Regular Paper

Detecting Research Fronts Using Neural Network Model for Weighted Citation Network Analysis

Hisato Fujimagari[1,a)]   Katsuhide Fujita[1,b)]

Abstract: The performance of different types of weighted citation networks for detecting emerging research fronts was investigated by a comparative study in the existing work. The citation networks are constructed and then divided into clusters to detect the research front. Additionally, some measures to weighted citations like difference in publication years between citing and cited papers and similarities of keywords between them, which are expected to be able to effectively detect emerging research fronts, were applied. However, the functions of deciding the edge's weight in the citation networks are decided based on the experiments. For deciding the effective weight's functions automatically depending on the characteristics of the dataset, a learning method is important. In this paper, we propose the novel learning method based on the Neural Networks for deciding the edge's weights for the citation networks. We have been evaluating our proposed method in three research domains including Gallium nitride, Complex Networks, and Nano-carbon. We demonstrate that our proposed method has the best performance of each approach by using the following measures of extracted research fronts: visibility, speed, and topological and field relevance than the existing methods.

Keywords: citation network analysis, neural network, research fronts

1. Introduction

Over the past several decades, the number of academic papers has increased exponentially [1], and each academic area has become specialized and segmented. Davidson et al. [2] describe the situation as follows: "For most of history, mankind has suffered from a shortage of information. Now, in just the infancy of the electronic age, we have begun to suffer from information excess." Therefore, it is hard for researchers to perceive their specialized fields as a whole, and segmentation occurs simultaneously with specialization, which brings a severe problem and also opportunity to find crucial knowledge by integrating different domains. Naturally, it is hard for researchers and managers to detect a research front in the early stages by human effort only. There is a strong need for computational tools of science mapping and emerging topic detection. Previous studies have established effective algorithms for creating academic landscapes and for detecting emerging topics for certain research fronts.

To support the detection of research fronts and visualization of academic landscapes, methods of science mapping by citation analysis have been proposed and developed [3], [4], [5]. Researchers have also focused on clustering and visualization. For example, Leydesdorff [6] made a large-scale investigation of a set of academic papers. Not only creating static academic landscapes, topological and semantic analysis of a citation network also helps us to focus on significant movements in research fronts and emerging research fields in a broad context [7]. By analyzing the relationships between the academic research and the social problems, the important path for solving the social problems can be detected [8]. The other approach is to detect emerging clusters of densely connected papers. Price et al. [1] employed the concept of a research front, that is, a research domain under development where papers cite each other densely. Scientists tend to cite the most recently published articles in their papers; therefore, the network belonging in a research front becomes very tight. In a given field, a research front refers to the body of articles that scientists actively cite. Researchers have been studying quantitative methods that can be used to identify and track a research front as it evolves over time. Small and Griffith [9] showed that activated scientific specialists generate clusters of co-cited papers. Braam et al. [10] also investigated the topics discussed in co-cited clusters by analyzing the frequency of indexing terms and classification codes occurring in these publications.

Fujita et al. [11] conducted the comparative studies for showing the characteristics of paper-paper weighted citation networks created by different citation patterns with different weight types. In particular, average publication year, similarities of citation information and similarities of keywords are effective information attributes for detecting research fronts. By introducing them as weights of links to the citation network, it is expected to detect research fronts compared with the non-weighted citation networks effectively. However, the weighting functions of edges are decided based on the human's knowledge in this existing work. In other words, the functions of deciding the edge's weight in the citation networks are decided based on the experiments in the

---

1   Department of Computer and Information Sciences, Faculty of Engineering, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184–8588, Japan
a)   fujimagari@katfuji.lab.tuat.ac.jp
b)   katfuji@cc.tuat.ac.jp

existing work. For deciding the effective weight's functions automatically depending on the characteristics of the dataset, a learning method is important. In addition, the combinations between four kinds of the weights proposed in the existing works (Frequency of citations, Publication years, Reference Similarity and Research Field Similarity) are effective for detecting the research fronts. Therefore, the learning technique is necessary to find the optimal combinations between four kinds of the weights.

In this paper, we propose a novel learning method based on the Neural Networks for deciding the edge's weights for the citation networks. By introducing the Neural Networks model to the weighted citation networks, the optimal combinations between four kinds of the weights (Frequency of citations, Publication years, Reference Similarity and Research Field Similarity) can be achieved. We also evaluate the performance of each method in detecting a research front by comparing visibility, speed, and topological and research field relevance of clustering. We evaluate the proposed method with the object of the best method for detecting the research front is the one that can detect a large, textually and topologically uniform cluster of papers at an earlier stage compared with the method without the learning. By considering the differences, we discuss which type of weight and citation patterns is most suitable for detecting emerging knowledge domains from diverse facets of evaluation.

The remainder of this paper is organized as follows. First, we give an overview of research domains analyzed in our comparative case study. Next, we describe the methodology based on the network clustering and network measures. Then, we present and discuss the performance of the types of weighted citation network for detecting emerging research fronts. Finally, we present our overall conclusions.

2. Research Domains and Core Papers

This paper studies the following three research domains. Gallium nitride (GaN) is widely recognized as a recent prominent innovation in the fields of applied physics and material science. Complex network (CNW) analysis is also recognized as pioneering a new research field after the leading works by physicists has received attention. Nano-carbon (carbon nanotube [CNT]) is well known as a recent prominent innovation in the fields of applied physics and material science. They are typical examples of recent remarkable innovations having somewhat different characteristics (e.g., breakthrough of the rapid development, material or model-based innovation). These three domains are the same with those selected in our previous studies [11], [12].

Core Papers are research papers that receive citations soon after publication, relative to other papers of the same field and age. Generally, papers reach their citation peak two, three, or even four years after publication. However, core papers are recognized very soon after publication, reflected by rapid and significant numbers of citations [13]. These papers are often key researches in their fields. In this paper, core papers are defined as highly cited papers published in the rapid-growth years expected for the review papers using Web of Science, which is a Web-based user interface of the Institute for Scientific Information's (ISI) citation databases. Rapid-growth years in each domain are as follows: Gallium ni-

tride, 1991-1994; CNW, 1998-2001; CNT, 1990-1994. A list of core papers in each domain, which opened a new research frontier, is shown as follows: In GaN, we define the core papers as (A-1) NAKAMURA S, 1991, JPN J APPL PHYS PT 2, V30, P1705 (Nakamura, 1991) and (A-2) NAKAMURA S, 1992, JPN J APPL PHYS PT 1, V31, P1258 (Nakamura 1992). In CNW, we define the core paper as (B-1) Watts DJ, 1998, NATURE, V393, P440 (Watts and Strogatz, 1998) and (B-2) Barabasi AL, 1999, SCIENCE, V286, P509 (Barabasi and Albert, 1999). In CNT, we define the core paper as (C) IIJIMA, S, 1991, NATURE, V354, P56 (Iijima, 1991).

3. Basic Methodology

3.1 Data Collection

First, we collected citation data from the Science Citation Index (SCI) and the Social Sciences Citation Index (SSCI), which maintains citation databases covering thousands of academic journals and offers bibliographic database services, because SCI and SSCI are two of the best sources for citation data. We used the Web of Science, which is a Web-based user interface of the ISI's citation databases. We searched the papers using the following terms as queries: *"GaN OR gallium nitride"* for the first domain, *"social networks OR social network OR random networks OR random network OR small-world OR scale-free OR complex networks"* for the second domain, and *"carbon AND (nano* OR micro*)"* for the third domain.

In this paper, queries were selected according to the following two steps: (a) the representative keyword, such as gallium nitride and social network, is selected and (b) if the definition of its domain is unclear, more keywords, such as random network, small-world, scale-free, and complex networks, were added. The second step is called as query expansion [14]. Our intention in using so many terms is to retain wide coverage of citation data in order to avoid omission of core papers. For example, we selected the seven search queries in CNW by the query expansion. After selecting the seven queries, we evaluated that these queries retain wide coverage of citation data with avoiding omission of core papers and stopped expanding the queries to eight or more.

The queries for each dataset explained in the previous paragraph are the same as those in the previous paper [11], [12], but retrieved data is not exactly the same because of the data expansion of bibliographic records registered in ISI's databases.

3.2 Creating Weighted Citation Networks

After collecting the data including the published year, title, author(s), abstract, author keywords, and citation based on the queries, we create weighted citation networks. We create citation networks by regarding papers as nodes and citations between papers directly as edges.

We define the citation graphs $G = (N, E, w)$ comprising a set N of nodes, with each node N_i representing a paper p_i and a set E of edges, with each edge E_{ij} directed from the citing node N_i to the cited node N_j, or from the citing node N_j to the cited node N_i. E_{ij} means the number of citations between p_i and p_j. Usually, the number of direct citations is one; however, the number of co-citations and bibliographic-couplings is more than one.

In other words, we will build the citation networks defined as a weighted non-directed graph, with each paper representing a node and three patterns of citations representing the edges in the graph. Each node (N_i) has several attributes: paper title, author(s), year of publication (y_i) and journal name, reference information (R_i), and Research Field (F_i). Research Field is a set by the authors when they write the papers.

The network is created in each year, enabling a time-series analysis of citation networks. When we create citation networks on year y, we use the data of papers published from 1970 to y. In this paper, only the largest-graph component is used because this paper focuses on the relationship among papers, and we should therefore eliminate papers that have no link with the largest-graph component.

The definitions of these weights are as follows:

(i) Frequency of citations: $w(E_{ij}) = E_{ij}$

(ii) Publication years: $w(E_{ij}) = \{(y_i + y_j)/2 - 1970\}$ if $(y_i + y_j)/2 < 1970$, $w(Eij) = 0$

(iii) Reference Similarity:
$w(E_{ij}) = Jaccard(R_i, R_j) + 1$

(iv) Research Field Similarity:
$w(E_{ij}) = Jaccard(F_i, F_j) + 1$

*$Jaccard(x, y) = x \cap y/x \cup y$ (Jaccard similarity is defined by Jaccard [15]).

By introducing four types of weights based on the attributes, we can detect the research fronts reflecting the important attributes, such as new research fronts growing rapidly.

3.3 Topological Clustering

After that, we divided the papers in the network into clusters. For dividing into clusters, a fast-modularity clustering proposed by Newman [16] is applied in order to discover tightly knit clusters with a high density of within-cluster edges, which enables the creation of a weighted graph consisting of a large number of nodes. The algorithm is based on the modularity Q, which is defined as follows:

$$Q = \sum_{s}(w_{ss} - a_s^2) = Tr(w) - \|w\|^2 \tag{1}$$

where w_{st} is the possibility of the weights of edges in the network that connected nodes in cluster s to those in cluster t, and $s_s = \sum_t w_{st}$. In the first part of the equation, $Tr(w)$ represents the sum of density of weights of edges within each cluster. A high value of this parameter means that nodes are densely connected within each cluster. The second part of the equation, $\|w\|$, represents the sum of density of weights of edges within each cluster when all edges are placed randomly.

In Newman's method, edges that connect clusters sparsely and extract clusters within which nodes are connected densely are cut. A high value of Q represents good community division where only dense edges remain within clusters and sparse edges between clusters are cut off, and $Q = 0$ means that a particular division gives no more within-community edges than would be expected by random chance. Then, the algorithm to optimize Q over all possible divisions to find the best structure of clusters is as follows. Starting with a state in which each node is the only

member of one of the n clusters, we repeatedly join clusters together in pairs, choosing at each step the joining that results in the greatest increase in Q. The change in Q upon joining two clusters is given by

$$\Delta Q = w_{st} + w_{ts} - 2a_s a_t \tag{2}$$

In this paper, we stop joining when $\Delta Q < 0$.

3.4 Topological Measures for Evaluating Citation Networks

For comparing the tendency of four types of weighted citation networks, visibility, speed, and topological and field relevance are calculated after clustering for each cluster to which these selected core papers belong. In this paper, we assume that the important research front is detected as a larger and more relevant cluster at an earlier stage.

When the normalized size of the cluster is larger, we can more easily distinguish the existence of emerging clusters from other clusters. When the average publication year of the cluster is younger, it means that the cluster can be speedily detected at its emerging stage. If there is a time lag during detection, the lack of methodologies' speed prevents us from finding the research fronts in the emerging stage. In other words, the lack of speed of emerging detections could fail to grow the seeds of innovations in the industry. Therefore, we consider the speed as the one of the most important measure for evaluating the methodologies. If the cluster is denser, we can check whether clustering is successful for dividing into clusters in the citation networks. If the cluster is more textually relevant, we can detect the textually similar clusters.

The size of a cluster is defined as normalized size to the relative in order to compare certain types of weights:

$$|N_i \in C|/|N| \times 100 \tag{3}$$

where N is the total number of entire nodes N and $|N_i \in C|$ is the number of nodes in cluster C.

The density is defined as follows:

$$|E_i \in C|/_{|N_i \in C|}C_2 \tag{4}$$

where $|E_i \in C|$ is the number of edges, both of the nodes are in cluster C, and $_{|N|}C_2$ is the number of combinations from N to 2.

The research field similarity between clusters are defined as follows:

$$\sum_{i \in C} \sum_{j \in C(i \neq j)} |F_i \cap F_j|/|F_i \cup F_j| \tag{5}$$

where $|F_i \cap F_j|$ is the number of overlapped fields between F_i and F_j, and $|F_i \cup F_j|$ is the number of fields in F_i or F_j.

In addition, the average of the publication year of all papers is defined as follows:

$$\sum_{i \in C} y_i/|C| \tag{6}$$

4. Learning Method for Weighted Citation Network Analysis Using Neural Network Model

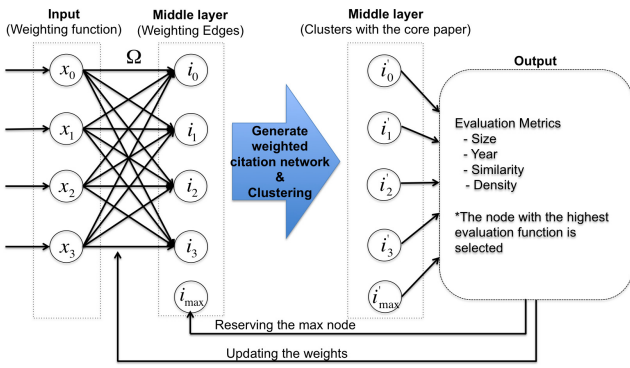The proposed approach is aimed at finding the most suitable
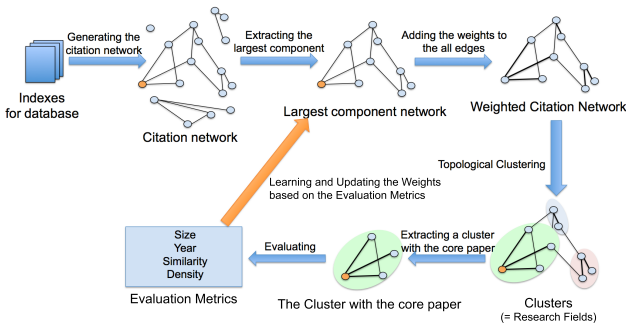
Fig. 1   Neural Network Model.

Fig. 2   Weighted Citation Network Analysis with Learning.

weighting function of the citation network's edges by combining a weighting function defined in the previous section. After that, we show the detailed descriptions of the proposed approach.

Neural Network Model

Figure 1 shows the Neural Network Model proposed in this paper. The Neural Network has the input layer, the middle layers and the output layer. The middle layer exists between the input layer and the output layer. Each middle layer has nodes with attributes and functions for linking the input and output layers. By changing the function for combining the values from the input layer, the values to the output layer become different. In other words, the function for combining the values from the input layer can be improved by repeating the procedures.

In our proposed model, we consider the four weighting functions for calculating the edges of citation networks written in the previous section as the nodes in the input layer. In the middle layer, the combination functions among the four weighting functions for calculating the edges of citation networks are regarded as the nodes.

$X = (x_0, \ldots, x_3)$ means the weighting functions of the edges of the citation network. Ω means the functions for combining every input values (X). $I = (i_0, \ldots, i_M)$ means the outputs by calculating the function Ω when the number of the nodes in the middle layer is M. We can get the $I' = (i'_0, \ldots, i'_M)$ by generating the weighted citation networks using I, and conducting the clustering. Finally, we get the size, year, similarity, and density written in the previous section, and select a node with highest evaluation value, which considers the size, year, similarity, and density.

Figure 2 shows the flow of the proposed method with learning. One of the big difference between the existing work [11] and our proposed method is the step of learning and updating the weights

based on the evaluation metrics. By repeating these procedures until finding the optimal combinations of the four weighting functions of deciding the weights of edges of the citation network, we can find the research fronts more effective. After that, we show the detail learning method using the Neural Network.

Weighting Functions between Initial Layer and Middle Layer

In the middle layer, the weights of the citations are calculated using the functions between the initial layer and the middle layer. In our method, the initial values of the functions for combining every input values are set as 0 or 1, randomly. The renewed function between the initial layer and the middle layer is defined as the Eq. (7).

$$i_n = \sum_{j \in I} x_j \omega_{jn} \tag{7}$$

i_n means the nth node in the middle layer. In the Eq. (7), i_n is calculated using the weighted summing. $i_0, \ldots, i_M$, which is the output from the functions between the input and middle layer have been normalized using the the highest and lowest limits in the renewing step.

Updating the functions in the middle layer

After defining the weighting functions in the middle layer, the cluster with the core paper is selected by generating the weighted citation networks and conducting the clustering. The clusters with the core paper generated by the functions are evaluated based on the size, year, density, and field similarity written in the previous section. The detailed procedures of updating the function for combining every input values are as follows:

We assume that $\vec{\omega}_0 = (\omega_{00}, \omega_{10}, \omega_{20}, \omega_{30}), \ldots, \vec{\omega}_3 = (\omega_{03}, \omega_{13}, \omega_{23}, \omega_{33})$ $(\vec{\omega}_0, \ldots, \vec{\omega}_3 \in \Omega)$. The updating vector for the function for combining every input values is $\vec{u}$.

(1) The highest nodes of the middle layer based on each topological measure are selected. The set of the four selected nodes are represented as $Select$. The elements of $Select$ are the nodes which are the highest values in the four topological measures (size, year, density, similarity). For example, when the i_0 is the highest in the size and the year measures, i_2 is the highest in the density measure, and i_3 is the highest in the similarity measure, $Select = (i_0, i_0, i_2, i_3)$.

(2) $\vec{u}$ for the function for combining every input values is defined as the average-sum of the functions for combining every input values which outputs are $Select$. For example, $\vec{u}$ is calculated as follows when $Select = (i_0, i_0, i_2, i_3)$: $\vec{u} = (\vec{\omega}_0 + \vec{\omega}_0 + \vec{\omega}_2 + \vec{\omega}_3)/4$.

(3) $\vec{\omega}_i^{(t)}$ is updated to $\vec{\omega}_i^{(t+1)}$ by the following function: $\vec{\omega}_k^{(t+1)} = (\vec{\omega}_k^{(t)} + \vec{u})/2$

These procedures are repeated until the conditions of stopping learning. After the one-time repeating, the best weighting functions of the edges of the citation networks are reserved for using the step 1. In other words, the highest nodes of the middle layer based on each topological measure are selected from I' as i_{max}.

Conditions of Ending the Learning Process

The conditions of ending the learning process are as follows:

(1) The functions in the middle layer aren't updated.

(2) All patters of the functions in the middle layer are finished.

Usually, the learnin g is stopped when the functions in the mid-

Table 1   Results of the topological measures and Q_{max} of each weighting function.

Research Field	Core Paper	Weighting function	Q_{max}	size	year	similarity	density
Gallium nitride	GaN-1	(i)	0.370599	9	1994.554	0.5625	0.0705
		(ii)	0.379776	32	1989.545	0.6354	0.0146
		(iii)	0.398590	28	1994.507	0.5926	0.0185
		(iv)	0.383212	26	1993.321	0.6064	0.0175
		(learning)	0.378652	30	1994.876	0.7590	0.0213
	GaN-2	(i)	0.370599	26	1994.570	0.5819	0.0172
		(ii)	0.379776	25	1994.496	0.6423	0.0178
		(iii)	0.398590	28	1994.507	0.5926	0.0185
		(iv)	0.383670	21	1994.919	0.6102	0.0236
		(learning)	0.379593	29	1994.901	0.7690	0.0210
	GaN-3	(i)	0.370599	31	1995.019	0.6269	0.0163
		(ii)	0.379776	25	1994.496	0.6423	0.0178
		(iii)	0.398590	21	1994.590	0.5879	0.0213
		(iv)	0.383670	21	1994.919	0.6102	0.0236
		(learning)	0.394264	31	1995.490	0.8910	0.0961
Complex Networks	CN-1	(i)	0.814976	58	2000.484	0.6779	0.0074
		(ii)	0.843881	59	2000.815	0.6640	0.0068
		(iii)	0.858989	49	2000.929	0.6510	0.0081
		(iv)	0.857758	47	2001.112	0.6405	0.0082
		(learning)	0.851434	59	2001.095	0.7130	0.0082
	CN-2	(i)	0.814976	58	2000.484	0.6779	0.0074
		(ii)	0.843881	59	2000.815	0.6640	0.0068
		(iii)	0.858989	49	2000.929	0.6510	0.0081
		(iv)	0.857758	47	2001.112	0.6405	0.0082
		(learning)	0.849351	52	2000.822	0.7020	0.0078
Carbon Nanotube	Carb-1	(i)	0.856571	10	1993.812	0.7675	0.0159
		(ii)	0.852119	10	1993.769	0.7706	0.0158
		(iii)	0.854688	9	1993.885	0.7695	0.0168
		(iv)	0.867004	9	1993.802	0.7800	0.0166
		(learning)	0.856809	10	1993.882	0.6798	0.0166

dle layer aren't updated because the updating functions lead to the maximum point of each topological measure.

5.   Experimental Results

In this experiment, we applied our proposed method to three research fields (Gallium nitride, Complex network, and Carbon Nanotube) to evaluate. We compared the Q_{max} and four topological measures (cluster size, average publication year, field similarity, density) among five types of the weighted citation networks. The number of the nodes in the middle layer is four in this experiment. As discussed in Sections 2 and 3, we have prepared the datasets for those three domains and applied our method to them.

Table 1 shows the results of each weighting function and the proposed approach. In Table 1, "(i)" means that the weighted function of the citation network is decide by "(i) Frequency of citations." "(ii)" means that the weighted function of the citation network is decide by "(ii) Publication years." "(iii)" means that the weighted function of the citation network is decide by "(iii) Reference Similarity." "(iv)" means that the weighted function of the citation network is decide by "(iv) Research Field Similarity." "(learning)" means our proposed method.

By comparing the proposed learning method with the existing weighted functions in each research field, our proposed method is higher in the similarity, year, and density than the other weighted functions in all core papers of the Gallium nitride field. In addition, our proposed method is higher in the size than the other weighted functions in the Gallium nitride field expected for the GaN-1. In the complex network field, the results of the comparisons between the learning method and others are almost same. Especially, our learning method can detect the research fronts

clearly compared with other methods because the similarity and the density are high in Gallium nitride and Carbon Nanotube.

On the other hand, the scores of our proposed learning method aren't the highest values than other methods in the Carbon Nanotube field. Our proposed method considers many kinds of the attributes of paper, therefore, it tries to improve all measures when there are trade-offs between the measures. However, this situation is dependent on the characteristics of the dataset. Therefore, the proposed learning method can improve for finding the research fronts in several kinds of the datasets compared with the methods without the learning.

6.   Conclusion

This paper proposed the novel Neural Network Model for defining the edge's weights for the citation networks. By introducing the Neural Network Model to the weighted citation networks, the optimal combinations between four kinds of the weights (Frequency of citations, Publication years, Reference Similarity and Research Field Similarity) can be achieved. We also evaluated the performance of each method in detecting a research front by comparing visibility, speed, and topological and research field relevance of clustering. Our proposed method could detect the research fronts is the one that can detect a large, textually and topologically uniform cluster of papers at an earlier stage compared with the method without the learning in the experiments.

Future works will address improvements of updating functions in the neural network model. Our learning method sometimes unspreads quickly despite that it isn't optimal. The method for solving this problem is to employ the random nature in updating

the functions in the middle layer. Another important future work is to address the temporal aspect of the analysis of a citation network such as the evolution and emergence of new clusters and their underlying knowledge domains using our proposed method.

References

[1] Price, D.: Networks of scientific papers, *Science*, Vol.149, pp.510–515 (1965).

[2] Davidson, G.S., Hendrickson, B., Johnson, D.K., Meyers, C.E. and Wylie, B.N.: Knowledge mining with VxInsight: Discovery through interaction, *Journal of Intelligent Information Systems*, Vol.11, pp.259–285 (1998).

[3] Klavans, R. and Boyack, K.W.: Identifying a better measure of relatedness for mapping science, *Journal of the American Society for Information Science and Technology*, Vol.57, No.2, pp.251–263 (2006).

[4] Klavans, R. and Boyack, K.W.: Toward a consensus map of science, *Journal of the American Society for Information Science and Technology*, Vol.60, No.3, pp.455–476 (2009).

[5] Boyack, K.W., Klavans, A.R. and Borner, K.: Mapping the backbone of science, *Scientometrics*, Vol.64, pp.351–374 (2005).

[6] Leydesdorff, L. and Rafols, I.: A global map of science based on the ISI subject categories, *Journal of the American Society for Information Science and Technology*, Vol.60, No.2, pp.348–362 (2009).

[7] Shibata, N., Kajikawa, Y., Takeda, Y. and Matsushima, K.: Detecting emerging research fronts based on topological measures in citation networks of scientific publications, *Technovation*, Vol.28, No.11, pp.758–775 (2009).

[8] Ittipanuvat, V., Fujita, K., Sakata, I. and Kajikawa, Y.: Finding linkage between technology and social issue: A Literature Based Discovery approach, *Journal of Engineering and Technology Management*, Vol.32, pp.160–184 (2013).

[9] Small, H. and Griffith, B.: The structure of scientific literatures: I. identifying and graphing specialties, *Science Studies*, Vol.4, pp.17–40 (1974).

[10] Braam, R.R., Moed, H.F. and van Raan, A.F.J.: Mapping of science by combined co-citation and word analysis. I. Structural aspects, *Journal of the American Society for Information Science*, Vol.42, No.4, pp.233–251 (1991).

[11] Fujita, K., Kajikawa, Y., Mori, J. and Sakata, I.: Detecting research fronts using different types of weighted citation networks, *Journal of Engineering and Technology Management*, Vol.32, pp.129–146 (2013).

[12] Shibata, N., Kajikawa, Y., Takeda, Y. and Matsushima, K.: Comparative Study on Methods of Detecting Research Fronts Using Different Types of Citation, *Journal of the American Society for Information Science and Technology*, Vol.60, No.3, pp.571–580 (2009).

[13] Nederhof, A.J.: Bibliometric monitoring of research performance in the Social Sciences and the Humanities: A Review, *Scientometrics*, Vol.66, No.1, pp.81–100 (2006).

[14] Kostoff, R., Eberhart, H. and Toothman, D.: Database tomography for information retrieval, *Journal of Information Science*, Vol.23, pp.301–311 (1997).

[15] Jaccard, P.: The Distribution of the Flora in the Alpine Zone, *New Phytologist*, Vol.11, No.2, pp.37–50 (1912).

[16] Newman, M.E.J. and Girvan, M.: Finding and evaluating community structure in networks, *Phys. Rev. E*, Vol.69, No.2, p.026113 (2004).

Katsuhide Fujita was born in 1985. He received his B.E. and M.E. and Doctor of Engineering from Nagoya Institute of Technology in 2008, 2010 and 2011, respectively. From 2010.4 to 2011.9, he was a research fellow of the Japan Society for the Promotion of Science (JSPS). From 2010.6 to 2011.5, he was a visiting student at Sloan School of Management, Massachusetts Institute of Technology. From 2011.10 to 2012.11, he was a project research at School of Engineering, The University of Tokyo. From 2012.12, he is an associate professor at Faculty of Engineering, Tokyo University of Agriculture and Technology. His main research interests include multi-issue negotiation, multi-agent systems, intelligent agents. He is a Member of IPSJ, IEICE, JSAI, IEEE-CS, and AAAI.

Hisato Fujimagari was born in 1992. He received his B.E. from Tokyo University of Agriculture and Technology in 2013 and has been engaged in the Information Processing Society of Japan since 2013. From 2014, he is a master student in Tokyo University of Agriculture and Technology. His research interest is citation network analysis and information retrieval. He is a student member of IPSJ.