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

Yasuhiro Yamashita*, Yoshiko Okubo**

*yaya@na.rim.or.jp
Research Institute, Mitsui Knowledge Industry
Harmony Tower 21F, 1-32-2 Honcho, Nakano-ku, Tokyo 164-8721, Japan.

**okuboy@attglobal.net
Laboratoire Stratégie & Technologie, Ecole Centrale Paris
362 bis rue de Vaugirard, 75015 Paris, France.

Abstract
The objective of the present research is to analyze patterns of international scientific cooperation by use of an indicator, Probabilistic Partnership Index (PPI). We will investigate inter-sectoral cooperation between Japan and France. We will examine diverse levels of collaborative networks - domestic, bilateral and multi-lateral relationships - established within France-Japan cooperation. We will present PPI and compare it with three existing indicators in order to bring to light specificity of the new indicator. We intend to compare and develop bibliometric methods for measuring collaborative strengths between partners.

1 Introduction
The objective of the article is to analyze international cooperation by use of *Probabilistic Partnership Index (PPI)*. We will investigate inter-sectoral cooperation between France and Japan for the period of 1981-2004, by classifying every "research institution" appeared in the data set into "sectors". We will examine international collaborative patterns of two countries and investigate domestic collaborative patterns and multi-lateral relationships established within France-Japan cooperation. We present PPI indicator, obtained by Monte-Carlo simulation, which will enable us to examine cooperation composed of multi-level networks. PPI provides a normalized standard to each domestic sub-network which is formed within international cooperation. Each exchange within an entire network can be projected to its standard and be explored individually. 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. We intend to provide another complementary tool to analyze international collaboration.

Methods

Basic data
We use data derived from the Japan-France Collaboration Citation Report (JFCCR) compiled by Thomson Scientific, Philadelphia. JFCCR contains 10,642 publications co-signed between institutions of Japan and France, for the period of 1981 to 2004, out of which four types of documents are investigated in the present study - articles, letters, notes and reviews (9,830 co-publications).

"Institutions" 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 institutions are classified into: (F)rance, (J)apan and (O)ther (C)ountries. The first letter of each category represents sectors and countries and they are combined as: JU, FN, FH etc. except for "Other Countries" (OC).

The whole count method is applied for counting institutions. Total of 39,324 sector-linkages were identified in 9,830 co-publication investigated.

Proceedings of ISSI 2005, pages 517–526.
Edited by P. Ingwersen and B. Larsen.

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. 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 number of links of sector i and j respectively. In the present study, we adopt SOa and JDa, since $n_i >= m_i$ 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 proved as follows:

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

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: Zitt et al. 2000), 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 $PAI = (1 - PAI^2) / (1 + PAI^2)$. 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: $PPI = (n_{ij} - E_a[n_{ij}])/\sigma$, where $E_a[n_{ij}]$ and σ 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 m_{ij} against the probability distribution of population 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 $n_{ij}/E_a[n_{ij}]$ 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 -∞ and +∞, values are renormalized into ranges between -1 and 1, in order to enable comparison with PAI:

re-normalized PPI = $2 / (1 + \exp(-0.183 \times PPI)) - 1$.

The coefficient 0.183 of logistic function is arbitrarily decided according to the criteria that the PPI value which deviates 6σ from the 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) take part in a randomly selected article among all of the Japan-France co-authored articles. This assumption requires the following two constraints: (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, probability that a participant of sector i take part in an article is not calculated from m_i/M, because of the constraint (1), where M is 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 overestimated. 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. 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. Statistical computing software "R" (R Development Core Team 2004) was used for executing Monte-Carlo simulation, and Mersenne Twister (Matsumoto & Nishimura 1998) was used as pseudo random number generator.

Results

Comparison of Collaboration Indexes

In this section, we will analyze the structure of the Japan-France collaboration by using three existing indexes and PPI. We will compare the characteristics of each index in order to apply them to practical uses.

Salton-Ochiai and Jaccard Indexes

Table 1 shows the values of Salton-Ochiai and Jaccard Indexes of Japan-France 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 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" do not necessarily appear in this data set. But the presence of other countries is extraordinary, suggesting that the Japan-France cooperation is largely performed within a multilateral framework.

Table 1: Salton-Ochiai and Jaccard Indexes of Japan–France inter-sectoral relationships 1981-2004

SOa / JDa	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 1, 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 all 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 shows tight linkages, it is important to point out the size effect of the sectors to 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 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?

We will next present expected values of both indexes to compare them with the observed values of Table 1. 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 2 shows expected values of Salton-Ochiai and Jaccard Indexes of Japan-France cooperation.

Table 2: The expected values of Salton-Ochiai and Jaccard Indexes
Japan–France inter-sectoral relationships calculated by Monte-Carlo simulation 1981-2004

JDa \ SOa	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.

All the large values in Table 2 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 1, 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 1 and the expected value presented in each cell of Table 2 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.

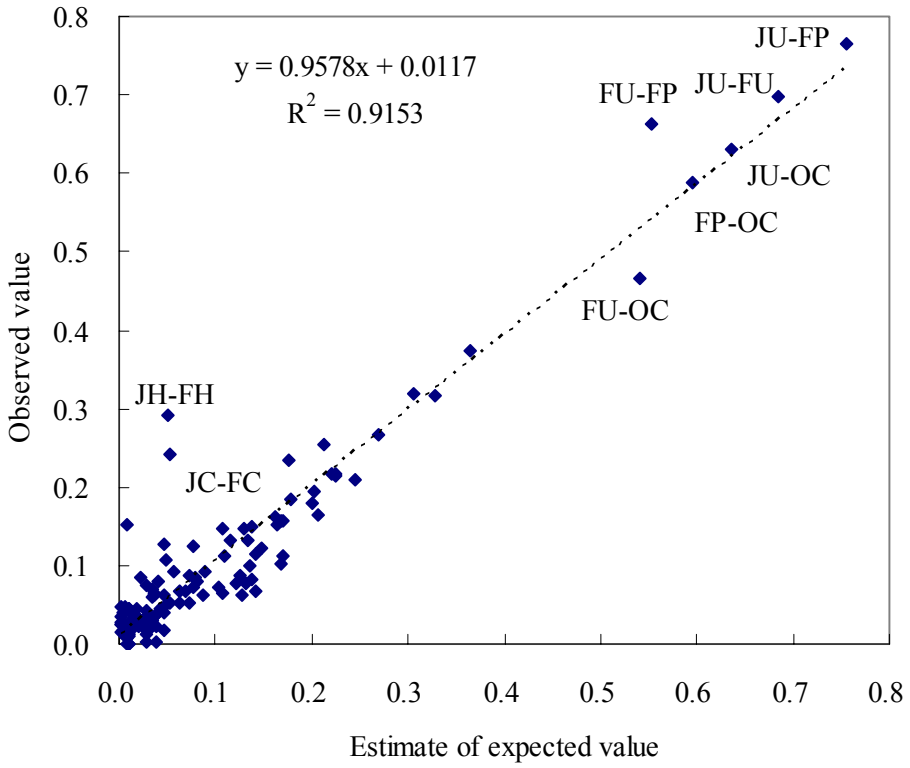

Figure 1： Estimate of expected and observed value of Salton-Ochiai Index

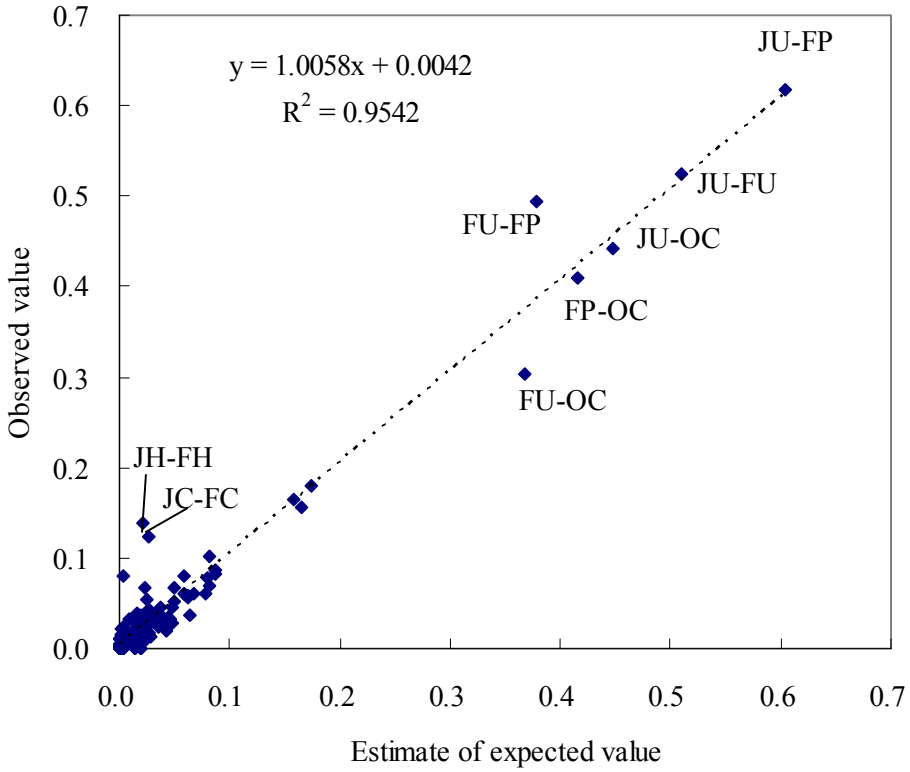

Figure 2： Estimate of expected and observed value of Jaccard Index

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.

3.1.2. 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 population. 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 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, because these sub-networks are analyzed as an integrated part of the entire set-up, they tend to be underestimated when there is a dense network in the whole connection. 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 contacts and analyzes networks individually, as it provides normalized forms to each one of them. 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 3. 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.

Table 3: 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.

The bias of 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 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.

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 France has 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.

The values scattered in the figure seem to be correlated, although coefficient of the determinant between them is only 0.34. 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 constrained 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.

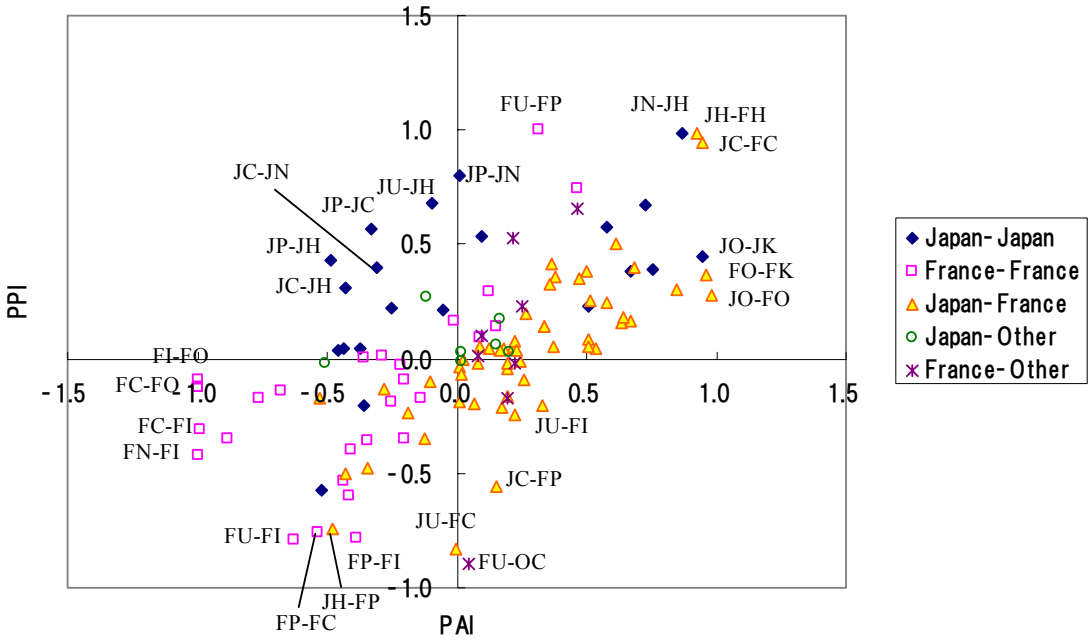

Figure 3: PAI and PPI indexes by type of linkage

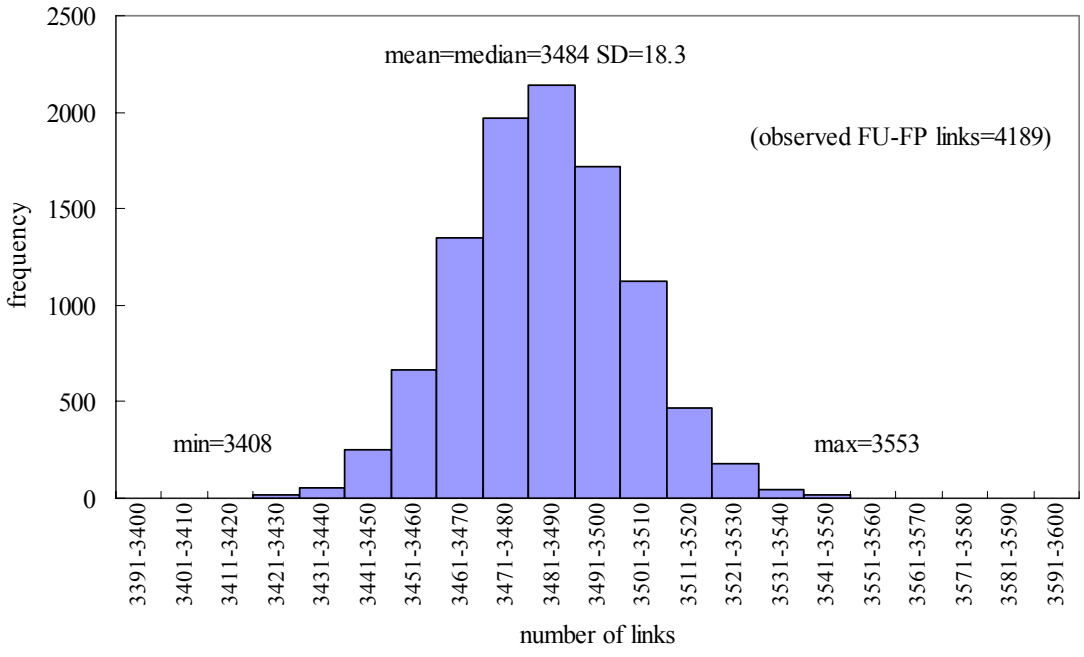

Figure 4: Frequency distribution of FU-FP links in the Monte-Carlo simulation

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 exists within a network of collaboration between countries and brings to light the way each participant (sectors) is situated in that network.

References

R Development Core Team (2004). R: A language and environment for statistical computing. Retrieved March 1, 2005 from: http://www.R-project.org, ISBN 3-900051-00-3.

Matsumoto, M. & Nishimura T. (1998). Mersenne Twister: A 623-dimesionally Equidistributed uniform Pseudo-Random number generator. *ACM Transactions on Modeling and Computer Simulation*, 8(1) 3-30.

Zitt, M., Bassecoulard, E. & Okubo, Y. (2000). Shadows of the Past in International Cooperation: Collaboration profiles of the top five producers of science. *Scientometrics* 47 (3) 627-657.