WILEY

Research trends in combinatorial optimization

Jann Michael Weinand[a],* , Kenneth Sörensen[b] , Pablo San Segundo[c],
Max Kleinebrahm[a] and Russell McKenna[d]

[a] *Chair of Energy Economics, Institute for Industrial Production, Karlsruhe Institute of Technology, Karlsruhe 76131, Germany*
[b] *Department of Engineering Management, University of Antwerp, Antwerp 2000, Belgium*
[c] *Centre for Automation and Robotics, Universidad Politécnica de Madrid, Madrid 28040, Spain*
[d] *Chair of Energy Transition, School of Engineering, University of Aberdeen, Aberdeen AB24 3FX, Scotland*
E-mail: jann.weinand@kit.edu [Weinand]; kenneth.sorensen@uantwerpen.be [Sörensen]; pablo.sansegundo@upm.es [San Segundo]; max.kleinebrahm@kit.edu [Kleinebrahm]; russell.mckenna@abdn.ac.uk [McKenna]

Abstract

Real-world problems are becoming highly complex and therefore have to be solved with combinatorial optimization (CO) techniques. Motivated by the strong increase in publications on CO, 8393 articles from this research field are subjected to a bibliometric analysis. The corpus of literature is examined using mathematical methods and a novel algorithm for keyword analysis. In addition to the most relevant countries, organizations, and authors as well as their collaborations, the most pertinent CO problems, solution methods, and application areas are presented. Publications on CO focus mainly on the development or enhancement of metaheuristics like genetic algorithms. The increasingly problem-oriented studies deal particularly with real-world applications within the energy sector, production sector, or data management, which are of increasing relevance due to various global developments. The demonstration of global research trends in CO can support researchers in identifying the relevant issues regarding this expanding and transforming research area.

Keywords: combinatorial optimization; bibliometric analysis; metaheuristics; genetic algorithms; exact algorithms; OR in energy

1. Introduction

Combinatorial optimization (CO) has always been of great interest in the scientific community (Cacchiani et al., 2018). In CO, problems are investigated that are characterized by a finite number of possible solutions (Bjorndal et al., 1995). While the discrete nature of these problems allows

*Corresponding author.

them to be solved in finite time by listing candidate solutions one by one and selecting the optimal solution, the number of such candidates typically grows rapidly with the input size, making many practical optimization problems insoluble for simple enumeration schemes (Cook, 2019). Researchers in CO explore the structural features of the problems and use these features to develop both precise and approximate general solution techniques. Usually, these CO problems are categorized based on their computational complexity. However, this worst-case evaluation does not always reflect the actual computational feasibility; the actual difficulty of the problems drives the development of solution approaches (Bjorndal et al., 1995). Through the development of effective methods and innovative approaches, as well as advances in software and hardware technology, hard real-world problems can already be solved more efficiently (Pardalos et al., 2013; Cacchiani et al., 2018). At the same time, new challenges are emerging, such as the consideration of uncertain conditions, the combination of hard problems, and the solving of problems in real time (Cacchiani et al., 2018).

In this context, the number of publications on CO problems has increased significantly in recent years (8393 in 2019, cf. Section 3.1), with over 150 reviews on the subject. However, these reviews cover only certain aspects of CO: for example, there are many reviews on specific CO problems such as the quadratic assignment problem (Loiola et al., 2007), the dynamic (Pillac et al., 2013) and the multiobjective (Jozefowiez et al., 2008) vehicle routing problem, the location-routing problem (Nagy and Salhi, 2007; Prodhon and Prins, 2014), or the minimum spanning tree problem (Pop, 2020). Furthermore, many studies review metaheuristics in CO in general (Blum and Roli, 2003; Gendreau and Potvin, 2005) or in comparison to each other regarding a specific problem (e.g., traveling salesman problem, TSP; Halim and Ismail, 2019). In addition, particular metaheuristics like ant colony optimization (Blum, 2005), other solution algorithms like Benders decomposition (Rahmaniani et al., 2017) as well as real-world applications of CO (e.g., sustainable supply chain network design; Eskandarpour et al., 2015) are reviewed.

Hence, a review of this growing body of literature, which thereby shows an overall picture of the CO research area, should be beneficial for CO researchers to understand and identify research and trends in CO in its entirety. For this purpose, different quantitative and qualitative approaches are employed to understand and organize the findings of existing studies. A bibliometric analysis is one of these approaches and has the potential to provide a systematic, reproducible, and transparent review process based on statistical measurements of research activities and researchers (Aria and Cuccurullo, 2017). This type of review can be classified as more objective and reliable than traditional review methods. Through a structured analysis of the extensive information on CO, the bibliometric review in this study can therefore achieve the following objectives (Aria and Cuccurullo, 2017): infer trends over time; show themes researched; identify shifts in the boundaries of the disciplines; detect the most prolific authors, organizations, and countries; and present an overview of the extant research. However, this paper should not be read as a complete and exhaustive list of all contributions in the field of CO, but rather as an insight into the current research focus and some of the most important challenges in the field, with the bibliometric study (including all of its limitations) used as a tool to highlight them.

These objectives are achieved with the present review study on CO. For this purpose, the paper is structured as follows. Section 2 describes the main methodological aspects of this study. Subsequently, the results and discussion of the bibliometric analysis are presented in Sections 3 and 4, respectively. Finally, Section 5 gives a summary and provides some conclusions.

Table 1
Search queries and resulting number of articles in the literature database Web of Science

Search name	Search query	Date	Number of studies
Optimization in general	(TS = "optimi*") AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article) Timespan: 1990–2019. Indexes: SCI-EXPANDED, SSCI.	23 March 2020	1,086,301
Initial search	(TS = "combinatorial optimi*") Timespan: 1990–2019. Indexes: SCI-EXPANDED, SSCI	23 March 2020	8769
Adjusted search	(TS = "combinatorial optimi*") AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article) Indexes = SCI-EXPANDED, SSCI Timespan = 1990–2019	23 March 2020	8393

2. Methodology

The research object of bibliometric analyses is a document system as well as the bibliometric characteristics of the individual documents. Thereby, the structure, features, and patterns of the underlying science are examined using mathematical and statistical methods (Weinand, 2020). The literature database *Web of Science*[1] and the web interface *biblioshiny* of the R-tool *bibliometrix* (Aria and Cuccurullo, 2017) are used to investigate the literature on CO. In Web of Science, the *adjusted search* query in Table 1 is used. In addition to the bibliometric analysis tool bibliometrix (Section 2.1), the statistical indicators *h*-, *g*-, and *m*-index (Section 2.2), an approach to measure trends (Section 2.3) and an algorithm for keyword and author analyses (Section 2.4) are explained in the following.

2.1. R-tool bibliometrix

The main part of the present analysis is based on the Web of Science analyzing tool and evaluations based on a newly developed algorithm (cf. Section 2.4). In addition, the R-tool bibliometrix is applied for the examination of the corpus of literature. Bibliometrix is an open-source tool for conducting comprehensive scientific mapping analyses. This tool has already been used in many bibliometric analyses. Due to its implementation in R, the package is flexible and facilitates integration with other statistical or graphical packages (Aria and Cuccurullo, 2017). One example for which bibliometrix has been used is determining the number of country collaborations (cf. Table 3).

---

[1]See http://apps.webofknowledge.com/WOS_GeneralSearch_input.do?product=WOS&search_mode=GeneralSearch &SID=D3JAVHgH6kUCRXvVCDb&preferencesSaved=.

Table 2
Arrow icons showing the trend of the increase in publications, using an example with a maximum average annual percentage increase in publications of 100 %

Interval of average annual percentage increase in publications (%)	Arrow icon
[0;20]	→
(20;40]	↗
(40;60]	↑
(60;80]	↑↑
(80;100]	↑↑↑

2.2. *Measures of influence:* h-*index,* m-*index, and* g-*index*

A variety of indices have been introduced in the past to measure research impact. The *h*-index was introduced to facilitate quantifying the cumulative impact and relevance of an individual's scientific output (Hirsch, 2005). Thereby an individual is associated to publications and can therefore be an author (cf. Table A1), country (cf. Table 3), organization (cf. Table 4), or source (cf. Table 5). The *h*-index reflects the number of *h* papers of an individual that have been cited at least *h* times. Together with the *h*-index, Hirsch (2005) also provided the *m*-index, which reflects the time period since the first publication of an individual by dividing the *h*-index by the number of years of scientific activity. Hirsch (2005) also classifies different values of the *m*-index with $m = 1$ being a "successful scientist", $m = 2$ being an "outstanding scientist," and $m = 3$ being a "truly unique individual." The *g*-index was introduced by Egghe (2006) as an alternative of the *h*-index. It represents the unique largest number of the top *g* most cited articles, which together received at least g^2 citations. This index therefore gives a higher weighting to highly cited articles than the *h*-index. In the present study, however, these indices cannot be used to evaluate a single individual's scientific activity but only for comparison with each other. This is because the publications on CO represent only a subset of an individual's total publications. As a result, for example, in relation to this subset the *m*-index is below 1 for all authors (cf. Table A1), whereas in relation to all publications by these authors could be well above 1.

2.3. *Measuring trends*

Important for the analysis of a research field are not only static measures but also the development of these indicators over time. Therefore, in some tables in Section 3, trends concerning the number of publications are indicated by arrow icons. To estimate the trends, publications from 2015 to 2019 are considered. The percentage increase in the number of publications per year is calculated, followed by the average from 2015 to 2019. The highest average value is then divided into five equal ranges. An example is used to show this measurement: it is assumed that the highest average annual percentage increase in publications of 20 countries between 2015 and 2019 is 100. In this case, an indication of the trend by means of arrow icons for different intervals as shown in Table 2 would be applied.
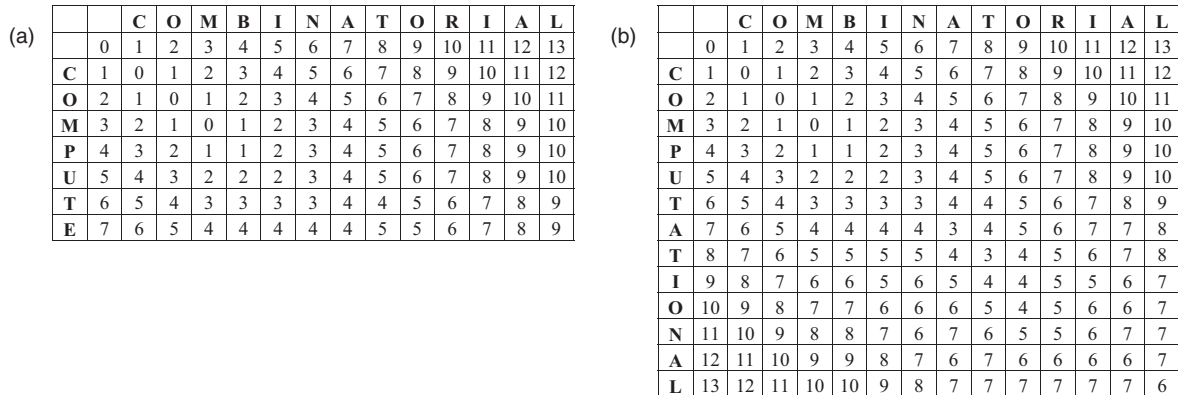
(a)

		C	O	M	B	I	N	A	T	O	R	I	A	L
	0	1	2	3	4	5	6	7	8	9	10	11	12	13
C	1	0	1	2	3	4	5	6	7	8	9	10	11	12
O	2	1	0	1	2	3	4	5	6	7	8	9	10	11
M	3	2	1	0	1	2	3	4	5	6	7	8	9	10
P	4	3	2	1	1	2	3	4	5	6	7	8	9	10
U	5	4	3	2	2	2	3	4	5	6	7	8	9	10
T	6	5	4	3	3	3	3	4	4	5	6	7	8	9
E	7	6	5	4	4	4	4	4	5	5	6	7	8	9

(b)

		C	O	M	B	I	N	A	T	O	R	I	A	L
	0	1	2	3	4	5	6	7	8	9	10	11	12	13
C	1	0	1	2	3	4	5	6	7	8	9	10	11	12
O	2	1	0	1	2	3	4	5	6	7	8	9	10	11
M	3	2	1	0	1	2	3	4	5	6	7	8	9	10
P	4	3	2	1	1	2	3	4	5	6	7	8	9	10
U	5	4	3	2	2	2	3	4	5	6	7	8	9	10
T	6	5	4	3	3	3	3	4	4	5	6	7	8	9
A	7	6	5	4	4	4	4	3	4	5	6	7	7	8
T	8	7	6	5	5	5	5	4	3	4	5	6	7	8
I	9	8	7	6	6	5	6	5	4	4	5	5	6	7
O	10	9	8	7	7	6	6	6	5	4	5	6	6	7
N	11	10	9	8	8	7	6	7	6	5	5	6	7	7
A	12	11	10	9	9	8	7	6	7	6	6	6	6	7
L	13	12	11	10	10	9	8	7	7	7	7	7	7	6

Fig. 1. Dynamic programming matrices for determining the Levenshtein distance between (a) "compute" and "combinatorial" and (b) "computational" and "combinatorial."

2.4. Keyword and author analysis

Keywords provide an important indication of the subjects covered in scientific studies. An analysis of these keywords can therefore provide an overview of the most important topics in a research field. These keywords could be examined with the help of bibliometrix. However, in its web interface biblioshiny, the exact strings x and y are compared to each other. If one character of a string is different from the other string, then these words are considered to be different keywords. For this reason, a separate keyword analysis algorithm was developed in MATLAB for this review. The algorithm considers similar strings as one keyword. The similarity of the strings is determined by the Levenshtein (1966) distance. The Levenshtein distance between x and y is the total cost of transforming x into y using the operations of inserting, deleting, and substituting a character. Thereby, the string distance problem is equivalent to the shortest path problem defined in a graph which is constructed as follows (Spiliopoulos and Sofianopoulou, 2007).

Given strings x and y of lengths m and n, respectively, the nodes are the points (i, j) in the grid $i = 0,…, m$ and $j = 0,…, n$, whereby the former string is put vertically and the latter string horizontally. There are three types of directed links from the nodes i and j (Spiliopoulos and Sofianopoulou, 2007):

- vertical links $(i, j) \rightarrow (i + 1, j)$, $i = 0,…,m − 1, j = 0,…, n$, with cost 1, to represent the deletion of x_{i+1};
- horizontal links $(i, j) \rightarrow (i, j + 1)$, $i = 0,…,m, j = 0,…, n − 1$, with cost 1, to represent the insertion of y_{j+1} after x_i (if $i = 0$, at the start of x);
- diagonal links $(i, j) \rightarrow (i + 1, j + 1)$, $i = 0,…,m − 1, j = 0,…, n − 1$, to represent the substitution of x_{i+1} by y_{j+1}. There is no cost involved if $x_{i+1} = y_{j+1}$, otherwise the cost is 1.

The distance between strings x and y is then given by the length of the shortest path between points $(0, 0)$ and (m, n) (Spiliopoulos and Sofianopoulou, 2007). Figure 1 shows two examples of determining Levenshtein distances in the so-called dynamic programming matrices. The top rows

of the matrices have 0 values, since the insertions at the start of the vertical string are not penalized. After filling in the matrices according to the above-mentioned links, the minimum value in the last row shows the distance. In the first example in Fig. 1a, there are three substitutions (P-U and E with B-I and A) and six insertions (N and O-R-I-A-L), whereas in the second example there are only six substitutions (P-U-T with B-I-N and I-O-N with O-R-I, cf. Fig. 1b).

In the developed algorithm, the Levenshtein distance is used to match equivalent strings under one keyword. Thus, the number of keyword occurrences can be determined more accurately. Strings of five or more characters are grouped for a Levenshtein distance up to one and strings of nine or more characters are grouped for a Levenshtein distance up to two. The latter case should ensure that, for example, the plural of a word and the simultaneous use of a hyphen is recognized as the same keyword (e.g. *metaheuristic* and *meta-heuristics*). Besides the number of keyword occurrences, the algorithm also determines the mean publication year and the mean citations of all articles with the respective keyword. In addition, the algorithm examines the keywords with regard to their simultaneous occurrence in the same articles (cf. Section 3.4). This latter function is also used to identify the collaboration of authors (cf. Fig. A1). The MATLAB script is applicable to any other bibliometric analysis and can be provided upon request.

3. Results

In the following, the main characteristics of the research field on CO are presented (cf. Section 3.1). Afterwards, an overview of contributions and collaborations of different countries and organizations is given in Section 3.2. Subsequently, the most relevant sources and studies are highlighted (cf. Section 3.3), before the most relevant topics on CO are discussed (cf. Section 3.4).

3.1. Development of the research field

The research field on CO includes 8393 articles, which have been published in 1415 different sources at the time of this analysis. A total of 14,423 authors were involved in the articles and the average number of citations per document is 23.55. This high citation rate is due to some highly cited publications (cf. Section 3.3); 55% of the articles are cited less than 10 times. The number of publications per year on CO has increased quite steadily over the years from 13 in 1990 to 533 in 2019. This could be related to the general increase in publications in the field of operations research. However, when comparing the share of articles on CO with the total number of articles on optimization in general (search query *optimization in general* in Table 1), a slight increase can be observed: in 1990, 15,868 articles on optimization in general and 76 (0.5%) on CO had already appeared, and in 2019, the share is 0.8% with 8456 in the total number of 1,100,191 publications.

3.2. Publication distribution and collaboration of countries, organizations, and authors

In total, authors from 85 countries have contributed to articles on CO. The fact that CO is only extensively researched in a limited number of countries is shown by the share of the top 20 relevant

countries in Table 3, which, with a number of 7157 articles, are involved in 85% of the publications on CO. In the top 20 are Australia, 11 countries from Europe, 5 from Asia, and 3 from America. By far the most articles were authored by researchers from the United States (23%), followed by China (14%), France (8%), Germany (7%), and Japan (7%). The development of the number of articles in these top five relevant countries in Fig. 2a is of interest: while authors in the United States published the most articles each year from 1990 to 2013, China was responsible for the fewest publications in the 1990s. However, the number of annual articles by Chinese authors has been increasing almost exponentially since then, and since 2014 most of the annual articles on CO are written by Chinese authors. Furthermore, as the arrow icons in Table 3 show, the annual percentage increase in the number of articles since 2015 among the top 20 countries is the highest in China (cf. Section 2.3). This could be related to the rapid growth of the Chinese economy and the associated increase in energy consumption (Zhang et al., 2017). As Section 3.4.3 shows, the production sector and the energy sector are the most important application areas of CO research.

However, these numbers of published articles do not provide any indication of the importance of CO research in the overall research of a country. For this reason, the total publications of a country are also shown in Table 3 in relation to the mean annual (between 2012 and 2018) gross domestic expenditure on R&D of the respective country (UNESCO, 2020). In the top five most productive countries, with the exception of France, CO seems to be of comparatively low importance. In contrast, CO research seems to be of great priority in Portugal, Iran, Spain, Belgium, and Canada. In addition, the United States has by far the highest number of citations (67,298) and the highest h-index (104), g-index (212), and m-index (3.47). However, in terms of average article citations, the United States is only in sixth position (35.1), behind Australia (41.5), Belgium (40.9), Canada (40.8), England (36.8), and Switzerland (35.2). Due to the high number of citations, the g-index in these countries is also comparatively high. In relation to the number of publications, the m-index is high in some countries, such as Iran, which published its first articles on CO in 2001.

Most collaborative publications have been produced by authors from United States and China (146), United States and Canada (73), United States and Germany (69), United States and France (60) as well as England and China (55). Using the corresponding author for assigning the articles to a country (cf. Table 3), the countries that frequently participate in international cooperation can be identified. Switzerland (48%) has the largest share of collaborative publications, ahead of Belgium (45%) and Austria (40%). In contrast, authors from Iran (88% single country publications) and Taiwan (85%) show the lowest share of collaborative publications. In general, the share of international or intranational collaboration is rather low, which is also reflected by the number of single-authored publications (1.087; 12%).

A total of 3539 different organizations from the 85 countries were participating in the publications on CO. The top 20 of the most relevant organizations are listed in Table 4. These organizations have been involved in 1696 publications (20%) so far. Most of the articles have originated from authors of the top five organizations *French National Centre for Scientific Research* (253, France), *University of California System* (162, United States), *Chinese Academy of Sciences* (118, China), *National Institute for Research in Digital Science and Technology* (115, France), and *Massachusetts Institute of Technology* (101, United States). However, it should be noted that some of these organizations are associations of several institutions, such as the *Chinese Academy of Sciences* or the *University of California System*. The annual publication volume of these five organizations in Fig. 2b shows a slightly rising trend in addition to annual fluctuations. The trend in the annual

Table 3
The top 20 of the most productive countries in terms of publications on combinatorial optimization

Country	Total publications		Trend	h-Index	g-Index	m-Index	TPGD (PPP, $ billion)$^{-1}$	Average article citations No.	Corresponding author's country		Single country publication		Collaborative publications	
	No.	%							No.	%	No.	%	No.	%
United States	1918	23	↗	104	212	3.47	4	35.1	1388	17	977	70	411	30
China	1197	14	↑↑	62	103	2.21	3	16.6	1068	13	812	76	256	24
France	676	8	↑	53	86	1.77	11	19.2	499	6	319	64	180	36
Germany	586	7	↑	49	76	1.63	5	17.4	432	5	279	65	153	35
Japan	547	7	↗	48	86	1.60	3	18.4	495	6	376	76	119	24
England	486	6	↑	55	123	1.83	11	36.8	330	4	205	62	125	38
Italy	425	5	↗	42	109	1.45	14	32.5	325	4	236	73	89	27
Canada	415	5	↑↑	52	124	1.79	15	40.8	254	3	162	64	92	36
Spain	398	5	↑↑	36	68	1.20	20	17.7	304	4	204	67	100	33
Brazil	308	4	↑↑	36	69	1.29	8	19.7	256	3	192	75	64	25
Australia	251	3	↑↑	36	100	1.24	11	41.5	166	2	115	69	51	31
India	251	3	↗	30	48	1.03	4	13.5	214	3	170	79	44	21
Netherlands	215	3	→	31	60	1.07	13	20.3	130	2	82	63	48	37
Taiwan	210	3	↑	42	67	1.45	n.a.	25.3	185	2	158	85	27	15
Belgium	199	2	↑↑	34	88	1.36	15	40.9	149	2	82	55	67	45
Turkey	179	2	↗	32	49	1.14	11	18.2	151	2	120	79	31	21
Austria	171	2	↑↑	32	60	1.10	13	24.1	120	1	72	60	48	40
Iran	163	2	↑	30	45	1.58	23	17.0	148	2	130	88	18	12
Portugal	155	2	↑	26	51	1.04	39	20.0	107	1	74	69	33	31
Switzerland	153	2		32	72	1.10	10	35.2	108	1	56	52	52	48

Note: The number for "total publications" does not have to sum up to 8393 or 100%, since more than one country could have contributed to a single publication. The percentages for "total publications" and "corresponding author's country" refer to the total number of 8393 articles, while for "single publication" and "collaborative publication" they refer to the number of publications of "corresponding author's country." TPGD is the total publications per gross domestic expenditure on R&D (UNESCO, 2020) expressed in the purchasing power parity of the respective country. As gross domestic expenditure on R&D, the mean annual value between 2012 and 2018 is used. PPP, purchasing power parity.
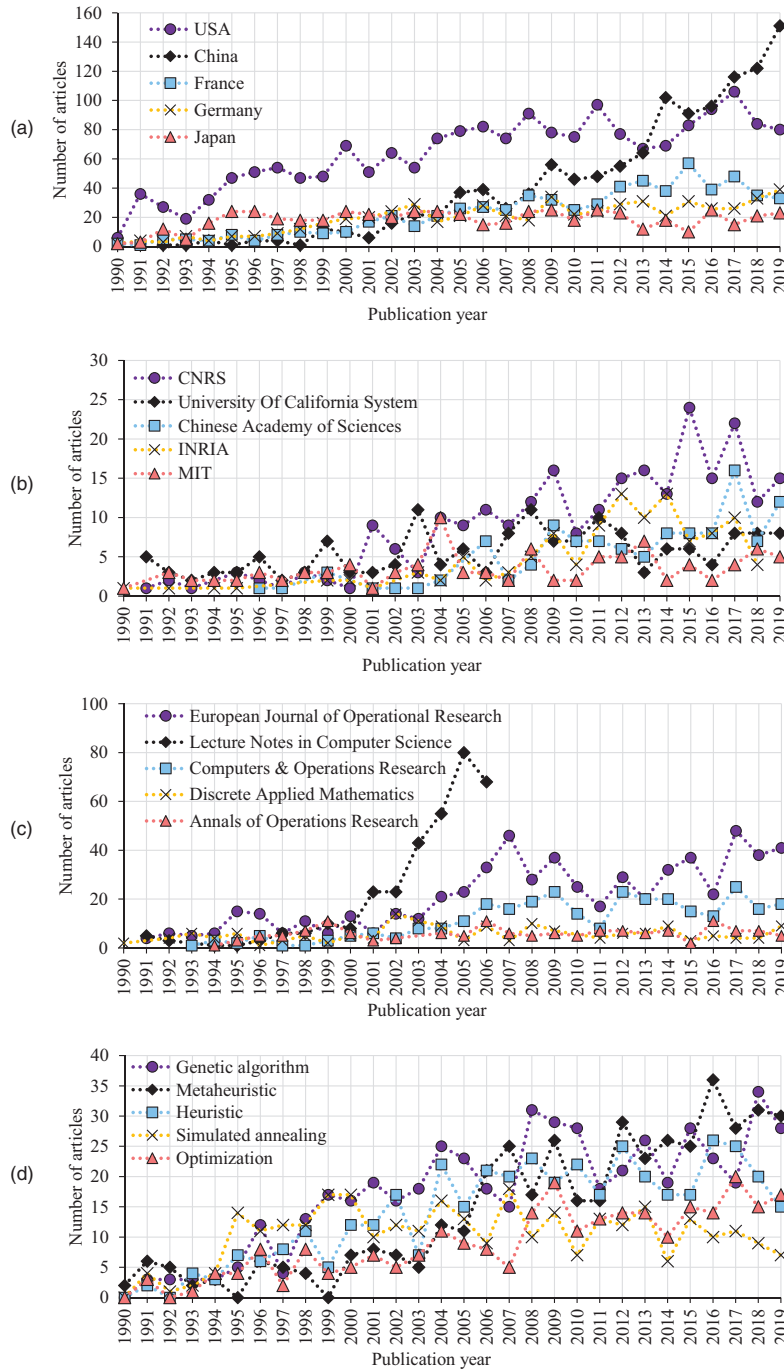
Fig. 2. Annual development of publications of the top five most productive countries (a), organizations (b), sources (c), and keywords (d). For reasons of clarity, zero values are not displayed. Web of Science stopped covering articles from the source *Lecture Notes in Computer Science* in 2007.

Table 4
The top 20 of the most productive organizations in terms of publications on combinatorial optimization

Organization	Country	Total publications		Trend	h-Index	g-Index	m-Index
		No.	%				
French National Centre for Scientific Research (CNRS)	France	253	3	↑	31	52	1.07
University of California System	United States	162	2	↑	37	62	1.28
Chinese Academy of Sciences	China	118	1	↑↑↑	20	41	0.83
National Institute for Research in Digital Science and Technology (INRIA)	France	115	1	↑	25	39	0.83
Massachusetts Institute of Technology (MIT)	United States	101	1	↗	30	70	1.00
State University System of Florida	United States	101	1	↑	25	41	1.04
University of Montreal	Canada	89	1	↑	27	67	0.93
Universite Libre de Bruxelles	Belgium	88	1	↑	22	75	0.88
University System of Georgia	United States	84	1	↑↑↑	25	53	0.89
University of Bologna	Italy	80	1	↑	22	46	0.88
University of Tokyo	Japan	78	1	↑↑↑	21	41	0.84
Carnegie Mellon University	United States	65	1	↑	23	48	0.79
Tsinghua University	China	64	1	↑	19	40	0.73
University of Texas System	United States	64	1	↑	17	45	0.63
Huazhong University of Science and Technology	China	63	1	↑↑↑	25	41	1.25
Indian Institute of Technology System (IIT System)	India	63	1	↑↑	15	29	0.54
Nanyang Technological University	Singapore	63	1	↑	27	41	1.08
Russian Academy of Sciences	Russia	63	1	↑↑	9	16	0.31
Polytechnic University of Catalonia	Spain	62	1	↑↑	19	47	0.79
Universidade de Lisboa	Portugal	62	1	↑↑	17	39	0.68

Note: The percentage values refer to the total of 8393 publications.

percentage increase in publications (cf. Table 4) shows a particularly strong increase for *Chinese Academy of Sciences*, *University System of Georgia*, *University of Tokyo*, and *Huazhong University of Science and Technology*. The publications on CO from the *University Of California System* have the highest *h*-index (37), followed by the *French National Centre for Scientific Research* (31) and the *Massachusetts Institute of Technology* (30). However, the order changes when the number of citations is taken into account for the most cited articles: for the *g*-index, the *Universite Libre de Bruxelles* (75) is at the top of the list, ahead of the *Massachusetts Institute of Technology* (70) and the *University of Montreal* (67). The publications of the *University of California System* and *Huazhong University of Science and Technology* are particularly relevant with regard to the first publication year, with *m*-indexes of 1.28 and 1.25, respectively. The majority of collaborations took place between organizations of the same country (cf. Fig. 3): most collaborations were conducted
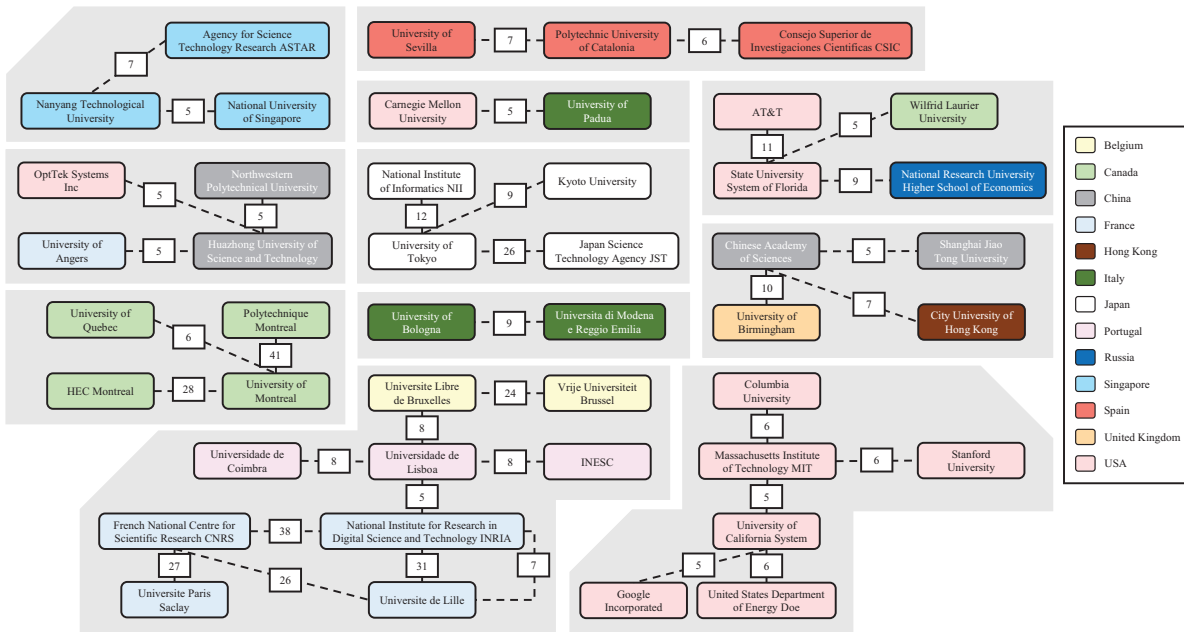
Fig. 3. Collaboration network of the top 20 most productive organizations together with their top three most productive collaborations. Collaborations with less than five collaborative publications are not shown.

by the *University of Montreal* and *Polytechnique Montreal* (41), the *French National Centre for Scientific Research* and the *National Institute for Research in Digital Science and Technology* (38) as well as the *National Institute for Research in Digital Science and Technology* and the *University of Lille* (31). The publication distribution and collaboration of the most relevant authors can be found in the Appendix.

3.3. Most relevant sources and articles

In addition to the authors of the articles, it is also important to analyze the sources in which the articles are published predominantly. The studies on CO have been published in 1415 different sources. A large proportion of the articles on CO (35%) is published in the top 20 most productive sources (cf. Table 5) *European Journal of Operational Research* (EJOR), *Lecture Notes in Computer Science* (LNCS), *Computers & Operations Research*, *Discrete Applied Mathematics*, and *Annals of Operations Research*. Most articles on CO were published in EJOR, with 7% of all publications. Therefore, it is not surprisingly that this source has the highest h-index (58) and g-index (99), followed by *Computers & Operations Research* regarding the h-index (46) and by LNCS regarding the g-index (87). The source *Expert Systems with Applications* has a particularly high m-index (2.07) due to a high h-index (31) in relation to the starting year of the first publication on CO (2005). In the most relevant source, EJOR, the share of articles on CO in all publications has increased

Table 5
The top 20 of the most productive scientific sources in terms of publications on combinatorial optimization

Source	Total publications (TP)		Trend	Average article citations	h-Index	g-Index	m-Index
	No.	%					
European Journal of Operational Research	615	7	↑	26	58	99	2.00
Lecture Notes in Computer Science	335	4	-	29	28	87	0.97
Computers & Operations Research	306	4	↗	8	46	44	1.70
Discrete Applied Mathematics	158	2	↗	25	23	58	0.77
Annals of Operations Research	149	2	↗	8	31	31	1.19
Journal of Combinatorial Optimization	138	2	↑↑	24	17	53	0.74
Mathematical Programming	125	1	→	18	30	43	1.03
Applied Soft Computing	120	1	↑↑↑	24	25	46	1.56
Expert Systems with Applications	120	1	↗	24	31	51	2.07
Computers & Industrial Engineering	110	1	↑↑	24	25	48	0.93
Journal of Heuristics	108	1	↗	22	30	47	1.36
Theoretical Computer Science	87	1	↑	17	19	33	0.73
International Journal of Production Research	82	1	↑↑	14	23	30	0.82
Operations Research Letters	79	1	↗	21	18	39	0.62
Journal of the Operational Research Society	76	1	↗	28	21	46	0.72
Journal of Global Optimization	70	1	↗	7	18	19	0.69
Discrete Optimization	66	1	↗	31	13	44	0.93
Information Sciences	65	1	↑	13	22	27	0.85
Algorithmica	61	1	↑	21	13	33	0.45
Computational Optimization and Applications	60	1	↗	11	19	23	0.83

Note: The percentage values refer to the total of 8393 publications. The trend in number of publications is not shown for the *Lecture Notes in Computer Science*, since this source is not indexed in Web of Science since 2007.

steadily, from 2% in 1990 to 6% in 2019. In addition to the findings from Section 3.1, this again demonstrates the increasing importance of CO in the research field of operations research.

However, when examining the development of the annual publications of the top five most relevant sources in Fig. 2c, it is striking that the annual publications of LNCS increased exponentially between 2001 and 2005. If this trend had continued, LNCS would have been in first rank among the most productive sources on CO. However, from 2007 onward, no publications by LNCS are listed

in Web of Science anymore. A reason for this could be that LNCS publishes mostly conference proceedings. The trend in the annual percentage increase in publications shows a strong upward trend for only a few sources, with *Applied Soft Computing* at the top (cf. Table 5).

The articles in a research field that are most cited can be assessed as most influential to that field. Table A2 lists the 20 globally most cited articles in the CO research area. Many of these articles are about newly developed metaheuristics like ant colony optimization (Dorigo et al., 1996; Dorigo and Gambardella, 1997), harmony search (Geem et al., 2001), or variable neighborhood search (Mladenović and Hansen, 1997). The study by Dorigo et al. (1996) on ant colony optimization is cited most frequently (5646). Furthermore, there are some articles in the top 20 that deal with the development of new algorithms (evolutionary algorithms, Yao et al., 1999; Han and Kim, 2002; approximation algorithms, Goemans and Williamson, 1995) or the enhancement of existing metaheuristics (MAX–MIN ant system, Stützle and Hoos, 2000). It is notable that the newly developed metaheuristics are mostly tested and benchmarked using the TSP (Dorigo et al., 1996; Dorigo and Gambardella, 1997; Mladenović and Hansen, 1997; Stützle and Hoos, 2000; Geem et al., 2001). The studies described in this paragraph also have by far the most local citations, that is, within the 8393 articles examined here. This means that these fundamental new methods are either applied or extended in many other articles. Ant colony optimization is the most common method with four studies in the top 20 of the most cited articles (Dorigo et al., 1996; Dorigo and Gambardella, 1997; Stützle and Hoos, 2000; Dorigo and Blum, 2005).

In addition to the development of these theoretical approaches, CO methods are applied in case studies in the top 20 articles, such as vision/image processing (Boykov and Kolmogorov, 2004; Boykov and Funka-Lea, 2006; Jaqaman et al., 2008), quantum computers (Knill et al., 2001), or atomic decomposition (Donoho and Huo, 2001). In contrast to the studies on developments of metaheuristics and algorithms described above, the local citations of these articles on specific applications are very low. Furthermore, only 2 of the top 20 most cited articles were published in the top 5 most relevant sources: one each in EJOR and *Computers & Operations Research* (cf. Table A2).

However, the most cited articles described above do not provide a complete picture of research trends in CO, as many are relatively old with the most recent article dating from 2008. Therefore, Table 6 shows the top 10 articles with the highest annual citation rates that have appeared since 2010. With 108 citations, Deng et al. (2019) is cited the most annually. Similar to the sixth most cited study by Nouiri et al. (2018), a new particle swarm optimization method is developed in Deng et al. (2019). An advanced method, in this case a search algorithm, is also introduced in Wen and Yin (2013). Some other studies are more application-oriented on phosphorus systems for white light emitting diodes (Xia et al., 2016), biological transport networks (Tero et al., 2010), or fatty acids production (Xu et al., 2013).

Furthermore, three survey studies are included in Table 6: first, on hyperheuristics (Burke et al., 2013), that is, heuristics that choose an appropriate solution heuristic depending on the problem. The other two publications are strongly linked regarding their content. Sörensen (2015) discusses the fact that many metaheuristics have been developed in recent years and that these are largely based on metaphors about natural or man-made processes, for example, ant or bee colony optimization. The author argues that besides some innovative studies of high quality, many papers have been published that are justified only by the fact that the developed metaheuristics are based on novel metaphors. Blum et al. (2011) also motivate their study on hybrid metaheuristics by the

Table 6
Articles among the scientific contributions on combinatorial optimization, which have the highest annual citation rate since 2010

Article title	Global citations		Publication year	Source
	No.	Per year		
A novel intelligent diagnosis method using optimal LS-SVM with improved PSO algorithm (Deng et al., 2019)	108	108	2019	*Soft Computing*
Recent developments in the new inorganic solid-state LED phosphors (Xia et al., 2016)	264	66	2016	*Dalton Transactions*
Quantum annealing with manufactured spins (Johnson et al., 2011)	533	59	2011	*Nature*
Hyperheuristics: a survey of the state of the art (Burke et al., 2013)	354	51	2013	*Journal of the Operational Research Society*
Metaheuristics—the metaphor exposed (Sörensen, 2015)	223	45	2015	*International Transactions in Operational Research*
An effective and distributed particle swarm optimization algorithm for flexible job-shop scheduling problem (Nouiri et al., 2018)	82	41	2018	*Journal of Intelligent Manufacturing*
Rules for biologically inspired adaptive network design (Tero et al., 2010)	375	38	2010	*Science*
Modular optimization of multigene pathways for fatty acids production in *E. coli* (Xu et al., 2013)	262	37	2013	*Nature Communications*
Hybrid metaheuristics in combinatorial optimization: a survey (Blum et al., 2011)	323	36	2011	*Applied Soft Computing*
A feasible method for optimization with orthogonality constraints (Wen and Yin, 2013)	225	32	2013	*Mathematical Programming*

fact that research on metaheuristics for CO problems was mostly algorithm-oriented in the past. Nevertheless, the authors see a trend that the focus of research on CO metaheuristics is shifting from this algorithm orientation to a problem orientation. As a result, metaheuristics are now often hybridized with other optimization techniques in order to find the best approaches to solving problems.

As already described, the article by Dorigo et al. (1996) has the most global citations (5646) among the articles examined here and is also very often locally cited (235) by these 8393 studies. However, five studies are even more frequently cited locally by the 8393 articles (cf. Table 7). The

Table 7
Five most frequently cited references in the articles on combinatorial optimization

Article title	Local citations	Publication year	Source
Computers and intractability: a guide to the theory of NP-completeness (Garey and Johnson, 2009)	911	1979	*A Series of Books in the Mathematical Sciences*
Optimization by simulated annealing (Kirkpatrick et al., 1983)	702	1983	*Science*
Genetic algorithms in search, optimization, and machine learning (Goldberg, 2012)	527	1989	—
Tabu search—part I (Glover, 1989)	276	1989	*ORSA Journal on Computing*
"Neural" computation of decisions in optimization problems (Hopfield and Tank, 1985)	266	1985	*Biological Cybernetics*

focus of the most local cited study is on the identification and handling of NP-complete (combinatorial optimization) problems, that is, problems that can be solved in polynomial time on a nondeterministic turing machine (Garey and Johnson, 2009). The other three publications in Table 7 deal with the introduction of new metaheuristics (simulated annealing, Kirkpatrick et al., 1983, and tabu search, Glover, 1989) or a review of a metaheuristic (genetic algorithm, Goldberg, 2012). These three metaheuristics are also the most frequently employed in the 8393 studies on CO (cf. Section 3.4).

3.4. Most relevant subjects

In this section, first the algorithm based on the Levenshtein distance is evaluated (cf. Section 3.4.1). Second, the most relevant keywords and thus topics in the research area of CO are shown in Section 3.4.2. Section 3.4.3 then attempts to determine the most relevant application areas of CO.

3.4.1. Evaluation of keyword algorithm

The developed algorithm led to a much better recognition of related keywords than was the case with the web interface *biblioshiny* of the R-tool bibliometrix. For example, different spellings with or without hyphen or in singular or plural are now combined. The keyword *combinatorial optimization*, which occurs most frequently in the articles with 3.051 times, serves as an example. The different spellings that have been grouped for this keyword are the following, whereby the change in spelling compared to the first keyword is shown in bold: *combinatorial optimization, combinatorial optimization**S**, combinatorial optimi**S**ation, combinatorial optimi**S**ation**S**, and combinatori**C** optimization*. Thus 7% more appearances of the keyword *combinatorial optimization* could be identified (cf. Fig. 4). For the seven most relevant keywords shown in Fig. 4, the greatest improvement was achieved for the keyword *genetic algorithm*: here the recognition rate increased by 99%. However, there are also examples that seem to be always stated in the same way: for *simulated annealing*, the developed algorithm yielded no improvement.
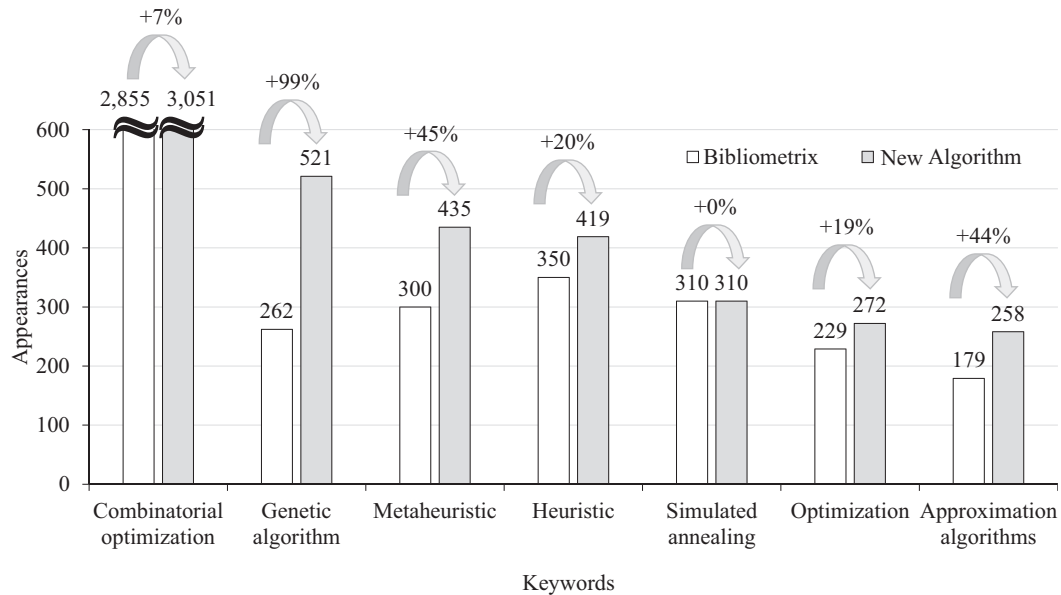
Fig. 4. Comparison of the appearance of the most relevant keywords determined on the one hand with the R-tool bibliometrix and on the other hand with the new algorithm developed for this study.

3.4.2. Combinatorial optimization problems and solution methods

In this section, the most relevant keywords on CO that appear in at least 1% of publications are discussed. As already shown in Section 3.4.1, the keyword *combinatorial optimization* occurs most frequently (in 3051 publications). This is not surprising since the keyword corresponds to the overall topic of this bibliometric analysis and thus the search query in Web of Science. Therefore, the keyword *combinatorial optimization* is excluded in the following analysis. Thus, the top five relevant keywords are *genetic algorithm* (521 occurrences), *metaheuristic* (435), *heuristic* (419), *simulated annealing* (310), and *optimization* (272). The annual occurrence of these keywords fluctuates, but in general there is an increasing trend (except for simulated annealing, cf. Fig. 2d). The fact that simulated annealing is less frequently covered in actual CO publications is further demonstrated by the low mean publication year (2005.7, cf. Fig. 5). In Fig. 5, the 25 keywords that appear in at least 1% of publications are divided into different categories and their mean publication year is shown. In addition, the correlation matrix in Fig. 6, which shows how often these keywords occur together in articles, is also relevant for the following analysis.

The most frequently addressed specific optimization problems are the TSP (225 appearances), integer programming (225), and the scheduling problem (219). While no clear trend can be identified for integer programming, the TSP is mainly investigated in connection with the metaheuristics ant colony optimization (20), genetic algorithms (18) as well as local search (16), and the scheduling problem with tabu search (13) and simulated annealing (10) (cf. Fig. 6). The TSP is examined so frequently since it is representative of CO problems. If an efficient (polynomial-time) algorithm could be found for the TSP, then efficient algorithms could also be found for all other NP-complete problems (Hoffman et al., 2013). Integer programming problems (e.g., assignment problem) are
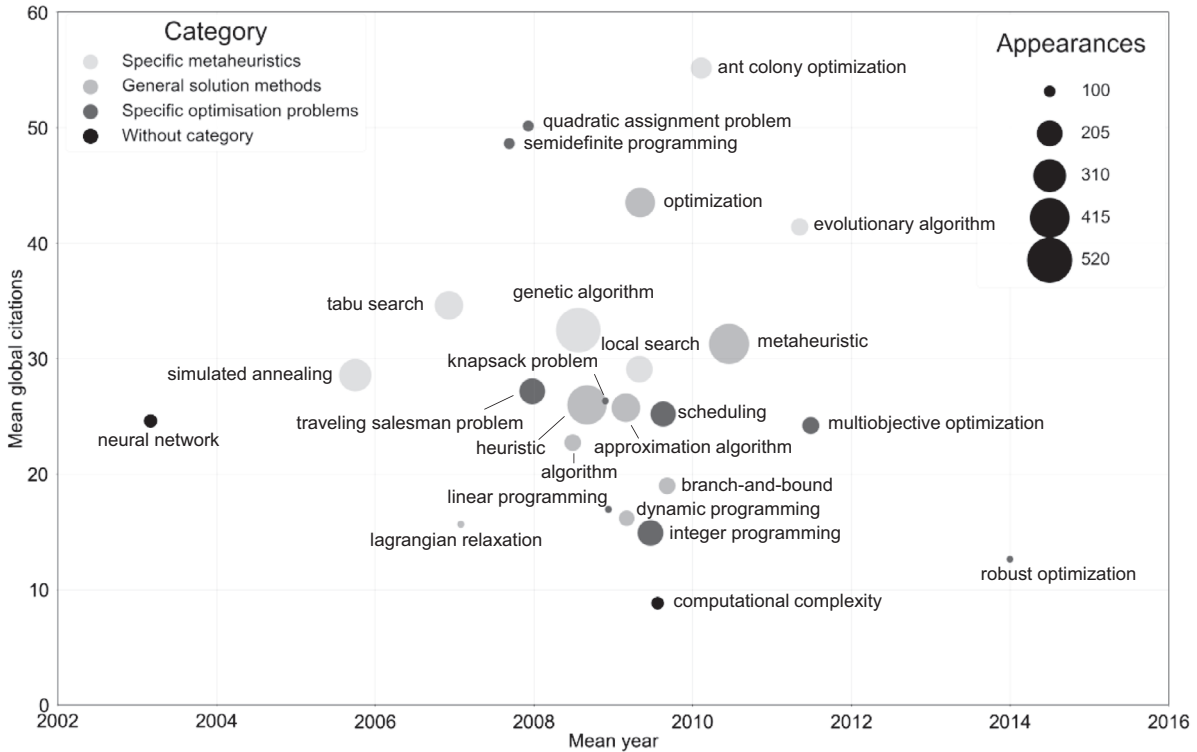
Fig. 5. The top 25 most relevant keywords on CO with their number of appearances, mean global citations, and mean publication year in the 8393 articles.

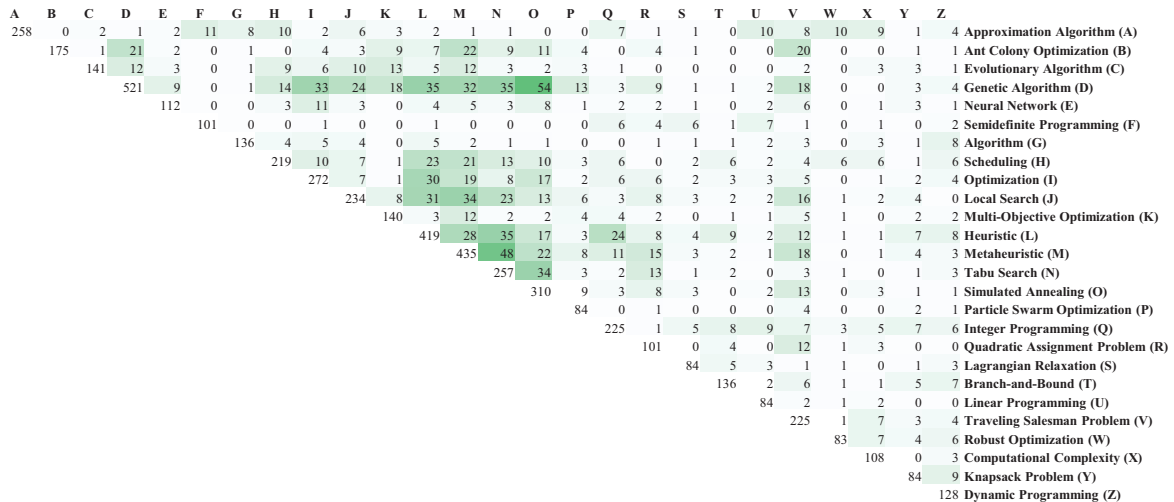| A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
258	0	2	1	2	11	8	10	2	6	3	2	1	1	0	0	7	1	1	0	10	8	10	9	1	4	Approximation Algorithm (A)
	175	1	21	2	0	1	0	4	3	9	7	22	9	11	4	0	4	1	0	0	20	0	0	1	1	Ant Colony Optimization (B)
		141	12	3	0	1	9	6	10	13	5	12	3	2	3	1	0	0	0	0	2	0	3	3	1	Evolutionary Algorithm (C)
			521	9	0	1	14	33	24	18	35	32	35	54	13	3	9	1	1	2	18	0	0	3	4	Genetic Algorithm (D)
				112	0	0	3	11	3	0	4	5	3	8	1	2	2	1	0	2	6	0	1	3	1	Neural Network (E)
					101	0	1	0	1	0	0	0	0	6	4	6	1	7	1	0	1	1	0	2		Semidefinite Programming (F)
						136	4	5	4	0	5	2	1	1	0	1	1	1	2	3	0	3	1	8		Algorithm (G)
							219	10	7	1	23	21	13	10	3	6	0	2	6	2	4	6	6	1	6	Scheduling (H)
								272	7	1	30	19	8	17	2	6	6	2	3	3	5	0	1	2	4	Optimization (I)
									234	8	31	34	23	13	6	3	8	3	2	2	16	1	2	4	0	Local Search (J)
										140	3	12	2	2	4	4	2	0	1	1	5	1	0	2	2	Multi-Objective Optimization (K)
											419	28	35	17	3	24	8	4	9	2	12	1	1	7	8	Heuristic (L)
												435	48	22	8	11	15	3	2	1	18	0	1	4	3	Metaheuristic (M)
													257	34	3	2	13	1	2	0	3	1	0	1	3	Tabu Search (N)
														310	9	3	8	3	0	2	13	0	3	1	1	Simulated Annealing (O)
															84	0	1	0	0	0	4	0	0	2	1	Particle Swarm Optimization (P)
																225	1	5	8	9	7	3	5	7	6	Integer Programming (Q)
																	101	0	4	0	12	1	3	0	0	Quadratic Assignment Problem (R)
																		84	5	3	1	1	0	1	3	Lagrangian Relaxation (S)
																			136	2	6	1	1	5	7	Branch-and-Bound (T)
																				84	2	1	2	0	0	Linear Programming (U)
																					225	1	7	3	4	Traveling Salesman Problem (V)
																						83	7	4	6	Robust Optimization (W)
																							108	0	3	Computational Complexity (X)
																								84	9	Knapsack Problem (Y)
																									128	Dynamic Programming (Z)

Fig. 6. Correlation matrix of the most relevant keywords in the research field of combinatorial optimization. The numbers indicate how often the keywords appear together in publications. The darker the fields in the matrix are colored green, the more often these keywords appear together.

also often closely related to CO (Conforti et al., 2014). However, not every CO problem can be formulated as an integer programming problem, if its feasible region is infinite (Ibaraki, 1976). The scheduling problem is often used for decision support, for example, in the case of project scheduling in project management (Hartmann and Briskorn, 2010), personnel scheduling (van den Bergh et al., 2013), or maintenance scheduling (Froger et al., 2016).

The keywords with the highest mean publication year represent the most recent topics in CO. The two most recent of the specific optimization problems are robust optimization (mean publication year: 2014.0) and multiobjective optimization (2011.4). This is also demonstrated by the trends in the annual percentage increase in publications on these topics since 2015. The difficulties, that several objectives have to be optimized simultaneously and that not all parameters are known in advance, are often encountered when applying optimization techniques to real-world problems (Schmidt et al., 2019). The method of robust optimization includes several approaches to protect a decision maker against parameter ambiguity and stochastic uncertainty. Thereby, the manager must determine what it means for him to have a robust solution. Based on worst-case analysis, a solution is evaluated using the realization of the most unfavorable uncertainty (Gabrel et al., 2014). Multiobjective optimization involves optimizing multiple objectives at the same time by selecting a (Pareto)efficient solution that cannot be improved in one objective without worsening it in another objective. In recent years, the concepts of both areas have been combined into multiobjective robust optimization (Schmidt et al., 2019).

The most prominent general solution techniques for solving these CO problems are metaheuristics (435 appearances), heuristics (419), optimizations (272), and approximation algorithms (258) (cf. Fig. 5). Real-world CO problems are usually large and exact solution procedures are mostly inadequate. Hence heuristics are mainly used in practice to solve complex CO problems (Hertz and Widmer, 2003). In the past, typically specialized heuristics were developed. However, this approach changed over the years: more general (metaheuristics) and less specialized solution approaches emerged. The motivation here is that applying a metaheuristic to a specific problem or problem class requires less effort than developing a specialized heuristic from scratch (Gendreau and Potvin, 2005). For heuristics and metaheuristics, for certain inputs good solutions (i.e., close to the optimal solution of a problem) are determined, but it is often uncertain why the heuristics work well. In this context, approximation algorithms are helpful, which bring mathematical rigor to the study of heuristics. Thus, it can be proven how well a heuristic performs on all instances and an idea of the types of instances on which a heuristic does not perform well can be given (Williamson and Shmoys, 2011).

Furthermore, many different specific metaheuristics are included in the 8393 studies (cf. Fig. 5). Various taxonomies can be found in the literature to distinguish metaheuristics (see Zäpfel et al., 2010). To give one example: Hertz and Widmer (2003) distinguish metaheuristics by two principles, namely local search and population search. The authors define the local search methods as an intensive exploration of the solution space by moving from the current solution to another promising solution in the neighborhood at each iteration. In comparison, the population search consists of maintaining a variety of good solutions and combining them to produce better solutions. The three metaheuristics most frequently found in the 8393 studies on CO are among the classic examples of population search and local search: genetic algorithm (521 occurrences; population search) as well as simulated annealing (310; local search), and tabu search (257; local search). These three metaheuristics are also often considered jointly in studies having at least 34 shared occurrences (c.f.

Fig. 6). The ant colony optimization, which is so frequently covered in the top 20 most cited articles, is also among the most relevant specific metaheuristics with 175 appearances. The articles on ant colony optimization are most cited on average (55, cf. Fig. 5) followed by articles on particle swarm optimization (52) and articles on quadratic assignment problems (50). For more information on the chronological development of the research field of metaheuristics please refer to Sörensen et al. (2018). Information on the most important topics of countries, organizations and sources can be found in Table A3.

3.4.3. Application areas for combinatorial optimization

In addition to the analysis of keywords, which can contain a multitude of words, an analysis with single words has also been conducted with the MATLAB algorithm. In this case, the keywords are divided into individual words, that is, *combinatorial optimization*, for example, into *combinatorial* and *optimization*. This is done in order to identify important application areas for CO methods. Table A4 lists the single keywords that appear in at least 1% of publications on CO. In this table, the words that are identified as specific real-world application areas of CO (and not specific problems or methods as in Fig. 5) are written in bold and the most relevant keywords and sources for the articles of these application areas are shown. The articles of a potential application area were checked by a manual inspection. For example, *scheduling*, *traveling*, or *vehicle* could represent application areas, but these keywords refer almost exclusively to the CO problems scheduling problem, TSP or the vehicle routing problem, respectively.

A total of 12 application areas were identified (cf. keywords written in bold in Table A4). The following four areas are covered most frequently (in 2% of all 8393 articles each): *production*, *data*, *power*, and *management*. The research field of *production*, which is about *production planning* (e.g., Shishvan and Sattarvand, 2015) or *assembly* (e.g., Becker and Scholl, 2009), for example, is also strongly linked to the application field *manufacturing* (e.g., Amen, 2006). Companies nowadays operate in global production networks (Lanza et al., 2019), which is the consequence of intense off-shoring, outsourcing, global procurement, and expansion into new international markets. Therefore, the global production network of a typical multinational manufacturing company today includes plants spread across the globe, each facing increasing pressure to coordinate its operations with one another and with the rest of the supply chain (Ferdows et al., 2016). Research in the field of CO on the application area of production also has a comparatively high mean publication year (2011.1), which demonstrates the increasing need for novel problem solutions in this area.

The research field of *data* is linked to *information* and these two application areas mainly focus on *data mining* (e.g., Brandner et al., 2013) or *information theory* (e.g., Braun et al., 2017), respectively. New approaches are needed in these areas, as the flood of data in recent decades have exceeded the ability to process, analyze, store, and understand the data sets. A good example is web pages whose number has increased from 1 million to 1 trillion between 1998 and 2008 alone (Fan and Bifet, 2013). The associated increasing use of data mining technologies also has a direct influence on the application field of information: for example, the increasingly emerging approaches of privacy-preserving data mining aim to protect sensitive information of individuals (Xu et al., 2014).

Big data is also related to the energy sector, which is increasingly using smart meters (sensor and measurement devices in smart grids) to collect data on real-time electricity consumption in order to better forecast and shift electrical loads (Wen et al., 2018). Energy sector related issues, which

are often NP-hard (Goderbauer et al. 2019), seem to be the most frequently represented subjects in publications on CO, with the keywords *power* and *energy*. The topic of *power* is mainly about the above-mentioned *smart grid* (e.g., Meskina et al., 2018) applications or *optimal power flow* (e.g., Abido, 2002) calculations. Relevant topics in the field of *energy* include *energy efficiency* (e.g., Al-harbi et al., 2019), and *energy consumption* (e.g., Weinand et al., 2019). In the course of the energy system transition to reduce anthropogenic greenhouse gas emissions, these issues are becoming increasingly important. This is also reflected by the mean publication year of the application area *energy* (2011.7), which is the most actual among the 12 identified application fields (cf. Table A4). Due to the developments in industrial production already discussed above, the greenhouse gas emissions increase further, therefore energy research is also increasingly connected with the application field of production. Due to improved technologies and companies, which are increasingly trying to make their production planning energy efficient, the emissions do not increase linear. Some reasons for this from a company's perspective are policy/legislation, scarcity of resources, rising energy prices and an increasing environmental awareness (Biel and Glock, 2016).

The last two major fields of application which are covered in more than 100 CO articles are *management* with many *supply chain management* studies (e.g., Mohammadi Bidhandi et al., 2009) and *decision* with mainly studies about *decision support systems* (e.g., Haastrup et al., 1998). Supply chain management studies are obviously strongly tied to production and decision support is actually needed in each of the application areas.

The analysis demonstrates the importance of CO methods, which are applied in many different fields and are used to solve many of the current global problems. In addition, it is notable that in many of these application areas, genetic algorithms are mainly used to solve the underlying problems. This metaheuristic already turned out to be the most relevant in the area of CO (cf. Section 3.4.2). Furthermore, this single-word analysis also confirms the presumptions that *robust optimization* is currently the most prominent topic on CO, since the two strongly related terms *robust* and *uncertainty* show the highest mean publication years 2013.8 and 2013.3, respectively (cf. Table A4). The consideration of uncertainties also becomes increasingly relevant for the real-world application areas discussed above like production planning (e.g., inhomogeneity of products, Mundi et al., 2019) or energy system analysis (e.g., stochastic nature of renewables and unknown future global energy and economy outlook, Mavromatidis et al., 2018).

4. Discussion

In this section, the results are reflected in terms of the challenges and prospects of the research field, separated for (meta)heuristics (cf. Section 4.1) and exact algorithms (cf. Section 4.2). Furthermore, the limitations of the study are discussed in Section 4.3.

4.1. Challenges and prospects of (meta)heuristics

One overarching conclusion from the data presented in this paper is that a large majority of papers in the field of CO are still of the "problem-algorithm-results" type. In these papers, an algorithm for a specific CO problem is developed and tested, and in most cases demonstrated to perform well by

comparing it to other algorithms for the same problem. Several authors (Barr et al., 1995; Hooker, 1995; Kendall et al., 2016) have warned that this type of research, especially the competitive testing aspect, yields very little scientific *knowledge* beyond the anecdotal. It wastes enormous amounts of research time on "development" activities (polishing code, compiler tuning, etc.) necessary to achieve top-notch performance. Papers in which attempts are made to draw some generalizable conclusions on heuristics and metaheuristics exist (e.g., Watson et al., 2003; Santini et al., 2018), but still represent a marginal phenomenon.

The practice of deciding which papers to publish based on competitive testing (which has been called the "horse race") has also resulted in a considerable *publication bias*. Mainly positive results (algorithm *X* works well for problem *Y*) appear in the literature and negative results, which demonstrate that some type of heuristic does *not* work for some problem (category), are few and far between (although some exist, e.g., Sörensen and Schittekat, 2013). Combined with the fact that there are very few generally accepted protocols for testing and reporting on algorithmic results, most academic journals do not even require simple statistical tests to demonstrate that a "better performance" is significant in the statistical sense. This makes one wonder how well many of the results published in the literature would stand the scrutiny of independent replication and testing (see, e.g., Sörensen et al., 2019, where an independent replication was not able to confirm the authors' original performance claims).

Another observation is that there seems to be a widening divide between the communities on metaphor-based metaheuristics (also called "nature-inspired" metaheuristics, even though the inspiration for this category of metaheuristics now comes from sources that have little to do with nature), and the more traditional metaheuristics that are not based on some metaphor. In the field of metaphor-based metaheuristics, the decision on whether an algorithmic idea is valuable or not seems to hinge on the novelty of the metaphor that inspired it, with ever more outlandish metaphors being proposed (interior decoration, Gandomi, 2014; the FIFA world cup, Razmjooy et al., 2016; "intelligent" water drops, Hosseini, 2009; and—in a demonstration of spectacular opportunism— the spread of covid-19, Martínez-Álvarez et al., 2020).[2] Without implying that the field of traditional metaheuristics does not have its issues, one can only conclude that the scientific standards in the subfield of metaphor-based metaheuristics are particularly low. As an illustration, it is remarkable that one of the top-cited papers in this field is the paper introducing the harmony search algorithm (Geem et al., 2001) (cf. Section 3.3). This metaheuristic is supposedly based on musicians playing music together (a solution in harmony search lingo is called a "melody," for example), even though this algorithm has been unequivocally demonstrated (Weyland, 2010, 2015) to be a special case of evolution strategies, a metaheuristic that predates it by 30 years. The community project "Metaheuristics in the Large" (MitL) has recently formulated a framework that enables combinatorial assembly and comparison of metaheuristics, and thereby also addresses issues of reproducibility and scalability (Swan et al. 2020).

Another trend, that is perhaps difficult to glean from the bibliometric results in this paper is that there is an increasing focus on "rich" problems (i.e., problems with complex formulations that involve many specific constraints and objectives), often based on real-life applications. Without doubt, heuristic, and metaheuristic ideas have penetrated into the mindsets of practitioners that

---

[2]The interested reader is referred to the "EC Bestiary," a satirical compilation of all metaphor-inspired metaheuristics published in the literature (Campelo and Aranha, 2019).

develop software for real-life optimization, and increasingly, (meta)heuristic research is finding its way to practical application. As an illustration, (software) companies like PTV and ORTEC increasingly participate in conferences and write out challenges for researchers to solve their real-life problems (Kheiri et al., 2019).

Finally, researchers in the field of (meta)heuristics are increasingly reaching out to related fields, incorporating ideas and techniques to develop better optimization algorithms. The combination of heuristics with exact methods (often called "matheuristics") has developed into a field of its own. Combinations with constraint programming and machine learning are also increasingly being found. For a concise review of such combinations, please refer to Talbi (2016).

4.2. Challenges and prospects of exact algorithms

The main keywords concerning publications on exact CO algorithms, according to the data reported in this work, are *branch-and-bound* and *integer programming*. Also from the data, the number of publications related to exact algorithms is considerably less than those related to approximation algorithms. One reason for this is that many hard combinatorial problems of practical interest are NP-hard and only approximation methods are able to provide *good* solutions. However, exact algorithms for NP-hard problems that are able to solve a number of real network instances with millions of nodes to proven optimality, have recently been described (e.g., San Segundo et al., 2016; Walteros and Buchanan. 2020).

Another explanation for the relatively small number of publications concerning exact CO algorithms might lie in the already mentioned "problem-algorithm-results" structure of the papers. Typically, this line of work focuses on the study of the specific structure of problem instances and develops specialized new algorithms, which are then compared extensively with the current state-of-the-art approaches. Only algorithms that show a significant improvement over state-of-the-art get to be published in the top CO journals. In addition, it has been argued that exact algorithms should also be *certifying*, that is, provide an easily verifiable proof that the solution is correct (Gocht et al., 2020).

Notwithstanding, this "problem-algorithm-results" stream of research has shown some exciting improvements for some fundamental NP-hard/NP-complete problems, such as the maximum independent set problem and the Boolean satisfiability (SAT) problem. Specifically, for the latter and during the past 20 years, the progress on the algorithmic methods has been reported to have at least the same impact as the advances in hardware (Fichte et al., 2020). In addition, research on specific NP-complete problems has also led to the study of effective transformations between problems. Among the many examples of recent successful transformations, SAT modules of state-of-the-art constraint programming solvers can be pointed out (e.g., Zhou et al., 2015) as well as the SAT-based bounding functions used by state-of-the-art maximum clique solvers (Li et al., 2018). Complementary to the development of new exact algorithmic techniques are the numerous CO challenges that are organized periodically, such as XCSP3, PACE, DIMACS, and SAT competitions, where algorithms are compared against benchmarks of practical interest and the state-of-the-art is settled.

From the data available, it is also worth mentioning the study of methods that aim at speeding up the convergence of exact CO algorithms by reducing the problem instance exploiting structural properties. These techniques can be applied just once, in a preprocessing phase known as

kernelization. Alternatively, they can be applied at every node of the branching tree, denoted the *branch-and-reduce* paradigm. Both techniques have recently contributed to solve to proven optimality some hard combinatorial large-scale problem instances (Akiba and Iwata, 2016; Hespe et al., 2020). Future research in exact CO algorithms capable of solving large-scale real-life problems will definitely continue over time. Exciting developments are to be expected from new algorithmic branch-and-bound and branch-and-reduce techniques, but one may also look forward to new algorithmic frameworks, of which the recent *branch-cut-and-prize* framework for vehicle routing and other related problems are good examples (Pessoa et al., 2020).

Since these challenges and prospects refer mostly to the results of Section 3, other emerging issues are likely to be neglected. One example would be machine learning based approaches, which show promising results in different applications by making decisions that were otherwise made by handcrafted expert knowledge-based heuristics in a more principled and optimized way (Bengio et al., 2020). For example, in the original AlphaGo paper (Silver et al., 2016), a machine learning algorithm is first trained based on expert knowledge and refined in a further step using a reward signal from games of self-play (reinforcement learning). In general, machine learning is used in both exact and heuristic frameworks (Bengio et al., 2020).

4.3. Limitations of this study

As with all review studies, the evaluation of the topic in this bibliometric analysis depends on the type of search query. Most probably, a large number of studies cover CO problems, but do not use combinatorial optimization, neither as a keyword, nor in the title, abstract, or article body. The search among the Keywords Plus increases the chance that a representative sample of articles on CO has been investigated because these keywords are generated from titles and phrases that appear in the references of an article. This means, that on the basis of a specialized algorithm, a search across disciplines is performed for all the articles that have cited references in common (Clarivate, 2021). However, if these publications are also not assigned to CO via Keywords Plus in Web of Science, they are not included in the present analysis. Thus, a 100% comprehensive picture of CO cannot be given. For example, the search query "optimi*" results in about 1 million articles, some of which might involve CO problems. On the other hand, the objective is to analyze studies that explicitly deal with CO as a methodology and the analyzed corpus of literature is assumed to yield a representative sample of the CO research field. In addition, the search query was limited to peer-reviewed original research articles, as is common in bibliometric analyses. As a consequence, some important contributions in the research field, such as the *Handbook of Combinatorial Optimization* (Pardalos et al., 2013) or the *Handbook of Heuristics* (Martí et al., 2018), are not covered.

Furthermore, some methods use parameters that are set in a more or less arbitrary way, and different parameters may yield different results. Examples are the number of years in the trend analysis, or the maximum distance set for the Levenshtein distance. It is not possible to recognize all relevant keywords during the grouping process using distances such as the Levenshtein distance, because this always involves some interpretation. For example, Fig. 5 shows the keyword traveling salesman problem, which is used 225 times. The higher occurrence (cf. Table A4) of the words traveling (384) and salesman (349), however, suggests that the TSP is covered in more than 225 articles. However, not all articles can be grouped by an algorithm if significantly different keywords

are used. For example, unlike simulated annealing, which is always known under that name, large neighborhood search is also known as destroy-and-repair and ruin-and-recreate, but the automatic analysis does not (and cannot) group these concepts. Therefore, the keyword analysis in Section 3.4 is only to be understood as an indication of trends.

5. Summary and conclusions

More and more real-world problems are becoming highly complex and have to be solved with combinatorial optimization techniques, which have always been of great interest to the scientific community. Consequently, the number of publications on combinatorial optimization has increased exponentially between 1990 and 2019, amounting to 8393 at the time of this analysis. This requires a study of the corpus of literature to show the status quo and trends in research on combinatorial optimization. The present study therefore uses a bibliometric analysis, supported by the literature database Web of Science, the R-tool bibliometrix as well as a novel algorithm developed for keyword analyses.

Among the 85 contributing countries, the United States is the most important contributor with 1918 articles and the highest *h*-index (104), followed by China (1197 articles), with the highest annual number of publications since 2014, and France (676 articles). In general, the share of collaborative publications on combinatorial optimization studies is rather low, with the most collaborations between United States and China (146). The most productive organizations are the *French National Centre for Scientific Research* (235 articles), the *University of California System* with 162 articles as well as the highest *h*-index (37) and the *Chinese Academy of Sciences* (118 articles). Due to the low proportion of cross-country collaborations, organizations from the same country tend to collaborate, with the *University of Montreal* and *Polytechnique Montreal* recording the highest number of collaborations among institutes (42). Core sources on combinatorial optimization are EJOR, LNCS, and *Computers & Operations Research*, which published around 15% of the 8393 articles.

The analysis of the most relevant publications and author keywords shows that the majority of studies focuses on the development, extension, and application of metaheuristics. While in most cases genetic algorithms predominate (6% of the 8393 publications), metaheuristics are mostly tested or applied to the TSP. Among the most globally cited publications are several articles on ant colony optimization. It appears that in the past the development of metaheuristics required only a new type of metaphor to justify its development. In this context, there seems to be a widening divide between the communities on metaphor-based metaheuristics (also called "nature-inspired" metaheuristics) and the more traditional metaheuristics that are not based on some metaphor. However, this trend of algorithm orientation in metaheuristics seems to have partly changed to a more problem-oriented approach in recent years. In this context, the analysis of the keywords showed that combinatorial optimization problems are particularly relevant in real-world application areas in the energy sector, production sector, and data management. This is due to the need to solve complex problems related to global production networks, the reduction of greenhouse gas emissions in the course of energy system transition, and the increasing amount of big data and its processing. The most current topics in the research area of combinatorial optimization are uncertainties and the associated increasingly relevant methodology of robust optimization, which is also becoming

more and more important in the aforementioned application areas. A discussion of the challenges and prospects of the field further reveals that most articles are still of the "problem-algorithm-results" type, which requires large amounts of research time on "development" activities necessary to achieve top-notch performance.

The present bibliometric analysis demonstrates global research trends in combinatorial optimization. This study can therefore support the scientific community as well as policy makers in identifying relevant issues regarding the expanding and transforming combinatorial optimization research area and its real-world applications.

Acknowledgments

References

Abido, M.A., 2002. Optimal power flow using particle swarm optimization. *International Journal of Electrical Power & Energy Systems* 24, 7, 563–571.

Akiba, T., Iwata, Y., 2016. Branch-and-reduce exponential/FPT algorithms in practice: a case study of vertex cover. *Theoretical Computer Science* 609, 211–225.

Alharbi, F., Tian, Y.-C., Tang, M., Zhang, W.-Z., Peng, C., Fei, M., 2019. An ant colony system for energy-efficient dynamic virtual machine placement in data centers. *Expert Systems with Applications* 120, 228–238.

Amen, M., 2006. Cost-oriented assembly line balancing: model formulations, solution difficulty, upper and lower bounds. *European Journal of Operational Research* 168, 3, 747–770.

Aria, M., Cuccurullo, C., 2017. bibliometrix: an R-tool for comprehensive science mapping analysis. *Journal of Informetrics* 11, 4, 959–975.

Barr, R.S., Golden, B.L., Kelly, J.P., Resende, M.G.C., Stewart, W.R., 1995. Designing and reporting on computational experiments with heuristic methods. *Journal of Heuristics* 1, 1, 9–32.

Becker, C., Scholl, A., 2009. Balancing assembly lines with variable parallel workplaces: problem definition and effective solution procedure. *European Journal of Operational Research* 199, 2, 359–374.

Bengio, Y., Lodi, A., Prouvost, A., 2020. Machine learning for combinatorial optimization: a methodological tour d'horizon. *European Journal of Operational Research* 290, 405–421.

Biel, K., Glock, C.H., 2016. Systematic literature review of decision support models for energy-efficient production planning. *Computers & Industrial Engineering* 101, 243–259.

Bjorndal, M.H., Caprara, A., Cowling, P.I., Della Croce, F., Lourenço, H., Malucelli, F., Orman, A.J., Pisinger, D., Rego, C., Salazar, J.J., 1995. Some thoughts on combinatorial optimisation. *European Journal of Operational Research* 83, 2, 253–270.

Blum, C., 2005. Ant colony optimization: introduction and recent trends. *Physics of Life Reviews* 2, 4, 353–373.

Blum, C., Puchinger, J., Raidl, G.R., Roli, A., 2011. Hybrid metaheuristics in combinatorial optimization: a survey. *Applied Soft Computing* 11, 6, 4135–4151.

Blum, C., Roli, A., 2003. Metaheuristics in combinatorial optimization. *ACM Computing Surveys* 35, 3, 268–308.

Boykov, Y., Funka-Lea, G., 2006. Graph cuts and efficient N-D image segmentation. *International Journal of Computer Vision* 70, 2, 109–131.

Boykov, Y., Kolmogorov, V., 2004. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26, 9, 1124–1137.

Brandner, H., Lessmann, S., Voß, S., 2013. A memetic approach to construct transductive discrete support vector machines. *European Journal of Operational Research* 230, 3, 581–595.

Braun, G., Jain, R., Lee, T., Pokutta, S., 2017. Information-theoretic approximations of the nonnegative rank. *Computational Complexity* 26, 1, 147–197.

Burke, E.K., Gendreau, M., Hyde, M., Kendall, G., Ochoa, G., Özcan, E., Qu, R., 2013. Hyper-heuristics: a survey of the state of the art. *Journal of the Operational Research Society* 64, 12, 1695–1724.

Cacchiani, V., Malaguti, E., Toth, P., 2018. Applied combinatorial optimization. *European Journal of Operational Research*. Available at https://www.journals.elsevier.com/european-journal-of-operational-research/call-for-papers/applied-combinatorial-optimization (accessed 28 March 2020).

Campelo, F., Aranha, C., 2019. EC-BESTIARY: a bestiary of evolutionary, swarm and other metaphor-based algorithms. Available at https://fcampelo.github.io/EC-Bestiary/ (accessed 19 July 2020).

de Castro, L.N, von Zuben, F.J., 2002. Learning and optimization using the clonal selection principle. *IEEE Transactions on Evolutionary Computation* 6, 3, 239–251.

Clarivate (2021) KeyWords Plus generation, creation, and changes. Available at https://support.clarivate.com/ScientificandAcademicResearch/s/article/KeyWords-Plus-generation-creation-and-changes?language=en_US (accessed on 3 February 2021).

Conforti, M., Cornuéjols, G., Zambelli, G. (eds), 2014. *Integer Programming*, Springer International Publishing, Cham.

Cook, W., 2019. Computing in combinatorial optimization. In Steffen, B., Woeginger, G. (eds) *Computing and Software Science: State of the Art and Perspectives*, Springer International Publishing, Cham, pp. 27–47.

Deng, W., Yao, R., Zhao, H., Yang, X., Li, G., 2019. A novel intelligent diagnosis method using optimal LS-SVM with improved PSO algorithm. *Soft Computing* 23, 7, 2445–2462.

Donoho, D.L., Elad, M., 2003. Optimally sparse representation in general (nonorthogonal) dictionaries via 1 minimization. *Proceedings of the National Academy of Sciences* 100, 5, 2197–2202.

Donoho, D.L., Huo, X., 2001. Uncertainty principles and ideal atomic decomposition. *IEEE Transactions on Information Theory* 47, 7, 2845–2862.

Dorigo, M., Blum, C., 2005. Ant colony optimization theory: a survey. *Theoretical Computer Science* 344, 2–3, 243–278.

Dorigo, M., Gambardella, L.M., 1997. Ant colonies for the travelling salesman problem. *Biosystems* 43, 2, 73–81.

Dorigo, M., Maniezzo, V., Colorni, A., 1996. Ant system: optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)* 26, 1, 29–41.

Egghe, L., 2006. Theory and practise of the g-index. *Scientometrics* 69, 1, 131–152.

Eskandarpour, M., Dejax, P., Miemczyk, J., Péton, O., 2015. Sustainable supply chain network design: an optimization-oriented review. *Omega* 54, 11–32.

Fan, W., Bifet, A., 2013. Mining big data. *ACM SIGKDD Explorations Newsletter* 14, 2, 1–5.

Feo, T.A., Resende, M.G.C., 1995. Greedy randomized adaptive search procedures. *Journal of Global Optimization* 6, 2, 109–133.

Ferdows, K., Vereecke, A., de Meyer, A., 2016. Delayering the global production network into congruent subnetworks. *Journal of Operations Management* 41, 1, 63–74.

Fichte, J.K., Hecher, M., Szeider, S., 2020. A time leap challenge for SAT-solving. In Simonis, H. (ed.) *Principles and Practice of Constraint Programming: 26th International Conference, CP 2020*, Louvain-la-Neuve, Belgium, September 7–11, pp. 267–285.

Froger, A., Gendreau, M., Mendoza, J.E., Pinson, É., Rousseau, L.-M., 2016. Maintenance scheduling in the electricity industry: a literature review. *European Journal of Operational Research* 251, 3, 695–706.

Gabrel, V., Murat, C., Thiele, A., 2014. Recent advances in robust optimization: an overview. *European Journal of Operational Research* 235, 3, 471–483.

Gandomi, A.H., 2014. Interior search algorithm (ISA): a novel approach for global optimization. *ISA transactions* 53, 4, 1168–1183.

Garey, M.R., Johnson, D.S., 2009. *Computers and Intractability: A Guide to the Tof NP-Completeness* (27th edn). Freeman, New York.

Geem, Z.W., Kim, J.H., Loganathan, G.V., 2001. A new heuristic optimization algorithm: harmony search. *Simulation* 76, 2, 60–68.

Gendreau, M., Potvin, J.-Y., 2005. Metaheuristics in combinatorial optimization. *Annals of Operations Research* 140, 1, 189–213.

Gerkey, B.P., Matarić, M.J., 2016. A formal analysis and taxonomy of task allocation in multi-robot systems. *The International Journal of Robotics Research* 23, 9, 939–954.

Glover, F., 1989. Tabu search—part I. *ORSA Journal on Computing* 1, 3, 190–206.

Gocht, S., McBride, R., McCreesh, C., Nordström, J., Prosser, P. and Trimble, J., 2020. Certifying solvers for clique and maximum common (connected) subgraph problems. In Simonis, H. (ed.) *Principles and Practice of Constraint Programming: 26th International Conference, CP 2020*, Louvain-la-Neuve, Belgium, September 7–11, pp. 338–357.

Goderbauer, S., Comis, M., Williamowski, F.J.L., 2019. The synthesis problem of decentralizes energy systems is strongly NP-hard. *Computers & Chemical Engineering* 124, 343–349.

Goemans, M.X., Williamson, D.P., 1995. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)* 42, 6, 1115–1145.

Goldberg, D.E., 2012. *Genetic Algorithms in Search, Optimization, and Machine Learning* (30th edn). Addison-Wesley, Boston, MA.

Haastrup, P., Maniezzo, V., Mattarelli, M., Mazzeo Rinaldi, F., Mendes, I., Paruccini, M., 1998. A decision support system for urban waste management. *European Journal of Operational Research* 109, 2, 330–341.

Halim, A.H., Ismail, I., 2019. Combinatorial optimization: comparison of heuristic algorithms in travelling salesman problem. *Archives of Computational Methods in Engineering* 26, 2, 367–380.

Han, K.-H., Kim, J.-H., 2002. Quantum-inspired evolutionary algorithm for a class of combinatorial optimization. *IEEE Transactions on Evolutionary Computation* 6, 6, 580–593.

Hartmann, S., Briskorn, D., 2010. A survey of variants and extensions of the resource-constrained project scheduling problem. *European Journal of Operational Research* 207, 1, 1–14.

Hertz, A., Widmer, M., 2003. Guidelines for the use of meta-heuristics in combinatorial optimization. *European Journal of Operational Research* 151, 2, 247–252.

Hespe, D., Lamm, S., Schulz, C., Strash, D., 2020. WeGotYouCovered: the winning solver from the PACE 2019 Challenge, vertex cover track. In Bücker, H. M., Li, X. S., (eds) *2020 Proceedings of the SIAM Workshop on Combinatorial Scientific Computing*, Philadelphia, PA. Society for Industrial and Applied Mathematics, Philadelphia, PA, pp. 1–11.

Hirsch, J.E., 2005. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America* 102, 46, 16569–16572.

Hoffman, K.L., Padberg, M., Rinaldi, G., 2013. Traveling salesman problem. In *Encyclopedia of Operations Research and Management Science*, Vol. 1, Springer US, New York, pp. 1573–1578.

Hooker, J.N., 1995. Testing heuristics: we have it all wrong. *Journal of Heuristics* 1, 1, 33–42.

Hopfield, J.J., Tank, D.W., 1985. "Neural" computation of decisions in optimization problems. *Biological Cybernetics* 52, 3, 141–152.

Hosseini, H.S., 2009. The intelligent water drops algorithm: a nature-inspired swarm-based optimization algorithm. *International Journal of Bio-Inspired Computation* 1, 1/2, 71.

Ibaraki, T., 1976. Integer programming formulation of combinatorial optimization problems. *Discrete Mathematics* 16, 1, 39–52.

Jaqaman, K., Loerke, D., Mettlen, M., Kuwata, H., Grinstein, S., Schmid, S.L., Danuser, G., 2008. Robust single-particle tracking in live-cell time-lapse sequences. *Nature Methods* 5, 8, 695–702.

Johnson, M.W., Amin, M.H.S., Gildert, S., Lanting, T., Hamze, F., Dickson, N., Harris, R., Berkley, A.J., Johansson, J., Bunyk, P., Chapple, E.M., Enderud, C., Hilton, J.P., Karimi, K., Ladizinsky, E., Ladizinsky, N., Oh, T., Perminov, I., Rich, C., Thom, M.C., Tolkacheva, E., Truncik, C.J.S., Uchaikin, S., Wang, J., Wilson, B., Rose, G., 2011. Quantum annealing with manufactured spins. *Nature* 473, 7346, 194–198.

Jozefowiez, N., Semet, F., Talbi, E.-G., 2008. Multi-objective vehicle routing problems. *European Journal of Operational Research* 189, 2, 293–309.

Kendall, G., Bai, R., Błazewicz, J., de Causmaecker, P., Gendreau, M., John, R., Li, J., McCollum, B., Pesch, E., Qu, R., Sabar, N., Berghe, G.V., Yee, A., 2016. Good laboratory practice for optimization research. *Journal of the Operational Research Society* 67, 4, 676–689.

Kheiri, A., Dragomir, A.G., Mueller, D., Gromicho, J., Jagtenberg, C., van Hoorn, J.J., 2019. Tackling a VRP challenge to redistribute scarce equipment within time windows using metaheuristic algorithms. *EURO Journal on Transportation and Logistics* 8, 5, 561–595.

Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. *Science (New York, N.Y.)* 220, 4598, 671–680.

Knill, E., Laflamme, R., Milburn, G.J., 2001. A scheme for efficient quantum computation with linear optics. *Nature* 409, 6816, 46–52.

Lanza, G., Ferdows, K., Kara, S., Mourtzis, D., Schuh, G., Váncza, J., Wang, L., Wiendahl, H.-P., 2019. Global production networks: design and operation. *CIRP Annals* 68, 2, 823–841.

Levenshtein, V.I., 1966. Binary codes capable of correcting deletions, insertions, and reversals. *Cybernetics and Control Theory* 10, 8, 707–710.

Li, C.-M., Fang, Z., Jiang, H., Xu, K., 2018. Incremental upper bound for the maximum clique problem. *INFORMS Journal on Computing* 30, 1, 137–153.

Loiola, E.M., Abreu, N.M.M., Boaventura-Netto, P.O., Hahn, P., Querido, T., 2007. A survey for the quadratic assignment problem. *European Journal of Operational Research* 176, 2, 657–690.

Martí, R., Pardalos, P. M., Resende, M. G. C. (eds), 2018. *Handbook of Heuristics*. Springer International Publishing, Cham.

Martínez-Álvarez, F., Asencio-Cortés, G., Torres, J.F., Gutiérrez-Avilés, D., Melgar-García, L., Pérez-Chacón, R., Rubio-Escudero, C., Riquelme, J.C., Troncoso, A., 2020. Coronavirus optimization algorithm: a bioinspired metaheuristic based on the COVID-19 propagation model. Available at https://arxiv.org/pdf/2003.13633 (accessed 19 July 2020).

Mavromatidis, G., Orehounig, K., Carmeliet, J., 2018. A review of uncertainty characterisation approaches for the optimal design of distributed energy systems. *Renewable and Sustainable Energy Reviews* 88, 258–277.

Meskina, S.B., Doggaz, N., Khalgui, M., Li, Z., 2018. Reconfiguration-based methodology for improving recovery performance of faults in smart grids. *Information Sciences*, 454–455, 73–95.

Mladenović, N., Hansen, P., 1997. Variable neighborhood search. *Computers & Operations Research* 24, 11, 1097–1100.

Mohammadi Bidhandi, H., Yusuff, Mohd.R., Megat Ahmad, M.M.H., Abu Bakar, M.R., 2009. Development of a new approach for deterministic supply chain network design. *European Journal of Operational Research* 198, 1, 121–128.

Mundi, I., Alemany, M.M.E., Poler, R., Fuertes-Miquel, V.S., 2019. Review of mathematical models for production planning under uncertainty due to lack of homogeneity: proposal of a conceptual model. *International Journal of Production Research* 57, 15–16, 5239–5283.

Nagy, G., Salhi, S., 2007. Location-routing: issues, models and methods. *European Journal of Operational Research* 177, 2, 649–672.

Nouiri, M., Bekrar, A., Jemai, A., Niar, S., Ammari, A.C., 2018. An effective and distributed particle swarm optimization algorithm for flexible job-shop scheduling problem. *Journal of Intelligent Manufacturing* 29, 3, 603–615.

Pardalos, P M., Du, D-Z; Graham, R L. (eds), 2013. *Handbook of Combinatorial Optimization*. Springer, New York.

Pessoa, A., Sadykov, R., Uchoa, E., Vanderbeck, F., 2020. A generic exact solver for vehicle routing and related problems. *Mathematical Programming* 183, 1–2, 483–523.

Pillac, V., Gendreau, M., Guéret, C., Medaglia, A.L., 2013. A review of dynamic vehicle routing problems. *European Journal of Operational Research* 225, 1, 1–11.

Pop, P.C., 2020. The generalized minimum spanning tree problem: an overview of formulations, solution procedures and latest advances. *European Journal of Operational Research* 283, 1, 1–15.

Prodhon, C., Prins, C., 2014. A survey of recent research on location-routing problems. *European Journal of Operational Research* 238, 1, 1–17.

Rahmaniani, R., Crainic, T.G., Gendreau, M., Rei, W., 2017. The Benders decomposition algorithm: a literature review. *European Journal of Operational Research* 259, 3, 801–817.

Razmjooy, N., Khalilpour, M., Ramezani, M., 2016. A new meta-heuristic optimization algorithm inspired by FIFA World Cup competitions: theory and its application in PID designing for AVR system. *Journal of Control, Automation and Electrical Systems* 27, 4, 419–440.

San Segundo, P., Lopez, A., Pardalos, P.M., 2016. A new exact maximum clique algorithm for large and massive sparse graphs. *Computers & Operations Research* 66, 81–94.

Santini, A., Ropke, S., Hvattum, L.M., 2018. A comparison of acceptance criteria for the adaptive large neighbourhood search metaheuristic. *Journal of Heuristics* 24, 5, 783–815.

Schmidt, M., Schöbel, A., Thom, L., 2019. Min-ordering and max-ordering scalarization methods for multi-objective robust optimization. *European Journal of Operational Research* 275, 2, 446–459.

Shishvan, M.S., Sattarvand, J., 2015. Long term production planning of open pit mines by ant colony optimization. *European Journal of Operational Research* 240, 3, 825–836.

Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., Hassabis, D., 2016. Mastering the game of go with deep neural networks and tree search. *Nature* 529, 7587, 484–489.

Sörensen, K., 2015. Metaheuristics-the metaphor exposed. *International Transactions in Operational Research* 22, 1, 3–18.

Sörensen, K., Arnold, F., Palhazi Cuervo, D., 2019. A critical analysis of the "improved Clarke and Wright savings algorithm. *International Transactions in Operational Research* 26, 1, 54–63.

Sörensen, K., Schittekat, P., 2013. Statistical analysis of distance-based path relinking for the capacitated vehicle routing problem. *Computers & Operations Research* 40, 12, 3197–3205.

Sörensen, K., Sevaux, M., Glover, F., 2018. A history of metaheuristics. In Martí, R., Pardalos, P. M., Resende, M. G. C. (eds) *Handbook of Heuristics*, Springer International Publishing, Cham, pp. 791–808.

Spiliopoulos, K., Sofianopoulou, S., 2007. Calculating distances for dissimilar strings: the shortest path formulation revisited. *European Journal of Operational Research* 177, 1, 525–539.

Stützle, T., Hoos, H.H., 2000. MAX-MIN ant system. *Future Generation Computer Systems* 16, 8, 889–914.

Swan, J., Adriaenson, S., Brownlee, A.E.I., Johnson, C.G., Kheiri, A., Krawiec, F., Merelo, J.J., Minku, L.L., Özcan, E., Pappa, G.L., Garcia-Sanchez, P., Sörensen, K., Voß, S., Wagner, M., White, D.R., 2020. Towards Metaheuristics "in the Large". Available at https://arxiv.org/abs/2011.09821 (accessed 30 November 2020).

Taillard, E., 1993. Benchmarks for basic scheduling problems. *European Journal of Operational Research* 64, 2, 278–285.

Talbi, E.-G., 2016. Combining metaheuristics with mathematical programming, constraint programming and machine learning. *Annals of Operations Research* 240, 1, 171–215.

Tero, A., Takagi, S., Saigusa, T., Ito, K., Bebber, D.P., Fricker, M.D., Yumiki, K., Kobayashi, R., Nakagaki, T., 2010. Rules for biologically inspired adaptive network design. *Science (New York, N.Y.)* 327, 5964, 439–442.

UNESCO (2020) *Science,technology and innovation: Gross domestic expenditure on R&D (GERD), GERD as a percentage of GDP, GERD per capita and GERD per researcher*. Available at http://data.uis.unesco.org/Index.aspx?DataSetCode=SCN_DS&lang=en (accessed 22 April 2020).

van den Bergh, J., Beliën, J., Bruecker, P., Demeulemeester, E., Boeck, L. (2013. Personnel scheduling: a literature review. *European Journal of Operational Research* 226, 3, 367–385.

Walteros, J.L., Buchanan, A., 2020. Why is maximum clique often easy in practice? *Operations Research* 68, 6, 1866–1895.

Watson, J.-P., Beck, J.C., Howe, A.E., Whitley, L.D., 2003. Problem difficulty for tabu search in job-shop scheduling. *Artificial Intelligence* 143, 2, 189–217.

Weinand, J.M., 2020. Reviewing municipal energy system planning in a bibliometric analysis: evolution of the research field between 1991 and 2019. *Energies* 13, 6, 1367.

Weinand, J.M., Kleinebrahm, M., McKenna, R., Mainzer, K., Fichtner, W., 2019. Developing a combinatorial optimisation approach to design district heating networks based on deep geothermal energy. *Applied Energy* 251, 113367.

Wen, L., Zhou, K., Yang, S., Li, L., 2018. Compression of smart meter big data: a survey. *Renewable and Sustainable Energy Reviews* 91, 59–69.

Wen, Z., Yin, W., 2013. A feasible method for optimization with orthogonality constraints. *Mathematical Programming* 142, 1–2, 397–434.

Weyland, D., 2010. A rigorous analysis of the harmony search algorithm. *International Journal of Applied Metaheuristic Computing* 1, 2, 50–60.

Weyland, D., 2015. A critical analysis of the harmony search algorithm—how not to solve Sudoku. *Operations Research Perspectives* 2, 97–105.

Williamson, D.P., Shmoys, D.B., 2011. *The Design of Approximation Algorithms*. Cambridge University Press, Cambridge.

Xia, Z., Xu, Z., Chen, M., Liu, Q., 2016. Recent developments in the new inorganic solid-state LED phosphors. *Dalton Transactions* 45, 28, 11214–11232.

Xiao, L., Boyd, S., 2004. Fast linear iterations for distributed averaging. *Systems & Control Letters* 53, 1, 65–78.

Xu, L., Jiang, C., Wang, J., Yuan, J., Ren, Y., 2014. Information security in big data: privacy and data mining. *IEEE Access* 2, 1149–1176.

Xu, P., Gu, Q., Wang, W., Wong, L., Bower, A.G.W., Collins, C.H., Koffas, M.A.G., 2013. Modular optimization of multi-gene pathways for fatty acids production in *E. coli*. *Nature communications* 4, 1409.

Yao, X., Liu, Y., Lin, G., 1999. Evolutionary programming made faster. *IEEE Transactions on Evolutionary Computation* 3, 2, 82–102.

Zäpfel, G., Braune, R., Bögl, M., 2010. Summarizing remarks on metaheuristics. In Zäpfel, G., Bögl, M., Braune, R. (eds) *Metaheuristic Search Concepts: A Tutorial with Applications to Production and Logistics*, Springer, Berlin, pp. 145–155.

Zhang, C., Zhou, K., Yang, S., Shao, Z., 2017. On electricity consumption and economic growth in China. *Renewable and Sustainable Energy Reviews* 76, 353–368.

Zhou, N.-F., Kjellerstrand, H., Fruhman, J., 2015. *Constraint Solving and Planning with Picat*. Springer, Cham.

Appendix

A.1. Publication distribution and collaboration of authors

Table A1 shows the top 20 of the most productive authors in the field of CO, with Panos M. Pardalos (50 publications), Gerhard J. Woeginger (46), and Fred Glover (40) at the top. The trend in the number of publications is particularly upward for F. Glover, J.K. Hao, K. Aihara, N. Mladenovic,

Table A1
The top 20 of the most productive authors in terms of publications on combinatorial optimization

| Author	Total publications	Trend	h-Index	g-Index	m-Index
P.M. Pardalos	50	↑	15	26	0.60
G.J. Woeginger	46	↗	13	21	0.48
F. Glover	40	↑↑↑	19	33	0.68
J.K. Hao	37	↑↑↑	16	30	0.67
M.G.C. Resende	36	↑	20	36	0.77
M. Hifi	27	↑	14	22	0.58
K. Aihara	25	↑↑↑	14	25	0.54
X. Yao	24	↑	12	24	0.41
S.M. Sait	23	↑↑	5	13	0.19
V.G. Deineko	22	↗	8	12	0.32
N. Mladenovic	22	↑↑↑	15	22	0.63
C.C. Ribeiro	22	↑	16	22	0.67
T. Stützle	22	↑↑	16	22	0.70
A.N. Letchford	21	↑↑	9	14	0.45
S. Onn	20	↑	8	13	0.30
W. J. Gutjahr	19	↑↑	15	19	0.71
A. Kasperski	19	↑	10	16	0.67
P.F. Stadler	19	↑↑↑	10	19	0.34
Z. Tang	19	↑↑	6	7	0.33
C. Blum	18	↑↑	10	18	0.56
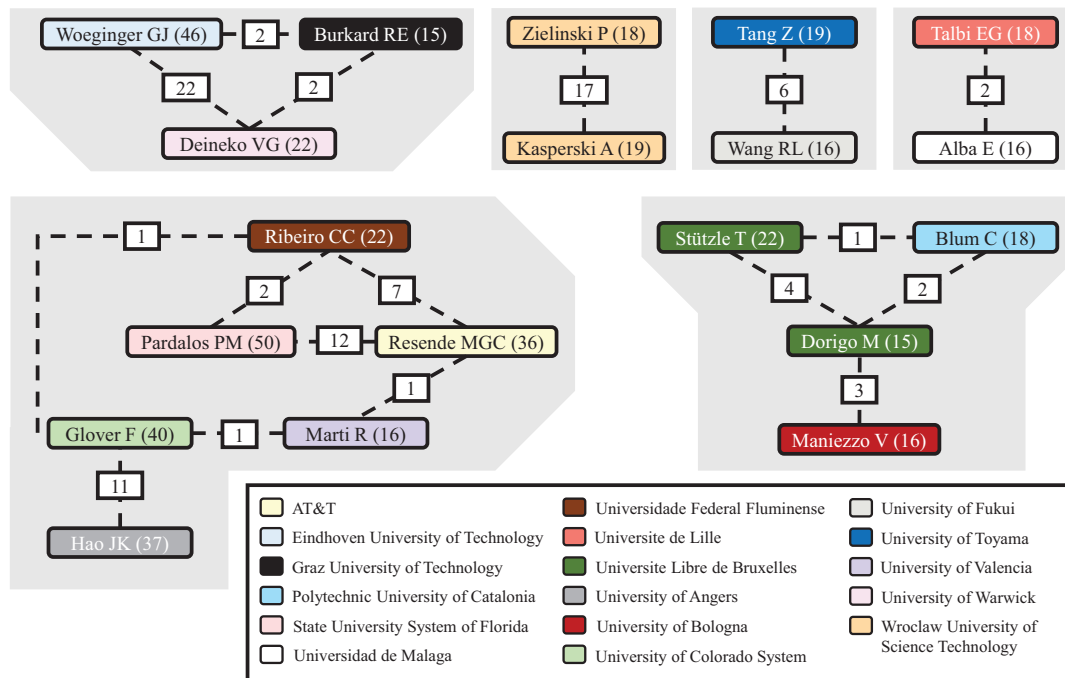
Fig. A1. Collaborations network of authors with at least 15 publications and at least one collaboration among these authors. In parentheses behind the authors is the total number of publications on CO, and in the boxes on the edges refer to the number of collaborations. The colors show the affiliation of the authors during the majority of their publications.

and P.F. Stadler. Interestingly, the fifth-ranked Mauricio G.C. Resende has the highest *h*-index, *g*-index, and *m*-index among the top 20, which means that his articles are cited a lot, that his most cited articles have very high citation rates, and that his articles are most cited in relation to the year of publication of the first article. Figure A1, which shows the collaborations of the top authors, also shows in which organizations the top authors have published most of their articles. It is evident that these authors are collaborating in the field of CO, albeit to a small extent. Most collaborations (22) took place between Gerhard J. Woeginger and Vladimir Deineko.

A.2. Most frequently cited articles

Please see Table A2.

A.3. Most important topics of countries, organizations, and sources

Table A3 shows the most common and therefore probably most important subject of each of the top five most relevant countries, organizations, and sources. These topics overlap almost entirely with the most relevant topics in all 8393 articles. Nevertheless, trends can be deduced for certain topics. In the CO publications from the United States and Germany, for example, the keyword approximation algorithm appears most frequently (in 5% of the publications in each case), while in

Table A2
Most frequently cited articles among the scientific contributions on combinatorial optimization

Article title	Global citations		Local citations	Publication year	Source
	No.	Per year	No.		
Ant system: Optimization by a colony of cooperating agents (Dorigo et al., 1996)	5646	235	242	1996	IEEE Transactions on Systems Man and Cybernetics Part B-Cybernetics
A scheme for efficient quantum computation with linear optics (Knill et al., 2001)	3485	183	0	2001	Nature
A new heuristic optimization algorithm: harmony search (Geem et al., 2001)	2662	140	47	2001	Simulation
An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision (Boykov and Kolmogorov, 2004)	2508	157	18	2004	IEEE Transactions on Pattern Analysis and Machine Intelligence
Evolutionary programming made faster (Yao et al., 1999)	1955	93	16	1999	IEEE Transactions on Evolutionary Computation
Variable neighborhood search (Mladenović and Hansen, 1997)	1680	73	143	1997	Computers & Operations Research
Optimally sparse representation in general (nonorthogonal) dictionaries via l(1) minimization (Donoho and Elad, 2003)	1530	90	4	2003	Proceedings of the National Academy of Sciences of the United States of America
Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming (Goemans and Williamson, 1995)	1503	60	156	1995	Journal of the ACM
MAX-MIN ant system (Stützle and Hoos, 2000)	1445	72	113	2000	Future Generation Computer Systems-the International Journal of Escience
Fast linear iterations for distributed averaging (Xiao and Boyd, 2004)	1371	86	2	2004	Systems & Control Letters

Continued

Table A2
Continued

Article title	Global citations		Local citations	Publication year	Source
	No.	Per year	No.		
Learning and optimization using the clonal selection principle (Castro and Zuben, 2002)	1287	72	21	2002	IEEE Transactions on Evolutionary Computation
Greedy randomized adaptive search procedures (Feo and Resende, 1995)	1197	48	164	1995	Journal of Global Optimization
Graph cuts and efficient N-D image segmentation (Boykov and Funka-Lea, 2006)	1143	82	3	2006	International Journal of Computer Vision
Benchmarks for basic scheduling problems (Taillard, 1993)	1135	42	89	1993	European Journal of Operational Research
Ant colonies for the travelling salesman problem (Dorigo and Gambardella, 1997)	1012	44	55	1997	Biosystems
Ant colony optimization theory: A survey (Dorigo and Blum, 2005)	1004	67	47	2005	Theoretical Computer Science
Uncertainty principles and ideal atomic decomposition (Donoho and Huo, 2001)	995	52	3	2001	IEEE Transactions on Information Theory
Robust single-particle tracking in live-cell time-lapse sequences (Jaqaman et al., 2008)	785	65	0	2008	Nature Methods
Quantum-inspired evolutionary algorithm for a class of combinatorial optimization (Han and Kim, 2002)	717	40	43	2002	IEEE Transactions on Evolutionary Computation
A formal analysis and taxonomy of task allocation in multi-robot systems (Gerkey and Matarić, 2016)	643	40	6	2004	International Journal of Robotics Research

Table A3
Most important topics in publications on combinatorial optimization in the top five relevant countries, organizations, and sources

Country, organization, or source	Main topics	Number of publications	
		No.	%
United States	Approximation algorithms	104	5
	Heuristics	97	5
	Genetic algorithms	81	4
China	Genetic algorithms	84	7
	Ant colony optimization	47	4
	Local search	43	4
France	Metaheuristics	52	8
	Heuristics	45	7
	Branch and bound	34	5
Germany	Approximation algorithms	29	5
	Integer programming	20	3
	Simulated annealing	17	3
Japan	Genetic algorithms	57	10
	Neural network	37	7
	Traveling salesman problem	23	4
French National Centre for Scientific Research (CNRS)	Metaheuristics	18	7
	Dynamic programming	17	7
	Scheduling	16	6
University of California System	Approximation algorithms	12	7
	Algorithms	10	6
	Dynamic programming	6	4
Chinese Academy of Sciences	Local search	7	6
	Inverse problem	4	3
	Evolutionary algorithms	4	3
National Institute for Research in Digital Science and Technology (INRIA)	Branch and bound	11	10
	Multiobjective optimization	8	7
	Integer programming/metaheuristics	7	6
Massachusetts Institute of Technology (MIT)	Approximation algorithms	13	13
	Heuristics/semidefinite programming/randomized algorithms	4	4
European Journal of Operational Research	Heuristics	93	15
	Metaheuristics	72	12
	Integer programming	64	10
Lecture Notes in Computer Science	Ant colony optimization/simulated annealing	6	2
	Linear programming/approximation algorithms	4	1
Computers & Operations Research	Metaheuristics	35	11
	Genetic algorithms	34	11
	Heuristics	21	7
Discrete Applied Mathematics	Approximation algorithms	13	8
	Integer programming	9	6
	Computational complexity	7	4

Continued

Table A3
Continued

Country, organization, or source	Main topics	Number of publications	
		No.	%
Annals of Operations Research	Metaheuristics	16	11
	Heuristics/integer programming	13	9
	Tabu search	12	8

Note: The percentage values refer to the total number of publications of the respective country, organization or source.

China and Japan genetic algorithms are most relevant (7% and 10%, respectively). Japanese authors also seem to be among the pioneers in the field of neural networks in CO, as they are involved in 33% of these studies. The research of the institutes logically shows similar tendencies as the research of their countries. It is also interesting to note that for the most relevant source EJOR a clear trend toward heuristics and metaheuristics can be seen (with at least 12% of the publications each) while the keywords in the second most relevant source *Lecture Notes in Computer Science* show no clear trend.

A.4. Single-word analysis

Table A4
The most relevant "single keywords" in publications on combinatorial optimisation divided into categories

Single word	Appearances		Mean year	Two most relevant keywords	Two most relevant journals
	No.	%			
Optimization	5239	62	2009.7		
Combinatorial	3654	44	2009.3		
Algorithms	2437	29	2009.9		
Problem	2318	28	2009.8		
Programming	1186	14	2009.6		
Search	1133	13	2009.5		
Network	1009	12	2008.4		
Scheduling	736	9	2009.9		
Genetic	683	8	2008.9		
Heuristic	674	8	2009.1		
Graph	600	7	2009.3		
System	557	7	2009.3		
Metaheuristic	542	6	2010.7		
And	470	6	2009.7		
Design	441	5	2009.4		
Annealing	426	5	2006.3		
Integer	422	5	2010.1		
Assignment	411	5	2008.8		

Continued

Table A4
Continued

Single word	Appearances		Mean year	Two most relevant keywords	Two most relevant journals
	No.	%			
Local	409	5	2010.2		
Analysis	395	5	2009.9		
Approximation	388	5	2009.4		
Traveling	384	5	2008.6		
Simulated	379	5	2006.2		
Methods	376	4	2008.9		
Computational	364	4	2010.0		
Constrained	364	4	2009.9		
Evolutionary	353	4	2010.7		
Multiobjective	351	4	2010.9		
Salesman	349	4	2008.2		
Of	344	4	2009.8		
Routing	318	4	2010.2		
Linear	316	4	2009.8		
Ant	307	4	2010.1		
Colony	307	4	2010.6		
Tabu	304	4	2007.5		
Quadratic	276	3	2009.7		
Model	276	3	2011.2		
Parallel	275	3	2007.5		
Dynamic	274	3	2010.1		
Neural	273	3	2004.7		
Functional	256	3	2009.5		
Set	249	3	2009.1		
Planning	247	3	2010.5		
Evolution	245	3	2011.5		
Theory	244	3	2008.8		
Knapsack	241	3	2010.2		
Complexity	238	3	2009.2		
Vehicle	228	3	2010.5		
Selective	215	3	2011.2		
Swarm	209	2	2010.7		
Computing	207	2	2010.5		
Production	202	2	2011.1	Genetic algorithm (23) and production planning (8)	*European Journal of Operational Research* (19) and *International Journal of Production Research* (10)
Relaxation	200	2	2008.5		
Optimal	198	2	2010.1		
Minimum	198	2	2010.6		
Data	191	2	2011.4	Data mining (31) and heuristics (12)	*European Journal of Operational Research* (14) and *Computers & Operations Research* (11)
Neighborhood	189	2	2010.8		
Learning	189	2	2011.6		

Continued

Table A4
Continued

Single word	Appearances		Mean year	Two most relevant keywords	Two most relevant journals
	No.	%			
Maximum	187	2	2009.1		
Stochastic	184	2	2009.7		
Tree	184	2	2009.7		
Bound	183	2	2009.1		
Allocation	180	2	2010.4		
Flow	172	2	2008.6		
Power	172	2	2009.3	Genetic algorithm (11) and smart grid (6)	*International Journal of Electrical Power & Energy Systems* (14) and *Electric Power Systems Research* (8)
Structural	170	2	2010.3		
Hybrid	168	2	2011.4		
Multiple	168	2	2010.1		
Distribution	166	2	2010.3		
Machine	165	2	2010.7		
Time	164	2	2010.0		
Location	150	2	2010.6		
Semidefinite	148	2	2008.4		
Particle	145	2	2010.5		
Path	145	2	2011.2		
Matching	144	2	2008.2		
Packing	142	2	2011.0		
Binary	141	2	2011.0		
Variable	141	2	2010.9		
Decomposition	141	2	2011.6		
Branch	135	2	2009.0		
Intelligent	135	2	2011.9		
Discrete	132	2	2010.8		
Management	130	2	2011.3	Genetic algorithm (15) and supply chain management (10)	*European Journal of Operational Research* (15) and *Computers & Operations Research* (7)
Spanning	129	2	2010.1		
Decision	125	1	2010.1	Decision support system (17) and decision making (17)	*European Journal of Operational Research* (16) and *Computers & Operations Research* (6)
Control	125	1	2011.2		
Shop	125	1	2010.8		
Energy	123	1	2011.7	Energy efficiency (14) and energy consumption (12)	*Applied Ocean Research* (6) and *IEEE Systems Journal* (5)
Assembly	123	1	2012.1		
Robust	122	1	2013.8		
Generation	122	1	2010.1		
Clustering	118	1	2010.2		
Artificial	116	1	2011.2		

Continued

Table A4
Continued

Single word	Appearances		Mean year	Two most relevant keywords	Two most relevant journals
	No.	%			
Lagrangian	113	1	2007.5		
Global	112	1	2008.0		
Random	111	1	2008.4		
Sequence	110	1	2010.0		
Mixed	109	1	2011.3		
Fuzzy	107	1	2007.4		
Resource	104	1	2012.4		
Strategic	103	1	2011.6		
Facility	101	1	2010.7		
Convex	99	1	2008.8		
Hopfield	99	1	2004.5		
Layout	99	1	2008.8		
Partitioning	99	1	2008.2		
Estimation	98	1	2011.3		
Modeling	96	1	2010.6		
Clique	94	1	2011.0		
Minimization	94	1	2011.6		
Adaptive	93	1	2010.3		
Matrix	93	1	2008.8		
Nonlinear	93	1	2009.1		
Information	93	1	2010.1	Information theory (9) and information entropy (5)	*European Journal of Operational Research* (2) and *IEEE Transactions on Wireless Communications* (2)
Cooperative	92	1	2011.0		
Image	92	1	2009.9	Image segmentation (19) and image processing (9)	*Journal of Visual Communication and Image Representation* (5) and *IEEE Transactions on Image Processing* (4)
Shortest	92	1	2011.1		
Memetic	91	1	2012.2		
Polynomial	91	1	2006.5		
Job	91	1	2012.0		
Chains	91	1	2010.4		
Polyhedral	90	1	2007.1		
Coloring	90	1	2010.2		
Transportation	89	1	2009.6	Transportation (22) and integer programming (10)	*European Journal of Operational Research* (12) and *Computers & Industrial Engineering* (4)
Branch-and-bound	88	1	2010.1		
Manufacturing	87	1	2009.8	Manufacturing (14) and genetic algorithm (13)	*European Journal of Operational Research* (13) and *International Journal of Production Research* (11)

Continued

Table A4
Continued

Single word	Appearances		Mean year	Two most relevant keywords	Two most relevant journals
	No.	%			
Vertex	87	1	2011.1		
Cover	87	1	2011.8		
Protein	86	1	2009.6	Protein design (19) and protein structure prediction (7)	*Journal of Computational Biology* (15) and *Journal of Computational Chemistry* (11)
Placement	86	1	2007.4		
Detection	86	1	2011.3		
Communication	85	1	2010.4	Telecommunications (9) and OR in telecommunications (8)	*European Journal of Operational Research* (8) and *IEEE Transactions on Wireless Communications* (7)
Differential	84	1	2012.5		
Cutting	84	1	2007.3		
Process	84	1	2009.2		
Uncertainty	83	1	2013.3		
Cut	83	1	2009.6		
Greedy	83	1	2008.5		
Automated	82	1	2010.9		
Matroid	82	1	2007.2		
Markov	81	1	2007.1		
Inverse	81	1	2010.6		
Balancing	81	1	2010.8		
Grasp	80	1	2009.4		
Simulation	80	1	2009.3		
Line	80	1	2011.7		
Technique	80	1	2008.9		

Note: The words that represent specific application areas of CO (and not problems or methods as in Figure 5) are written in bold and the two most relevant keywords and journals for the articles of these application areas is shown. The percentage values refer to the total of 8393 publications.