Science and technology fields and higher education institutions with mathematically trained contributors: Metadata analysis of IEEE papers

Takashi Ikegawa
The University of Tokyo/Waseda University
Tokyo, Japan
tikegawa@ms.u-tokyo.ac.jp

***Abstract*—Recently, the shortage of mathematically trained individuals who can integrate mathematical knowledge and skills with other types of knowledge and skills and thereby contribute to innovation has become apparent. To find clues for solving this problem, we identify scientific and technology fields (STFs) that require mathematically trained contributors and identify higher education institutions (HEIs) that provide students an opportunity to develop the mathematical knowledge needed in the STFs. We perform this by analyzing a large amount of metadata of papers in IEEE journals, which is easily accessible via the Scopus API. The metadata contains author's affiliation information, which is useful for identifying HEIs with a mathematical faculty. For example, if an author's affiliation information contains terms related to mathematics such as "Department of Mathematics" or "Dept. of Math.," the author can be considered to have had advanced mathematical training. We demonstrate that 1) the contributions of mathematically trained individuals to research fields covered by published IEEE papers have continued to increase, 2) their contributions are particularly important to micro-STFs such as information theory, reliability, fuzzy systems, and neural networks, and 3) the degree of their contributions to the various STF-focused journals differs between HEIs.**

***Index Terms*—Survey data collection, mathematical training, IEEE papers, scientific and technology field, higher education institution, Scopus API.**

I. Introduction

Mathematics (or mathematical science) plays a key role in innovation by providing solutions to many challenging problems in industrial as well as scientific fields. For example, mathematics has greatly contributed to such industrial fields as information security, image analysis, data mining, and theoretical economics [1]–[3].

However, it has become apparent that there is a shortage of mathematically trained individuals who can integrate (advanced) mathematical knowledge and skills with other types of knowledge and skills and thereby contribute to innovation [3], [4].

A procedure for solving this problem is herewith presented:

Step 1: Identify scientific and technology fields (STFs) that require innovative thinking from mathematically trained individuals,

Step 2: Identify higher education institutions (HEIs) that provide students an opportunity to develop the mathematical knowledge that is needed in the identified STFs,

Step 3: Investigate the educational practices (e.g., curriculum, project-based learning, and industrial internship programs) provided by the identified HEIs that produce mathematically trained graduates who are active in the identified STFs, and

Step 4: Promote widely such excellent educational practices to other HEIs.

In this paper, we focus on the solutions of Steps 1 and 2 among aforementioned Steps 1 to 4. We present a method comprising Steps 1 and 2 that is based on bibliometric analysis (i.e., analysis of a large number of written publications) [5]. We also investigate the STFs and HEIs with mathematically trained contributors by analyzing a large amount of metadata of IEEE papers.

Major databases that have been standard for bibliometric analysis include Scopus, the Web of Science (WoS), and Google Scholar [6]. We used the Scopus database because its application programming interface (API) [7] is RESTful [8] and provides a variety of tools including search and retrieval for various target items [9].

The data obtained using the Scopus API, which can be treated as the traditional "big" data [10], has missing values. Then, we developed methods for imputing the missing values using rule-based machine learning.

The main contributions of this paper are as follows.

1) We present a method that uses raw data from the Scopus database to identify journals that publish papers from mathematically trained contributors and to identify their HEI affiliations.
2) We introduce a measure referred to as the "contribution rate of mathematically trained author(s) (CRMT)" that quantitatively represents the degree of contribution by mathematically trained author(s) to a journal's published papers.
3) The results obtained by applying this method to published IEEE papers reveal that
   - the contributions of the mathematically trained contributors to the research fields covered by the IEEE papers have continued to increase (i.e., the CRMT for all journals has been increasing for the past decade),

- the STFs that substantially benefited from the contributions of mathematically trained contributors include information theory, reliability, fuzzy systems, and neural networks (i.e., main topics covered by journals with a very high CRMT), and
- the degree of the contributions by mathematically trained contributors to the identified STF-focused journals differs between HEIs.

The rest of the paper is organized as follows. Section II explains the target data acquired using the Scopus API and the reasons for focusing on IEEE papers. Section III describes the procedure used to to obtain lists of journals and HEIs ranked by CRMT. Sections IV and V respectively explain the procedures used to extract the organization/(sub-)department names and organization type, which are essential items of information for data analysis. Section VI describes the imputation of missing values. Section VII discusses the STFs that require mathematically trained contributors and the HEIs of the contributors. Finally, Section VIII summarizes the key points and mentions future work.

II. Target data acquired

We used a bibliometric approach, which has been used for providing quantitative analysis of written publications, to data collection and analysis. We focused on IEEE papers for the following three reasons in particular.

- New information and communication technologies (ICTs) (or digital technologies in a broad sense) including artificial intelligence and big data analytics are playing a key role in the fourth industrial revolution, which has emerged over the past decade. Mathematics is a well-known contributor to the emergence and advancement of such new ICTs. Hence, analysis of IEEE journals with a scope related to the research fields relevant to the ICTs (see Appendix) should enable us to efficiently and effectively identify the various STFs that require innovative thinking from mathematically trained individuals.
- IEEE journals continue to maintain rankings at the top of their fields according to journal citation reports based on scientometric indices such as the Journal Impact Factor.[1] Analysis of the metadata of IEEE papers should thus enable us to efficiently and effectively identify individuals with "advanced" mathematical training.
- The metadata of IEEE papers can be easily acquired using the Scopus API.

We used the Scopus API to acquire the raw data required to obtain the results of interest.

Digital object identifier (DOI) is an identifier permanently assigned to an object associated with content including scholarly material (e.g., papers and e-books) [11]. It enables an IEEE paper for which metadata has been acquired to be uniquely identified.

Paper metadata is descriptive information about an IEEE paper, including the journal name, paper title, publication date, group of author names, and group of author affiliations.

Organization data is descriptive information about each author's affiliated organization(s).

TABLE I
HIERARCHICAL ELEMENT STRUCTURE OF PAPER METADATA ASSOCIATED WITH JAVASCRIPT OBJECT NOTATION (JSON) OBJECTS IN RAW DATA.

Element	JSON object in raw data
• *Journal name*	"sourcetitle"
• *Paper title*	"citation-title"
• *Publication date*	"publicationdate"
• *Author information group*	"author-group"
– *Author name*	"author"
– *Affiliation information*	"affiliation"
∗ *Country name*	"country"
∗ *Organization name*	"organization"
∗ *Organization ID*	"affiliation-id"
∗ *Organization source information*	"ce:source-text"

TABLE II
ELEMENT STRUCTURE OF ORGANIZATION DATA ASSOCIATED WITH JSON OBJECTS IN RAW DATA.

Element	JSON object in raw data
• *Organization name*	"affiliation-name"
• *Alternative organization name*	"name-variant"
• *Organization type*	"org-type"

III. Data analysis

Figure 1 shows the procedure used to obtain lists of journals and HEIs ranked by CRMT. It comprises three stages: raw data acquisition, author information data extraction, and result production. We describe them in the following subsections.

A. Raw data acquisition

In this stage, the DOI data, paper metadata, and organization data are created using JSON-formatted [12] raw data acquired using the Scopus API:

Digital object identifier (DOI): The DOIs of all papers for every IEEE journal are acquired using the following steps:

1) The International Standard Serial Number (ISSN) [13] of each IEEE journal is acquired in advance by using an ISSN search engine[2] with a keyword from the target journal name.
2) The DOIs of all papers in the target journals are acquired using the Scopus Search API[3] with a parameter value of the corresponding ISSN.

Paper metadata: Raw data for each IEEE paper are acquired using the Abstract Retrieval API[4] with a parameter value of the corresponding DOI. Appropriate elements from the raw data are then extracted to create the paper metadata.

Table I lists the hierarchical element structure of paper metadata associated with JSON objects in the raw data. Note that we represent a name for data elements in *italics* (except for a JSON-formatted name, which is denoted in sans-serif font).

[1] https://www.ieee.org/publications/subscriptions/journal-citations.html

[2] https://portal.issn.org/

[3] https://dev.elsevier.com/documentation/ScopusSearchAPI.wadl

[4] https://dev.elsevier.com/documentation/AbstractRetrievalAPI.wadl

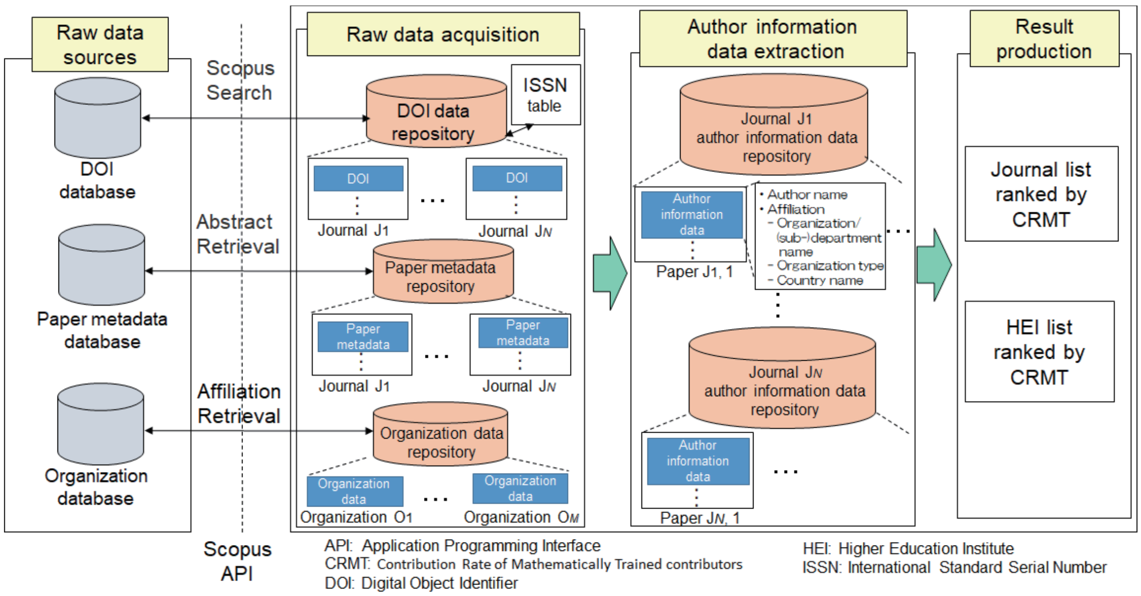

Fig. 1. Procedure to obtain lists of journals and HEIs ranked by CRMT.

A paper usually has multiple authors. Then, an element of the *author information group*, which has a JSON object name of "author-group", has a structure containing multiple items of author information (see Table I).

For example, Fig. 2 shows the JSON-formatted paper metadata for a paper by Ikegawa *et al.* [14].

Organization data: Raw organization data is acquired using the Affiliation Retrieval API[5] with a parameter value of the corresponding *Organization ID*. The appropriate elements are then extracted from the raw data to create the organization data.

Table II lists the element structure of the organization data.

B. Author information data extraction

In this stage, the group of author's information (referred to as "author information data") is created for each paper from the paper metadata and organization data. It has the following structure.

- *Author name*
- *Affiliation information*
  - *Organization name*
  - *Department name*
  - *Sub-department name*
  - *Organization type*
  - *Country name*

Remark 1 We found that the raw data acquired using the Scopus API occasionally had missing values, including

- *Affiliation information*, *Organization name*, and *Organization source information* in the paper metadata, and
- *Organization type* in the organization data.

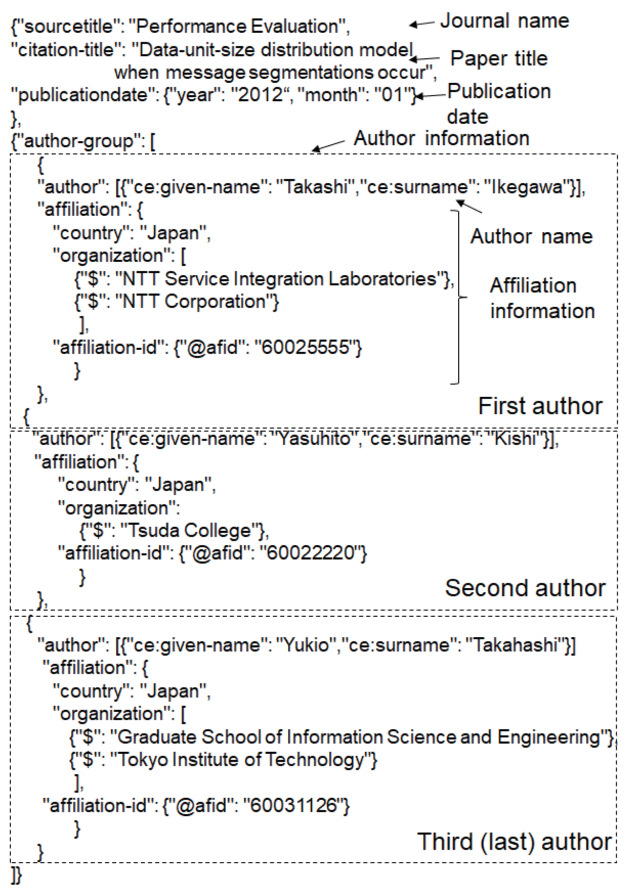

Fig. 2. JSON-formatted paper metadata for paper by Ikegawa *et al.* [14].

[5] https://dev.elsevier.com/documentation/AffiliationRetrievalAPI.wadl

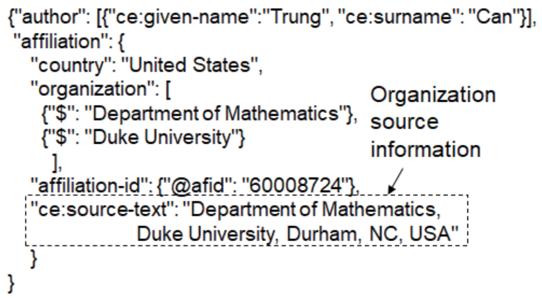

Fig. 3. JSON-formatted paper metadata including *Organization source information* (object "ce:source-text").

For example, the paper metadata shown in Fig. 2 has missing values of object "ce:source-text" (element *Organization source information* equivalently), whereas the paper metadata shown in Fig. 3 has these values.

We imputed the missing values using rule-based machine learning (see Section VI) and obtained results with high accuracy. ■

C. Result production

In this stage, lists of journals and HEIs ranked by CRMT are produced using the author information data. These lists are used for identifying the STFs that require innovative thinking from mathematically trained contributors and identifying the contributor HEIs (see Section VII).

IV. Procedure for Extracting Values of Organization and (Sub-)Department Names

In this section, we explain the method used to extract the values of elements for *Organization name*, *Department name*, and *Sub-department name* in the author information data from the values of object "organization" (element *Organization name* equivalently) in the paper metadata.

As shown in Fig. 2, the number of entries comprising object "organization" is variable.[6] For the example shown in Fig. 2, the number of entries for the second author's "organization" is one whereas that for the first and third (last) authors' "organization" is two.

Table III shows examples of patterns for entries comprising object "organization". From this table, we find that the arrangement of entries comprising "organization" has the rule (i.e., sequential pattern), denoted by $\mathtt{R}_{\mathtt{org}}$:

Rule of entry arrangement for object "organization" $\mathtt{R}_{\mathtt{org}}$

- The last entry value is the name of the author's organization (if the number of entries is one, the author's organization name is identified with this value, and the (sub-)department name, it any, is unknown).
- If object "organization" consists of two entries, the first entry's value is the name of the author's department, and the second (last) entry's value is the name of the author's organization.
- If object "organization" consists of three entries, the values of the first and second entries are the names of the author's sub-department and department, respectively, and the third (last) entry's is the name of the author's organization.

Hence, rule $\mathtt{R}_{\mathtt{org}}$ enables us to extract values of elements for *Organization name*, *Department name*, and *Sub-department name* in author information data from values of object "organization" in paper metadata.

[6]If object "organization" is composed of multiple entries, it is represented in an array structure. The number of entries equals the degree of the array.

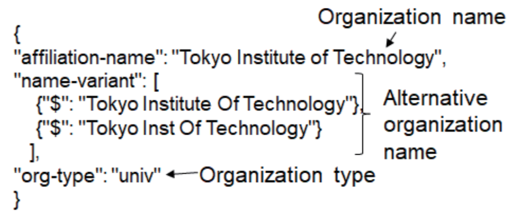

Fig. 4. JSON-formatted organization data for *Organization ID* 60031126, which is assigned to "Tokyo Institute of Technology."

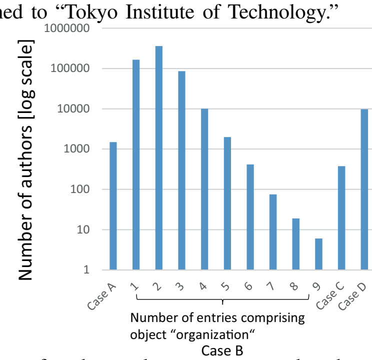

Fig. 5. Number of authors whose paper metadata has missing values for different missing-value patterns (according to Table V) for papers published during 2011–2020.

V. Procedure for Extracting Values of Organization Type

In this section, we explain the method used to extract the values of an element of *Organization type* in the author information data from that in the organization data.

Table IV lists organization classification based on values of object "org-type" in the organization data. From this table and the value of object "org-type" in the organization data, we can identify the organization type such as HEI and private enterprise (values of element *Organization type* equivalently) of the target organization.

Figure 4 shows JSON-formatted organization data for *Organization ID* 60031126, which is assigned to "Tokyo Institute of Technology." From this figure and Table IV, we find that "Tokyo Institute of Technology" is an HEI.

VI. Imputation of Missing Values

This section briefly explains the methods used for imputing missing values (as described in Remark 1).

A. Imputation of missing values for names of organization and (sub-)department when there is no Organization name

We classify the missing-value patterns for elements within element *Author information group* (i.e., object "author-group") in the paper metadata (see Table V).

TABLE III

EXAMPLES OF PATTERNS FOR ENTRIES COMPRISING OBJECT "organization".

No. of entries	Pattern*	Examples
1	{Organization name}	{University of Portsmouth}, {National Statistics Institute}
2	[{Department name}, {Organization name}]	[{Department of Mathematics}, {Southeast University}] [{Assembly Technology Development Group}, {Intel Corporation}]
3	[{Sub-department name}, {Department name}, {Organization name}]	[{Department of Automation}, {School of Electronic and Information Engineering}, {Xi'An Jiaotong University}]

∗ Square brackets "[ ]" represent an array structure.

TABLE IV

ORGANIZATION CLASSIFICATION BASED ON VALUES OF OBJECT "org-type" IN ORGANIZATION DATA.

Category		Values of object "org-type"*	Examples
Higher education institution (HEI)	Large	univ	Beijing Institute of Technology, Purdue University
	Small	coll, coll|ngov, coll|resi	Thunderbird School of Global Management
	Medical	meds, hosp|meds	Indiana University School of Medicine
	Research	resi|univ	Tata Institute of Fundamental Research
Private enterprise	Corporation	comp, comp|govt, comp|resi, comp|ngov, comp|lawf	AT&T Inc., Apple Computer, DuPont, Intel Corporation, Motorola, Mozilla Corporation, Toyota Motor Corporation
	Law firm	lawf	Fish & Richardson P.C., Steptoe & Johnson LLP
Medical institution	–	hosp, hosp|ngov, hosp|resi	Dana-Farber Cancer Institute, Toronto General Hospital
Public research institution	–	resi, ngov|resi	Agricultural Research Corporation Sudan
Government-affiliated organization	–	govt, govt|hosp, govt|resi, poli	Australian Bureau of Statistics, Australian Institute of Sport
Military organization	–	milo, milo|resi	Finnish Defence Research Agency
Non-profit organization	–	fund, fund|govt, ngov	The Eye & Ear Foundation of Pittsburgh
Museum	–	museum	Adler Planetarium, Van Gogh Museum

∗ The definition of values for object "organization" is not explicitly described in Scopus documentation. The classification used here was thus generated by the author.

Figure 5 shows the number of authors whose paper metadata has missing values for different missing-value patterns (according to Table V) for papers published during 2011–2020.

We imputed the missing values for Case C (i.e., missing values of element for *Organization name*) using values of object "ce:source-text", which is expressed in comma-separated value (CSV) format, using rule-based machine learning as follows.[7] The basic arrangement rule of CSVs extracted from object "ce:source-text" is as follows (see Fig. 3 for example): sub-department name → department name → organization name → address information.

Therefore, this simple rule with the value of element *Organization name* and the value of element *Alternative organization name* enables us impute the missing values.

B. Imputation of missing values of Organization type when there is no Organization name

Figure 6 shows the ratios of value of element *Organization type*, which is represented by the names categorized in Table IV for papers published during 2011–2020. From this figure, we found that the ratio of missing values of *Organization type* exceeds 50%. We imputed these missing values in accordance with the following rules:

Higher education institution (HEI): If the value of element *Organization name* contains terms such as

`University, Univ., Université, Universität, Polytechnic, Politecnico, and Politécnica,`

the organization is treated as an HEI.

Private enterprise: If the value of element *Organization name* contains terms such as

`Corporation, Corp., Incorporation, Inc., Company, Co., Compagnie, Limited, and Ltd.,`

the organization is treated as a private enterprise.

Application of this simple method resulted in 65% of the missing values being imputed for the example shown in Fig. 6.

TABLE V

CLASSIFICATION OF MISSING-VALUE PATTERNS FOR ELEMENTS WITHIN *Author information* (OBJECT "author-group") IN PAPER METADATA.

Case	Affiliation information (object "affiliation")	Organization name object ("organization")	Organization source information (object "ce:source-text")
A	Missing	–	–
B	Existing	Existing	Existing/Missing
C	Existing	Missing	Existing
D	Existing	Missing	Missing

[7]Missing values for Cases A and D could not be imputed due to lack of other information.

VII. RESULTS AND DISCUSSION

In this section, we discuss the STFs that require innovative thinking from mathematically trained individuals and the institutions with which they are affiliated.

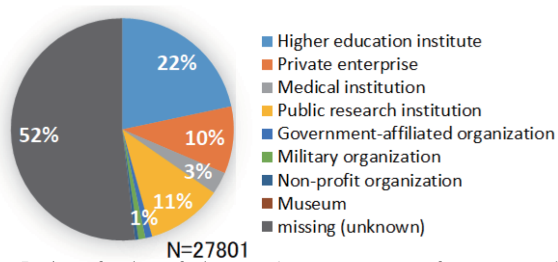

Fig. 6. Ratios of value of element *Organization type* for papers published during 2011–2020.

The publications issued by the IEEE are mainly classified into two categories: journals (or transactions) and magazines. We focused on journals that publish papers reporting novel work. The “IEEE Access” journal[8] was excluded because it covers most of the research fields handled by the IEEE.

A. Treatment of papers with unresolved missing data

We found that if certain values are missing, the reliability of the results was reduced despite the imputation of missing data as described in the preceding section. These values are

> **D1:** If the value of element *Organization type* is missing, it is difficult to identify the HEIs.
>
> **D2:** If the value of element *Organization type* for an author exists and the author’s affiliation is classified as an HEI and the values of elements *Department name* and/or *Sub-department name* are missing, it is difficult to identify the mathematically trained contributors.

Such papers were treated as invalid and excluded from the analysis.

B. Definition of “mathematically trained”

We considered an author to be mathematically trained if the author is in the mathematics (sub-)department of an HEI; i.e., the name of the department contains at least the following words:[9]

```
Mathematics or ``Math.'', Statistics or
``Stat.'', and Operations Research or
``OR.''
```

Table VI lists example names of (sub-)departments that are primarily responsible for mathematics education.

C. Contribution rate of mathematically trained author(s) (CRMT)

As mentioned, we used CRMT to quantitatively represent the degree of contribution by mathematically trained authors to a journal’s published papers. We define a paper satisfying the following two conditions as one to which a mathematically trained author contributed:

- At least one of the authors was mathematically trained and
- The paper does not contain incomplete data; i.e., it is not missing D1 and/or D2.

1) Trend in number of papers with contribution from mathematically trained author(s): Figure 7 shows the total number of papers published, the number of papers without missing data (D1 and/or D2), and the number of papers with mathematically trained author(s) for IEEE journals published during 1990–2020. It shows that the rate of growth in the number of papers with mathematically trained authors was larger than that of all published papers.

We examined this characteristic quantitatively using CRMT as the metric. It is defined as the number of papers with mathematically trained author(s) divided by the number of papers without incomplete data (D1 and/or D2).

[8]https://ieeeaccess.ieee.org/learn-more-about-ieee-access/

[9]Mathematically trained people should include individuals who received the Ph.D. degree in mathematics (or the related STFs). This extension is a further study because we need additional information such as an author’s biography.

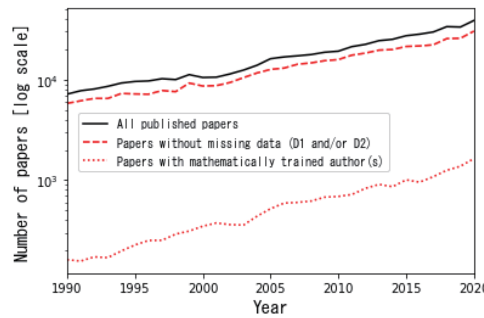

Fig. 7. Total number of papers published, number of papers without missing data (D1 and/or D2), and number of papers with mathematically trained author(s).

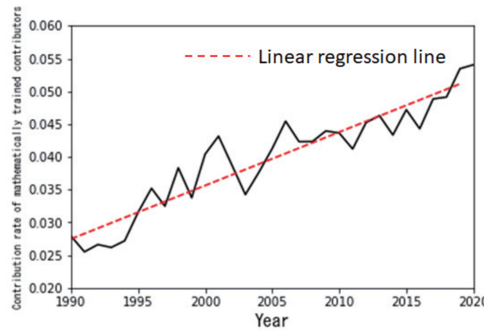

Fig. 8. Contribution rate of mathematically trained author(s) (CRMT).

The CRMT along with the linear regression line is shown in Figure 8. It shows that the contribution of mathematically trained authors to STFs has been increasing year by year.

2) STFs that require mathematically trained contributors: In our analysis of STFs that require mathematically trained contributors, we focused on papers published after 2011 to investigate the most recent trend.

The keywords of a journal’s name can be used for rough identification of the journal’s scope. We identified the STFs that require mathematically trained contributors by using these keywords.

Figure 9 shows the CRMT by journal. The y-axis shows the total number of journal editions during 2011–2020, and the x-axis shows the contribution rate. The total number of journals was 140. The mean, standard deviation, median, and third quartile values of CRMT were 0.053, 0.053, 0.040, and 0.072, respectively.

Table VII lists the journals ranked by CRMT for papers published during 2011–2020 and includes the research fields (areas) assigned by the WoS [15] to each journal.

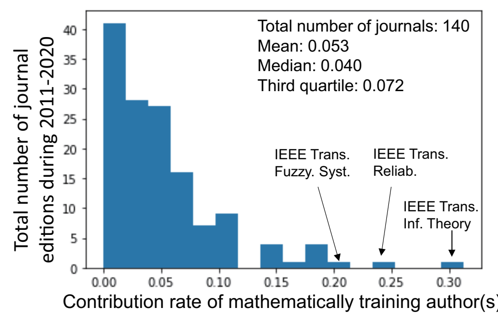

Fig. 9. Histogram of CRMT for papers published during 2011–2020.

TABLE VI
EXAMPLE NAMES OF (SUB-)DEPARTMENTS THAT ARE PRIMARILY RESPONSIBLE FOR MATHEMATICS EDUCATION.

	Examples	**Note**
Mathematics	Department of Mathematics, Department of Mathematics and Statistics, Dept. of Math., Center for Applied Math., Dept. of Pure Math.	English
	Département de Mathématiques	French
	Fakultät für Mathematik	German
Statistics	Department of Statistics, Dept. of Stat.	English
	Inst. fur Medizinische Informatik und Statistik	German
Operations Research	Department of Data Science and Operations Research	English

TABLE VII
JOURNALS RANKED BY CRMT FOR PAPERS PUBLISHED DURING 2011–2020.

Rank	**Journal name**	**No. of papers without missing data (D1 and/or D2)**	**No. of papers with mathematically trained author(s)**	**CRMT**	**Research fields assigned by WoS***
1	IEEE Transactions on Information Theory	4092	1278	0.31	CS, EN
2	IEEE Transactions on Reliability	833	211	0.25	CS, EN
3	IEEE Transactions on Fuzzy Systems	1388	298	0.21	CS, EN
4	IEEE Transactions on Neural Networks/IEEE Transactions on Neural Networks and Learning Systems	2162	446	0.20	CS, EN
5	IEEE/ACM Transactions on Computational Biology and Bioinformatics	1313	254	0.19	BM, CS, **MA**
6	IEEE Transactions on Automatic Control	3696	682	0.18	AC, EN
7	IEEE Transactions on Network Science and Engineering	302	55	0.18	EN, **MA**
8	IEEE Transactions on Systems, Man, and Cybernetics, Part B/IEEE Transactions on Cybernetics	1896	304	0.16	AC, CS
9	IEEE Transactions on Computational Imaging	109	15	0.14	EN, IP
10	IEEE Control Systems Letters	533	78	0.14	–
11	IEEE Transactions on Systems, Man, and Cybernetics, Part A/IEEE Transactions on Systems, Man, and Cybernetics: Systems	1225	157	0.13	CS
12	IEEE Transactions on Pattern Analysis and Machine Intelligence	1564	179	0.11	CS, EN
13	IEEE Transactions on Control of Network Systems	574	63	0.11	AC, CS
14	IEEE Transactions on Signal Processing	2568	275	0.11	EN
15	IEEE Transactions on Medical Imaging	1787	191	0.11	CS, EN, IP
16	IEEE Transactions on Computational Social Systems	341	36	0.11	CS, EN
17	IEEE Transactions on Image Processing	3198	330	0.10	–

∗ See Table IX for the definition of abbreviations.

From Table VII, we find that

- mathematically trained individuals contributed to the fields of computer science and engineering in the broad categories specified by the WoS (see Table IX), and
- from the ten top-ranked journals, in particular, they contributed to such micro-STFs as information theory, reliability, fuzzy systems, neural networks, bioinformatics, automatic control, network science, cybernetics, and image processing.

Note that the journal with the highest CRMT, i.e., "Transactions on Information Theory," covers a wide range of topics including coding theory, data compression, signal processing, pattern recognition, cryptography, and quantum information theory.[10] This means that future work includes identification of more micro-STFs by focusing on the topics of the papers rather than the journal names.

D. HEIs with which mathematically trained contributors are affiliated

Table VIII lists the four top-ranked HEIs with respect to CRMT with which mathematically trained contributors are affiliated. Table VIII (a) is for journals with a CRMT exceeding the third quartile, whereas Tables VIII (b), (c), and (d) are for the three top-ranked journals ("IEEE Transactions on Information Theory," "IEEE Transactions on Reliability," and "IEEE Transactions on Fuzzy Systems," respectively).

From Tables VIII (b), (c), and (d), we see that the HEIs with which mathematically trained contributors are affiliated differ by journal (the STFs equivalently).

The identified HEIs are well suited for developing educational programs that produce mathematically trained graduates who are active in fields other than mathematics. Therefore, future work includes investigating the educational programs provided by the identified HEIs. [11]

VIII. CONCLUSION AND FUTURE WORK

We have identified the fields of science and technology (STFs) to which mathematically trained individuals contribute and the higher educational institutions (HEIs) with which they are affiliated. To collect and analyze the metadata of IEEE papers efficiently, we used the Scopus API to collect the raw data and used rule-based machine learning to extract the

[10] https://ieeeittrans.ee.technion.ac.il/

[11] Mathematics is generally considered to comprise pure (fundamental) and applied mathematics. Surveys of educational programs in the field of applied mathematics have been performed (e.g., [16]). A survey with respect to "pure" mathematics is required.

TABLE VIII
FOUR TOP-RANKED HEIS WITH RESPECT TO CRMT WITH WHICH MATHEMATICALLY TRAINED CONTRIBUTORS ARE AFFILIATED.

Rank	Organization name	Department name	Country
1	Southeast University	School of Mathematics	China
2	Nanyang Technological University	School of Physical and Mathematical Sciences	Singapore
3	Chinese Academy of Sciences	Key Lab of Mathematics Mechanization	China
4	University of California	Department of Mathematics	United States

(a) Journals with CRMT exceeding third quartile value

Rank	Organization name	Department name	Country
1	Nanyang Technological University	School of Physical and Mathematical Sciences	Singapore
2	National University of Singapore	Department of Mathematics	Singapore
3	Zhejiang University	Department of Mathematics	China
4	Stanford University	Department of Statistics	United States

(b) Journal "IEEE Transactions on Information Theory"

Rank	Organization name	Department name	Country
1	McMaster University	Department of Mathematics and Statistics	Canada
2	King Abdulaziz University	Department of Statistics	Saudi Arabia
3	Tamkang University	Department of Mathematics	Taiwan
4	University of Isfahan	Department of Statistics	Iran

(c) Journal "IEEE Transactions on Reliability"

Rank	Organization name	Department name	Country
1	Ghent University	Department of Applied Mathematics, Biometrics, and Process Control	Belgium
2	University of Glamorgan	Department of Computing and Mathematical Sciences	United Kingdom
3	Liaoning University of Technology	Department of Mathematics	China
4	North China Electric Power University	Department of Mathematics and Physics	China

(d) Journal "IEEE Transactions on Fuzzy Systems"

appropriate data. We demonstrated that the contributions of mathematically trained individuals to STF-focused journals has been increasing and that the STFs receiving the greatest contributions include information theory, reliability, fuzzy systems, and neural networks. Furthermore, we identified the HEIs with which a large number of mathematically trained authors are affiliated.

This paper focused on "mathematically" trained contributors for electronic and electrical engineering fields (i.e., mathematicians with the other STF's knowledges and skills). Hence, the proposed approach in this paper will be available for developing educational practices to produce engineers with the multiple of specialized knowledges and skills.

Future work includes evaluating the accuracy of the implemented rule-based machine learning, expanding the survey scope to include leading journals other than IEEE journals, analyzing more micro-STFs using paper abstract information, and identifying the excellent educational programs provided by the HEIs with which mathematically trained authors are affiliated.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number 19K02871.

APPENDIX

This appendix discusses the range of research fields (areas) covered by IEEE journals.

There are several kinds of classification for research fields. We used the classification specified by the WoS [15]. This classification has 153 research fields categorized into five broad classes.

Table IX lists the WoS research fields covered by IEEE journals, and Table X lists the research fields covered by IEEE journals for the five classes. These tables show that IEEE journals cover a wide range of research fields.

TABLE IX
WOS RESEARCH FIELDS COVERED BY IEEE JOURNALS.

Category	Fields	Abbreviation
Life Sciences & Bio-medicine	Biochemistry & Molecular Biology	BM
	Mathematical & Computational Biology	–
	Medical Informatics	–
	Neurosciences & Neurology	–
	Radiology, Nuclear Medicine & Medical Imaging	–
	Rehabilitation	–
Physical Sciences	Geochemistry & Geophysics	–
	Mathematics	**MA**
	Oceanography	–
	Optics	–
	Physics	–
Social Sciences	Business & Economics	–
	Education & Educational Research	–
Tech-nology	Acoustics	–
	Automation & Control Systems	AC
	Computer Science	CS
	Energy & Fuels	–
	Engineering	EN
	Imaging Science & Photographic Technology	IP
	Instruments & Instrumentation	–
	Materials Science	–
	Nuclear Science & Technology	–
	Remote Sensing	–
	Telecommunications	TEL
	Transportation	–

TABLE X
RANGE OF RESEARCH FIELDS COVERED BY IEEE JOURNALS.

Category	No. of fields specified by WoS	No. of fields covered by IEEE journals	Coverage*
Arts & Humanities	14	0	0
Life Sciences & Biomedicine	76	6	0.08
Physical Sciences	17	5	0.29
Social Sciences	25	2	0.08
Technology	21	12	0.57

$*$ Coverage is defined as number of fields covered by IEEE journals divided by number of fields specified by WoS.

REFERENCES

[1] SIAM, "Mathematics in Industry Report," Accessed May 1, 2021. [Online]. Available: http://www.siam.org/Portals/0/Publications/Reports/Mathematics_Industry_2012.pdf

[2] National Research Council, The Mathematical Sciences in 2025, The National Academies Press. 2013.

[3] P. Bond, "The Era of Mathematics: An Independent Review of Knowledge Exchange in the Mathematical Sciences," Accessed May 1, 2021. [Online]. Available: https://epsrc.ukri.org/newsevents/pubs/era-of-maths/

[4] The Conference Board, "US Labor Shortages: Challenges and Solutions," Accessed May 1, 2021. [Online]. Available: https://conference-board.org/topics/labor-shortages/us-labor-shortages-report-2020

[5] O. Ellegaard and J.A. Wallin, "The bibliometric analysis of scholarly production: How great is the impact?," *Scientometrics*, vol. 105, pp. 1809–1831, 2015.

[6] A-W. Harzing and S. Alakangas, "Google Scholar, Scopus and the Web of Science: a longitudinal and cross-disciplinary comparison," *Scientometrics*, vol. 106, pp. 787–804, 2016.

[7] Elsevier, "API Interface Specification," Accessed May 1, 2021. [Online]. Available: https://dev.elsevier.com/api_docs.html

[8] C. Pautasso, O. Zimmermann, and F. Leymann, "RESTful web services vs. "big" web services: making the right architectural decision," in *Proceedings of the 17th International Conference on World Wide Web (WWW '08)*, pp. 805—814, 2008.

[9] M. E. Rose and J. R. Kitchin, "Pybliometrics: Scriptable bibliometrics using a Python interface to Scopus," *SoftwareX*, vol. 10, p. 100263, 2019.

[10] X. Wu, X. Zhu, G. Wu, and W. Ding, "Data mining with big data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 26, no. 1, pp. 97–107, Jan. 2014.

[11] ISO, "Information and documentation–Digital object identifier system," ISO 26324:2012, 2012.

[12] T. Bray, "The JavaScript Object Notation (JSON) Data Interchange Format," RFC 8259, Dec. 2017.

[13] ISO, "Information and documentation – International standard serial number (ISSN)," ISO 3297:2020, 2020.

[14] T. Ikegawa, Y. Kishi, and Y. Takahashi, "Data-unit-size distribution model when message segmentations occur," *Perform. Evaluation*, vol. 69, no. 1, pp. 1–16, Jan. 2012.

[15] Clarivate Analytics, "Research Areas (Categories / Classification)," Accessed on May. 1, 2021. [Online]. Available: https://images.webofknowledge.com/images/help/WOS/hp_research_areas_easca.html

[16] R. Levy, E. Chow, B. Kwon, K. Socha, M. McCarthy, and P.R. Turner, "SIAM Education Subcommittee Report on Undergraduate Degree Programs in Applied Mathematics," *SIAM Review*, vol. 59, no. 1, pp. 199–204, 2017.