Computers & Industrial Engineering 110 (2017) 395–403

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

Short Communication

A bibliometric analysis of Genetic Algorithms throughout the history

Son Duy Dao *, Kazem Abhary, Romeo Marian

School of Engineering, University of South Australia, Mawson Lakes Campus, Mawson Lakes, SA 5095, Australia

ARTICLE INFO

Article history:
Received 2 April 2017
Accepted 6 June 2017
Available online 19 June 2017

Keywords:
Genetic algorithms
Bibliometric analysis
Publication statistics
Survey

ABSTRACT

In this article, a bibliometric analysis of Genetic Algorithms (GA) throughout the history is conducted. A big picture of publications associated with GA is created. A number of dominant statistics of GA publications by years, research fields, document types, source titles, countries, institutions and authors are provided herein. In addition, some insights as well as future perspectives of publications associated with GA are discussed.

1. Introduction

Genetic Algorithms (GA) is a popular solution method, often used to optimise solutions to problems in many fields such as engineering, computer science, economic management, and supply chain management (Aguilar-Rivera, Valenzuela-Rendón, & Rodríguez-Ortiz, 2015; Cheng & Chang, 2007; Lee, Kim, & Joo, 2012; Su, Huang, Fan, & Mak, 2015). GA has a number of advantageous characteristics compared to other meta-heuristics such as flexibility in defining constraints as well as quality measures, capability of working with both continuous and discrete variables, capability of handling large search space, capability of providing multiple optimal/good solutions, and great potential for applying parallel computing techniques to shorten the processing time (Fahimnia, Luong, & Marian, 2008). Since being introduced over four decades ago, significant developments in various components of GA such as chromosome encoding (Dao, Abhary, & Marian, 2014; Zhong & Chen, 2002), crossover (Qing-dao-er-ji & Wang, 2012; Wang & Zheng, 2002), mutation (Tang & Tseng, 2013; Wang, Yin, & Wang, 2009), evaluation (Chang, Hsieh, & Wang, 2007; Hyun, Kim, & Kim, 1998), selection (Stern, Chassidim, & Zofi, 2006; Wu, Chu, Wang, & Yan, 2007) as well as algorithm structure (Dao, Abhary, & Marian, 2015; Zhou, Zheng, Yan, & Wang, 2011) have been made. These developments make today's GA much more powerful than the traditional GA. It should be noted that GA is a searching philosophy only; there is no standard GA. Generally speaking, when using GA, users must customise some GA components, e.g. chromosome encoding, crossover or mutation. Due to the increasing computing power, application of GA has been expanded to many different fields. It would be interesting to see the big picture of the developments of GA throughout its history.

In this article, a bibliometric analysis of the developments of GA throughout its history is conducted. First, the source of statistical data is described and then statistical data as well as analysis of publication quantity associated with GA, published by year, field, document type, source title, country, university and author, are presented. Finally, discussion and conclusion are given.

2. Source of statistical data

Scopus, managed by Elsevier, is the largest abstract and citation database of peer-reviewed journals, books and conference proceedings in the fields of science, technology, medicine, social sciences, arts as well as humanities (Scopus, 2015). In this article, the Scopus database is used to do the bibliometric analysis of the developments of GA throughout its history. With the search engine "*Title, Abstract, Keywords*" in the Scopus database, all documents associated with GA, published from the beginning of GA history to 2014, are searched. Statistical data and analysis related to different aspects of publications associated with GA will be presented in the subsequent Sections.

3. Latest statistics of publications associated with Genetic Algorithms

3.1. Statistics by year

With the keywords "Genetic Algorithm" OR "Genetic Algorithms", 124,799 publications associated with GA were found

* Corresponding author at: Room M2-14, Building M, Mawson Lakes Campus, University of South Australia, Mawson Lakes, SA 5095, Australia.
E-mail address: son.dao@mymail.unisa.edu.au (S.D. Dao).

http://dx.doi.org/10.1016/j.cie.2017.06.009
0360-8352/

in the Scopus database (accessed on 12 November 2015). The trend in quantity of the publications is shown in Fig. 1. As can be seen from Fig. 1, the first documents associated with GA were published in 1972 and the quantity of the publications has been remarkably increasing since 1992. In 2014 alone, 11,020 documents associated with GA were published. From 1992 to 2014, on average, there was an increase of 475.7 publications per year; the biggest increase with 1774 publications happened in 2004 and there was a decrease of 147 publications in 2012. The annual increase of publications associated with GA from 1992 to 2014 is shown in Fig. 2. The trends in Figs. 1 and 2 clearly show that the number of publications associated with GA will continue to increase in the future.

3.2. Statistics by field

Majority of 124,799 publications associated with GA, found in Section 3.1, are in the fields of engineering, computer science and mathematics. The publications in different fields are detailed in Fig. 3. As can be seen in Fig. 3, there are 66,366, 54,241 and 21,230 publications associated with GA in engineering, computer science and mathematics, respectively. In other fields, each has less than 9000 publications. In terms of percentage, as shown in Fig. 4 and 53.2% of the publications are in the field of engineering; while 43.5% and 17.0% are in computer science and mathematics, respectively. It is noted that the summation of all values in Fig. 4 exceeds 100% because one publication may be classified in more than one field.

3.3. Statistics by document type

Quantity of publications associated with GA in different document types, e.g. journal article, conference paper, book, book chapter, note, report, etc., is shown in Table 1. As can be seen, majority of the publications are journal articles and conference papers. More specifically, there are 63,797 journal articles and 56,093 conference papers associated with GA in Scopus database. Following the article and conference paper, it is conference review (critical response to a conference) with 2054 documents. For other types of publications like book, book chapter, review, editorial, letter, note, and abstract report, each has less than 1300 documents as shown in Table 1. More than half (51.1%) of 124,799 publications associated with GA in Scopus database are journal articles, followed by conference papers with 44.9%, as shown in Fig. 5. In addition, the trends of the conference and journal papers

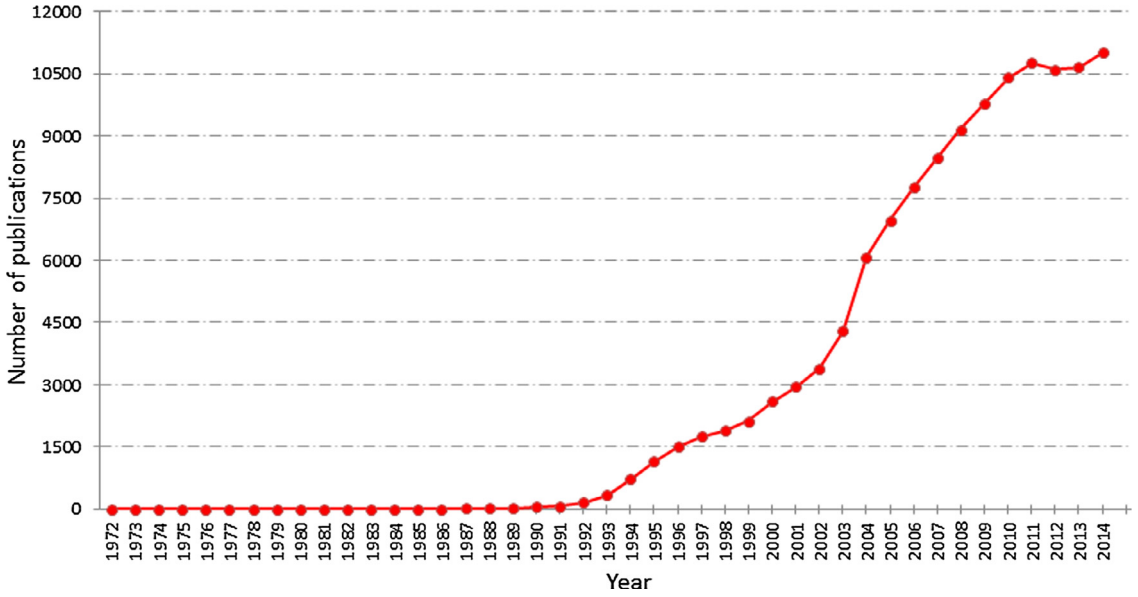

Fig. 1. Number of publications associated with GA.

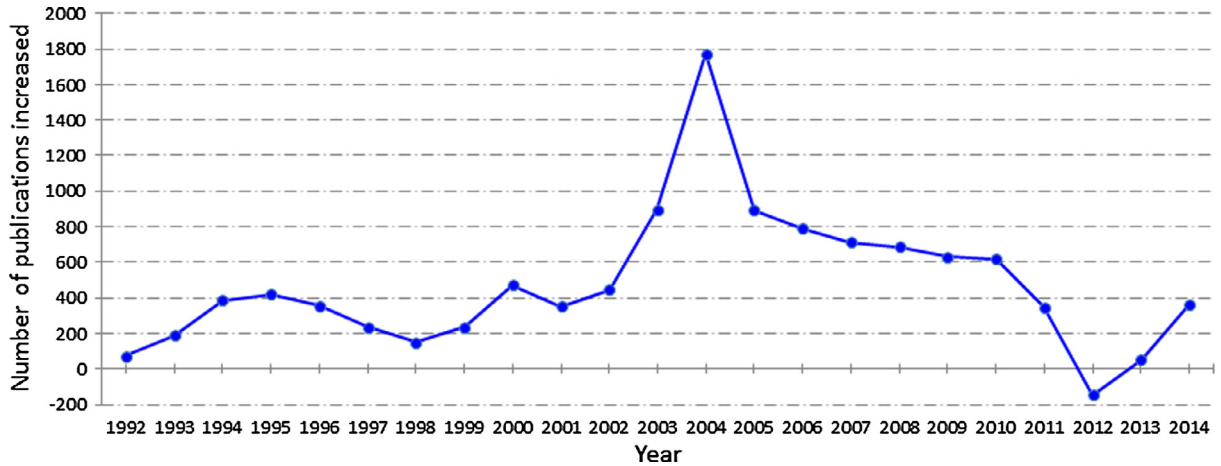

Fig. 2. Annual increase of publications associated with GA.

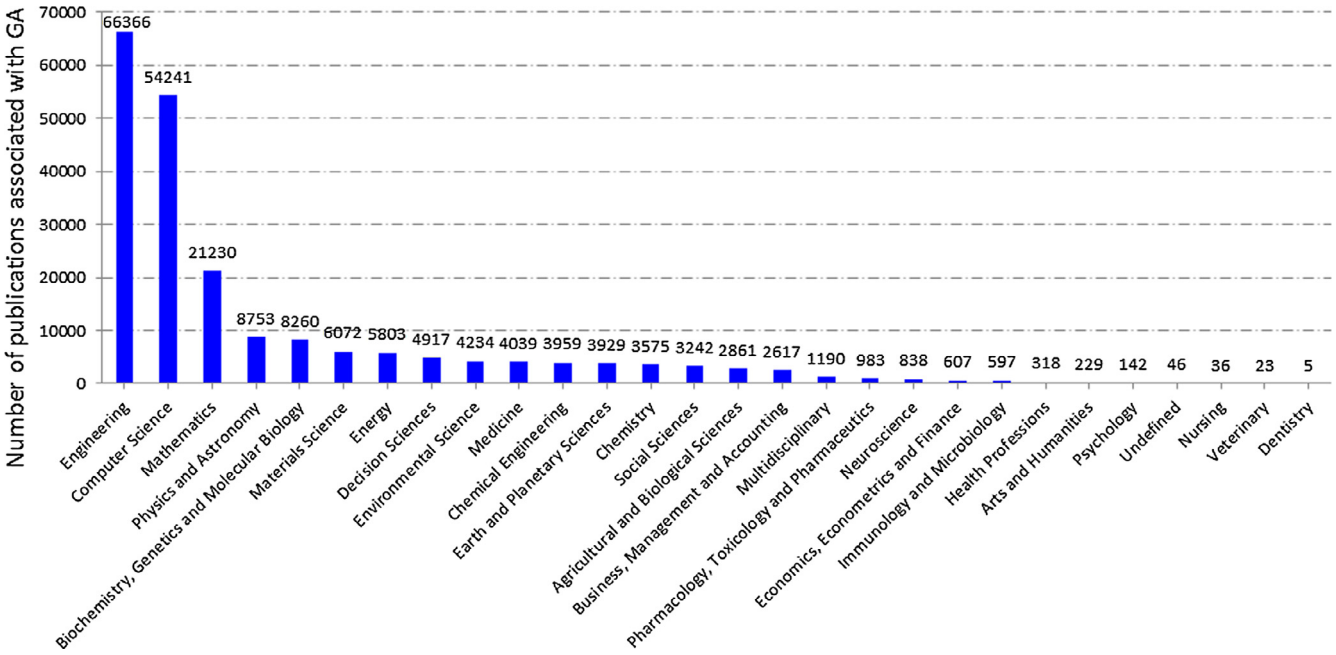

Fig. 3. Number of publications associated with GA in different fields.

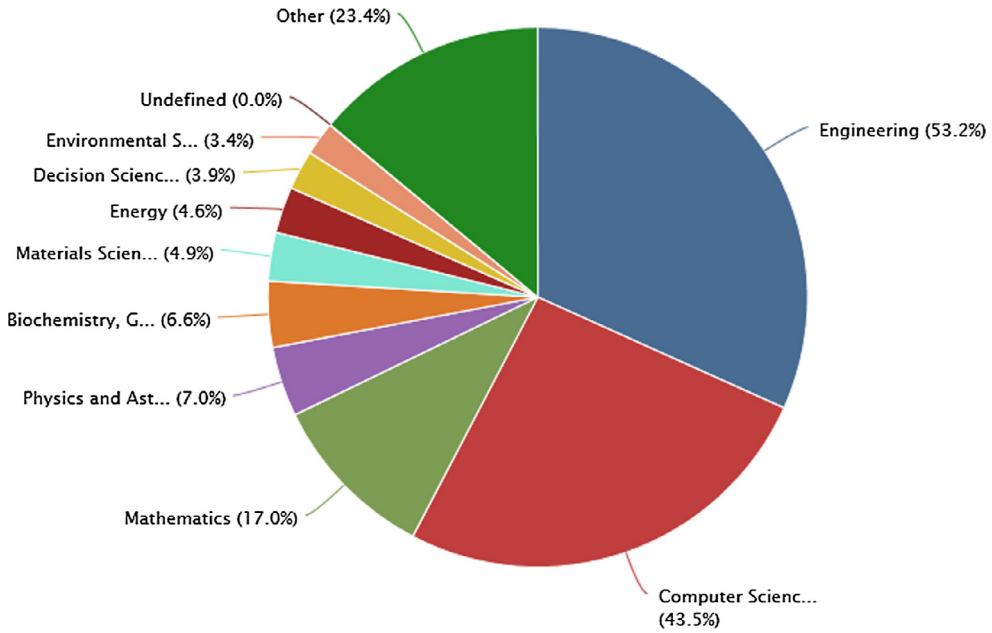

Fig. 4. Percentage of publications associated with GA in different fields.

published over the years are shown in Fig. 6. Obviously, up to 2009, annual quantities of both conference and journal papers associated with GA increased continuously. In 2010, number of journal articles slightly decreased, while number of conference papers continued to increase. From 2011 to 2014, the trends of journal and conference publications were opposite, going up and down, respectively.

3.4. Statistics by source title

The top ten source titles publishing documents associated with GA in Scopus database are shown in Table 2 where it can be seen that seven out of the ten source titles are conference proceedings, book series, lecture notes or edited volumes. Only three out of the ten are standard journals, i.e. Expert Systems with Applications, Bioinformatics and International Journal of Advanced Manufacturing Technology. It should be noted that Applied Mechanics and Materials, and Advanced Materials Research are not considered as standard journals, since they only publish complete volumes on given topics, proceedings and complete special topic volumes, and do not publish stand-alone papers by individual authors (TTP, 2015a, 2015b). In addition, Studies in Computational Intelligence as well as Communications in Computer and Information Science are two book series published by Springer.

Table 1
Number of publications associated with GA in different document types.

Document type	Quantity
Article	63,797
Conference paper	56,093
Conference review	2054
Review	1252
Book chapter	602
Article in press	352
Editorial	144
Letter	121
Note	105
Book	92
Erratum	71
Short survey	51
Undefined	50
Report	11
Abstract report	4
Total	124,799

Table 2
Top ten source titles published documents associated with GA.

No	Source	Quantity
1	Applied Mechanics and Materials	1446
2	Proceedings of SPIE the International Society for Optical Engineering	1446
3	Advanced Materials Research	1339
4	Expert Systems with Applications	874
5	Bioinformatics	712
6	International Journal of Advanced Manufacturing Technology	634
7	Proceedings of the World Congress on Intelligent Control and Automation WCICA	627
8	Proceedings of the IEEE International Conference on Systems Man and Cybernetics	618
9	Studies in Computational Intelligence	501
10	Communications in Computer and Information Science	478

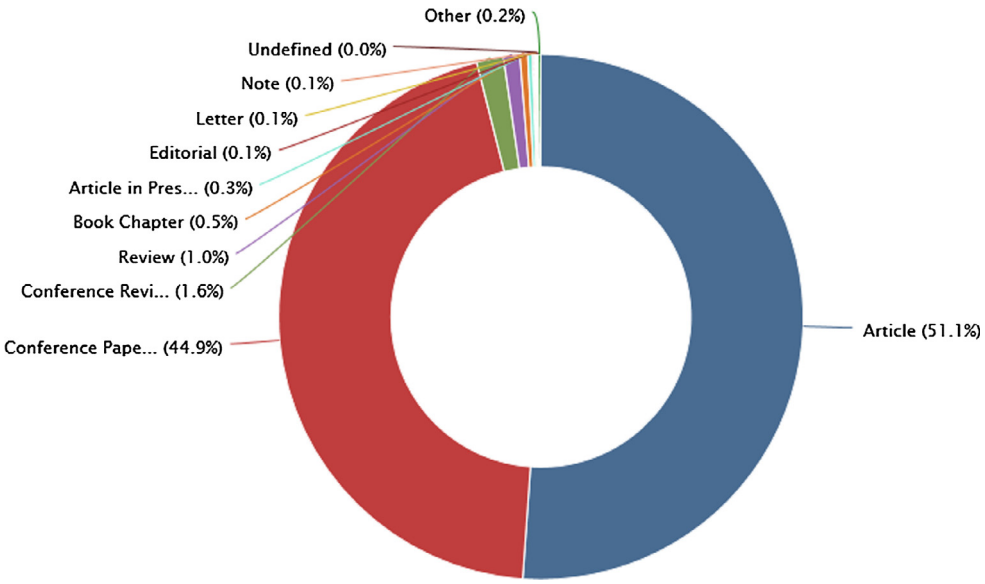

Fig. 5. Percentage of publications associated with GA in different document types.

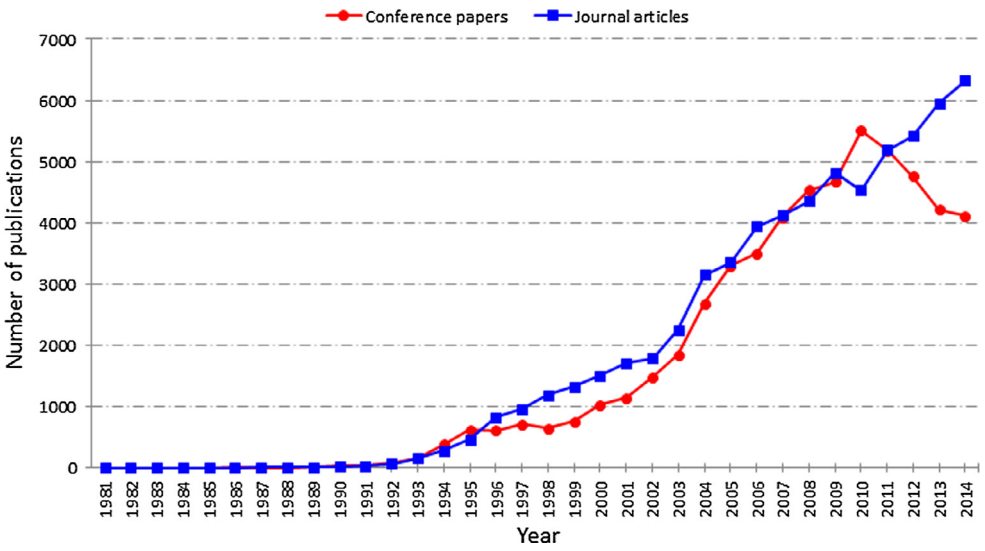

Fig. 6. Conference and journal papers associated with GA published over the years.

Statistics in Table 2 shows that for the top two source titles, namely Applied Mechanics and Materials, and Proceedings of SPIE the International Society for Optical Engineering, each has published 1446 documents associated with GA. Following the two source titles, it is Advanced Materials Research, with 1339 documents. The journal that has published largest number of articles associated with GA is Expert Systems with Applications, with 874 articles. The other two journals, namely Bioinformatics and International Journal of Advanced Manufacturing Technology, have published 712 and 634 articles, respectively. For the rest of the source titles, each has published less than 630 documents.

Annual quantity of documents associated with GA, published by different source title, is visualised in Fig. 7. Overall, the annual quantities increased over time, except the quantity published by Bioinformatics. After 2010, number of the documents associated with GA, published by Bioinformatics, significantly decreased. In contrast, the quantities of the documents, published by Applied Mechanics and Materials, and Advanced Materials Research, remarkably increased. Fig. 7 clearly shows that two source titles, i.e. Applied Mechanics and Materials, and Advanced Materials Research, have dominated other source titles in the last five years.

3.5. Statistics by country

The top 15 countries publishing documents associated with GA are shown in Fig. 8. China has published the largest number of documents associated with GA, followed by United States. More specifically, China has published 35,105, while United States has published 17,850 documents associated with GA. Following China and United States, there are India, Japan and United Kingdom, with 7527, 6923 and 6753 publications respectively. The rest of the top 15 countries are Iran, Taiwan, Canada, Spain, South Korea, Italy, France, Germany, Brazil and Australia, with numbers of publications associated with GA as shown in Fig. 8. It can be seen that China and United States dominate the other countries in terms of publishing research documents associated with GA. More interestingly, the quantity published by China is nearly twice as many as the quantity published by United States.

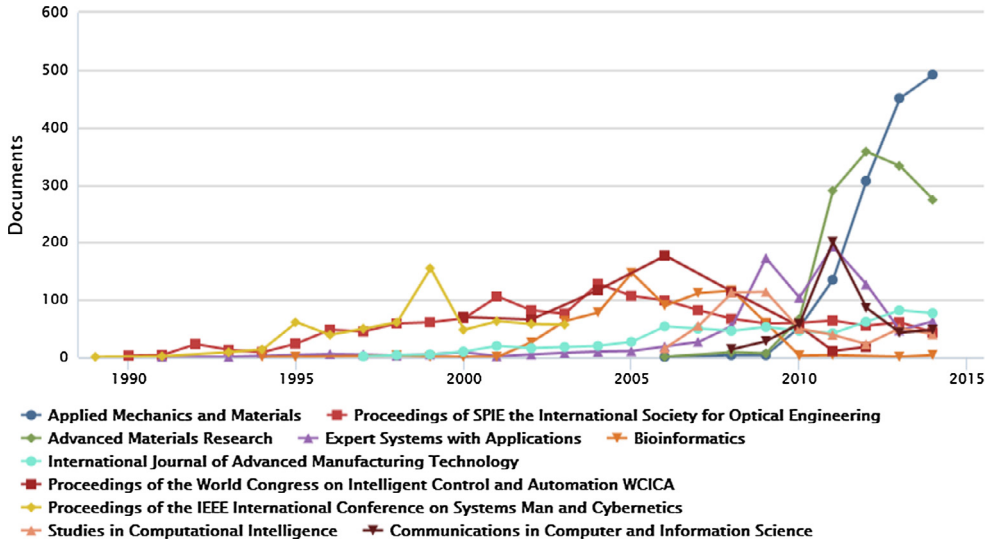

Fig. 7. Annual quantity of documents associated with GA published by top ten source title.

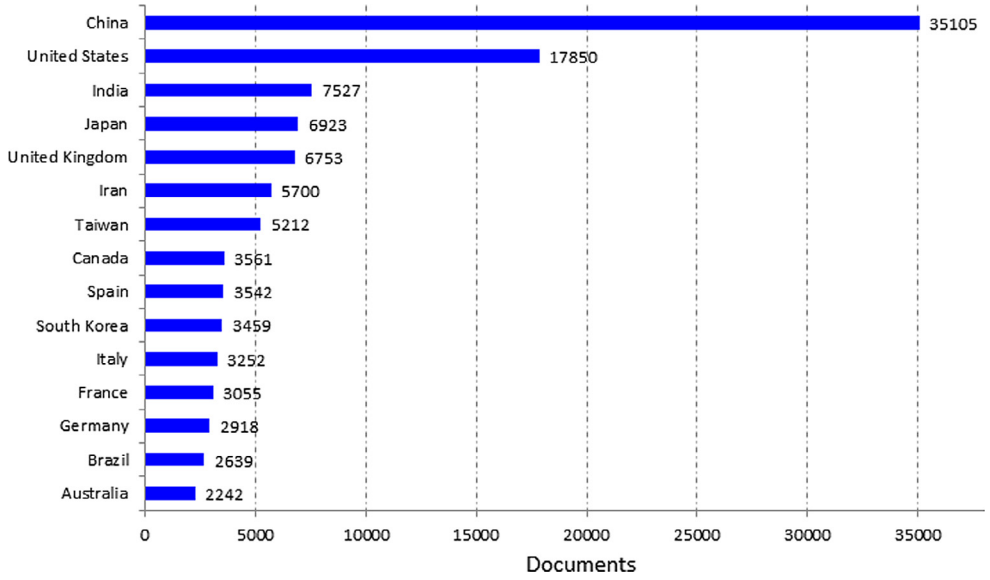

Fig. 8. Top 15 countries publishing documents associated with GA.

3.6. Statistics by institution

The top five institutions that have published the largest numbers of research documents associated with GA are shown in Table 3. As can be seen from Table 3, Tsinghua University is the world leading institution in terms of publishing research documents associated with GA. Tsinghua University has published 1214 research documents associated with GA. Following Tsinghua University, it is Shanghai Jiaotong University, with 1105 publications. The other three institutions, i.e. Harbin Institute of Technology, Zhejiang University and Islamic Azad University, have published 1045, 963 and 952 research documents associated with GA, respectively. Interestingly, the top four institutions are based in China and the fifth institution is located in Iran.

3.7. Statistics by author

The top 15 authors who have published the largest numbers of research documents associated with GA are shown in Fig. 9. Mitsuo Gen, from Fuzzy Logic Systems Institute, Iizuka, Japan, is the author with the largest number of publications associated with GA, 219 publications. Following Mitsuo Gen, there are Castillo, O. (from Tijuana Institute of Technology, Tijuana, Mexico) and Pedrycz, W. (from Systems Research Institute of the Polish Academy of Sciences, Warsaw, Poland), who have published 154 and 153 documents, respectively. Other dominant authors with more than 100 publications associated with GA are Werner, D.H. (from Pennsylvania State University, Department of Electrical Engineering, State College, United States), Melin, P. (from Tijuana Institute of Technology, Tijuana, Mexico), Goldberg, D.E. (from University of Illinois at Urbana-Champaign, Urbana, United States), Jiao, L. (from Xidian University, International Research Center for Intelligent Perception and Computation, Xi'an, China), Oh, S.K. (from Suwon University, Department of Electrical Engineering, Hwasong, South Korea), Deb, K. (from Michigan State University, East Lansing, United States), Abraham, A. (from Machine Intelligence Research Labs - Scientific Network for Innovation and Research Excellence, Auburn, United States) and Alba, E. (from Universidad de Malaga, Departamento de Lenguajes y Ciencias de la Computación, Malaga, Spain). For the last four authors in the top 15, i.e. Herrera, F. (from Universidad de Granada, Department of Computer Science and Artificial Intelligence, Granada, Spain), Kubota, N. (from Tokyo Metropolitan University, Faculty of System Design, Hachioji, Japan), Bandyopadhyay, S. (from Indian Statistical Institute, Kolkata, Kolkata, India) and Fukuda, T. (from Beijing Institute of Technology, Intelligent Robotics Institute, Beijing, China), each has published nearly 100 documents associated with GA.

It is interesting to note that United States has four authors in the top 15 authors who have published the largest numbers of research documents associated with GA, followed by Japan (two), China (two), Spain (two), Mexico (two), Poland (one), South Korea (one) and India (one). Obviously, Mitsuo Gen, from Fuzzy Logic Systems Institute, Iizuka, Japan, is the world leading author in terms of quantity of publications associated with GA.

Table 3
Top five institutions publishing research documents associated with GA.

No	Institution	Quantity
1	Tsinghua University	1214
2	Shanghai Jiaotong University	1105
3	Harbin Institute of Technology	1045
4	Zhejiang University	963
5	Islamic Azad University	952

4. Discussions

GA is a universal optimisation solution method, with the history of about four decades. As can be seen from Section 3.1, a significant number of research documents associated with GA (124,799 documents) have been published, up to 2014. Majority of these publications are in the fields of engineering and computer science, as shown in Figs. 3 and 4. In addition, Fig. 5 indicates that most of the publications are journal articles or conference papers. Number of publications associated with GA has significantly increased in the last 10 years, as indicated in Fig. 1. Based on the trend of the publications in Fig. 1, it can be predicted that the number of publications associated with GA will continue to increase in the future.

Since journal publications generally have greater impact than other kinds of publications like conference papers, book chapters,

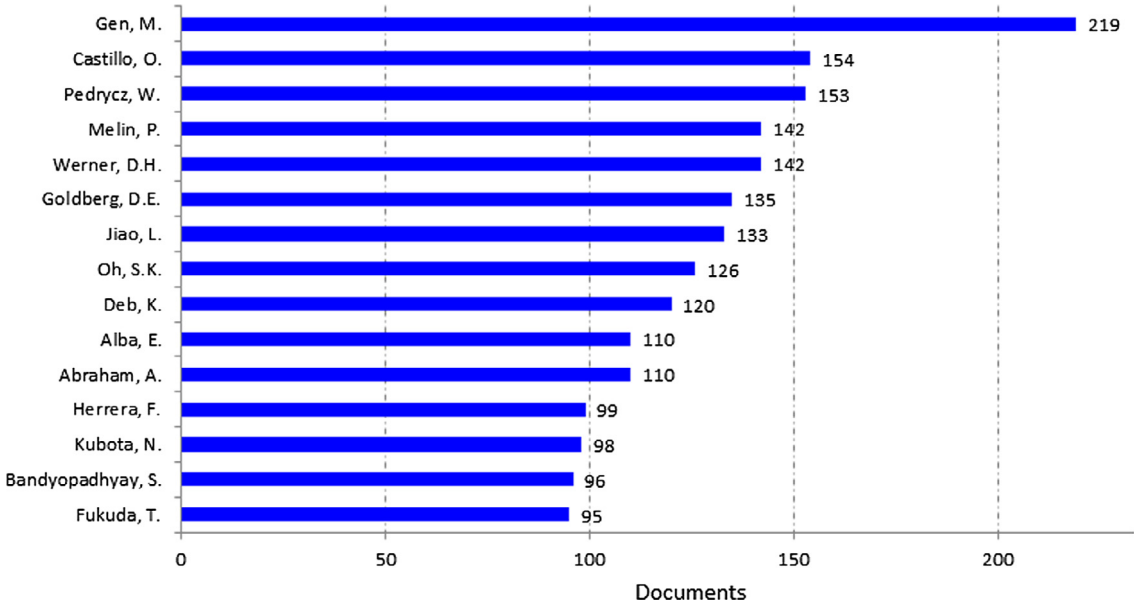

Fig. 9. Top 15 authors with the largest numbers of publications associated with GA.

letters, notes, etc., the rest of the discussions will address different aspects of journal publications only. The top three journals that have published the largest numbers of research documents associated with GA are Expert Systems with Applications, Bioinformatics, and International Journal of Advanced Manufacturing Technology, as indicated in Table 2. The annual number of research articles associated with GA, published by the top three journals, are visualised in Fig. 10. As can be seen from Fig. 10, the overall trends of number of articles associated with GA, published by Expert Systems with Applications and International Journal of Advanced Manufacturing Technology, have been going up. However, this trend of Bioinformatics is quite different, going up before 2010 and then going down significantly. Finding out the root cause of the sudden decrease of quantity of articles associated with GA, published by Bioinformatics in the last 5 years, is beyond the scope of this article.

Going further into journal publications associated with GA, the annual publications by top three countries are discussed. As can be seen from Fig. 8, the top three countries that have published the largest numbers of research documents associated with GA are China, United States and India. The annual journal publications by the top three countries are shown in Fig. 11. It can be seen that overall trends of the annual journal publications have gone up; however the publication rate of China has gone up much faster than those of United States and India in the last decade. In addition, the journal publication rate of India has continually increased during the last decade, while the rate of United States has been stable, at about 600 articles per year. Interestingly, India has kept up with United States since 2014, with about the same number of journal publications associated with GA, as indicated in Fig. 11. Nevertheless, United States and India were still far behind China in the number of journal publications associated with GA in 2014, about 600 compared to more than 2100 publications. Based on the trends in Fig. 11, it can be predicted that the numbers of journal articles associated with GA, published by top three countries: China, United States and India, will continue to increase in the future, especially China.

As indicated in Table 3, top three institutions that have published the largest numbers of research documents associated with GA are Tsinghua University, Shanghai Jiaotong University and Harbin Institute of Technology. Numbers of journal articles associated with GA, annually published by the three institutions, are shown in Fig. 12. Generally speaking, patterns of the journal publication rates are quite similar, going up before 2006, then going down, and finally being stable after 2010. The peaks of the annual publication rates of the top three institutions were in the period of 2004–2008, with about 70 articles; and since 2010 the publications rates have been roughly stable at about 45 articles per year. It

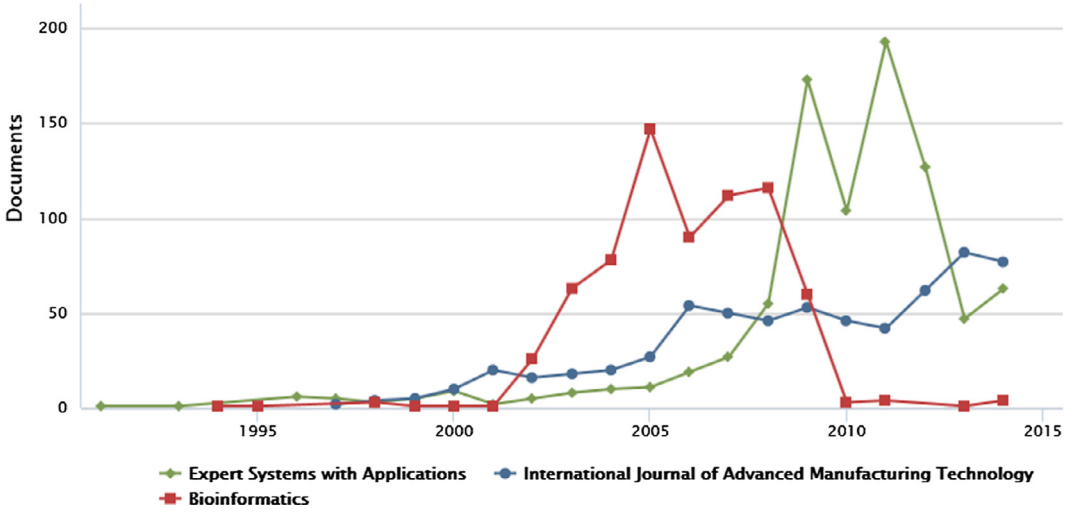

Fig. 10. Annual quantities of articles associated with GA published by top three journals.

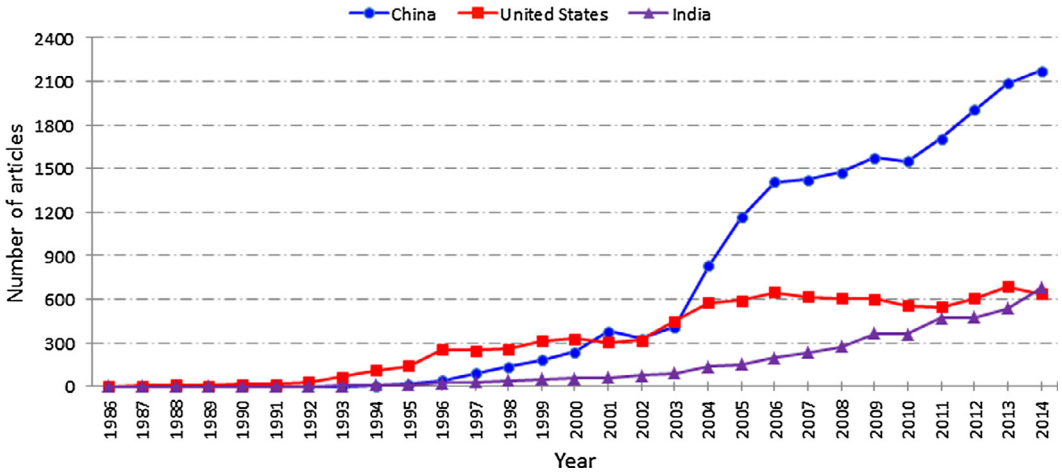

Fig. 11. Annual journal publications associated with GA by top three countries.

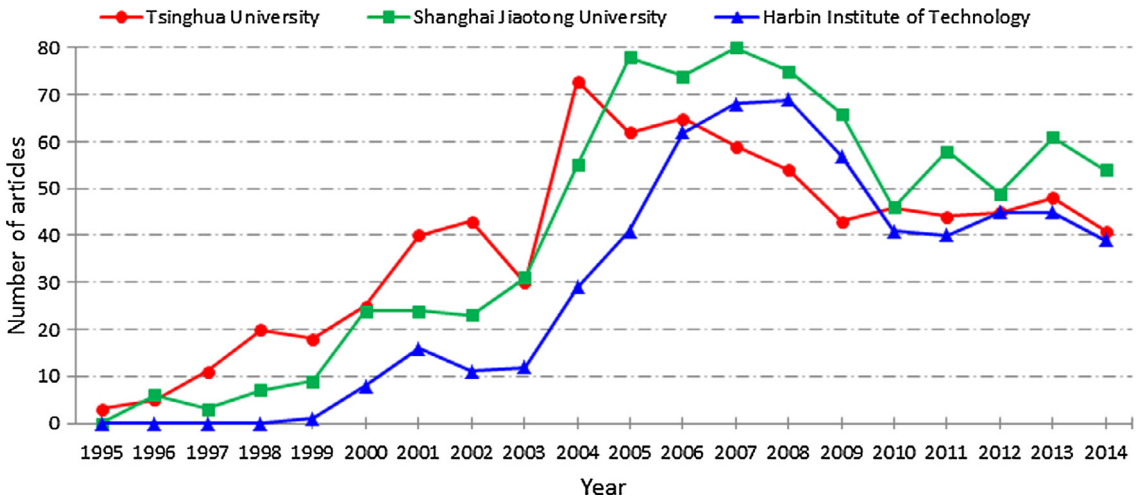

Fig. 12. Annual journal publications associated with GA by top three institutions.

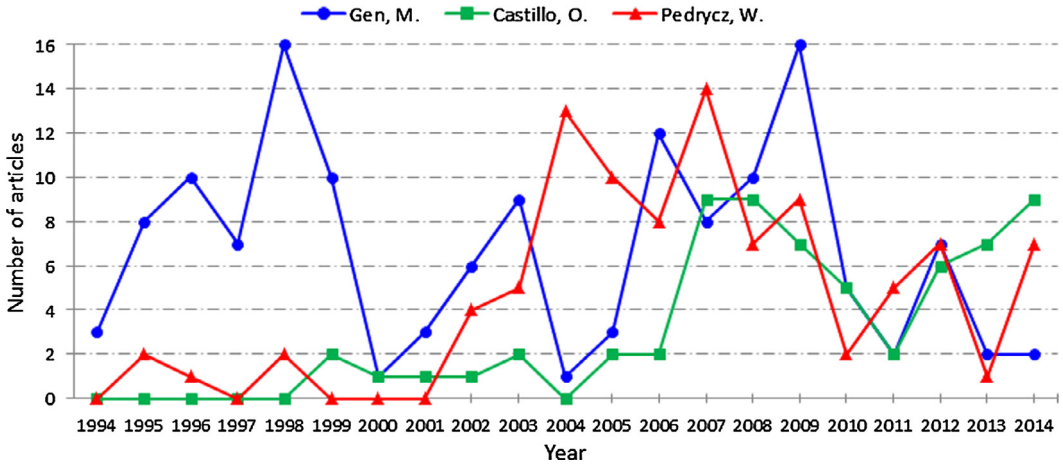

Fig. 13. Annual journal publications associated with GA by top three authors.

should be noted that Shanghai Jiaotong University has been the world leading institution in terms of number of journal publications associated with GA since 2005. Based on the trends in Fig. 12, it can be predicted that in the future, the annual journal publications associated with GA, by the top three institutions, will be stable as in the last five years.

Finally, journal publication rates of some top authors are discussed. Annual numbers of journal articles associated with GA, published by the top three authors, are shown in Fig. 13. As can be seen from Fig. 13, Gen, M. has been being active in publishing journal articles associated with GA since 1994; while Castillo, O. and Pedrycz, W. have been being active since 1999 and 1995, respectively. Up to 2014, Gen, M. has published 141 journal articles associated with GA; while those numbers of Castillo, O. and Pedrycz, W. are 65 and 97, respectively. The author with the most impressive publication rate is Gen, M., with, on average, 6.7 journal articles associated with GA per year, followed by Pedrycz, W. with 4.9 and then Castillo, O. with 4.1 publications. More interestingly, in 1998 and 2009, each year, Gen, M. published 16 journal articles associated with GA. The two most productive years of Pedrycz, W. were 2004 and 2007, with 13 and 14 journal publications associated with GA, respectively. The trends in Fig. 13 indicate that Castillo, O. has been becoming more and more active recently, with 9 journal publications associated with GA in each of the following years: 2007, 2008 and 2014.

5. Conclusions

In this article, a big picture of publications associated with GA has been created. With Scopus, the largest abstract and citation database of peer-reviewed publications, 124,799 research documents associated with GA, published from the beginning of GA history to 2014, have been found. The first documents associated with GA were published in 1972 and the number of the publications has been remarkably increasing since 1992. From 1992 to 2014, on average, there was an increase of 475.7 publications per year. In 2014 alone, 11,020 research documents associated with GA were published. Majority of 124,799 publications associated with GA are in the fields of engineering (66,366 documents), computer science (54,241 documents) and mathematics (21,230 documents). In addition, majority of the publications are journal articles (51.1%) and conference papers (44.9%). The journal that has published largest number of articles associated with GA is Expert Systems with Applications (874 articles), followed by Bioinformatics with 712 articles. China has published the largest number of documents associated with GA (35,105 publications), followed by United States (17,850 publications). Tsinghua University is the world leading institution in terms of publishing research documents associated with GA (1214 documents), followed by Tsinghua University (1105 publications). Gen, M. (from Japan) is the author with the largest number of publications associated with GA (219 publications),

followed by Castillo, O. (from Mexico) with 154 documents. Moreover, annual journal publications associated with GA, published by (1) top three journals, (2) top three countries, (3) top three institutions and (4) top three authors, have been discussed. Finally, the future perspectives regarding some aspects of publications associated with GA have been predicted.

In future work, the authors would focus on more detailed aspects of GA publications, e.g. developments of chromosome encoding, crossover, mutation, evaluation, selection as well as hybrid GA.

References

Aguilar-Rivera, R., Valenzuela-Rendón, M., & Rodríguez-Ortiz, J. J. (2015). Genetic algorithms and Darwinian approaches in financial applications: A survey. *Expert Systems with Applications, 42*(21), 7684–7697.

Chang, P. C., Hsieh, J. C., & Wang, C. Y. (2007). Adaptive multi-objective genetic algorithms for scheduling of drilling operation in printed circuit board industry. *Applied Soft Computing, 7*(3), 800–806.

Cheng, B. W., & Chang, C. L. (2007). A study on flowshop scheduling problem combining Taguchi experimental design and genetic algorithm. *Expert Systems with Applications, 32*(2), 415–421.

Dao, S. D., Abhary, K., & Marian, R. (2014). Optimisation of partner selection and collaborative transportation scheduling in Virtual Enterprises using GA. *Expert Systems with Applications, 41*(15), 6701–6717.

Dao, S. D., Abhary, K., & Marian, R. (2015). An adaptive restarting genetic algorithm for global optimization. In *Proceedings of the world congress on engineering and computer science, October 21–23, 2015, San Francisco, USA* (pp. 455–459).

Fahimnia, B., Luong, L., & Marian, R. (2008). Optimization/simulation modeling of the integrated production-distribution plan: an innovative survey. *WSEAS Transactions on Business and Economics, 3*(5), 52–65.

Hyun, C. J., Kim, Y., & Kim, Y. K. (1998). A genetic algorithm for multiple objective sequencing problems in mixed model assembly lines. *Computers & Operations Research, 25*(7–8), 675–690.

Lee, K., Kim, B. S., & Joo, C. M. (2012). Genetic algorithms for door-assigning and sequencing of trucks at distribution centers for the improvement of operational performance. *Expert Systems with Applications, 39*(17), 12975–12983.

Qing-dao-er-ji, R., & Wang, Y. (2012). A new hybrid genetic algorithm for job shop scheduling problem. *Computers & Operations Research, 39*(10), 2291–2299.

Scopus (2015). *About scopus*. Elsevier <https://www.elsevier.com/solutions/scopus> Viewed 03 November 2015.

Stern, H., Chassidim, Y., & Zofi, M. (2006). Multiagent visual area coverage using a new genetic algorithm selection scheme. *European Journal of Operational Research, 175*(3), 1890–1907.

Su, W., Huang, S. X., Fan, Y. S., & Mak, K. L. (2015). Integrated partner selection and production–distribution planning for manufacturing chains. *Computers & Industrial Engineering, 84*, 32–42.

Tang, P. H., & Tseng, M. H. (2013). Adaptive directed mutation for real-coded genetic algorithms. *Applied Soft Computing, 13*(1), 600–614.

TTP (2015a). *Advanced materials research*. Trans Tech Publications Inc. <http://www.ttp.net/1022-6680.html> Viewed 15 November 2015.

TTP (2015b). *Applied mechanics and materials*. Trans Tech Publications Inc. <http://www.ttp.net/1660-9336.html> Viewed 15 November 2015.

Wang, Y., Yin, H., & Wang, J. (2009). Genetic algorithm with new encoding scheme for job shop scheduling. *The International Journal of Advanced Manufacturing Technology, 44*(9–10), 977–984.

Wang, L., & Zheng, D. Z. (2002). A modified genetic algorithm for job shop scheduling. *The International Journal of Advanced Manufacturing Technology, 20*(1), 72–76.

Wu, X., Chu, C. H., Wang, Y., & Yan, W. (2007). A genetic algorithm for cellular manufacturing design and layout. *European Journal of Operational Research, 181*(1), 156–167.

Zhong, T. X., & Chen, J. C. (2002). A hybrid-coded genetic algorithm based optimisation of non-productive paths in CNC machining. *The International Journal of Advanced Manufacturing Technology, 20*(3), 163–168.

Zhou, W., Zheng, J., Yan, J., & Wang, J. (2011). A novel hybrid algorithm for assembly sequence planning combining bacterial chemotaxis with genetic algorithm. *The International Journal of Advanced Manufacturing Technology, 52*(5–8), 715–724.