

Analysis of Chinese Patents Associated with Incremental Clustering Algorithms: A Review

Archana Chaudhari^{1*}, Preeti Mulay², Amit Kumar Tiwari³

¹ Department of Computer Engineering, Dr. D. Y. Patil Institute of Technology, Pimpri, Pune

² Symbiosis Institute of Technology (SIT), Symbiosis International (Deemed University) (SIU), Pune, India

³ Ex-Head - Intellectual Property, Symbiosis Centre for Research and Innovation (SCRI), Symbiosis International (Deemed University) (SIU), Pune, India

Corresponding author: * chaudhari.archana12@gmail.com

Received Date: 12 January 2022

Accepted Date: 13 February 2022

Published Date: 30 March 2022

HIGHLIGHTS

- The patent landscape of China as it is emerging in the field of incremental learning via incremental clustering.
- Knowledge investment in the field of intellectual property will bring more prosperity.
- Artificial Intelligence-related inventions are booming, shifting from theory to commercial application.

ABSTRACT

With the advent of Internet-of-Things (IoT) and the overall Information-Technology world, an enormous amount of data is getting generated dynamically and in real-time mode, in almost all domains of research and application systems. Such huge data has embedded patterns and hidden information to extract and learn. This learning is incremental in nature for all involved entities and users, as the data is growing exponentially in real-time. To achieve learning from such dynamic data sources, incremental clustering algorithms are used mandatorily. This mandate has given rise to increased patents related to the incremental clustering concept, which is primarily a significant part of the Machine Learning field. In this paper, we contribute to the in-progress discussion on the use of intellectual property resources, particularly patents related to machine learning, incremental clustering, incremental learning with a special focus on China. Due consideration of the prior art search, the author found that China, the country of registration of the application extensively contributes to the intellectual property related to incremental clustering domain hence felt the need to undertake this detailed patent analysis about this topic. We hope all readers, research scholars will benefit with the latest research presented in this paper pertaining to various patents in the advanced areas of computer engineering.

Keywords: *patent analytics, artificial intelligence, machine learning, incremental clustering*



INTRODUCTION

Artificial Intelligence (AI) is progressively driving essential developments in technology and business, from medical diagnosis to advanced manufacturing. AI is new electricity, every industry transformed by AI. AI-related inventions are booming, shifting from theory to commercial application. AI is an amalgamation of Machine Learning (ML), Deep Learning (DL), etc. Some areas of AI-related patenting are developing more quickly as compared to others. ML is the leading AI technique disclosed in patents. ML-related patents acquire more than one-third of all identified inventions. ML techniques are categorized into Supervised, Unsupervised, and Reinforcement learning. Clustering is the subcategory of unsupervised learning — however, traditional clustering lacks in the concept of new learning. Incremental learning via incremental clustering algorithm is the novel concept in ML.

As per the World Intellectual Property Indicators (WIPO) – 2018 (Indicators, 2018) report the Chinese patent application worldwide increased by 43.6% in 2017. Also, WIPO Technology Trends 2019: Artificial Intelligence (WIPO, 2019) says that "Companies dominate the patenting activity are mostly from Japan, the United States of America (U.S.) and China." Therefore, the present researchers are motivated to study the patent landscape of China as it is emerging in the field of AI. Moreover, patents are the most abundant source of scientific information and facts (Raturi, Sahoo, Mukherjee, & Tiwari, 2010). Due consideration of the prior art search, the author found that Chinese Universities extensively contribute to the intellectual property related to incremental learning via incremental clustering domain.

Incremental Clustering

The incremental clustering makes effective use of the new information that is being evolved and the existing knowledge base to have accurate decision-making. On the influx of new labelled or unlabelled data, incremental clustering either updates the existing clusters or forms a new cluster as shown in figure 1.

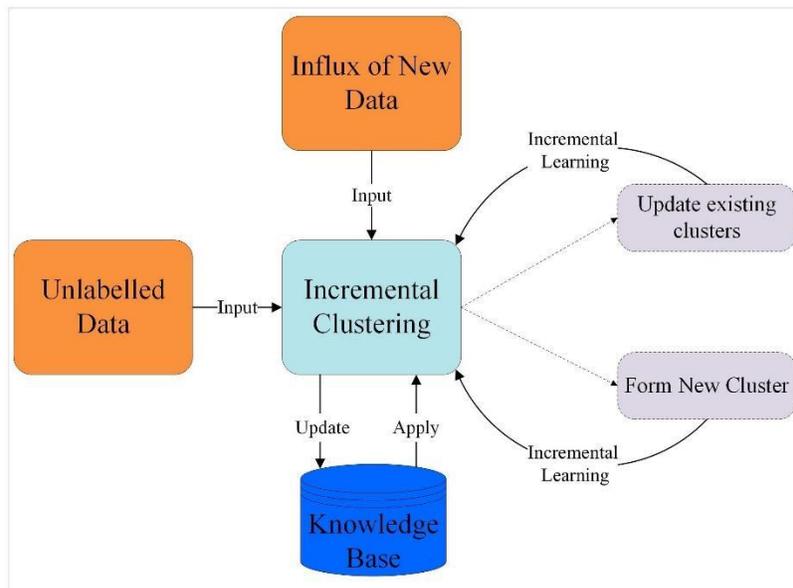


Figure 1: Working Principle of an Incremental Clustering



The Incremental-Clustering (IC) is an elite category of clustering algorithms. IC generates basic clusters from the available dataset and then on the influx of new data either append existing clusters or form new clusters. Some of the clustering methods are of incremental manners such as BIRCH(T. Zhang, R. Ramakrishnan, & Livny, 1997), COBWEB(T. Zhang et al., 1997), DBSCAN(Martin Ester, Hans-Peter Kriegel, Jörg Sander, & Xu, 1996), and Incremental K-Means(Zhu, Yu, Kan, & Liu, 2018). However, most of the algorithms required the input from the user, such as the number of clusters, radius, number of minimum point, etc. If the end-user selects/gives the wrong input(s), then it hampers the quality of clusters. To address these issues and to design a parameter-free algorithm to accommodate the arrival of new data effectually the 3 - phase algorithm called CFBA (Closeness Factor-Based Algorithm) is developed (P. Mulay & Kulkarni, 2013a, 2013b). It is further extended by TBCA (Threshold-based Clustering Algorithm) to analyze diabetic patients clinical parameters (P. Mulay, Joshi, & Chaudhari, 2018; Preeti Mulay et al., 2017). TBCA uses the threshold value 0 to 1 only. However, the threshold range from -1 to +1. Correlation-Based Incremental Clustering Algorithm (CBICA) (Dr. Preeti Mulay & Shinde, 2017), a new variant of CFBA has taken shape. CBICA uses Pearson's coefficient of correlation similarity measures. Personalization in diabetic patient care is successfully achieved using CBICA. One more variant of distributed CFBA is there i.e. Log likelihood based gradational clustering algorithm (A. Chaudhari & Mulay, 2019a). It is further extended to Nearness factor based incremental clustering algorithm (A. Y. Chaudhari & Mulay, 2020). The detail introductions of the incremental clustering and distributed incremental clustering methods are available in literature (A. Chaudhari, Joshi, Mulay, Kotecha, & Kulkarni, 2019; A. Chaudhari & Mulay, 2019b; Preeti Mulay, Joshi, & Chaudhari, 2020). The IC is also applied in the field of smart metres (Kuralkar, Mulay, & Chaudhari, 2020). By looking at all these professional IC algorithm's summary, it's clear that there is a need for such algorithms today and in future too, to effectively handle ever-growing data from varied domains.

This paper aims at in-progress discussions on the use of "Intellectual-Property-Resources" particularly the patents related to incremental clustering, machine learning, and data mining and incremental learning, with special focus on China the country of application. Incremental Learning easily achieved via Incremental Clustering is the most important and growing field of study and hence felt the need to undertake this detailed study about this topic.

This paper is organised in various sections; section 2 depicts a journey through the "Chinese Patent System," followed by bibliometric analysis in section 3. Conclusive summary and future research direction described as a part of section 4.

A JOURNEY THROUGH THE CHINESE PATENT SYSTEM

According to China's AI Development Report 2018 (Xue Lan, 2018), "The market value of China's AI industry reached 23.7 billion yuan (about USD 3.5 billion) in 2017, with the growth rate expected to reach 75 percent in 2018."

McKinsey Global Institute says, "AI will give the Chinese economy a productivity boost that will add 0.8 to 1.4 percentage points to annual GDP growth in the short term"(Henry, February 28, 2019). According to the World Economic Forum report, China plans for AI to be worth 1 trillion yuan (the US \$146 billion) to the country by 2030.

Figure 2 shows the facts and figures of the Chinese patent system. The patent application is directly proportional to the grant. Patent grant is equivalent to patent enforce. Last five year of statistical profiles of China's patent system have been growing continuously. Its growth rate is almost double in the patent



application, grant, and patent in force. With this study, the researcher found that in China applicants are interested in enforcing their rights, and it appears that the trends increases.

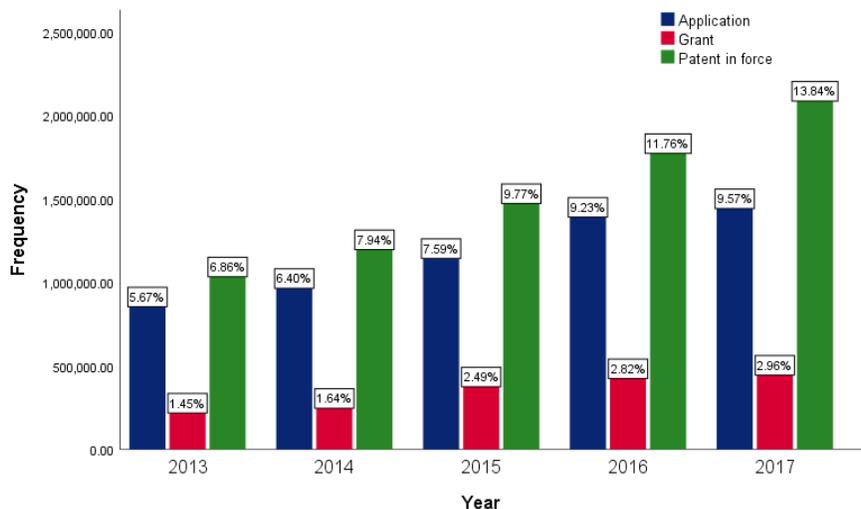


Figure 2: Fact and figure of the Chinese patent system (source (database, 2018))

METHODOLOGY

Filing a patent application is a necessity for the protection of an invention. A patent helps in restricting others from copying the patentee's invention/technology. But if the patent application has not been filed in a professional manner, it would not be an easy task to have a patent granted based on the patent application filed.

It is mandatory for the researchers to have in-depth knowledge about the ongoing invention in incremental learning via the incremental clustering field and the inventors who contribute to the Intellectual Property (IP) domain. In the present study, bibliometric analysis was conducted using the Espacenet patent database. Espacenet is a multinational patent database offered by European patent authority free of charge to the public. It is offering easy-to-use features for searching patent information. Espacenet offers free access to more than 110 million patent documents from over 120 patent authorities. Espacenet allows the user to search the query in three different filter options-smart search, advanced search, and classification search. Table 1 highlights the prominent features of Espacenet.

Table 1: prominent characteristics of Espacenet Database

Available search interface	Number Search
	Form Search
	Command-line search
Boolean operator	AND, OR, NOT
Wildcards	Any length
	Precisely one character
	Right truncation
	Proximity operations
Maximum search terms per field	10 terms
Maximum search terms per mask	20 terms



Maximum number of result	500
Keyword highlighting	possible
Full document download	Yes (With CAPTCHA)
Patent machine translation	Yes (Google)
Link to legal status	Yes(INPADOC)
Nr of max. exportable patents records	500
Nr of max. fields in export	24
Format	CSV, Xls

Various Internationally agreed Numbers for the Identification of (bibliographic) Data (INID) codes are used in the present study. The query for the Espacenet database was “(txt = incremental and txt = clustering) and txt = algorithm” of country China for six years during 2015-2019. The research is all about the new thing to do, and the incremental clustering domain has become more popular since 2015, hence author took current five years duration for a research study — preliminary investigation through planned keywords search tactic generated in all 2725 patents from Espacenet search engine. The retrieved data were analyzed from structured patent data, such as patent number, filing date, and assignees and unstructured data like text segments, keywords or key-phrases, abstracts or descriptions of the inventions (Vincent, Singh, Chakraborty, & Gopalakrishnan, 2017). The additional information required for the bibliometric analysis collected from journals, magazines, government websites, white papers, books, and other accessible databases. Figure 3 shows the process flow diagram for the comprehensive literature review (Aristodemou & Tietze, 2018).

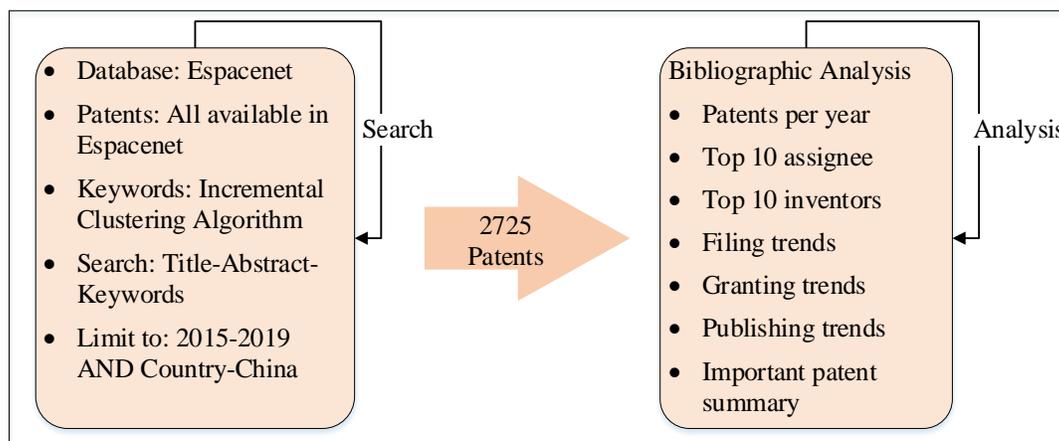


Figure 3: Process flow of the search and analysis of the patent literature review
 Source: <https://worldwide.espacenet.com> (accessed on 25 July 2020)

Bibliometric Analysis

This study analyses patent application filing, publication, and grant trends. Also identifies the prominent assignee, prominent inventors, legal status and citation analysis of patents in incremental learning via incremental clustering sector. Further, the future technological trends for R&D and investment activities could be predicted based on the analysis done (Vincent et al., 2017).



Patent Filing Trends

Patents are territorial in nature. They provide protection only in the jurisdiction where an application is filed and a patent granted. Figure 4 illustrates the annual patents filing trend in the Incremental Clustering Algorithm (ICA) domain. The research reveals that the highest number of patent applications was filed in the Chinese patent office in the year 2016. These filing explosions may be consistent with substantial advancements and corporate prioritisation of ICA seen in recent years.

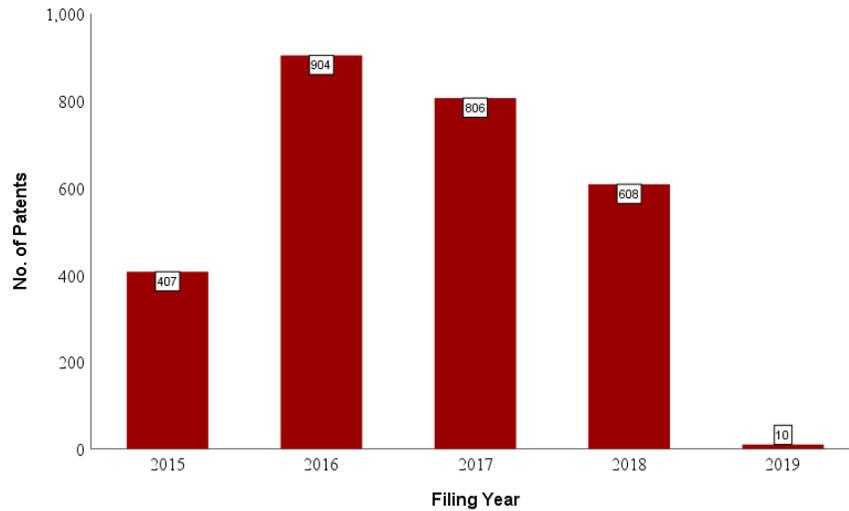


Figure 4: Number of a patent filed in the era ICA
Source: <https://worldwide.espacenet.com> (accessed on 25 July 2020)

Patent Publication Trends

Patent filing is directly proportional to patent publications not necessarily. Figure 5 shows the trends in ICA patent applications published from 2015 to 2019. It has observed that patents related to ICA have growing trends in publication. The present study reveals that, patent publication rate increased by 75.17% in the year 2018.

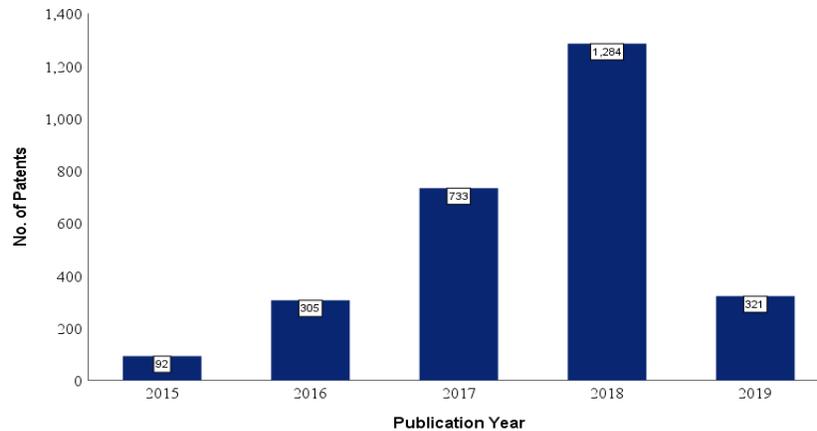


Figure 5: Patent application publishing trend in the era ICA
Source: <https://worldwide.espacenet.com> (accessed on 25 July 2020)



Patent Granting Trends

Figure 6 shows year after year changing patent grant count in the era of ICA. In the last five-year highest number of patents granted in the year 2018. The study reveals that patent granting are great increases from the year 2015 to 2018.

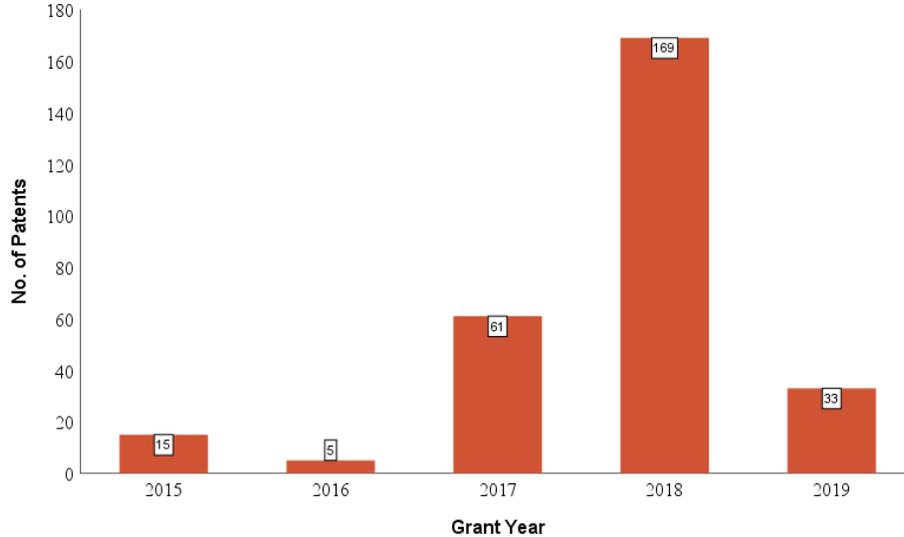


Figure 6: Patent application publishing trend in the era ICA
Source: <https://worldwide.espacenet.com> (accessed on 25 July 2020)

Figure 7 also compares patent filings, publication, and granting trends since the year 2015. The study reveals that the boom in ICA is almost doubling between 2015 and 2019.

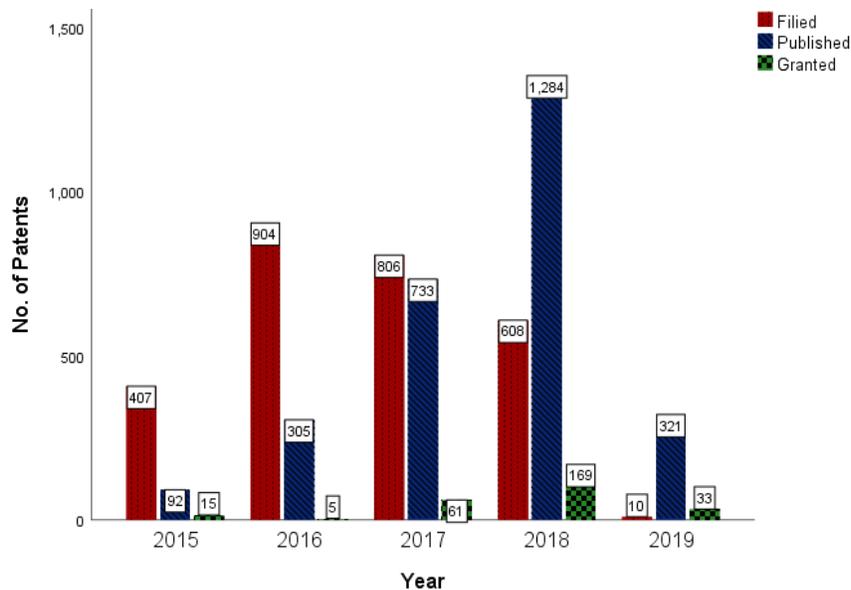


Figure 7: Comparison of a patent application filed, published, and granted in the era ICA
Source: <https://worldwide.espacenet.com> (accessed on 25 July 2020)



Patent Assignee Statistics

Figure 8 indicates the top ten contributing universities/industry affiliations of China in the field of ICA. The extracted information demonstrates that research and development, patenting activities are rising in public and private organisations in China. The INTEL is a leading assignee in the global perspective. Table 2 shows yearly patent filing trends of the leading assignee. The present study found that in the last five years, INTEL filed the highest number of patents in 2018 and contributed to the development of the country.

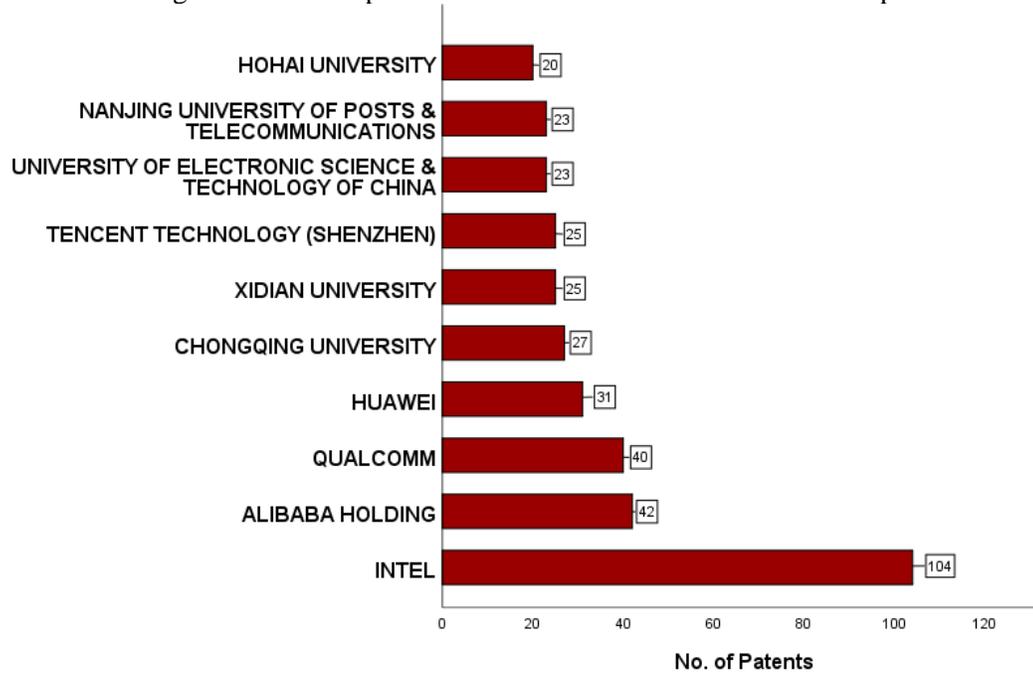


Figure 8: Top 10 assignee

Source: <https://worldwide.espacenet.com> (accessed on 25 July 2020)

Table 2: Metadata of top five assignee

Assignee	2015	2016	2017	2018	2019
INTEL	3	34	12	51	4
ALIBABA HOLDING	10	13	09	07	3
QUALCOMM	2	19	19	0	0
HUAWEI	6	16	6	2	1
CHONGQING UNIVERSITY	6	5	11	4	1

Patent Inventor Statistics

Figure 9 depicts the top ten leading inventors contributing in the area of ICA to understand the influence of particular inventors. The inventor Wang Yue is the highest patent filler in the last five years. The statistic indicates that a few inventors have taken the lead in the conceptual development of this emerging domain.



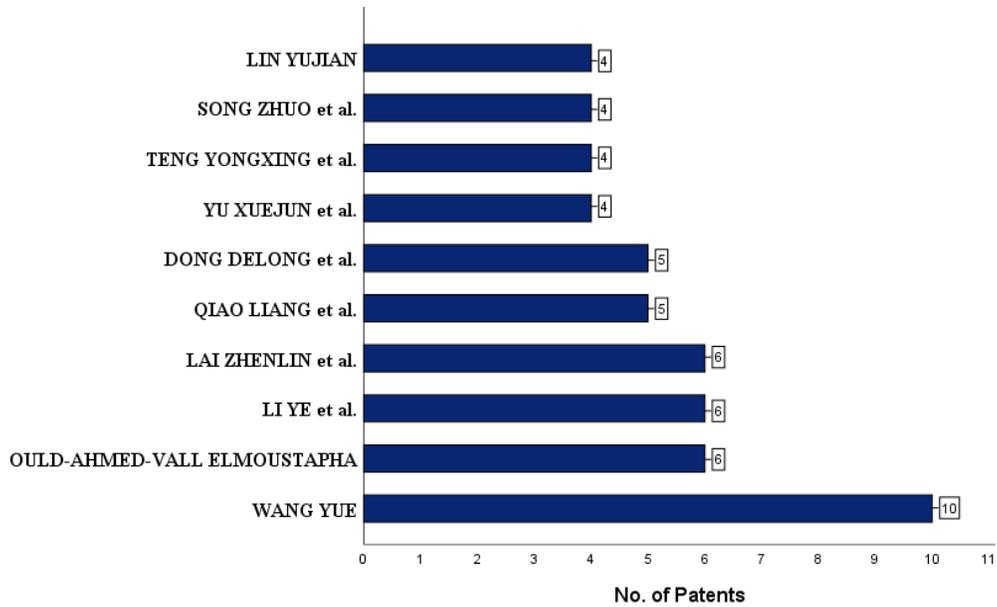


Figure 9: Key contributing inventors
 Source: <https://worldwide.espacenet.com> (accessed on 25 July 2020)

Legal Status of Patents

The legal status of the ICA patent analysed (Figure 10). It has been observed that pending patent applications are comparatively higher than the granted patents. It is clear that a significant number of patents lapsed in the field of ICA.

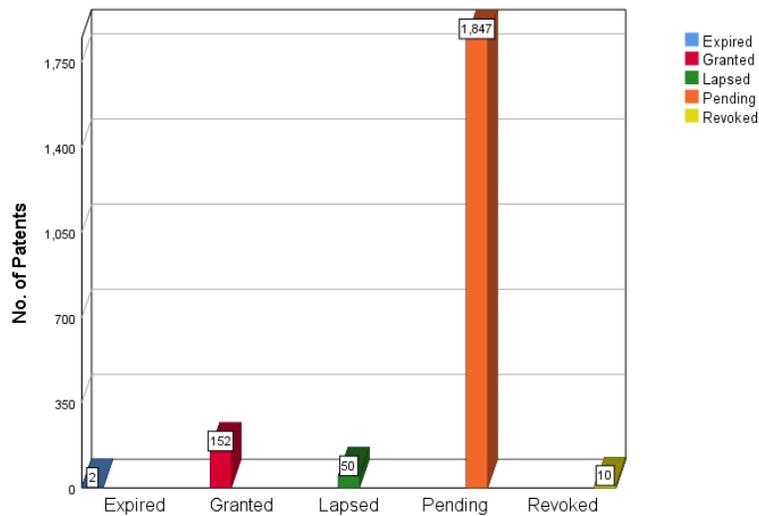


Figure 10: Legal status of patents in the era ICA
 Source: <https://worldwide.espacenet.com> (accessed on 25 July 2020)



Technology Domain

Figure 11 illustrates the ICA patent categorization by subject area. The maximum percentage of patents are distributed in the area of computer technology (53.17%), digital communication (15.35%), and telecommunication (5.05%), followed by biotechnology (1.24%).

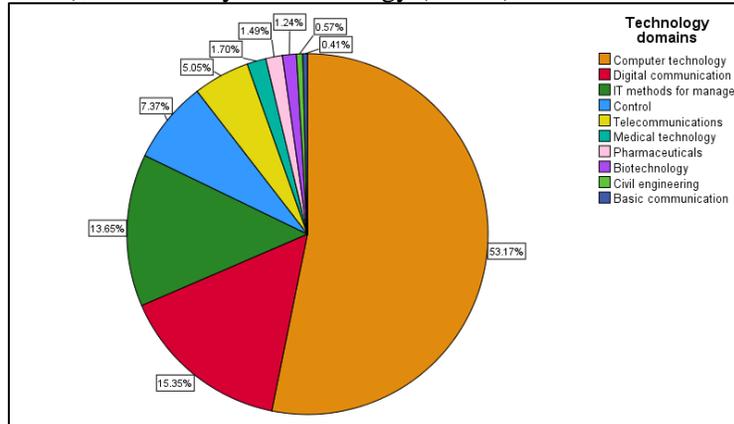


Figure 11: Percentage of a patent by subject area
 Source: <https://worldwide.espacenet.com> (accessed on 25 July 2020)

International Patent Classification (IPC) in the era ICA

IPC established by the Strasbourg Agreement 1971 ((WIPO)). IPC is a hierarchical system of patent classification. It classifies the patent according to the area of technology they pertain to. IPC provides language-independent symbol (A: Human Necessities; B: Performing Operations, Transporting; C: Chemistry, Metallurgy; D: Textiles, Paper; E:Fixed Constructions; F: Mechanical Engineering, Lighting, Heating, Weapons, Blasting; G: Physics; H:Electricity) ((WIPO)). The present study falls into the IPC classification code G06F (Electrical Digital Data Processing: Computer systems based on specific computational models). Figure 12 depicts 23 IPC codes under which the ICA patents are classified. The top five IPC code with the highest number of ICA patents are G06F-017/30; G06F-017/27; G06F-009/50; G06F-017/50; and G06F-009/30.

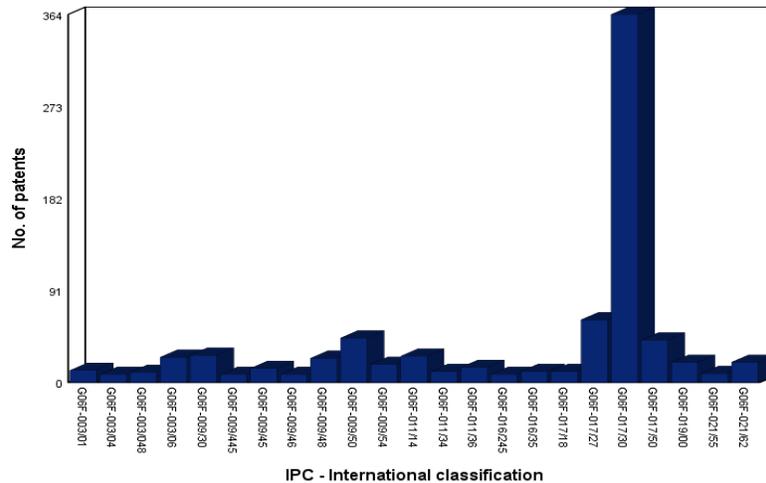


Figure 12: IPC classification in the era of ICA
 Source: <https://worldwide.espacenet.com> (accessed on 25 July 2020)



Citation Analysis

Citation Analysis could be used to track the impact of the patent, including which technologies & products are highly valuable. In general, two types of citation used in patent terminology: Backward citations are patents that are cited by a specific patent and forward citations are patents that cite a particular patent. Table 3 demonstrates the most frequently cited patent from 2015 to 2019. This citation trend reveals that, if the patent is cited multiple times, then it has positive economic value.

Table 3: A citation analysis of the top five patent in ICA

Sr. No	Patent No.	Backward Citation	Foreword Citation
1	CN105320646	7	2
2	CN105590032	0	2
3	CN106225788	5	1
4	CN108089572	4	1
5	CN106169115	3	1

Qualitative Analysis of Relevant Patent

This section further discusses briefly a qualitative analysis of important patents use today.

CN109067612 (A): An online traffic identification method based on incremental clustering algorithm

This patent provides an online traffic identification method based on an incremental clustering algorithm with good real-time performance. It comprises the following steps:

- Step 1. The offline identification phase extracts the pre-processed data set from the pre-prepared training data set to complete the pre-processing of the data set.
- Step 2. Perform a preliminary clustering and mapping work on the pre-processed data set by using the semi-supervised learning flow method based on the improved K-means algorithm to obtain a preliminary classified data set.
- Step 3. Perform real-time feature extraction on the newly added data stream to obtain incrementally clustered data objects.
- Step 4. Based on the formed data set that completes the clustering and mapping, the incremental clustering is used to determine the network application type of traffic, thereby achieving the purpose of traffic identification.
- Step 5. Periodically delete the data points that do not belong to the cluster group corresponding to the network application type to be identified or the data objects further from the center of each cluster, obtain a new data set, and go to step 3 to achieve continuous identification of online traffic.

The beneficial effects of the invention are:

The network traffic identification method based on incremental clustering is based on machine learning technology. By constructing a suitable recognition model to learn the prepared data, the online traffic can be incrementally clustered in real-time, combined with the prepared training set. The preliminary semi-supervised classification can realize online identification of network traffic, and is not subject to new network technologies such as dynamic port technology and encryption technology, and has excellent real-time performance and high recognition rate.



CN109145180 (A): Enterprise hot event mining method based on incremental clustering

The technical solution adopted by the present invention to solve the technical problem thereof is to provide an enterprise hotspot event mining method based on incremental clustering, which comprises the following steps:

- Step 1. Obtain corporate network news through web crawlers
- Step 2. Noise reduction and storage of the crawled corporate network news into the database
- Step 3. Read corporate news from the database for pre-processing
- Step 4. Cluster analysis of pre-processed corporate news
- Step 5. The center of mass news is extracted from the cluster of hot event clusters obtained by cluster analysis and stored in the database
- Step 6. Get the required enterprise hotspot events from the database for display

The beneficial effects of the invention are:

The present invention can provide enterprises or individuals with hot events generated by enterprises in the past, and the method based on incremental clustering can Extracting corporate hotspot events more quickly and accurately is an unsupervised method for automatically generating enterprise hotspot events, which is more efficient than manually searching, excluding, and selecting hotspot events.

CN106339495 (A): Topic detection method and system based on hierarchical incremental clustering

The invention obtains standard features of texts by high threshold incremental clustering, improves the accuracy of topic clustering, and reduces the program. Time complexity.

Comprising the following steps:

- Step 1. Obtaining text information, and dividing the text information by using a word segmenter;
- Step 2. Composing each of the words into a word sequence according to a segmentation order; respectively determining a word frequency of each of the words - inverting a file frequency and a hash An integer value;
- Step 3. The integer value constructs the text word vector, and determining whether a class center vector of a class is pre-stored, and calculating a class center vector in which a class is pre-stored a similarity between the text word vector and each class center vector;
- Step 4. Comparing each of the similarities with a preset categorization threshold;
- Step 5. If it is determined that there is a similarity greater than the categorization threshold, then The category corresponding to the similarity of the classification threshold is determined as a candidate set; the category corresponding to the maximum similarity in the candidate set is determined, and the text information is classified into the category corresponding to the maximum similarity;
- Step 6. Further, if it is determined that each of the similarities is less than or equal to the categorization threshold, the text information is determined to be a new category.

The beneficial effects of the invention are:



The invention obtains the common features of the text through high threshold incremental clustering, improves the accuracy of topic clustering, and reduces the time complexity of the program.

CN105740371 (A): Density-based incremental clustering data mining method and system

The aims of the present study are to solve the problem of low efficiency of data mining for dynamic data sets in the prior art.

The embodiment of the present invention is implemented in the following manner:

- Step 1. The original data set is clustered by using the DBSCAN algorithm to obtain data with class labels;
- Step 2. When the new data is added to the data of the class tag, the Incremental DBSCAN algorithm is used to perform incremental clustering processing on the tagged data;
- Step 3. The results of the two clustering processes are superimposed to form a final data mining result.

The beneficial effects of the invention are:

The technical solution provided by the invention can avoid waste of computing resources caused by repeated clustering, can improve the efficiency of incremental clustering, can enhance the timeliness of data mining, and improve the efficiency of data mining.

CN105320646 (A): Incremental clustering-based news topic mining method and apparatus thereof

The invention provides a news topic mining method and device based on incremental clustering, and the invention improves the accuracy of news mining, comprising the following steps:

- Step 1. Pre-processing the input text; extracting features from the pre-processed text to establish a text representation model;
- Step 2. The calculated text represents the similarity between the models, and the topic clustering is performed by the similarity;
- Step 3. Ranking the clustering results to obtain Chinese ranking results of hot topics;
- Step 4. Combine English to Chinese machine translation to get the English ranking results of hot topics;
- Step 5. Weight Chinese ranking results and English ranking results to get the final ranking of hot topics.

The beneficial effects of the invention are:

The broad application topic detection and tracking technology of the present invention, such as network public opinion information monitoring, internet financial analysis, network forum information monitoring and network information security, etc., can collect numerous sources from various information sources. The dazzling information, after analyzing and processing the collected information through data mining, can form a clear-cut, hot topic that attracts the public and identify emergencies, helping online news users solve information overload problems for the Internet regulatory department. The security decision-making provides information basis and is conducive to promoting the long-term development and progress of society.



An outlook of the study

This research work perhaps laid the firm groundwork that will prompt inventive, creative ability and enlighten the significance of an incremental clustering algorithm process to bring in the change through improvement in their processes. The present study implies a significant research gap is consequently providing critical space to analysts around the globe to concentrate research in real-time data analysis (Archana Chaudhari & Mulay, 2019), artificial intelligence, database systems, learning system, probability, and neural network. These domains have a more prominent scope and therefore needs to be investigated.

Conclusion

This present study revealed the detailed analysis of Intellectual-Property-Resources, particularly patents related to incremental clustering and incremental learning with a special focus on China's research during the period 2015-19. Incremental Clustering is the niche area of research related to Advanced Machine Learning. It is easy to carry out incremental learning via incremental clustering related to substantial data sets of varied domains in a collaborative form. The present researchers are motivated to study the patent landscape of China, the country of registration of the application as it is emerging in the field of incremental learning via incremental clustering. The patent database Espacenet is a rich source of patent information and various functionalities, including translation framework, is being utilised in this study. The analysis reveals that the patent filings are increasing in the last 5 years. In the year 2016 the highest number of patents were filed for the selected theme of the present work. The publications are also increasing gradually, and in the year 2018 it was highest among the last 5 years. The grant of the patents are also increasing, and in the year 2018 was a most successful year for various assignees as in this year the patents are granted most as compared to the whole data of last 5 years. Another interesting finding of the present work is mostly the Chinese Universities and filing more patents collectively. However, INTEL is leading among all. The trends suggest that there is an excellent scope for a non-Chinese organisation to file patents in the field of ICA. Since the datasets are very recent, the citation counts will go up with the years, and many more new patents lead towards economic development. Knowledge investment in the field of intellectual property will bring more prosperity.

ACKNOWLEDGEMENT

The authors appreciate the reviewers for their contributions towards improving the quality of this research.

CONFLICT OF INTEREST DISCLOSURE

All authors declare that they have no conflicts of interest to disclose.

REFERENCES

- (WIPO), W. I. P. O. International Patent Classification (IPC). Retrieved from <https://www.wipo.int/classifications/ipc/en/>
- (WIPO), W. I. P. O. (2019). WIPO Technology Trends 2019 - Artificial Intelligence. *WIPO Technology Trends 2019 - Artificial Intelligence*. Retrieved from https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf



- Archana Chaudhari, & Mulay, P. (2019). SCSi: Real-Time Data Analysis with Cassandra and Spark. In M. Mittal, V. E. Balas, L. M. Goyal, & R. Kumar (Eds.), *Big Data Processing Using Spark in Cloud* (pp. 237-264). Singapore: Springer Singapore.
- Aristodemou, L., & Tietze, F. (2018). The state-of-the-art on Intellectual Property Analytics (IPA): A literature review on artificial intelligence, machine learning and deep learning methods for analysing intellectual property (IP) data. *World Patent Information*, 55, 37-51. doi:<https://doi.org/10.1016/j.wpi.2018.07.002>
- Chaudhari, A., Joshi, R. R., Mulay, P., Kotecha, K., & Kulkarni, P. (2019). Bibliometric Survey on Incremental Clustering Algorithms. *Library Philosophy and Practice*, 1-23. doi:<https://digitalcommons.unl.edu/libphilprac/2762>
- Chaudhari, A., & Mulay, P. (2019a). Algorithmic analysis of intelligent electricity meter data for reduction of energy consumption and carbon emission. *The Electricity Journal*, 32(10), 106674. doi:<https://doi.org/10.1016/j.tej.2019.106674>
- Chaudhari, A., & Mulay, P. (2019b). A bibliometric survey on incremental clustering algorithm for electricity smart meter data analysis. *Iran Journal of Computer Science*. doi:10.1007/s42044-019-00043-0
- Chaudhari, A. Y., & Mulay, P. (2020). Cloud4NFICA-Nearness Factor-Based Incremental Clustering Algorithm Using Microsoft Azure for the Analysis of Intelligent Meter Data. *International Journal of Information Retrieval Research (IJIRR)*, 10(2), 21-39.
- database, W. s. (2018). Statistical Country Profiles China. Retrieved from https://www.wipo.int/ipstats/en/statistics/country_profile/profile.jsp?code=CN
- Dr. Preeti Mulay, & Shinde, K. (2017). CBICA: Correlation Based Incremental Clustering Algorithm. *IJCTA*, 10(9), 575-582. Retrieved from <https://www.semanticscholar.org/paper/Cbica%3A-Correlation-based-incremental-clustering-a-Shinde-Mulay/49156b7374fc017ce495c5c140a72567a97bfe62>
- Henry, M. (February 28, 2019). ARTIFICIAL INTELLIGENCE: AN OVERVIEW OF THE PATENT LANDSCAPE *Henry patent law firm* (Vol. 2019). TEXAS.
- Indicators, W. I. P. (2018). World Intellectual Property Indicators 2018. Retrieved from https://www.wipo.int/edocs/pubdocs/en/wipo_pub_941_2018.pdf
- Kuralkar, S., Mulay, P., & Chaudhari, A. (2020). Smart Energy Meter: Applications, Bibliometric Reviews and Future Research Directions. *Science & Technology Libraries*, 39(2), 165-188. doi:10.1080/0194262X.2020.1750081
- Martin Ester, Hans-Peter Kriegel, Jörg Sander, & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise *KDD-96 Proceedings*, 226-231. Retrieved from accessed on <https://dl.acm.org/citation.cfm?id=3001507>
- Mulay, P., Joshi, R., & Chaudhari, A. (2018, 30 Nov.-2 Dec. 2018). *Mapping of Six Sigma to Threshold Based Incremental Clustering Algorithm*. Paper presented at the 2018 IEEE Punecon, Pune, India, India.
- Mulay, P., Joshi, R., & Chaudhari, A. (2020). Distributed Incremental Clustering Algorithms: A Bibliometric and Word-Cloud Review Analysis. *Science & Technology Libraries*, 39(3), 289-306. doi:10.1080/0194262X.2020.1775163
- P. Mulay, & Kulkarni, P. A. (2013a). Evolving Systems using incremental clustering approach. *Evolving Systems*, 4, 70-85. doi:10.1007/s12530-012-9068-z.
- P. Mulay, & Kulkarni, P. A. (2013b). Knowledge augmentation via incremental clustering: new technology for effective knowledge management. *International Journal Business Information Systems*, 12(1), 68-87. Retrieved from <https://dl.acm.org/citation.cfm?id=2407446>
- Preeti Mulay, Joshi, R. R., Anguria, A. K., Gonsalves, A., Deepankar, D., & Ghosh, D. (2017). *Threshold Based Clustering Algorithm Analyzes Diabetic Mellitus*. Paper presented at the In Proceedings of



- the 5th International Conference on Frontiers in Intelligent Computing: Theory and Applications Singapore.
- Raturi, M. K., Sahoo, P. K., Mukherjee, S., & Tiwari, A. K. (2010). Patinformatics – An Emerging Scientific Discipline Available at SSRN: <https://ssrn.com/abstract=1566067>. Retrieved from <http://dx.doi.org/10.2139/ssrn.1566067>
- T. Zhang, R. Ramakrishnan, & Livny, M. (1997). BIRCH: A New Data Clustering Algorithm and Its Applications. *Data Mining and Knowledge Discovery*, 1(2), 141-182. doi:<https://doi.org/10.1023/A:100978382>
- Vincent, C. L., Singh, V., Chakraborty, K., & Gopalakrishnan, A. (2017). Patent data mining in fisheries sector: An analysis using Questel-Orbit and Espacenet. *World Patent Information*, 51, 22-30. doi:<https://doi.org/10.1016/j.wpi.2017.11.004>
- Xue Lan, L. Z., Dai Yixin, Deng Xinghua, Li Daitian, Yu Zhen, Yang Fangjuan. (2018). *China's AI Development Report 2018*. Retrieved from China: http://www.sppm.tsinghua.edu.cn/eWebEditor/UploadFile/China_AI_development_report_2018.pdf
- Zhu, W., Yu, W., Kan, B., & Liu, G. (2018). *Smart Meter Data Analytics Based on Modified Streaming k-Means*. Paper presented at the 2018 IEEE 12th International Conference on Compatibility Power Electronics and Power Engineering (CPE-POWERENG) Chengdu, China. <https://ieeexplore.ieee.org/document/8113085>

