PAPER

Predicting Research Trends Identified by Research Histories via Breakthrough Researches

Nagayoshi YAMASHITA[†a)], *Nonmember*, Masayuki NUMAO[††], *and* Ryutaro ICHISE[†††], *Members*

SUMMARY Since it is difficult to understand or predict research trends, we proposed methodologies for understanding and predicting research trends in the sciences, focusing on the structures of grants in the Japan Society for the Promotion of Science (JSPS), a Japanese funding agency. Grant applications are suitable for predicting research trends because these are research plans for the future, different from papers, which report research outcomes in the past. We investigated research trends in science focusing on research histories identified in grant application data of JSPS. Then we proposed a model for predicting research trends, assuming that breakthrough research encourages researchers to change from their current research field to an entirely new research field. Using breakthrough research, we aim to obtain higher precision in the prediction results. In our experimental results, we found that research fields in Informatics correlate well with actual scientific research trends. We also demonstrated that our prediction models are effective in actively interacting research areas, which include Informatics and Social Sciences.
key words: *scientometrics, data mining, grant application analysis*

1. Introduction

Understanding trends in scientific research is important in research fields ranging from classical fields, such as Mathematics and Philosophy, to interdisciplinary fields, such as Environmental Science and Genome Science. Although new discoveries in science often occur through combinations of existing disparate research, it is difficult to understand or predict research trends, because such research fields are subdivided and highly focused. We therefore propose methodologies for understanding and predicting research trends in the sciences, focusing on the structures of grants in the Japan Society for the Promotion of Science (JSPS), a Japanese funding agency. Grant applications are suitable for predicting research trends because these are research plans for the future, different from papers, which report research outcomes in the past.

First, to understand trends in scientific research, we focused on research histories. We assumed that if many researchers change their research themes within the same period, the scientific research trends change. We extracted research histories from "research fields" in grant application data of JSPS. A "research field" is selected according to the content of the research project, as identified by a predefined "List of Research Fields," a classification table showing research areas selected for screening.

Second, we proposed a model for predicting research trends based on breakthrough research. Breakthrough research radically changes the understanding and approach to an existing scientific concept and often leads to paradigm shifts or the creation of new paradigms or fields of science. We assumed that breakthrough research encourages researchers to change from their current research field to either a new approach or an entirely new research field. If many researchers in the same research field apply their research to a new research field as breakthrough research, then scientific research trends change from past to new research fields. We extracted every pair of research fields in the work authored by the same researchers in the "Scientific Research," "Grants-in-Aid for Young Scientists," and "Challenging Exploratory Research" categories, which are the three main categories in the JSPS. The "Scientific Research" and "Grant-in-Aid for Young Scientists" categories represent individual research, whereas the "Challenging Exploratory Research" is a category for breakthrough research. These pairs of research fields provide an indicator of a shift to breakthrough research. Using breakthrough research, we aim to obtain higher precision in the prediction results.

In this paper, first, we briefly explain related works and grants in the JSPS. Next, we describe and evaluate a model for understanding scientific research trends using research histories. We then introduce and test a model for predicting the research trends by focusing on breakthrough research. Finally, we discuss our results and conclude our paper with a brief summary and directions for future work.

2. Related Work

Several papers have presented methods for investigating research trends in science, including co-authorship [1], [2], co-citation [3]–[8], and text information [9]–[12]. These approaches are summarized below.

Co-authorships are used to investigate collaborative research trends. Pham investigates the collaborative and citation behavior of journals and conferences by analyzing the properties of their co-authorship and citation subgraphs [2].

Co-citation is a popular similarity measure used to establish a subject similarity measure between two papers. In general, if papers A and B are both cited by paper C, then

Manuscript received December 9, 2013.
Manuscript revised August 25, 2014.
[†]The author is with the Japan Society for the Promotion of Science, Tokyo, 102–0083 Japan.
[††]The author is with the Institute of Scientific and Industrial Research, Osaka University, Ibaraki-shi, 567–0047 Japan.
[†††]The author is with the Principles of Informatics Research Division, National Institute of Informatics, Tokyo, 101–8430 Japan.
a) E-mail: nagayoshi3@gmail.com
DOI: 10.1587/transinf.2013EDP7435

A and B are related to one another even though they do not directly reference each other. Typically, the number of papers citing both papers A and B affects the strength of the similarity measure between papers A and B. Co-citation can be used for visualizing author relationships [3]. The change of co-citation relationships can be used to track research trends [4]. Jiam uses a matrix of co-citations as an information source, and then transforms this matrix into an FP-tree for visualization [5]. Abercrombie examines a scientometric model that tracks the emergence of an identified technology from initial discovery (via original scientific and conference literature) to critical discoveries (via original scientific, conference literature, and patents) [6]. The co-citation relationships in each journal have been used for visualizing structures in chemistry [7]. Finally, a "Science Map" is proposed for mapping sets of papers into two dimensions on the basis of co-citations [8].

Text information is used to visualize relationships between research fields. Okuoka uses word-count co-occurrences to visualize the relationships between research fields [9]. Woo provides indicators and visualization methods for measuring the latest research trends and aspects underlying scientific and technological documents for researchers and policy planners using "co-word analysis" [10]. Co-word analysis reveals patterns and trends by measuring the association strength of term representatives of relevant publications or other texts. Text information is also used for visualizing research fields in Information Science via co-citations. This method of incorporation is called the hybrid method. Frizo applies a hybrid clustering method based on Fisher's inverse chi-square approach to integrate full-text with citations and provide a mapping for the field of Information Science [11]. Chaomei enhances contemporary co-citation network analysis by enabling analysts to identify co-citation clusters of cited references intuitively, synthesize thematic contexts in which these clusters are cited, and trace how research focus evolves over time [12].

3. Research on Grant Application Data

In our research, we investigate scientific research trends by using grant application data. This is beneficial for several reasons. First, the format of such data is well-organized. Second, there are numerous grant applications in the JSPS ranging from the Humanities and Social Sciences to Medicine. Third, each application has a set of predefined attributes, such as various research fields, keywords, and the research institute of each researcher. Therefore, to evaluate relationships between research fields, Satoh uses the relationships of collaborative research between principal investigators and co-researchers [13]; similarly, Herr II uses titles and abstracts of grant applications [14].

Since papers have various formats and do not always have organized categories of research themes, collecting papers from and summarizing research themes proves to be difficult. To solve these problems effectively, the selection of papers is mainly based on papers with specific research theme or highly cited papers. On the contrary, the format of numerous grant application data and categories of research themes in the JSPS are well organized. Therefore, we can identify the relationships between interdisciplinary research themes, e.g., between Informatics and research themes such as Breeding Science, Brain Science, and Psychology.

We investigate research histories by using research fields of grant application data, assuming that these histories represent scientific research trends [15]. While previous studies only use research trends from past years for predicting scientific research trends, our proposed method uses a combination of trends identified from past years and breakthrough research. Using breakthrough research, we aim to obtain higher precision in the prediction results.

4. Grants-in-Aid for Scientific Research, JSPS

The JSPS provides Grants-in-Aid for Scientific Research for publicly recruited and screened applications. These are the most popular grants in Japan and are competitive funds intended to significantly develop all scientific research (research based on the free ideas of the researcher), covering basic to applied research in all fields, ranging from Humanities and Social Sciences to the Natural Sciences [16].

Grant applications include attributes such as research fields, keywords, research institutes of applicants, co-researchers, and categories of grants. On the basis of the content of the research project, authors select one research field from the "List of Categories, Areas, Disciplines and Research Fields for Grants-in-Aid for Scientific Research," a classification table showing the predefined areas for screening. In this table, categories are subdivided into areas; areas are subdivided into disciplines, which are further subdivided into research fields. Finally, research fields are subdivided into keywords. A keyword is selected from the "Table of Keywords" depending on the most closely related keyword that describes the content of the applicant's research project [16].

The main categories of Grants-in-Aid for Scientific Research are "Scientific Research (S)(A)(B)(C)," "Challenging Exploratory Research," and "Grant-in-Aid for Young Scientists (A)(B)." Scientific Research is a grant category for creative and pioneering researches by one researcher or multiple researchers. Challenging Exploratory Research is a grant category for early-stage researches based on a unique, challenging concept with a high set goal. Grant-in-Aid for Young Scientists is a category for researches by one researcher below 40 years. Note that Challenging Exploratory Research is allowed to be funded in conjunction with the other two categories [16].

5. A Model for Understanding Research Trends Using Research Histories

In this section, we propose a model for understanding scientific research trends using research field histories. We define an accepted migration rate for extracting research field histories for each researcher. We also define research field networks that represent the relationships of research fields using these accepted migration rates. We then evaluate whether the research field networks in Informatics represent actual scientific research trends.

5.1 Extracting Research Histories and Building Research Field Networks

Accepted migration rate $M^{a}_{xy}(n)$ from research field x to research field y in year n is the ratio of researchers who applied for grants in research field y in year $n+1$ to those who completed their projects in research fields x at the end of year n. This is defined as follows:

$$M^{a}_{xy}(n) = \frac{|F^{a}_{x}(n) \cap N_{y}(n+1)|}{|F^{a}_{x}(n)|} \times 100, \qquad (1)$$

where $F^{a}_{x}(n)$ is the set of researchers who completed their projects supported by Grants-in-Aid for Scientific Research in research field x at the end of year n, and $N_{y}(n+1)$ is the set of researchers who applied for new grants in research field y in year $n+1$. These are extracted from "Scientific Research" and "Grant-in-Aid for Young Scientists." As an example, the accepted migration rate from "Media Informatics/Database" to "Intelligent Informatics" in 2010 is the ratio of researchers who applied for new grants in "Intelligent Informatics" in 2011 to those who completed their projects in "Media Informatics/Database" at the end of 2010.

5.2 Construction of Research Field Networks

The research field networks in year n are constructed by creating a directed graph in which nodes represent research fields and links represent pairs of nodes with an accepted migration rate exceeding α in year n.

Research field networks of each α in 2005 are constructed, and then the average path lengths and clustering coefficients of these networks are calculated as shown in Fig. 1. Note that isolated nodes are excluded in calculating the average path length. It is desirable to have high values of threshold α, average path lengths, and clustering coefficients because threshold α is selected such that many nodes are connected by fewer links. A network with fewer links is easy to understand. As shown in the figure, the average path length is the highest when α is 4; it rapidly decreases when α is more than 5. Also, the clustering coefficient decreases as α increases. Thus, we set α to 5.

We created a research field network for Informatics by including any research fields within Informatics that have links between pairs of research fields with at least one node within Informatics. The resulting network has 91 research fields, including 17 research fields in Informatics. The research field network for 2009 is illustrated in Fig. 2. In the figure, node colors represent areas and node sizes represent the number of proposals in each research field. Each link has a weight that represents the accepted migration rate. These

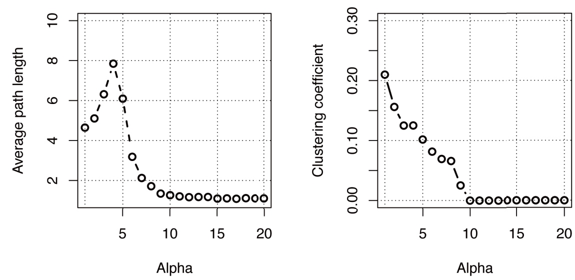

Fig. 1 Average path length and clustering coefficient in each α.

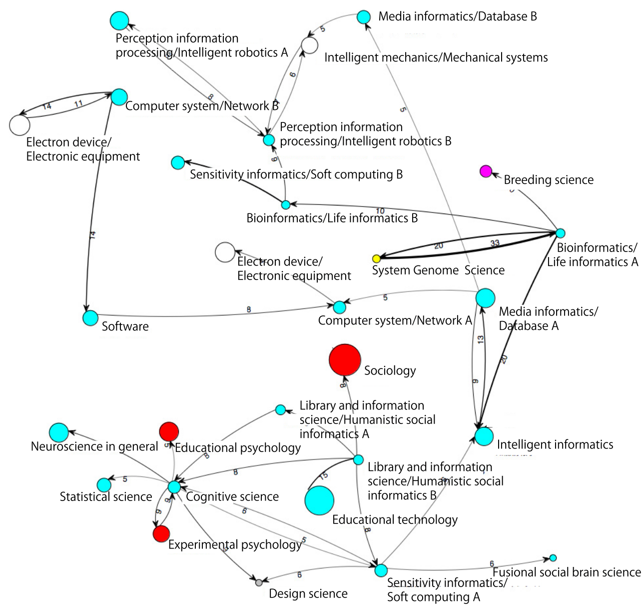

Fig. 2 Research field network for Informatics (2009).

networks are constructed for each year in the range 2003–2010. As necessary, we investigate the relationships of research fields in detail by keywords. Keywords are predefined in the "List of Research fields" and define the scope of each research field. Keywords in different research fields are not regarded as the same even if the same keywords are in different research fields.

5.3 Evaluation of Research Field Networks in Informatics

We evaluated whether the research field networks in Informatics represent actual scientific research trends. Five research fields—"Media Informatics/Database B," "Perception Information Processing/Intelligent Robotics B," "Bioinformatics/Life Informatics A," "Bioinformatics/Life Informatics B," and "Cognitive Science"—are selected to be compared with actual research trends. To determine the actual research trends, we studied web pages of academic societies and "Google trends" [17]. Google trends shows how often a particular word is retrieved relative to the total search-volume. The horizontal axis represents time and the vertical is how often a word is retrieved for relative to the total number of searches. Research trends were shown quantitatively by retrieving each Japanese research field name in Google trends.

Human interface research in **"Media informatics / Database B"** has been applied to various fields. In the research field networks, "Media Informatics/Database B" has links with "Perception Information Processing/Intelligent Robotics A" (2004–2008). "Media informatics/Database B" has links with "Cognitive Science" (2003, 2004), "Media Informatics/Database A" (since 2008), "Sensitivity Informatics/Soft Computing A" (since 2007), and "Rehabilitation Science/Welfare Engineering B" (2008).

Analyzing keywords used by researchers, 16 researchers who chose "Human Interface" also selected other research fields, such as "Image Processing," "Speech Processing," "Pattern Recognition" in "Perception Information Processing/Intelligent Robotics A," and "Sensitivity Interface in "Sensitivity Informatics/Soft Computing A." The website of "Special Interest Groups-Human Computer Interaction Academic Society" also stated that research goals in human interface have been changing from interfaces between humans and technology to interactions between humans and technology [18].

"Robot Research" in "**Perception Information Processing / Intelligent Robotics B**" has been applied to various research fields, especially sensor research. In our constructed research field networks, "Perception Information Processing/Intelligent Robotics B" has a link with "Bioinformatics/Life Informatics B" (since 2006). "Perception Information Processing/Intelligent Robotics B" has links with "Rehabilitation Science/Welfare Engineering B" (since 2006), and "Biomedical Engineering/Biological Material Science A" (2007). Eight researchers changed research fields between "Perception Information Processing/Intelligent Robotics B" and "Perception Information Processing/Intelligent Robotics A." These researchers selected keywords related to sensors, including "Sensor Fusion" and "Sensing Device Systems." The website of "Network Robot Technical Group in the Institute of Electronics" also states that as one of the three network robot technologies, unconscious robots coordinated with sensors embedded in an environment, wearable computers and actuators are introduced [19]. Google trends also showed that "Robot + Sensor" starts to be retrieved since 2012 and "Robot+Welfare" starts to be retrieved since 2010 in Fig. 3.

"**Bioinformatics/Life Informatics**" has been related to other research fields in Informatics. In "**Bioinformatics/Life Informatics A**," three out of 13 researchers who completed their projects changed research fields from "Bioinformatics" to "Knowledge Discovery and Data Mining" and "Knowledge Bases and Knowledge Systems" in "Intelligent Informatics". 12 researchers are applied for grants in these two research fields successively (since 2007). In "**Bioinformatics/Life Informatics B**," six researchers changed research fields from "Neural Information Processing" to "Intelligent Robot" in "Perception Information Processing/Intelligent Robotics B" or from "Bioinformatics" to "Neural Networks" in "Sensitivity Informatics/Soft Computing B." According to the website of "Special Interest Groups, Bioinformatics and Genomics in Information Processing Society of Japan," by the end of 20th century, the development of science and technology lies in Information Science and Bioscience. There are innumerable computers connected to networks and there has been progress in genome sequencing technologies. In the 21st century, these two research fields would combine to solve the problems of life and contribute to the progress of mankind [20].

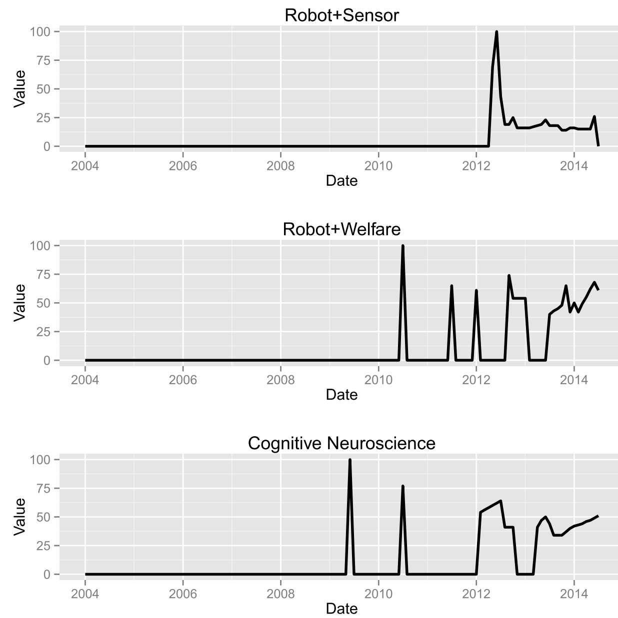

Fig. 3 Research field names in Google Trends.

"**Cognitive science**" has been related with neuroscience. In our research field networks, two out of 24 researchers who completed their projects changed research fields from "Comparative Cognitive Psychology" in Cognitive science to "Neural Information Processing" or "Cognitive Neuroscience" in neuroscience. 18 researchers are applied for grants in these two research fields successively (since 2007). The website of the Japanese Society for Cognitive Psychology also states that cognitive science has been developing to peripheral domains, and that cognitive neuropsychology has emerged in neuropsychology (2002) [21]. Google trends also showed that "Cognitive Neuroscience" starts to be retrieved since 2009 and has been retrieved regularly since 2012 in Fig. 3.

These experimental and comparative results show that research field networks do correlate to actual scientific research trends when compared with websites of academic societies.

6. A Model for Predicting Research Trends Based on Breakthrough Research

In addition to the methods described above, we propose a method for predicting scientific research trends by focusing on breakthrough research. In this section, we define duplicate application for predicting such trends. This is calculated by identifying each pair of research fields proposed by the same researchers in the "Scientific Research," "Grants-in-Aid for Young Scientists," and "Challenging Exploratory Research" categories. We explain how to calculate duplicate application and migration number, where the latter represents histories of research fields applied by researchers. We use this attribute in our prediction models.

6.1 Extracting Duplicate Application and Migration Number

Duplicate application $D_{xy}(n)$ from research field x to research field y in year n is the number of researchers who applied for research field y in "Challenging Exploratory Research" and those who applied for research fields x in "Scientific Research" or "Grant-in-Aid for Young Scientists" in year n. Accepted duplicate application $D^a_{xy}(n)$ from research field x to research field y in year n is the number of researchers who applied for research field y in "Challenging Exploratory Research" and had continuing grant of research field x in "Scientific Research" or "Grant-in-Aid for Young Scientists" in year n. These are defined by the following equations:

$$D_{xy}(n) = |S_x(n) \cap C_y(n)|, \tag{2}$$

$$D^a_{xy}(n) = |S^a_x(n) \cap C_y(n)|. \tag{3}$$

In these equations, $S_x(n)$ is the set of researchers who applied for research field x or had continuing grants in year n; $S^a_x(n)$ is the set of researchers whose projects in research field x were supported by "Scientific Research" or "Grant-in-Aid for Young Scientists" in year n; and $C_y(n)$ is the set of researchers who applied for research field y in "Challenging Exploratory Research" in year n. As an example, a duplicate application from "Media Informatics/Database" to "Intelligent Informatics" in 2010 is the number of researchers who applied for "Intelligent Informatics" in "Challenging Exploratory Research" and those who applied for "Media Informatics/Database" in "Scientific Research" or "Grant-in-Aid for Young Scientists" in 2010. Cumulative duplicate in year n is the sum of duplicate applications from year $n-2$ to n. Accepted cumulative duplicate in year n is the sum of accepted duplicate applications from year $n-2$ to n.

Next, we define migration number and accepted migration number. These values are extracted from the "Scientific Research" or "Grant-in-Aid for Young Scientists" categories. Migration number $N_{xy}(n)$ from research field x to y in year n is the number of researchers who applied for research field y in year $n+1$ and research field x in year n. Accepted migration number $N^a_{xy}(n)$ from research field x to y in year n is the number of researchers who applied for research field y in year $n+1$ and completed their projects in research field x at the end of year n. These are defined by the following equations:

$$N_{xy}(n) = |S_x(n) \cap N_y(n+1)|, \tag{4}$$

$$N^a_{xy}(n) = |F^a_x(n) \cap N_y(n+1)|. \tag{5}$$

For example, the accepted migration number from "Media Informatics/Database" to "Intelligent Informatics" in 2010 is the number of researchers who applied for "Intelligent Informatics" in 2011 and completed their projects in "Media Informatics/Database" at the end of 2010. Accepted migration numbers are updated less often than migration numbers because accepted migration numbers are updated after projects are completed.

Cumulative migration in year n is the sum of migration numbers from year $n-2$ to n. Accepted cumulative migration in year n is the sum of accepted migration number from year $n-2$ to n.

7. Experiments

In this section, we describe how to build prediction models and predict scientific research trends in areas that are actively interacting with each another. In our prediction models, the following four attributes are used: cumulative duplicate, cumulative migration, accepted cumulative duplicate, and accepted cumulative migration. Finally, we evaluate our prediction models and prediction results.

7.1 Constructing Prediction Models

A prediction model is constructed by applying logistic regression to a training set. Logistic regression is used for predicting a binary response in which the response variable takes only the values zero and one. Since our subject in this paper is predicting whether researchers apply their research outcomes to the other research fields or not, logistic regression was chosen.

The prediction model is represented by a regression line, and the training sets consist of sets of the four attributes obtained from the 2005 grant application data. These data are labeled based on grant application data from 2006 to 2007. A prediction result is obtained by applying the prediction model to a test set, which consists of sets of the four attributes obtained from the 2007 grant application data. Again, these data are labeled based on grant application data from 2008 to 2009. Positive instances of these labels are pairs of research fields that have more than two accepted migration numbers for the subsequent two years. Negative instances of these labels are the other pairs of research fields. Accepted migration numbers are used as labels since experimental results in the previous section showed that research field networks correlate to actual scientific research. If a response variable in a pair of research fields is over 0.5, our proposed method predicts that researchers would apply their research outcomes to these research fields in the future.

7.2 Finding Actively Interacting Areas

Clustering coefficients of research field networks in each

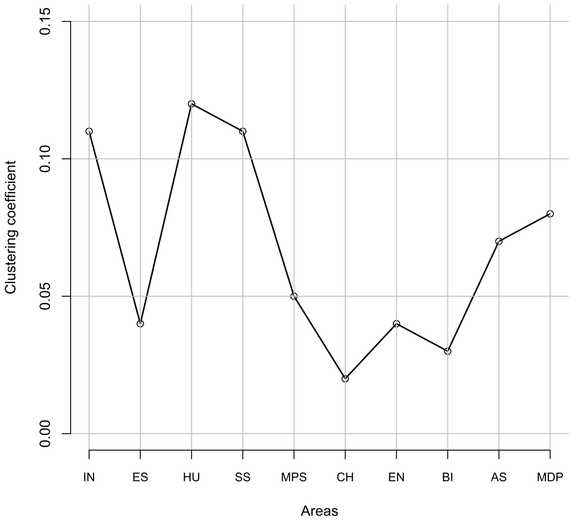

Fig. 4 Clustering coefficients of research field networks.

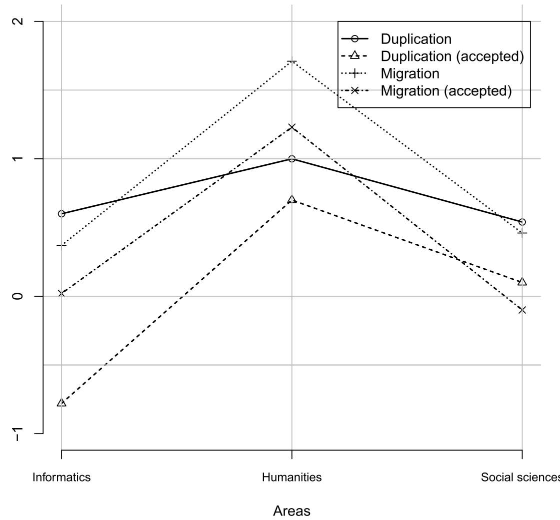

Fig. 5 Regression coefficients of prediction models.

Table 1 Number of research fields in each area.

Areas	Informatics	Humanities	Social sciences
size	228	104	156

area are calculated. Clustering coefficients are used for selecting research areas in which a lot of researchers are applying research outcomes to the other research fields frequently because we focused on researchers who applied their research outcomes to another research fields to predict research trends.

Research field networks in an area include research fields in the area and those with more than 1.0 accepted migration rate to those in the area. All pairs of research fields are included; however, both research fields are included in the other areas. Areas refer to Humanities (HU), Social Sciences (SS), Mathematical and Physical Sciences (MPS), Chemistry (CH), Engineering (EN), Biology (BI), Agricultural Sciences (AS), Medicine/Dentistry and Pharmacy (MDP), Informatics (IN), and Environmental Sciences (ES). Because clustering coefficients in Informatics, Humanities, and Social Sciences are more than 0.1, research field networks in these areas are actively interacting, as shown in Fig. 4. We therefore apply our proposed method to Informatics, Humanities, and Social Sciences. The number of research fields in these areas is shown in Table 1.

7.3 Evaluation of Prediction Models

The prediction models are evaluated in these areas. Figure 5 shows the regression coefficients in these prediction models. In the figure, duplicate refers to cumulative duplicates and migration refers to cumulative migration. The horizontal axis represents areas, while the vertical axis represents the values of regression coefficients.

Cumulative duplicates in Informatics and Social Sciences have the largest influence on prediction models. Conversely, in Humanities, cumulative migration has the largest influence.

7.4 Evaluation of Prediction Results

We evaluate the prediction results by precision, recall, and F-measure. These are calculated by comparing positive and negative instances between prediction results and actual results from 2008 to 2009.

Precision is the ratio of positive instances in actual results to the pairs of research fields predicted as positive instances. Recall is the ratio of pairs of research fields predicted as positive instances to all positive instances in the actual results. F-measure is the harmonic mean of precision and recall. Figure 7 illustrates the prediction results in these areas. The horizontal axis represents the areas, while the vertical axis represents the corresponding values of precision, recall, and F-measure.

F-measures in Informatics and Social Sciences are high, while that in Humanities is much lower. Thus, our proposed method, which focuses on breakthrough research, is effective in Informatics and Social Sciences.

8. Discussion

From the results, in Humanities cumulative duplicates have the smallest influence on prediction models and the F-measure is not high. We investigated the reason by using the ratio of researchers who applied for "Challenging Exploratory Research" to those who have continuing grants in "Scientific Research" or "Grant-in-Aid for Young Scientists." If the ratio is small, Cumulative duplicate is not sufficient to predict in the areas. In Fig. 6, "the same research fields" represents the ratio of researchers who applied for the same research fields in "Scientific Research," "Grant-in-Aid for Young Scientists," and "Challenging Exploratory Research." Also in the figure, "different research fields" represents the ratio of researchers who applied for different research fields in these categories.

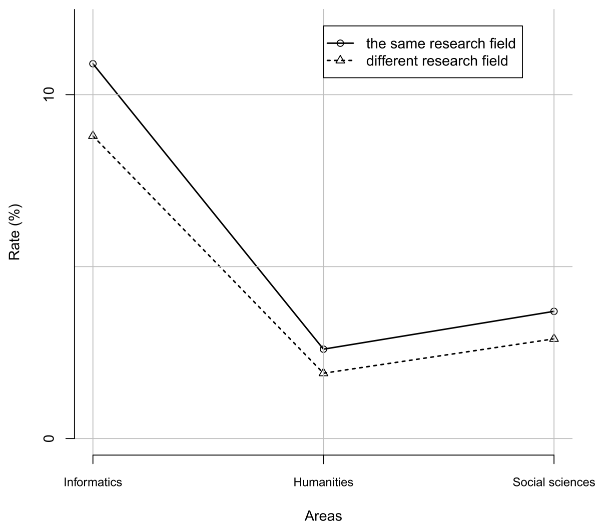

Fig. 6 Ratio of researchers with duplicate applications.

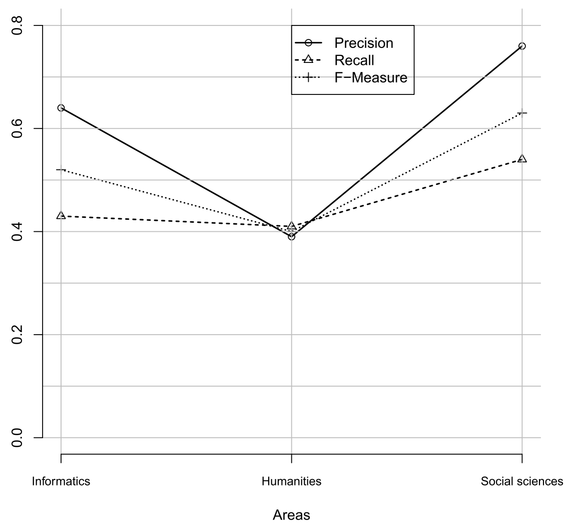

Fig. 7 Precision, recall and F measure in each area.

The ratio of researchers who applied for different research fields is 1.9% in Humanities and 2.9% in Social Sciences. These ratios for fields such as Economics and Psychology in Social sciences are higher than other research fields. On the contrary, there are no research fields in Humanities with comparatively high ratios. More specifically, no research fields in Humanities exceed four in duplicate applications. Thus, the number of researchers who applied for "Challenging Exploratory Research" is not sufficient to predict in Humanities. This is why Cumulative duplicate has the smallest influence and the F-measure of prediction results is low in Humanities (as shown in Fig. 7). Therefore, in areas with high clustering coefficients and sufficient duplicate applications, our proposed method is effective.

9. Conclusion and Future Work

In this study, we proposed methods for comprehending and predicting scientific research trends. Our experimental results showed that sets of research histories are identified with actual scientific research trends. Also, our proposed method, which focuses on breakthrough research, is effective in areas that are actively interacting with one another.

Future work includes the investigation of Research development process in the same research fields and networks of collaborative research. If several common properties in the network are found, we propose that the precision of predictions would be improved.

Acknowledgments

We are deeply grateful to the program officers in JSPS, Research Center for Science Systems.

References

[1] M.E.J. Newman, "Coauthorship networks and patterns of scientific collaboration," Proc. Natl. Acad. Sci. USA, pp.5200–5205, 2004.
[2] M.C. Pham, R. Klamma, and M. Jarke, "Development of computer science disciplines — A social network analysis approach," CoRR, pp.321–340, 2011.
[3] C. Chen and R.J. Paul, "Visualizing a Knowledge Domain's Intellectual Structure," IEEE Computer Society Press, pp.65–71, 2001.
[4] H. Small, Tracking and Predicting Growth Areas in Science, Scientometrics, pp.595–610, 2006.
[5] Z. Jian, C. Chaomei, and L. Jiexun, "Visualizing the intellectual structure with paper-reference matrices," IEEE Trans. Vis. Comput. Graphics, vol.15, pp.1153–1160, 2009.
[6] R.K. Abercrombie, A.W. Udoeyop, and B.G. Schlicher, "A study of scientometric methods to identify emerging technologies via modeling of milestones," Scientometrics, pp.1–16, 2012.
[7] K.W. Boyack, K. Börner, and R. Klavans, "Mapping the structure and evolution of chemistry research," Scientometrics, pp.45–60, 2009.
[8] A. Saka, M. Igami, and T. Kuwahara, Science Map 2008 — Study on Hot Research Areas (2003-2008) by bibliomethod — (in Japanese), NISTEP REPORT No.110, 2010.
[9] S. Okuoka, D. Katagami, and K. Nitta, "Analysis on change of research trend by research network and evaluation (in Japanese)," Proceedings of Knowledge Based Systems, pp.21–26, 2009.
[10] L. Woo, "How to identify emerging research fields using scientometrics: An example in the field of information security," Scientometrics, pp.503–525, 2008.
[11] J. Frizo, G. Wolfgang, and D.M. Bart, "A hybrid mapping of information science," Scientometrics, pp.607–631, 2008.
[12] C. Chen, J. Zhang, and M.S. Vogeley, "Visual analysis of scientific discoveries and knowledge diffusion," Proceedings of ISSI 2009 - 12th International Conference of the International Society for Scientometrics and Informetrics, pp.874–885, 2009.
[13] K. Satoh, R. Ichise, S. Kurihara, A. Aizawa, and M. Numao, "An analysis of research grant application data (in Japanese)," Proceedings of 25th Fuzzy System Symposium (CD-ROM), 2009.
[14] B.W. Herr II, E.M. Talley, G.A.P.C. Burns, D. Newman, and G. Larowe, "The NIH visual browser: An interactive visualization of biomedical research," Proceedings of International Conference on Information Visualisation, pp.505–509, 2009.
[15] N. Yamashita, M. Numao, and R. Ichise, "Relation extraction of research fields using proposal transitions of research fields in research grants and its prediction (in Japanese)," IPSJ SIG Technical Report, vol.2010-ICS-161, no.2, 2010.
[16] Japan Society for the Promotion of Science, Application Procedures for Grants-in-aid for Scientific Research-kakenhi

fy2012, 2011 (http://www.jsps.go.jp/j-grantsinaid/03_keikaku/data/h24/download/e/fullpage_e.pdf), (viewed Feb. 21 2012).
[17] Google Inc., Google Trends (https://www.google.co.jp/trends/), (viewed July 28 2014).
[18] K. Nakakoji, From Interfaces to Interactions (in Japanese) (http://www.sighci.jp/afiles/view/20), (viewed July 14 2011).
[19] M. Doi, The Institute of Electronics, Information and Communication engineers, network robot technical group (in Japanese) (http://www.irc.atr.jp/ieice_nwr/about.html), (viewed July 14 2011).
[20] Information Processing Society of Japan, The Special Interest Groups BIoinformatics and Genomics, Charter (in Japanese), (http://www.ipsj.or.jp/katsudou/sig/sighp/bio/tbio.html#effect), (viewed July 14 2011).
[21] Japanese Society for Cognitive Psychology, Charter (in Japanese), (http://cogpsy.jp/setsuritsu.html), (viewed July 14 2011).

Nagayoshi Yamashita received his Ph.D. degree in information science from Osaka University Osaka, Japan, in 2009. From 2009 to 2012, he was working in the Japan Society for the Promotion of Science (JSPS). He is currently working in System Department at GMO research. His research interests include machine learning and data mining.

Masayuki Numao is a professor in the Department of Architecture for Intelligence, the Institute of Scientific and Industrial Research, Osaka University. He received a bachelor of engineering in electrical and electronics engineering in 1982 and his Ph.D. in computer science in 1987 from Tokyo Institute of Technology. He was working in the Department of Computer Science, Tokyo Institute of Technology from 1987 to 2003, and was a visiting scholar at CSLI, Stanford University from 1989 to 1990. His research interests include Artificial Intelligence and Machine Learning. He is a member of Information Processing Society of Japan, Japanese Society for Artificial Intelligence, Japanese Cognitive Science Society, Japan Society for Software Science and Technology, and the American Association for Artificial Intelligence. Division of Information and Quantum Sciences, The Institute of Scientific and Industrial Research (ISIR), Osaka University, 8–1 Mihogaoka, Ibaraki, Osaka, 567–0047, Japan.

Ryutaro Ichise received his Ph.D. degree in computer science from Tokyo Institute of Technology, Tokyo, Japan, in 2000. From 2001 to 2002, he was a visiting scholar at Stanford University. He is currently an associate professor in the Principles of Informatics Research Division at the National Institute of Informatics in Japan. His research interests include machine learning, semantic web, and data mining.