, Anderberg, Simple Matching, Roger & Tanimoto, Ochiai II and Russel & Rao – and described that Salton-Ochiai and Jaccard Indexes (as well as Sorensen-Dice and Anderberg) have relatively similar characteristics due to their exclusion of negative co-occurences. In this article we will concentrate on Salton-Ochiai and Jaccard Indexes to investigate research collaboration, since the other six coefficients seem to have rarely been used in scientometric studies.

Here, let m_{ij}, m_i and m_j be the number of co-authored articles of sector i and j, and number of articles of sector i and j respectively, Salton-Ochiai Index for articles (SOa) is m_{ij} divided by geometric mean of m_i and m_j: $SOa = m_{ij} / (m_i \times m_j)^{1/2}$.

Jaccard Index (JDa), on the other hand, is the number of articles co-authored between sectors i and j divided by the union of articles of both sectors: $JDa = m_{ij} / (m_i + m_j - m_{ij})$.

Salton-Ochiai and Jaccard Indexes for collaborative links (SOl and JDl) are formulated as follows: $SOl = n_{ij} / (n_i \times n_j)^{1/2} = m_{ij} / (n_i \times n_j)^{1/2}$, $JDl = n_{ij} / (n_i + n_j - n_{ij}) = m_{ij} / (n_i + n_j - n_{ij})$, where n_{ij} be the number of links between sector i and j (=co-authored articles m_{ij}), and n_i, n_j be the number of links of sector i and j respectively. In the present study, we adopt SOa and JDa, since $n_i \geq m_i$ for any sector i SOa ≥ SOl and JDa ≥ JDl are valid because of dense linkage of the data.

Any value of Jaccard is below that of Salton-Ochiai due to the fact that the dominator of Jaccard is greater than or equal to that of Salton-Ochiai. This is proven as follows:

Given that $m_i \geq = m_j > 0$,

$$m_i + m_j - m_{ij} - (m_i \times m_j)^{1/2} \geq m_i + m_j - m_{ij} - (m_i \times m_i)^{1/2} = m_j - m_{ij} \geq 0.$$

The equal sign is valid only when $m_i = m_j = m_{ij}$ and in the condition SOa=JDa=1.

Probabilistic Affinity Index (PAI). Probabilistic Affinity Index (PAI),[13] a ratio of observed and expected number of links, indicates relative tendency of co-authorship and is formulated as follows: $PAI = n_{ij} / E[n_{ij}] = n_{..} \times n_{ij} / (n_i \times n_j)$.

Let $E[n_{ij}]$ be expected value of n_{ij} and $n_{..}$ be sum of all cells of the contingency matrix (diagonal be discarded). From the different point of view, PAI is an "activity index" of collaborative links. The value more than 1 indicates that there are more collaborative links than expected, and contrary, less than 1 indicates less links. For convenience, we re-normalized PAI into value between –1 to 1 using the following equation: Re-normalized PAI = $(PAI^2 - 1) / (PAI^2 + 1)$. PAI hereafter will indicate the re-normalized value.

Probabilistic Partnership Index (PPI). Probabilistic Partnership Index (PPI), standardized difference of observed number of links from expected, provides another view of deviation from expected value. PPI is formulated as follows: $PPI = (n_{ij} - \hat{E}_a[n_{ij}])/\hat{\sigma}$, where $\hat{E}_a[n_{ij}]$ and $\hat{\sigma}$ are expected value and standard deviation of distribution of number of links between sector i and j under the constraint of the number of articles and current participants, which is estimated by the Monte-Carlo method we will mention in next section. PPI is a standard score of n_{ij} against the probability distribution of current participants without any preference for collaborating partners. PPI is based on the idea that if all participants in Japan-France co-operation could randomly choose articles in which they contribute, how would collaborative links be distributed, and to what extent the actual distribution of links differs from such model?

PPI=0 indicates that the observed number of links equals with that of the expected value; PPI>0 indicates that the observed value is greater than the expected value and vice versa for PPI<0. By normalizing by the standard deviation, the effect of sector sizes can be reduced. For example, if $n_{ij}/\hat{E}_a[n_{ij}]$ is used as an indicator, index values would be close to 1 in the case of large sectors with more than half of all articles, because the number of articles would be much greater than the differences between observed and expected values. Although PPI ranges between $-\infty$ and $+\infty$, values are renormalized into ranges between –1 and 1, in order to enable comparison with PAI: re-normalized PPI = $2 / (1 + \exp(-0.183\times\text{PPI})) - 1$.

The coefficient 0.183 of logistic function is arbitrarily decided according to the criteria that the PPI value which deviates $6\hat{\sigma}$ from the estimated expected value corresponds to the re-normalised index value 0.5. PPI hereafter will indicate re-normalized value.

Estimation of expected values of collaborations. Expected value of number of articles of Japan-France international-intersectoral co-authorship is used to assess the difference from the actual co-authorship pattern. For the estimation of the expected values, we assumed that each participant (a subgroup composed of researchers within a sector) takes part in a randomly selected article among all of the Japan-France co-authored articles. This assumption requires the following two conditions: (1) co-existence of more than one same sector within an article is prohibited, since it is assumed that sub-groups in a same sector is organized in advance; (2) all articles must be attributed to at least one sector of Japan and France.

In the model, the probability that a participant of sector i takes part in an article is not calculated from m_i/M, because of the constraint (1), where M is the number of articles. When a sector i appears at probability of m_i/M, articles which are not attributed to a sector appear at probability of $(1-m_1/M)\times(1-m_2/M)\times\ldots\times(1-m_k/M)$, where k is the number of sectors of a country. In the case that gross number of sectors is not much greater than the number of articles, m_i/M is the overestimation of the probability. Thus, we employed the Monte-Carlo method for evaluating the expected values and standard deviations of the number of links of all patterns of cooperation. Monte-Carlo method is appropriate to solve complex problems which can not be solved analytically, such as a growth of researchers' personal networks in relation to the increase of the number of published articles in a discipline.[40]

In each trial of simulation, every participant is attributed to an article randomly selected at same probability from all the articles. Number of links for all patterns of cooperation is counted in each trial of the simulation, and then their means and standard deviations of 10000 trials are calculated.

For reducing computing time, 2-steps of attribution process were executed; (i) all articles were attributed to one participant randomly selected from all sectors in Japan-France collaboration; (ii) each of the rest of the participants was assigned to a randomly selected article that had not been attributed to the same sector. This process assumes a situation that the principal participant of each article (which does not always correspond to the first author, because the order of authors in each article is ignored in present study) randomly selected chooses their partners randomly. Statistical computing software "R"[41] was used for executing Monte-Carlo simulation, and Mersenne Twister[42] was used as pseudo random number generator.

Results

Basic observations

Table 1 illustrates the distribution of sectors in cooperation between Japan and France for the period of 1981–2004.

It shows a contrasting pattern of scientific performers participating in cooperation, and the difference seems to reflect the characteristics of the scientific systems of the two countries. The principal participants of Japan in the collaboration are universities, while in France, they are universities and public research institutions, mainly owing to the mixed laboratories as mentioned above. One can also see the difference in the degree of participation of corporation; Japanese firms participate more than three times in cooperative activity than do French firms. Intensive participation of firms in the overall R&D activity is one of the prominent features of Japanese knowledge production system, which is considerably different from that of France.

The development of sectors in the Japan-France cooperation is depicted in Tables 2a and 2b. In 24 years, Japanese public institutions doubled their share from 8.8% in 1981–1985 to 19.5% in 2000–2004. Counterbalanced by such increase, universities decreased their share, from 82.2% to 67.8%, although in absolute term their contribution increased by seven during the same period. All in all, Japanese participants have diversified: corporations, non-profits, hospitals and others all increased their shares.

Table 1. Distribution of sectors in France-Japan cooperation 1981–2004

Sector	France %	Japan %
University	36.8	69.9
Public	44.9	16.3
Corporation	2.3	7. 4
Non-profit	4.0	3.3
Hospital	5.7	2.3
Others	0.3	0.2
Unknown	2.1	0.7
International	3.9	
	100.00	100.00

Table 2a. Distribution of Japanese sectors, 1981–2004
Percent of sectors as of the total Japanese sectors in France-Japan cooperation

Years	Univ.	Public	Corp.	Non-profit	Hospital	Other	Unknown
1981–1985	82.2	8.8	5.1	2.0	1.2	0.0	0.6
2000–2004	67.8	19.5	5.9	3.3	2.4	0.2	0.7

Table 2b. Distribution of French sectors, 1981–2004
Percent of sectors as of the total French sectors in France-Japan cooperation

Years	Univ.	Public	Corp.	Non-profit	Hospital	Internat'l	Other	Unknown
1981–1985	37.4	45.5	1.8	5.0	4.5	2.9	0.2	2.7
2000–2004	38.0	44.1	2.3	2.9	5.5	4.3	0.3	2.5

In contrast, the degree of participation of French sectors remained fairly stable over the period studied: universities' share is around 38% in each period, while the share of public institutions slightly decreased, from 46% to 44%, between the period 1981–1985 and 2000–2004. Non-profits somewhat decreased, while international organizations in France increased their share from 2.9% to 4.3%. During this period, French sector-participation increased nine times in absolute term.

One of the greatest changes in the collaboration during the two decades is the increase of the share of third countries' participation, from 11.4% (1981–1985) to 16.2% (2000–2004). Publication analysis, on the other hand, reveals that 67% of all the Japan-France co-publications carry affiliations of third countries. The collaboration between Japan and France is therefore far from a bilateral undertaking spontaneously established between scientists, and this multi-nationality has manifestly developed during the decades examined.

Comparison of Collaboration Indexes

Salton-Ochiai and Jaccard Indexes. Table 3 shows the values of Salton-Ochiai and Jaccard Indexes of Japan-France inter-sectorial relationships during 1981–2004. In the table, cells above diagonal denote Salton-Ochiai (SOa) and those below diagonal denote Jaccard (JDa). The mean of Salton-Ochiai cells is 0.11, which is more than twice that of Jaccard (0.05).

At a glance, we can observe that the values of domestic links are smaller than those of Japan-France or Japan/France-Other Country relationships. The means of the domestic relationships are only 0.060 (SOa) and 0.020 (JDa) for Japan, and 0.078 (SOa) and 0.035 (JDa) for France, whereas they are 0.118 (SOa) and 0.055 (JDa) for Japan-France relationships, 0.213 (SOa) and 0.107 (JDa) for Japan-Other Countries, and 0.258 (SOa) and 0.129 (JDa) for France-Other Countries. Two indexes thus reveal that stronger ties are created multilaterally than bilaterally, i.e. between France-Japan-3rd countries than between only France and Japan. As the data we use are the co-authorships containing Japan-France linkages extracted from the SCI, the linkages of Other Countries (OC) do not necessarily appear in this data set. But here, the presence of other countries is extraordinary, which implies that the Japan-France cooperation is largely performed in a multilateral framework.

Table 3. Salton-Ochiai and Jaccard Indexes of Japan–France inter-sectoral relationships 1981–2004

JDa \ SOa	JU	JP	JC	JN	JH	JO	JK	FU	FP	FC	FN	FH	FI	FO	FK	OC
JU		*0.20*	*0.10*	*0.09*	*0.12*	*0.02*	*0.04*	*0.70*	*0.77*	*0.11*	*0.22*	*0.27*	*0.21*	*0.05*	*0.15*	*0.63*
JP	0.08		*0.07*	*0.08*	*0.04*	*0.02*	*0.04*	*0.32*	*0.38*	*0.08*	*0.11*	*0.08*	*0.15*	*0.03*	*0.08*	*0.32*
JC	0.03	0.03		*0.04*	*0.03*	*0.03*	*0.02*	*0.22*	*0.21*	*0.24*	*0.09*	*0.06*	*0.05*	*0.03*	*0.05*	*0.16*
JN	0.02	0.03	0.02		*0.15*	*0.03*	*0.04*	*0.12*	*0.16*	*0.06*	*0.11*	*0.09*	*0.06*	*0.03*	*0.06*	*0.15*
JH	0.02	0.01	0.01	0.08		*0.01*	*0.05*	*0.09*	*0.08*	*0.04*	*0.08*	*0.29*	*0.02*	*0.04*	*0.07*	*0.13*
JO	0.00	0.00	0.00	0.01	0.00		*0.03*	*0.02*	*0.03*	*0.01*	*0.02*	*0.02*	*0.01*	*0.04*	*0.01*	*0.03*
JK	0.00	0.01	0.01	0.02	0.02	0.01		*0.07*	*0.07*	*0.04*	*0.03*	*0.03*	*0.03*	*0.05*	*0.04*	*0.07*
FU	0.53	0.16	0.08	0.03	0.02	0.00	0.01		*0.66*	*0.07*	*0.12*	*0.16*	*0.07*	*0.04*	*0.07*	*0.47*
FP	0.62	0.18	0.07	0.04	0.02	0.00	0.01	0.49		*0.06*	*0.16*	*0.18*	*0.10*	*0.04*	*0.08*	*0.59*
FC	0.02	0.03	0.12	0.03	0.02	0.00	0.01	0.02	0.01		*0.02*	*0.04*	*0.00*	*0.00*	*0.03*	*0.13*
FN	0.06	0.05	0.05	0.05	0.04	0.00	0.01	0.04	0.04	0.01		*0.13*	*0.00*	*0.00*	*0.04*	*0.18*
FH	0.09	0.04	0.03	0.04	0.14	0.00	0.01	0.06	0.06	0.02	0.07		*0.02*	*0.02*	*0.06*	*0.26*
FI	0.06	0.07	0.03	0.03	0.01	0.00	0.01	0.02	0.03	0.00	0.00	0.01		*0.00*	*0.01*	*0.24*
FO	0.00	0.01	0.01	0.01	0.02	0.02	0.02	0.00	0.00	0.00	0.00	0.01	0.00		*0.01*	*0.05*
FK	0.03	0.03	0.02	0.03	0.04	0.00	0.01	0.02	0.02	0.01	0.02	0.03	0.01	0.00		*0.15*
OC	0.44	0.16	0.06	0.04	0.03	0.00	0.01	0.30	0.41	0.03	0.06	0.10	0.08	0.00	0.04	

The first character of each title denotes J(apan), F(rance) and O(ther countries), and the second character denotes U(niverstiy), P(ublic), C(ompary), H(ospital), I(nternational organization), O(ther) and (un)K(nown), except for the case of "OC" which denotes "Other Countries". The cells above diagonal are values of Salton-Ochiai Index and below diagonal are that of Jaccard Index.

As to individual relationships in Table 3, two indexes show that JU-FP (SOa=0.77, JDa=0.62), JU-FU (SOa=0.70, JDa=0.53), JU-OC (SOa=0.63, JDa=0.44) and FP-OC (SOa=0.59, JDa=0.41) are the strongest international relationships. Furthermore, Japan-Other Countries relationship shows that the link values are greater than SOa=0.10 (except for JO-OC and JK-OC), whereas most of the France-Other Countries' links exceeded SOa=0.10, except for FO-OC. On the other hand, while Japanese domestic relationships do not exceed SOa=0.2, FU-FP (SOa=0.66, JDa=0.49) shows a remarkably strong linkage among French domestic relationships, followed by FP-FH (SOa=0.18), FP-FN and FU-FH (SOa=0.16 respectively) indicated by Salton-Ochiai, and FN-FH (JDa=0.07), FP-FH and FU-FH (JDa=0.06 respectively) indicated by Jaccard. Despite the overall small values of the domestic linkages of both countries, FU-FP shows an exceptionally high value, reflecting the existence of numerous mixed structures between university and public sector in France.

As to Japan-France bilateral relationship, Salton-Ochiai Index shows a strong linkage of more than 0.20 for JU-FP (0.77), JU-FU (0.70), JP-FP (0.38), JP-FU (0.32), JH-FH (0.29), JU-FH (0.27) and JC-FC (0.24), while Jaccard Index shows a higher value of more than 0.10 for JU-FP (0.62), JU-FU(0.53), JP-FP (0.18), JP-FU (0.16), JH-FH (0.14) and JC-FC (0.12). These observations indicate that not only the linkages established between universities and public institutions are strong, but also those created between the same sectors of both countries (such as JC-FC, JN-FN and FH-FH) are particularly strong.

In spite of the fact that above results show tight linkages, it is important to consider the size of sectors and its effect on these links. JU (appear in 80.5% of total co-authorships), FP (70.9%), FU (57.2%) and OC (50.7%) participated in more than half of total Japan-France co-publications during 1981–2004. It seems rather rationale that when a researcher seeks a partner(s) in other sectors, he will select them according to their scientific size, in addition to his preference. If this is the case, researchers in larger sectors would have greater possibility of becoming partners of those in other sectors, which will result in larger Salton-Ochiai and Jaccard values. The question we must ask here is: to what extent large values of indexes reflect the strength of collaboration? or to what extent the size of sectors affects the index values? In order to answer these questions, we will next present expected values of both indexes to compare them with the observed values of Table 3. The expected values of Salton-Ochiai and Jaccard Indexes are calculated by 10000 times trials of Monte-Carlo simulation, in which index values are calculated in each trial and mean values of all trials are adopted as estimated expected values. Table 4 shows expected values of Salton-Ochiai and Jaccard Indexes of Japan-France cooperation.

All the large values in Table 4 are found in the linkages of above-mentioned large-sized sectors (JU, FU, FP and OC), and the order is the same as the observed values: JU-FP (SOa=0.76, JDa=0.60), JU-FU (SOa=0.68, JDa=0.51), JU-OC (SOa=0.64, JDa=0.45), FP-OC (SOa=0.60, JDa=0.42), FU-FP (SOa=0.55, JDa=0.38) and FU-OC (SOa=0.54, JDa=0.37). These values are all close to observed values except for FU-FP. By contrast, 3 links between small sectors, JC-FC, JN-FN and JH-FH, which show outstanding values in Table 3, are not as large. These observations suggest that the outstanding results revealed by the two indexes are mainly due to the size of partners.

By plotting observed and expected values of both indexes into scatter diagram, correlation between them can be clearly grasped. The observed value presented in each cell of Table 3 and the expected value presented in each cell of Table 4 are plotted in Figures 1 (Salton-Ochiai) and 2 (Jaccard). Links (dots) located near y = x means that links have neutral strength and the index value would mainly be decided by their size, whereas, links (dots) distant (above or below) from y = x have relatively strong/weak links. For instance, though four links between JU-FP, JU-FU, JU-OC and FP-OC mark large index values as measured by both Salton-Ochiai and Jaccard Indexes, as the dots indicating them are located near the y=x, these links should be considered as neutral and not strong. By contrast, despite the relatively smaller values of JH-FH and JC-FC, they are relatively strong as they are located above and far from y = x.

Table 4. The expected values of Salton-Ochiai and Jaccard Indexes
Japan–France inter-sectoral relationships calculated by Monte-Carlo simulation 1981–2004

SOa / JDa	*JU*	*JP*	*JC*	*JN*	*JH*	*JO*	*JK*	*FU*	*FP*	*FC*	*FN*	*FH*	*FI*	*FO*	*FK*	*OC*
JU		*0.20*	*0.13*	*0.09*	*0.08*	*0.02*	*0.04*	*0.68*	*0.76*	*0.17*	*0.22*	*0.27*	*0.22*	*0.06*	*0.16*	*0.64*
JP	0.09		*0.04*	*0.02*	*0.02*	*0.00*	*0.01*	*0.33*	*0.37*	*0.08*	*0.11*	*0.13*	*0.11*	*0.03*	*0.08*	*0.31*
JC	0.04	0.02		*0.01*	*0.01*	*0.00*	*0.01*	*0.22*	*0.25*	*0.05*	*0.07*	*0.09*	*0.07*	*0.02*	*0.05*	*0.21*
JN	0.02	0.01	0.01		*0.01*	*0.00*	*0.00*	*0.15*	*0.16*	*0.04*	*0.05*	*0.06*	*0.05*	*0.01*	*0.04*	*0.14*
JH	0.01	0.01	0.01	0.00		*0.00*	*0.00*	*0.13*	*0.14*	*0.03*	*0.04*	*0.05*	*0.04*	*0.01*	*0.03*	*0.12*
JO	0.00	0.00	0.00	0.00	0.00		*0.00*	*0.03*	*0.03*	*0.01*	*0.01*	*0.01*	*0.01*	*0.00*	*0.01*	*0.03*
JK	0.00	0.00	0.00	0.00	0.00	0.00		*0.07*	*0.08*	*0.02*	*0.02*	*0.03*	*0.02*	*0.01*	*0.02*	*0.06*
FU	0.51	0.16	0.08	0.04	0.03	0.00	0.01		*0.55*	*0.11*	*0.14*	*0.17*	*0.14*	*0.04*	*0.10*	*0.54*
FP	0.60	0.17	0.08	0.04	0.03	0.00	0.01	0.38		*0.13*	*0.17*	*0.20*	*0.17*	*0.05*	*0.12*	*0.60*
FC	0.04	0.03	0.03	0.02	0.02	0.00	0.01	0.03	0.03		*0.03*	*0.04*	*0.03*	*0.01*	*0.02*	*0.13*
FN	0.06	0.05	0.04	0.02	0.02	0.00	0.01	0.04	0.05	0.01		*0.05*	*0.04*	*0.01*	*0.03*	*0.18*
FH	0.09	0.06	0.05	0.03	0.02	0.00	0.01	0.06	0.07	0.02	0.02		*0.05*	*0.01*	*0.03*	*0.21*
FI	0.06	0.05	0.04	0.02	0.02	0.00	0.01	0.04	0.05	0.01	0.02	0.02		*0.01*	*0.03*	*0.18*
FO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		*0.01*	*0.05*
FK	0.03	0.03	0.02	0.02	0.02	0.00	0.01	0.02	0.03	0.01	0.01	0.02	0.01	0.00		*0.13*
OC	0.45	0.16	0.08	0.04	0.03	0.00	0.01	0.37	0.42	0.03	0.06	0.08	0.06	0.00	0.03	

Abbreviations are the same as Table 1. Cells above diagonal denote estimates of expected values of Salton-Ochiai and below diagonal that of Jaccard Index.

Coefficient between the observed and expected Salton-Ochiai is 0.92 for all cells and 0.74 when the 6 strongest links are excluded (JU-FP, JU-FU, JU-OC, FP-OC, FU-FP and FU-OC), whereas, Jaccard's coefficient is 0.95 (for all cells) and 0.71 (excluding above 6 links). These observations suggest that the sector size is one of the most prominent factors that determine index values.

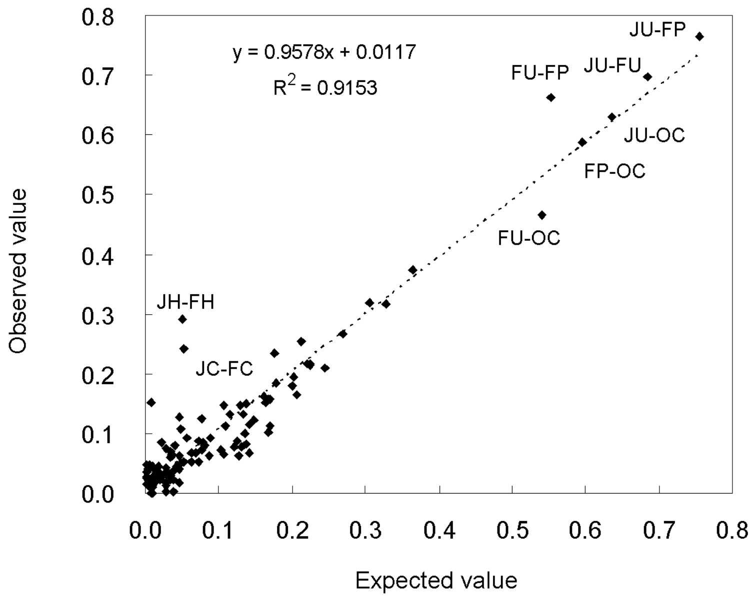

Figure 1. Expected and observed value of Salton-Ochiai Index

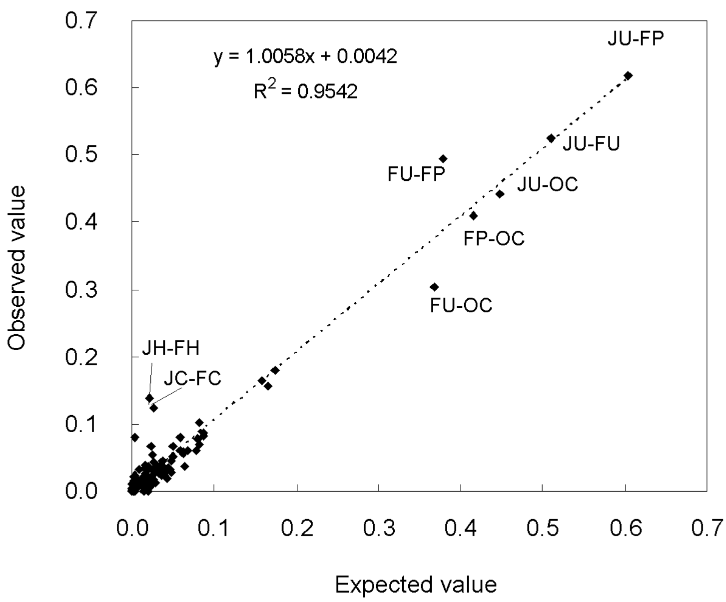

Figure 2. Expected and observed value of Jaccard Index

Probabilistic Affinity Index and Probabilistic Partnership Index. Probabilistic Affinity Index (PAI) measures the relative strength of each co-operative link in comparison with the total linkage, while Probabilistic Partnership Index (PPI) measures the rareness of occurrence of observed value, in comparison with an assumed distribution of links generated by randomly distributing participants within the current articles. PPI considers that all participants randomly select partners and they do not have particular preference in choosing their counterpart. PPI is calculated by attributing each participant to sub-network according to its situation, and the index values are established by comparing the situation to the average situation of the sub-network. This means that PPI is useful in analyzing a set of connection which consists of multi-level networks. It is also helpful in rating linkages according to their probability of occurrence under a given situation.

When PAI is applied to examine relatively sparse sub-networks (within a dense network), as they are analyzed as an integrated part of the entire set-up, they tend to be underestimated when there is a dense network. The sub-networks calculated by PAI are therefore affected by the denseness or sparseness of one another and provides a relative density of sub-networks. PPI, on the contrary, takes into account the difference of each set of exchanges, and, as it provides normalized forms to each one of the relationships, networks can be analyzed individually. It thus enables the comparison of all networks without being affected by a particularly dense relationship.

PAI and PPI values of Japan-France relationships are shown in Table 5. As a whole, PAI has larger absolute mean values than PPI, for example, in French domestic (PAI –0.39 vs. PPI –0.16), Japan-France (PAI 0.27 vs. PPI 0.04), Japan-Other Countries (PAI 0.08 vs. PPI 0.04) and France-Other Countries (PAI 0.20 vs. PPI 0.05) cooperation. Japanese domestic linkages (PAI 0.05 vs. PPI 0.34) are the exception. These results may be due to the fact that PPI is calculated for each sub-network, whereas PAI is calculated for the entire network.

The exceptionally strong linkages observed in Japanese domestic network when measured by PPI, is mainly due to a relatively large proportion of multi-sectoral linkages within the domestic relationship. JO and JK appear at extremely small frequency, while JN-JH linkage is quite strong (PAI 0.9 vs. PPI 1.0). This is partially due to the fact that some of the major Japanese non-profit organizations, such as the Cancer Institute, have hospitals (classified into both non-profit organization and hospital) and when such hospitals publish research articles they produce numerous JN-JH links.

Concerning the French domestic linkages, both indexes show that most of the links are negative, except for FU-FP and FN-FH which are largely above average.

Table 5. PAI and PPI value of Japan–France inter-sectoral relationships, 1981–2004

PPI \ *PAI*	*JU*	*JP*	*JC*	*JN*	*JH*	*JO*	*JK*	*FU*	*FP*	*FC*	*FN*	*FH*	*FI*	*FO*	*FK*	*OC*
JU		*–0.4*	*–0.5*	*–0.5*	*–0.1*	*–0.4*	*–0.4*	*0.4*	*0.4*	*0.0*	*0.2*	*0.2*	*0.3*	*0.1*	*0.2*	*0.3*
JP	–0.2		*–0.3*	*0.0*	*–0.5*	*–0.1*	*0.1*	*0.2*	*0.3*	*0.2*	*0.1*	*–0.4*	*0.5*	*0.2*	*0.2*	*0.2*
JC	–0.6	0.6		*–0.3*	*–0.4*	*0.7*	*–0.3*	*0.3*	*0.2*	*0.9*	*0.3*	*–0.2*	*0.0*	*0.4*	*0.2*	*0.0*
JN	0.0	0.8	0.4		*0.9*	*0.8*	*0.6*	*–0.1*	*0.1*	*0.6*	*0.6*	*0.4*	*0.3*	*0.6*	*0.5*	*0.1*
JH	0.7	0.4	0.3	1.0		*0.5*	*0.7*	*–0.3*	*–0.5*	*0.2*	*0.5*	*0.9*	*–0.5*	*0.8*	*0.7*	*0.0*
JO	0.0	0.2	0.4	0.4	0.2		*0.9*	*–0.3*	*–0.1*	*0.5*	*0.5*	*0.2*	*0.0*	*1.0*	*0.5*	*0.0*
JK	0.0	0.5	0.2	0.6	0.7	0.4		*0.0*	*0.0*	*0.7*	*0.2*	*0.1*	*0.2*	*1.0*	*0.6*	*0.0*
FU	0.4	–0.2	–0.1	–0.3	–0.5	–0.1	0.0		0.3	–0.4	–0.3	–0.3	–0.6	–0.2	–0.4	0.0
FP	0.4	0.2	–0.6	0.0	–0.7	–0.1	–0.1	1.0		*–0.5*	*–0.1*	*–0.2*	*–0.4*	*–0.2*	*–0.4*	*0.2*
FC	–0.8	0.0	0.9	0.2	0.0	0.0	0.2	–0.5	–0.8		*–0.7*	*–0.3*	*–1.0*	*–1.0*	*0.1*	*0.2*
FN	–0.2	0.0	0.1	0.5	0.3	0.1	0.1	–0.4	–0.2	–0.1		*0.5*	*–1.0*	*–1.0*	*0.0*	*0.1*
FH	0.0	–0.5	–0.2	0.3	1.0	0.0	0.1	–0.2	–0.3	0.0	0.7		*–0.9*	*0.2*	*0.1*	*0.2*
FI	–0.2	0.4	–0.2	0.1	–0.2	0.0	0.0	–0.8	–0.8	–0.3	–0.4	–0.3		*–1.0*	*–0.8*	*0.5*
FO	–0.2	0.0	0.0	0.2	0.3	0.3	0.4	0.0	–0.1	–0.1	–0.1	0.1	–0.1		*–0.4*	*0.1*
FK	–0.2	0.1	0.0	0.3	0.4	0.1	0.2	–0.4	–0.6	0.1	0.2	0.3	–0.2	0.0		*0.3*
OC	–0.1	0.2	–0.5	0.2	0.2	0.0	0.0	–0.9	–0.2	0.0	0.1	0.5	0.7	0.0	0.2	

Abbreviations denote same as Table 1. Cells above diagonal denote PAI and below denote PPI.

Three linkages between same sectors of each country show particularly strong ties in Japan-France relationship: JC-FC, JN-FN and JH-FH. Furthermore, JU-FU, JP-FP and JO-FO are also strong, all above zero. On the contrary, JH-FP and JP-FH are shown to be below zero by both indexes.

Japan and France's relationships with Other Countries show different values by two indexes. While all PAI values are nearly or greater than zero, PPI indicates four linkages below zero with particularly low value of JC (PPI –0.5) and FU (PPI –0.9). Among all collaborative linkages of France and Japan created with Other Countries respectively, two indexes show that FI-OC linkage is much greater than average (PAI 0.5 vs. PPI 0.7), the greatest of all the FI values measured by PAI and PPI. This result indicates that international organizations in France have a particularly strong preference to collaborate with Other Countries, even in the context of Japan-France co-operation.

We will next examine the similarities and differences of two indexes. Scatter diagram shows an overview of relationships between PAI and PPI (Figure 3). If dots are located in the first or the third quadrant, the values of the two indexes are the same, whereas if they are located in the second or the fourth quadrant, the values are opposite.

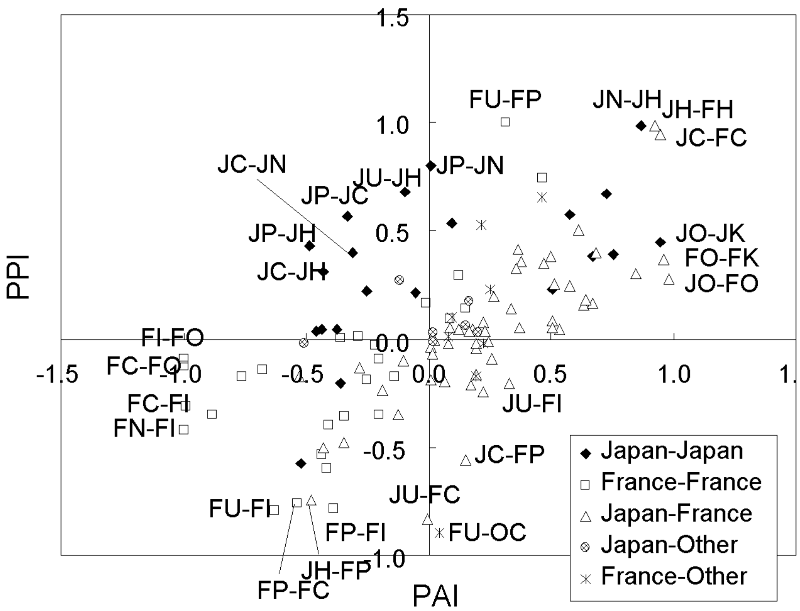

Figure 3. PAI and PPI indexes by type of linkage

The values scattered in the figure seem to be weakly correlated (Spearmans's rank correlation coefficient $\rho = 0.63$). As most of the French domestic relationships are located in the first and the third quadrants, the two indexes show almost identical values for these linkages. However, the French domestic relationships are distributed in the region of PPI > PAI in the first quadrant as well as in two regions where PPI < PAI and near PAI = –1 in the third quadrant. These positions of French domestic linkages seem to be affected by the difference of size of sectors which affects more the absolute values of PAI than those of PPI. Expected values of linkages between small sectors tend to be small, and often less than 1. In such situation, if only 1 or 2 co-operative links exist, ratios of expected values to observed values tend to be large, and the absolute PAI also tend to be large. However this is not always the case with PPI owing to the normalization by standard deviation.

In contrast, there are many Japanese domestic linkages located in the second quadrant. Japanese domestic relationships are mainly distributed in the first and second quadrants with only two exceptions in the third quadrant. Though, most of Japan-France linkages have positive correlations between PAI and PPI, some of them, such as JU-FI and JC-FP, fall into the fourth quadrant.

What are the factors that make such difference? One answer could be the difference of the nature of two indexes. Two indexes are based on the idea of expected number of links, but the concepts are not identical. Expected number of links for PAI is based on the ratio of current links unconstrained by participants, whereas PPI is based on the average links which could be affected by participants. As such, PPI is conditioned by current participants and the number of articles.

Furthermore, PPI uses standard deviation of distribution of the number of links created by random participants which compensates the size effect. PPI therefore demonstrates the rareness of occurrence. For example, the FU-FP linkage is positive measured by both indexes, but PPI value is much larger than that of PAI. In this case, the expected number of links being 3016 and the observed 4189, un-renormalized PAI is 1.4 (4189/3016). This process and value seem rather natural and intuitively clear, but information on how rare the observed value occur under the expected value is not considered in the PAI index. Expected number of links and the standard deviation of obtaining PPI estimated by Monte-Carlo simulation, on the other hand, are 3483.9 and 18.3, respectively, therefore un-renormalized PPI (standard score) is 38.5 (Figure 4). This value suggests that this situation does not occur accidentally under the premise that no preference exists between FU and FP.

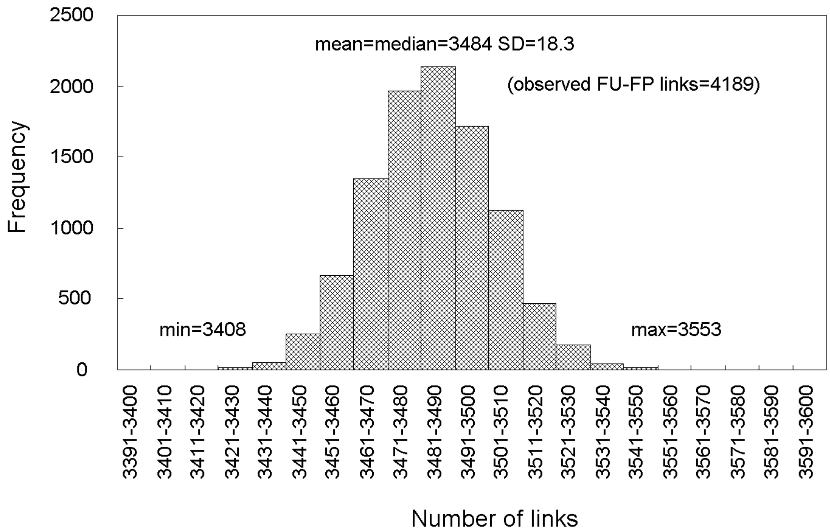

Figure 4. Frequency distribution of FU-FP links in the Monte-Carlo simulation

In fact, none of the participants are independent in the model assumed in PPI. One sector's preference, especially those between large-sized sectors (such as FU-FP), would affect other sectors' distribution and biases all articles regardless of their true preference. Thus, we should bear in mind that information on deviation from estimated expected value of PPI means deviation from field in which all participants does not have any preference for partners.

In such sphere, PPI provides a new way of observing international relationship. It highlights the importance of sub-networks that exist within a network of collaboration between countries and brings to light the way each participant (sector) is situated in that network.

Discussion and conclusions

We have used co-authorships to examine research collaboration between Japan and France over a period of 24 years. We have tried to classify research units according to sector in a way that the sectors of two countries with considerably different scientific structure can be compared "symmetrically". The factor which hindered the straight comparison of the two systems is the mixed laboratory structure which exists in France but not in Japan. We have tried to cope with this divergence by applying whole count participation to every sector belonging to one unit. This required detailed micro-level scrutiny of the classification of institutions.

Our second objective was to present an indicator – Probabilistic Partnership Index (PPI) – for use in measuring scientific linkages. This indicator is based on the Monte-Carlo simulation which provides a standard model to each network established in collaboration between two countries. Any relationship that occurs within a (whole) network can be projected to a standard model respectively and thus PPI enables all sub-networks to be investigated individually.

PPI differs from other similarity indexes such as the Salton-Ochiai and Jaccard Indexes which measure the strength of linkages between partners without taking into account the scientific size of two partners. While Jaccard and Salton-Ochiai Indexes provide important information on relative overlaps of linkages, one of their characteristics is that they are partially determined by the size of sectors, and this feature is especially prominent in collaboration between two large-sized partners. When the size of a sector differs greatly from one country to the other, these indexes run the risk of a skewed analysis of inter-sectoral cooperation. It is therefore essential to complement these indexes by other measures.

PAI is an index to measure collaborative linkages taking into account the size difference of the two counterparts. It indicates, like PPI, deviation from expected values of collaborative linkages. PAI calculates the expected value from the number of links, whereas PPI is calculated from the current participants. The signification of expected values is somewhat different between the two indexes. The expected value of PPI is affected by the condition of the current scientific size of each sector, while that of PAI is affected by the condition of the current links which are separated from articles (thus scientific size of each sector is not assumed to be constant).

PPI is a yardstick which gauges collaborative relationships by comparing each exchange with a model relationship created by ten thousand simulations. This model signifies the most frequent form of relationship and offers "a standard" of how an "ideal" collaboration pattern would be. The divergence from the model can be observed for each relationship.

The inter-sectoral collaboration pattern of two countries largely reflects the characteristics of the scientific systems of the two countries. Japanese private firms, for example, participate in international undertakings three times more readily than French firms, which reflects features of the overall S&T activity of Japan and France. As to the actors of scientific cooperation, the degree of participation of various sectors hardly changed between 1981–1985 and 2000–2004 in France, whereas it has by and large diversified in Japan with institutions other than universities and public research institutes increasing their shares in international activity. This trend also reflects the intensive policy of developing internationalization in Japan in the late 1980s to 1990s. One common feature is that strong interactions seem to occur between same sectors.

One of the most prominent characteristics of Japan-France collaboration is that it has developed primarily in a multilateral setting rather than in a purely bilateral relationship. In this context, stronger domestic networks have been created in France than in Japan. In addition to French universities and public institutions becoming intensely mixed, French sectors tend to construct a tight domestic network within international activity, whereas Japanese sectors tend to collaborate with international partners independently from their colleagues. All in all, collaboration policy does not seem to be clearly defined between the two countries, where 67% of collaborative activity is undertaken in a multi-lateral framework. The majority of co-authorship created between France and Japan is rather the result of research conducted under international programs than of relations spontaneously established among scientists, and this multi-nationality has manifestly developed during the decades examined.

Collaboration between France and Japan needs therefore to be studied within the context of multilateral cooperation. PPI is useful in examining individual networks within complex exchanges. It can bring to light the entire communication pattern of "a relationship" by taking into account every local network.

We have presented the usefulness of PPI for analyzing multi-level cooperation networks, but some aspects of the index need further investigation. First, the links under the condition of randomized actors are implicitly assumed to follow normal distribution.

However, the result of 10000 trials of the simulation showed a skewed distribution for many of the links. Further detailed study concerning the shapes of link distribution seems necessary.

Secondly, we adopted –0.183 as the coefficient of renormalization according to arbitrary determined criteria which suggested that the un-renormalized value of $6\hat{\sigma}$ corresponded to a renormalized value of 0.5. Further examination of the criteria for determining the coefficient in accordance with the conditions of the cooperative network in each study will be necessary.

Thirdly, it was also perceived that large deviations from the expected values of links between large sectors affect other links of PPI values, regardless of their real tendency. A method for eliminating effects of other links should be developed to deal with such situation. Exploring these avenues and answering these questions will be the focus of future work.

*

Authors are grateful to Thomson Scientific for their technical assistance.

References

1. D. J. DE SOLLA PRICE, D. BEAVER, Collaboration in an invisible college, *American Psychologist,* (1966) 1011–1018.
2. D. DE BEAVER, R. ROSEN, Studies in scientific collaboration, Part I – The professional origins of scientific co-authorship, *Scientometrics,* 1 (1) (1978) 65–84.
3. J. D. FRAME, M. CARPENTER, International research collaboration, *Social Studies of Science*, 9 (1979) 481–497.
4. *Science & Engineering Indicators,* National Science Board, Virginia (1987 and every two years onwards).
5. G. MELIN, Impact of national size on research collaboration – a comparison between Northern European and American universities, *Scientometrics*, 46 (1) (1999) 161–170.
6. R. J. W. TIJSSEN, H. F. MOED, *Science and Technology Indicators,* DSWO Press, Leiden, 1989.
7. J. F. MIQUEL, Y. OKUBO, N. NARVAEZ, L. FRIGOLETTO, Les scientifiques sont-ils ouverts à la coopération internationale? *La Recherche,* 20 (1989) 116–118.
8. A. SCHUBERT, T. BRAUN, International collaboration in sciences, *Scientometrics*, 19 (1–2) (1990) 3–10.
9. Y. OKUBO, J. F. MIQUEL, L. FRIGOLETTO, J. C. DORÉ, Structure of international collaboration in science; typology of countries through multivariate techniques using a link indicator, *Scientometrics* 25 (2) (1992) 321–351.
10. T. LUUKKONEN, R. J. W. TIJSSEN, O. PERSSON, G. SIVERSTEN, The measurement of international scientific collaboration, *Scientometrics,* 28 (1) (1993) 15–36.
11. J. F. MIQUEL, Y. OKUBO, Structure of international collaboration in science, Part II: Comparisons of profiles in countries using a link indicator, *Scientometrics* 29 (2) (1994) 271–297.
12. W. GLÄNZEL, C. DE LANGE, Modeling and measuring multilateral co-authorship in international scientific collaboration, Part II. A comparative study on the extent and change of international scientific collaboration links, *Scientometrics*, 40 (3) (1997) 605–626.
13. M. ZITT, E. BASSECOULARD, Y. OKUBO, Shadows of the past in international cooperation: Collaboration profiles of the top five producers of science, *Scientometrics*, 47 (3) (2000) 627–657.

14. J. M. RUSSELL, Publishing patterns of Mexican scientists: Differences between national and international papers, *Scientometrics,* 41 (1–2) (1998) 113–124.
15. B. MELY, M. ABD EL KADER, G. DUDOGNON, Y. OKUBO, Scientific publications of China in 1994: Evolution or revolution? *Scientometrics,* 42 (1) (1998) 3–16.
16. Y. OKUBO, J. F. MIQUEL, International scientific collaboration of Japan: Co-authorship analysis, *Journal of Science Policy and Research Management,* (Kenkyu Gijutu Keikaku Gakkai) 6 (4) (1991) 261–280.
17. P. S. NAGPAUL, Transnational linkages of Indian science: A structural analysis, *Scientometrics,* 46 (1) (1999) 109–140.
18. B. STEFANIAK, International cooperation of Polish researchers with partners from abroad: A scientometric study, *Scientometrics,* 41 (1–2) (1998) 155–167.
19. R. J. W. TIJSSEN, E. VAN WIJK, The global science base of information and communication technologies, *Scientometrics,* 42 (1) (1998) 41–60.
20. M. MEYER, O. PERSSON, Nanotechnology–Interdisciplinarity, patterns of collaboration and differences in application, *Scientometrics,* 42 (2) (1998) 195–205.
21. M. PEREIRA FRIEDRICH, P. DOS SANTOS RODRIGUES, Looking at science in Brazilian universities: The case of the Instituto de Biofisica Carlos Chagas Filho, *Scientometrics,* 42 (2) (1998) 247–258.
22. M. ABD EL KADER, T. OJASOO, J. F. MIQUEL, Y. OKUBO, J. C. DORÉ, Hierarchical author networks: An analysis of European Molecular Biology Laboratory (EMBL) publications, *Scientometrics,* 42 (3) (1998) 405–421.
23. E. YURTSEVER, S. GULGOZ, The increase in the rate of publications originating from Turkey, *Scientometrics,* 46 (2) (1999) 321–335.
24. S. KYVIK, O. PERSSON, Scientometric research in the Nordic countries – Introduction, *Scientometrics,* 49 (1) (2000) 3–6.
25. P. G. DASTIDAR, Ocean science & technology research across the countries: Global scenario, *Scientometrics,* 59 (1) (2004) 15–27.
26. I. GOMEZ, M. T. FERNANDEZ, J. SEBASTIAN, Analysis of the structure of international scientific cooperation networks through bibliometric indicators, *Scientometrics,* 44 (3) (1999) 441–457.
27. F. NARIN, E. S. WHITLOW, *Measurement of Scientific Co-peration and Co-authorship in CEC-related Areas of Science*, Commission of the European Communities, Brussels, 1990.
28. H. MOED, R. E. DE BRUIN, ANTON J. NEDERHOF, ROBERT J. W. TIJSSEN, International scientific cooperation and awareness within the European Community: Problems and perspectives, *Scientometrics,* 21 (3) (1991) 291–311.
29. M. LECLERC, Y. OKUBO, L. FRIGOLETTO, J.-F. MIQUEL, Scientific co-operation between Canada and the European Community, *Science and Public Policy*, 19 (1) (1992) 15–24.
30. Y. OKUBO, M. ZITT, Searching for research integration across Europe: A closer look at international and inter-regional collaboration in France, *Science and Public Policy,* 31 (3) (2004) 213–226.
31. O. H. THORSTEINSDOTTIR, External research collaboration in two small science systems, *Scientometrics,* 49 (1) (2000) 145–160.
32. J. S. KATZ, Geographical proximity and scientific collaboration, *Scientometrics,* 31 (1994) 31–34.
33. J. F. MIQUEL, Y. OKUBO, Structure of international collaboration in science, Part II: Comparisons of profiles in countries using a link indicator, *Scientometrics,* 29 (2) (1994) 271–297.
34. H. KRETSCHMER, Analysis of social relations in coauthorship networks, *JISSI,* 2 (2–3) (1996) 75–89.
35. P. JACCARD, Nouvelles recherches sur la distribution florale, *Bulletin de la Societe Vaudoise des Sciences Naturelles,* 44 (1908) 223–270.
36. A. OCHIAI, Zoogeographical studies on the soleoid fishes found in Japan and its neighbouring regions – II, *Bulletin of the Japanese Society of Scientific Fisheries,* 22 (1957) 526–530.
37. LABINTEL, UNIPS-CNRS, In: *Composante 5: Recherche Publique*, Report FutuRIS, ANRT, Paris, 2004 p. 3.
38. K. WARTA, C. MORICEAU, S. BUSSILLET, *Evaluation de la politique de contractualisation avec les universities – contractualisation et recherche*, Final Report submitted to CNE, Technopolis France, Paris, July, 2003.

39. A. DA S. MEYER, A. A. F. GARCIA, A. P. DE SOUZA, C. L. DE SOUZA JR., Comparison of similarity coefficients used for cluster analysis with dominant markers in maize (Zea mays L), *Genetics and Molecular Biology*, 27 (2004) 83–91.
40. F. YOSHIKANE, K. KAGEURA, Comparative analysis of coauthorship networks of different domains: The growth and change of networks, *Scientometrics,* 60 (3) (2004) 433–444.
41. R DEVELOPMENT CORE TEAM, *R: A Language and Environment for Statistical Computing.* Retrieved March 1, 2005 from: http://www.R-project.org, ISBN 3-900051-00-3 (2004).
42. M. MATSUMOTO, T. NISHIMURA, Mersenne Twister: A 623-dimesionally equidistributed uniform pseudo-random number generator, *ACM Transactions on Modeling and Computer Simulation,* 8 (1) (1998) 3–30.