Wiley
Complexity
Volume 2025, Article ID 7605463, 17 pages
https://doi.org/10.1155/cplx/7605463

WILEY

Research Article

The Influence Maximization in Complex Networks: Significant Trends, Leading Contributors, and Prospective Directions

Elaf Adel Abbas,[1] **Raaid Alubady**,[2,3] **Aqeel Sahi**,[2,4,5] **Mohammed Diykh**,[2,6,7] **and Shahab Abdulla**[6]

[1]*College of Computer Science and Information Technology, University of Karbala, Karbala, Iraq*
[2]*Department of Cybersecurity, Technical College of Engineering, Al-Ayen Iraqi University, Nasiriyah 64001, Thi-Qar, Iraq*
[3]*College of Information Technology, University of Babylon, Babil, Iraq*
[4]*School of Mathematics, Physics and Computing, University of Southern Queensland, Toowoomba, Queensland, Australia*
[5]*College of Engineering, Al-Shatrah University, Thi-Qar, Iraq*
[6]*UniSQ College, University of Southern Queensland, Toowoomba, Queensland, Australia*
[7]*College of Education for Pure Science, University of ThiQar, Nasiriyah 64001, Thi-Qar, Iraq*

Correspondence should be addressed to Aqeel Sahi; aqeel.sahi@unisq.edu.au

Received 30 April 2025; Revised 2 October 2025; Accepted 15 November 2025

Academic Editor: Atila Bueno

Influence maximization (IM) is a concept in social network analysis and data science that focuses on finding the most influential nodes (people, users, etc.) in a network to maximize the spread of information, behavior, or influence. IM studies have become more crucial due to the quick uptake of social media and networking technologies, which have revolutionized communication and information sharing. Using information from the Scopus database, this study conducts a thorough bibliometric analysis of the literature on instant messaging from 2006 to 2024 to investigate publishing trends, significant contributors, and developing themes. The three primary issues the study attempts to answer are finding the most productive journals, nations, and scholars in IM research; assessing the growth and influence of publications; and predicting future research trends. The results show that IM research is dominated by China and the US, with significant contributions from organizations like the Department of Computer Science and Microsoft Research Asia. The development of the field toward scalable algorithms and practical applications is highlighted by highly cited articles, such as Chen's (2009) work on successful instant messaging. The investigation also shows the possibility of incorporating AI into future advancements and points out shortcomings in behaviorally informed techniques. This study offers a valuable summary of information management research for academics and professionals trying to understand this ever-evolving topic.

Keywords: bibliometric analysis; complex networks; influence maximization (IM); research trends; social network analysis

1. Introduction

The adoption of social media and various networking technologies available today has changed how people go about their lives and communicate with each other [1, 2]. Various networking platforms serve different purposes, such as sharing news updates or messages with others online for publicizing events or expressing viewpoints on common interests. Enabling users to interact and create connections online can lead to the spread of ideas and news items among communities, while influencing opinions effectively through these user-generated networks [3]. It has been widely recognized that not all users in a network hold importance or influence levels due to their activities or connections within the network. Some users may have a higher level of engagement and impact based on their interactions or friends in their network circle [4].

Influence maximization (IM) involves determining a group of individuals within a network who can effectively and widely impact users. This task is recognized as intricate since it necessitates a measure to quantify influence [5]. At first, Domingos and Richardson viewed the question of IM from the standpoint of an algorithmic formulation. The problem was first formulated as a discrete optimization problem by Kempe, Kleiberg, and Tardos [6], Domingos and Richardson (2001) [7], and Richardson and Domingos (2002) [8]. Given the fact that IM through social networks has great significance in the present times, although attempts have been made to resolve the IM problem, there are significant glaring gaps and niches in the domains of the problem that seek to know "who" or "who" are likely to be sensitive receivers of the message [9, 10] and/or who influences the recipients [11, 12] continue to be appreciated by scholars.

One of the most common methods for forming a social network representation relies on the graph's structure, where nodes represent users and edges reflect their interactions. After that, a user's influence can be quantified in numerous ways. Still, some scholars claim that this propensity for influence is rooted in the geometric nature of a comprised network, which is why some individuals are defined as influential based on their position within the structure [13].

Other studies also consider user behavior traits as preferences or trust and look for them within the network structure. While the former can be regarded as behavior-agnostic, the latter can be referred to as behavior-aware. One of the main concerns with behavior-agnostic approaches is that they lack the specificity of identifying target individuals based on behavioral patterns and instead employ a graph-based structural approach to locate target users. They have power (and additional complexity) as to how, in everyday situations, users' behavior can be regarded and used in matching them with various systems [14].

As far as the authors are aware, up to this study, there has been descriptive statistical analysis of the IM theme in the literature as cited, although there have been several recent works focusing on this aspect [15–17]. Additional bibliometric analysis is very appealing to scholars as it not only provides an overview of trends in the development of the discipline but also enables them to identify possible directions for future research. Besides, instead of being daunted by the number of published papers, junior researchers might be able to look for a niche to begin their explorations with the help of bibliometric analysis [18]. As a result, this study aims to conduct a bibliometric analysis of IM by addressing the following questions:

- Q1: Which journals have published the largest share of research articles on IM?
- Q2: Which countries are the most productive in IM research, and what are the most highly cited studies and most influential scholars in the field?
- Q3: What are the anticipated research directions and emerging trends in IM over the next decade?

2. Method

2.1. Data Source and Search Strategy. The study's objectives include understanding the research lines and exploring the IM problem in social network research. Due to its accuracy in defining and evaluating scientific publications, bibliometrics was utilized [19]. The Scopus database was also employed to collect data. There are two basic and primary sources for carrying out bibliometric analyses: Scopus and Web of Science (WoS). However, Scopus has some advantages, such as better accessibility to the public, more exhaustive coverage of content, and profiles of all authors and institutions, as well as serial sources. In addition, several studies have confirmed that Scopus provides more extensive coverage than the other and indexes more unique sources than WoS covers [20].

A total of 1288 publications regarding IM problems were collected. Articles published in multidisciplinary journals that mentioned IM problems using direct messaging but not related to computer science concepts or methodologies were excluded. Following the search, 1446 items were taken out for additional examinations.

Next, a search string was developed that comprised all documents where "IM problem" appeared anywhere in the title, abstract, or keywords as well as "influence propagation," "seed selection," "viral marketing," and "information diffusion." The primary keywords should now be sufficiently represented to attain the goal. They are not overly generalist or overemphasizing. The primary purpose is to treat IM in its appropriate sense, concept, and approach.

It is worth noting that during the first phase, 2734 IM-based studies were found in the Scopus databases, mainly from computer sciences and engineering. Other than the retrieved items, the documents covered publications from the year 2006 to the year 2024 (18 years), focusing on accountancy, mathematics, and computer science. In addition, other types of nonjournal papers, such as book chapters, trade reports, and conference proceedings, among others, have not been accounted for in this study. With this information, we set our parameters and define our search strings. We first limited our search to studies that discussed the issue of IM, and second, attempts were made to include studies that were purely relevant to IM. Nonetheless, it is disturbing that only 1446 items were returned, even though the flowchart in Figure 1 depicts the data collection procedure. We selected Scopus database as the primary data source for this bibliometric evaluation due to its broader journal coverage, particularly in the fields of computer science, engineering, and technology, which are dominated to IM research. Compared to Web of Science, Scopus contains a larger number of peer-reviewed journals, conference proceedings, and publications from emerging regions and disciplines. This guarantees a more comprehensive and inclusive dataset for detecting key trends, influential authors, and collaborative networks within the rapidly evolving landscape of complex network studies.

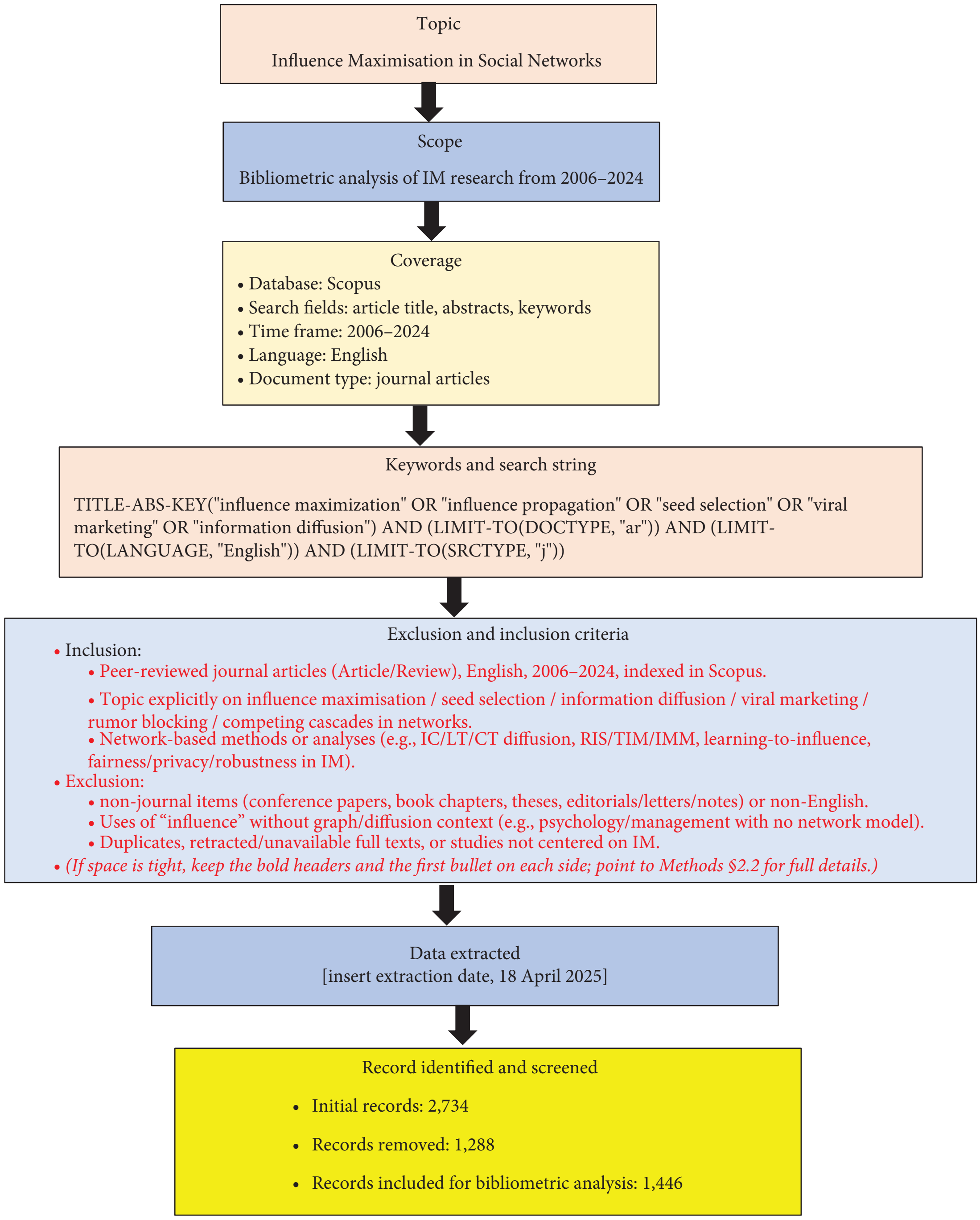

Figure 1: Overview of the data collection and screening process used in the bibliometric study on influence maximization in social networks, covering journal articles from 2006 to 2024.

3. Data Analysis

The first step involved applying bibliometric measures to the data in terms of quantity and quality [21]. While other measures like total citations (TC), number of cited publications (NCP), average citations per publication (C/P), and average citations per cited publication (C/CP) assessed the quality dimension, the total number of publications (TP) assessed the quantity dimension of publication output [22]. In addition, rather than using standard measurements, the g-index (g) and h-index (h) have also been included in bibliometric measures that determine the prospective forecast of future accomplishments. The indicators are used at several levels, including author, journal, organization, and country levels. Based on the RIS, the Scopus database was accessed to extract relevant research data, which was then processed and analyzed using Harzing's Publish or Perish (PoP) software. The second stage aimed at investigating the relationships within the constituents of the research, comprising authors and their affiliations (organizations), nations, their citations, and sketching keywords; hence, the research network was visualized. Only the term co-occurrence has been used in analyzing the potential of past, present, and future in the context of the IM problem. While our research is a bibliometric analysis rather than a systematic review of clinical or experimental tests, we used several aspects of preferred reporting items for systematic review and meta-analysis (PRISMA) to improve scientific transparency.

3.1. An Overview of the Collected Literature. The terms were analyzed based on their clusters, edges, and frequency. Two or more clusters containing various study themes are formed due to the arrangement of edges (the connections between keywords) and nodes (keywords) [23]. In the visual styles of the keywords, the bigger nodes are the ones that were most frequent, while the stronger associations are shown by thicker edges. At the same time, in the study's cluster analysis, it is possible to determine the combination of research interest and keywords within the group by grouping words with comparable meanings and differentiating them from those with different meanings in distinct groups. The scope of the work is further limited to the works that study IM and fall into the period between 2006 and 2024. For subsequent analysis, the study excluded books, editorials, review articles, and any publications except journal articles. Finally, 1446 papers have been taken up by the study. Nevertheless, the search encompassed only articles published in English. The papers concerned had 29,606 citations, or 1644 citations a year, and 18 for each published piece on average. Table 1 displays the citation metrics for each article.

In addition to IM as the most popular keywords as a main keyword were "social networks (online)" (14.5%), "economic and social impacts" (5.1%), "approximation algorithms" (5.0%), "maximization problem" (4.0%), "viral marketing" (3.0%), and others. Table 2 contains a list of the top keywords. Typically, high-frequency terms are well-liked within a specific subject [24]. The listed keywords are used together to maximize dissemination with social network research.

Table 1: Statistical information for published papers.

Metric	Value
Number of papers	1446
Number of years	18
Number of citations	29,606
The yearly average of citations	1644
The mean number of citations in each paper	20.47
g-index	149
h-index	75

Table 2: Top recurrent keywords and their corresponding total publication percentages across the dataset.

Number	Keywords	Total publication	Percentage
1	SOCIAL networking (online)	797	13.80%
2	Economic and social effects	284	4.90%
3	Approximation algorithms	277	4.80%
4	Maximization problem	220	3.80%
5	Viral marketing	209	3.60%
6	Marketing	208	3.60%
7	Greedy algorithms	177	3.00%
8	Heuristic algorithms	172	2.90%
9	Information diffusion	166	2.80%
10	Budget control	152	2.60%
11	Commerce	142	2.40%
12	Online social networks	132	2.20%
13	Influential nodes	126	2.10%
14	Optimization	126	2.10%
15	Information dissemination	118	2.00%
16	Diffusion	115	2.00%
17	Diffusion model	103	1.70%
18	Social influence	102	1.70%
19	Complex networks	97	1.60%
20	Algorithms	89	1.50%
21	Data mining	87	1.50%
22	Heuristic methods	87	1.50%
23	NP-hard	85	1.40%
24	Efficiency	71	1.20%
25	Linear threshold models	66	1.10%
26	Cascade model	65	1.10%
27	Real-world datasets	64	1.10%
28	Information propagation	63	1.00%
29	Propagation modeling	63	1.00%
30	Artificial intelligence	62	1.00%

3.2. Research Growth. Regardless of the rise and fall in research production during the year, it is noticeable that research productivity has increased over the years. We find that in the first four years from 2006 to 2010, publishing was limited and reached only 22 research papers, whereas in 2010, the total number of published research papers was only 22 research papers. The number of published research papers doubled in 2014 to about six times (reaching 66 research papers) and continued to increase until 2021, reaching 166 research papers. We also notice a decrease in the level of research production after this year, which is in the years 2022, 2023, and 2024. The average number of citations per publication was highest for publications published in 2007 (196.75). Nonetheless, the years 2017 (h-index 30) and 2018–2019–2020 (h-h-index 24) had the highest h-index, indicating a significant cumulative impact of the articles based on their quantity and quality. Since citation counts have been rising over time, low citations per publication in recent years were anticipated. Figure 2 shows the average number of citations per publication as well as the publishing trend.

3.3. Leading Nations, Institutions, and Authors in IM Research. With 673 and 353 articles, respectively, the United States and China were the most productive nations. The study found that the distribution of impact

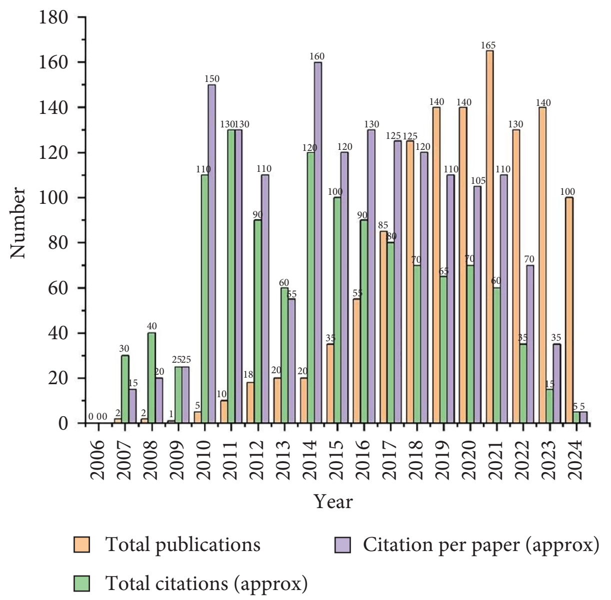

Figure 2: Annual trends in publications and citations related to influence maximization (2006–2024).

maximization research was roughly even between Asia and America, except for Europe, where research was unevenly distributed among the rest of the continents. The top 10 countries (see Table 3) included one North American country, six Asian countries, one African country, Australia, and the United States. The United States and China, the two most productive nations, both produced high-caliber publications with average citation counts per publication of 32.52 and 16.44, respectively. With average citations per publication of 84.23 and 93.81, respectively, Singapore and Canada published the fewest. Based on a rating of the top 10 most productive nations, Asian nations had the fewest citations: India (14.62) and Iran (13.94) were at the bottom of the list.

With the highest h-index and g-index and 74 published publications, the Department of Computer Science was the most productive institution (see Table 4). The number of publications exceeded that of the second-most productive department, the Department of Computer Science and Engineering (51). However, the top 10 list showed that Microsoft Research Asia had the highest average citations per publication with 580.12, followed by the School of Information Systems (229) and Microsoft Research (206). Three American universities were ranked as the most productive in the nation, accounting for 35.8% of all US articles. This indicates the distribution of the publications to other US institutions. Nevertheless, 51 articles were published from the Department of Computer Science and Engineering for the Indian Institute of Technology Bombay (45.5% of the country's total articles).

The two most prolific institutions carried out studies on distinct topics. Figure 3 shows that the Department of Computer Science's use influences maximization in numerous areas, such as "viral market," "community detection," "social network analysis," "seed selection," "information propagation," and others. Nonetheless, the Department of Computer Science and Engineering research mostly involved tourism; the keywords were "differential evaluation," "information dimension," and "information diffusion."

Table 5 shows the most productive authors who have published the most articles in the field of IM. The author with the most publications, Wu Weili, was the most prolific, followed by Chen Wei (32) from Singapore and Guo Jianxiong (23) from the United States. Although Chen Wei did not have the most research in the field of influence maximization, he had the highest number of citations as well as an h-index. Wu Weili and Guo Jianxiong, as the most productive authors, have similar research interests, which are in online social networks. Based on the study database, they had written 20 articles together using IM in social networks. Liu Wei and Chen Ling are experts in positive IM in social networks.

3.4. Top Active Journals. The journals were distributed among three publishing houses. Specifically, four of the 12 source titles were Elsevier journals; three were published by Springer Nature, three by IEEE, and one by Conference. Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics) were the most active source with 131 articles, followed by information sciences and IEEE Access with 38 and 33 articles, respectively. The average number of citations per publication is calculated as follows, Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (c/p = 177.07), followed by IEEE Transactions on Knowledge and Data Engineering (c/p = 53.82) and knowledge-based systems (c/p = 32.68). Table 6 lists the most active sources with research on social network IM together with their impact score (2023 SCImago Journal Rank (SJR) Citation Score).

3.5. Highly Cited Articles. Chen (2009) [25] published the most top-cited article in IM titled "Efficient IM in Social Networks." The publication had 1900 citations. The article also discussed another use of instant messaging, which is described in the section on "most active journals" of the organizational study. The study found that most of the top 15 articles were about social networks. At the same time, some papers also looked at the use of instant messaging in marketing, recommendation systems, and other areas. Table 7 lists the top 15 highly cited articles.

3.6. Co-Citation and Bibliographic Coupling Analysis. The VOSviewer was used to create the author-level and reference-level co-citation network, along with a bibliographic coupling analysis, to supplement keyword co-occurrence analysis. These visualizations advance topics influencing maximization research and offer a multilayered perspective on intellectual structure. A co-citation network of authors was constructed using VOSviewer. In Figure 4, each node in the graphic represents an author, the shapes of

Table 3: Bibliometric performance of top contributing countries in influence maximization research (2006–2024).

	Country	TC	TP	Cited publications	Average citations per publication	Average citations per cited publication	h-index	g-index
1	United States	8619	265	237	32.52	36.36	38	87
2	China	8140	495	388	16.44	20.97	39	79
3	Singapore	3201	38	34	84.23	94.14	19	38
4	Canada	2064	22	21	93.81	98.28	11	22
5	India	1638	112	92	14.62	17.80	24	38
6	Australia	1102	44	40	25.04	27.55	15	33
7	Japan	1059	34	30	31.14	35.30	15	32
8	Iran	1018	73	63	13.94	16.15	19	29
9	South Korea	815	27	25	30.18	32.60	11	27
10	Hong Kong	719	34	32	21.14	22.46	13	26

Note: Total publications (TP) refers to the number of publications analyzed for each country or entity. Cited publications denotes the number of publications that received at least one citation. Total citations (TC) denotes to the sum of all citations received. Average citations per publication is calculated as TC divided by TP. Average citations per cited publication is computed as TC divided by the number of cited publications. h-index is the number of publications (h) that have received at least h citations. g-index is the highest number (g) such that the top g articles received at least g^2 citations in total.

the nodes correspond to the frequency of co-citation, and the colors depict local intellectual communities by highlighting groups of authors whose works are frequently quoted together.

The study revealed a wide variety of groups: cluster 1 (red): theoretical model and algorithm efficiency (Wang X., Zhang J., Liu J., Liu X., and Shang J); cluster 2 (yellow): applied optimization and socioeconomic reference (Liy., He Z., Tan K.-L., Cai Q., and Sellis T); cluster 3 (green): integrated principle and application (Chen W., Xiao X., Thai M.T., and Song L); cluster 4 (blue): scalable algorithm for large-scale networks (Tang I., Motoda H., Muller E., Mahajan V., and Cordasco G); cluster 5 (purple): AI and multiagent application (Kleinberg J., Richardson M., and Walker D). This reflects the historical intellectual backbone of the network sector and highlights bridges to writers such as Chen W. and Tang J. Co-citation interactions between the primary references are depicted in Figure 5, where nodes stand in for the works, colors frequently depict groups of works that have argued together, and link thickness denotes co-element power. The analysis identifies three intellectual columns: the founding theoretical functions in social network analysis and influence propagation (Leskovec et al., Granovetter's threshold model, Barabasi–Albert network theory, and Goldenburg et al.'s diffusion study), as well as the studies conducted in the field of algorithms for maximizing the effect and the possibility of developing them, such as the studies that focus on efficiency, conducted by Chen W., Wang C., and Wang Y, and the algorithm optimization framework (Kempe, Kleinberg, Tardos).

The bibliographic coupling network's texts are shown in Figure 6, where the node size indicates the extent of shared contexts in recent publications and color groups authors that share book list profiles. According to this analysis, the largest cluster (Wu Veli, Tang Zueyan, and Guo Ziaxian) combined social network analytics and algorithm innovation, while the second cluster (Chen Wei and Tang Shogi) concentrated on big data references. Bridging figures like Tambe Milind connected AI, decision principle, and priority. When taken as a whole, these visualizations not only chart the region's historical underpinnings, remarkable functions, and thematic alignments but also show the multidisciplinary connections that are developing and the likely direction of future research.

4. Discussion

In this section, the findings were discussed as follows:

- The purpose of the study was to examine IM research on social network topics. The number of articles generally showed an upward trend, with a notable increase since 2005. Another notable increase was found in 2017, the second highest increase in the last 20 years, rising 50.60% from the previous year. This study demonstrated that a bibliometric examination of IM research delivered a similar tendency, with IM publications steadily rising since 2006. The highest average citations per publication were found in publications published in 2010 and 2014, according to quality measures. Furthermore, most citations in organizational science were found in many works published during those years. Chen's (2009) research shaded alight on the IM problem in social networks. Authors in that study addressed the problem by selecting a small group of users (seed nodes) to maximize the impact on users as much as possible. They proposed a new algorithm that achieved better results than greedy algorithms in speed, making it a unique solution for maximizing diffusion in complex networks.
- Chen (2009b) presents the MIA algorithm as a faster and more efficient alternative to greedy algorithms for maximizing influence in large-scale social networks. It relies on local influence structures (arborescence) instead of calculating global influence, presenting a faster model than traditional methods while achieving close to the same influence. The study

Table 4: Most influential institutions in the field of influence maximization (IM), based on total publications, citations, and bibliometric impact indicators between 2006 and 2024.

#	Institution	Country	TC	TP	Cited publications	Average citations per publication	Average citations per cited publication	h-index	g-index
1	Department of Computer Science	United States	2761	74	68	37.31	40.60	25	52
2	Department of Computer Science and Engineering	India	988	51	44	19.37	22.45	19	30
3	Department of Computer Engineering	United States	437	17	17	25.70	25.70	12	17
4	Microsoft Research Asia	China	4641	8	8	580.12	580.12	8	8
5	University of British Columbia	Canada	1100	7	7	157.14	157.14	6	7
6	Key Laboratory of Machine Perception	Singapore	1116	6	5	186.00	223.2	5	6
7	Dept. of Computer Science	Canada	870	5	4	174.00	217.5	3	5
8	Microsoft Research	United States	824	4	4	206.00	206	4	4
9	School of Computer Engineering	Singapore	1331	3	3	443.66	443.66	2	3
10	School of Information Systems	Singapore	458	2	2	229	229	2	2

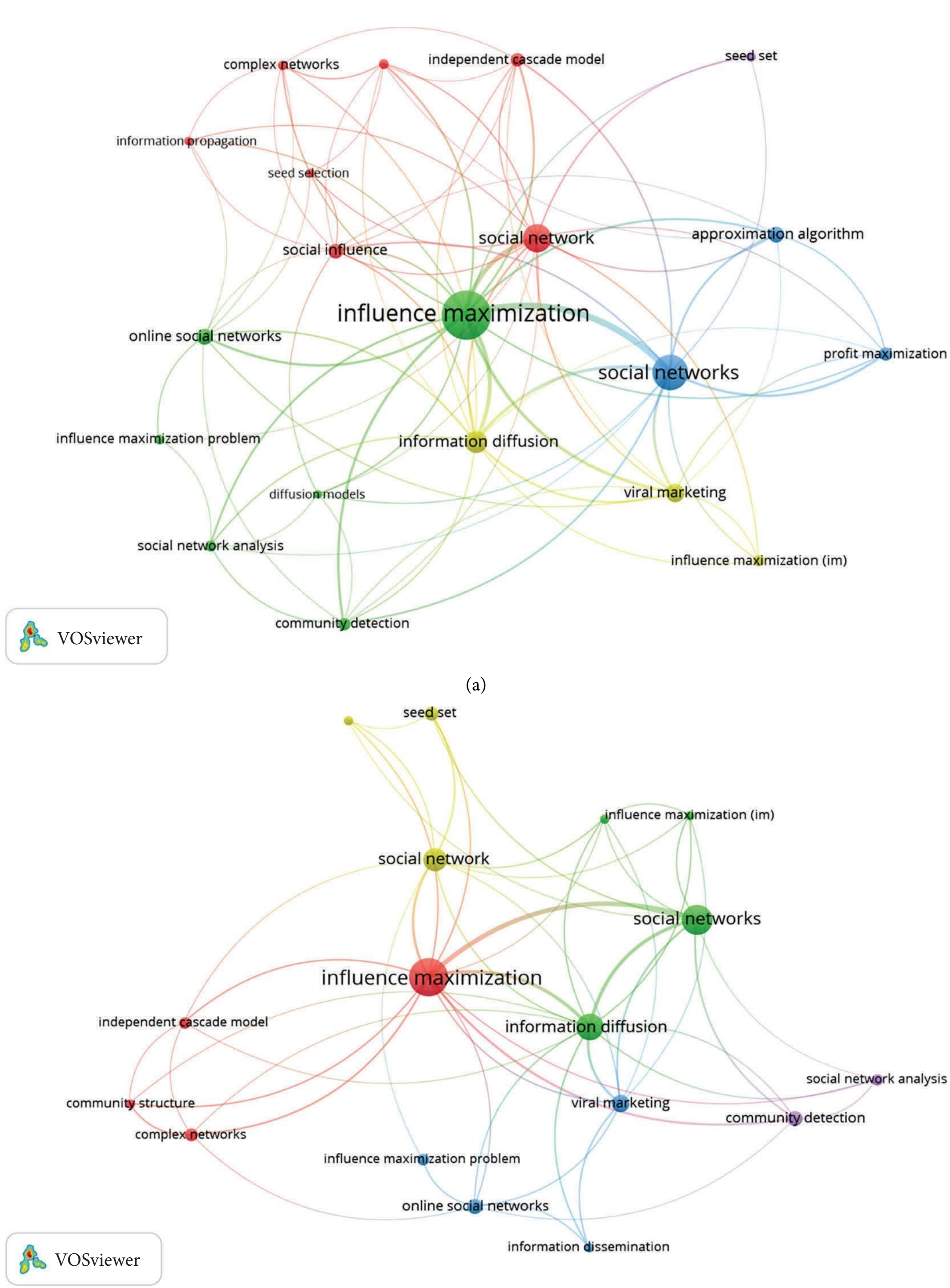

Figure 3: The topic keywords comparison between (a) Department of Computer Science and (b) Department of Computer Science and Engineering.

revealed Canada, Singapore, and the United States as the most productive countries. Furthermore, they discovered that Canada had led IM research, with Singapore coming in second and the US third. Asian institutions, including those in Japan, South Korea, Hong Kong, China, Taiwan, and Iran, scored lower than other institutions based on the average number of citations (18.80). For authors from Asian institutions, higher caliber and more significant research are required.

TABLE 5: Most productive authors in the field of influence maximization (2006–2024).

No.	Author	Affiliation	Country	Documents publication	Citations	h-index	g-index
1	Wu Weili	University of Texas at Dallas	United States	51	874	17	17
2	Chen Wei	School of Computer Science and Engineering	Singapore	32	6403	32	35
3	Guo Jianxiong	The University of Texas at Dallas	United States	23	316	13	13
4	Li Deying	University of Electronic Science and Technology of China	United States	18	140	7	7
5	Qiu Liqing	College of Computer Science and Engineering	United States	17	120	7	7
6	Thai My T.	University of Florida	United States	16	824	16	24
7	Liu Wei	College of Information Engineering	China	15	160	10	10
8	Tambe Milind	Harvard University	United States	15	259	15	15
9	Yang Wenguo	University of Chinese Academy of Sciences	China	15	82	5	5
10	Chen Ling	College of Information Engineering of Yangzhou University	China	14	151	10	10

Table 6: Most active sources published.

Sources	Documents	Citations	Average citations	Publishers	CiteScore 2023	SJR 2023
Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	131	1634	12.47328	Springer Nature	2.6	0.16
Information Sciences	38	1196	31.47368	Elsevier	14	2.238
IEEE Access	33	436	13.21212	IEEE	9.8	0.96
IEEE Transactions on Computational Social Systems	32	460	14.375	IEEE	10	1.716
Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining	27	4781	177.0741	Conference proceeding	7.5	1.004
Knowledge-Based Systems	25	817	32.68	Elsevier	14.8	2.219
IEEE Transactions on Knowledge and Data Engineering	23	1238	53.82609	IEEE	11.7	2.867
Physica A: Statistical Mechanics and Its Applications	22	563	25.59091	Elsevier	7.2	0.661
Social Network Analysis and Mining	22	307	13.95455	Springer Nature	5.7	0.667
Theoretical Computer Science	21	150	7.142857	Elsevier	2.6	0.57
Communications in Computer and Information Science	21	94	4.47619	Springer Nature	1.1	0.203

Table 7: The highly cited articles.

Authors	Titles	Citations
Chen (2009) [25]	Efficient Influence Maximization in Social Networks	1900
Chen (2010b) [26]	Scalable Influence Maximization for Prevalent Viral Marketing in Large-Scale Social Networks	1455
Goyal (2011c) [27]	CELF++: Optimizing the Greedy Algorithm for Influence Maximization in Social Networks	789
Chen (2010a) [28]	Scalable Influence Maximization in Social Networks Under the Linear Threshold Model	774
Borgs (2014) [29]	Maximizing Social Influence in Nearly Optimal Time	686
Tang (2014) [29]	Influence Maximization: Near-Optimal Time Complexity Meets Practical Efficiency	676
Tang (2015) [30]	Influence Maximization in Near-Linear Time: A Martingale Approach	654
Wang (2010) [31]	Community-Based Greedy Algorithm for Mining Top-K Influential Nodes in Mobile Social Networks	472
Goyal (2011a) [32]	SIMPATH: An Efficient Algorithm for Influence Maximization Under the Linear Threshold Model	468
Li (2018c) [33]	Influence Maximization on Social Graphs: A Survey	442
Bharathi (2007) [34]	Competitive Influence Maximization in Social Networks	405
Jung(2011) [35]	IRIE: Scalable and Robust Influence Maximization in Social Networks	353
He(2012b) [36]	Influence Blocking Maximization in Social Networks Under the Competitive Linear Threshold Model	349
Goyal (2011b) [37]	A Data-Based Approach to Social Influence Maximization	324
Nguyen (2011) [38]	Stop-and-Stare: Optimal Sampling Algorithms for Viral Marketing in Billion-Scale Networks	305

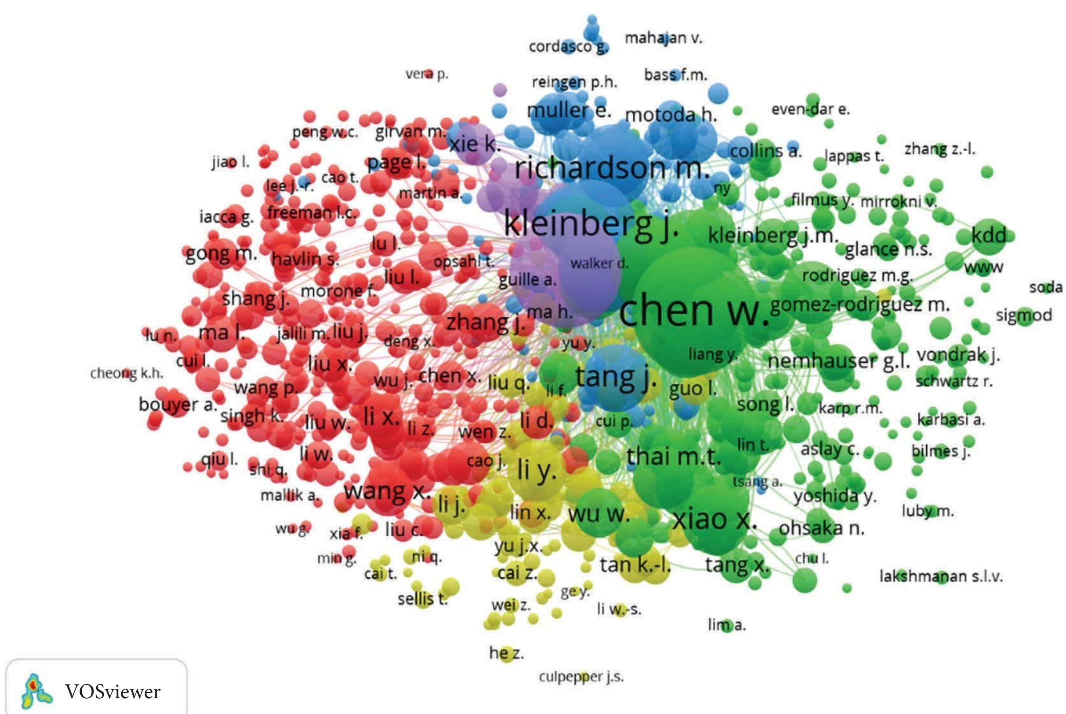

Figure 4: Author co-citation analysis, showing the frequency with which pairs of authors are cited together in the same documents.

- In the VOSviewer map shown in Figure 5, there were three tightly knit citation communities. A foundational cluster (green/yellow) is anchored by Kempe, Kleinberg, Tardos, and classic submodularity work (Nemhauser, Wolsey) which provides the theoretical basis for IM. A red, highly connected core around Leskovec, Krause, links diffusion modeling and network science (Barabási, Albert) with marketing diffusion studies (Goldenberg, Libai–Muller). To the right, a blue cluster dominated by Chen–Wang and collaborators reflects scalable IM algorithms (e.g., RIS/TIM/IMM, sketch-based methods), showing strong within-cluster cohesion and directed ties back to the theory core. Node size (citation weight) and edge thickness (co-citation strength) indicate that these groups form the intellectual backbone of the field and

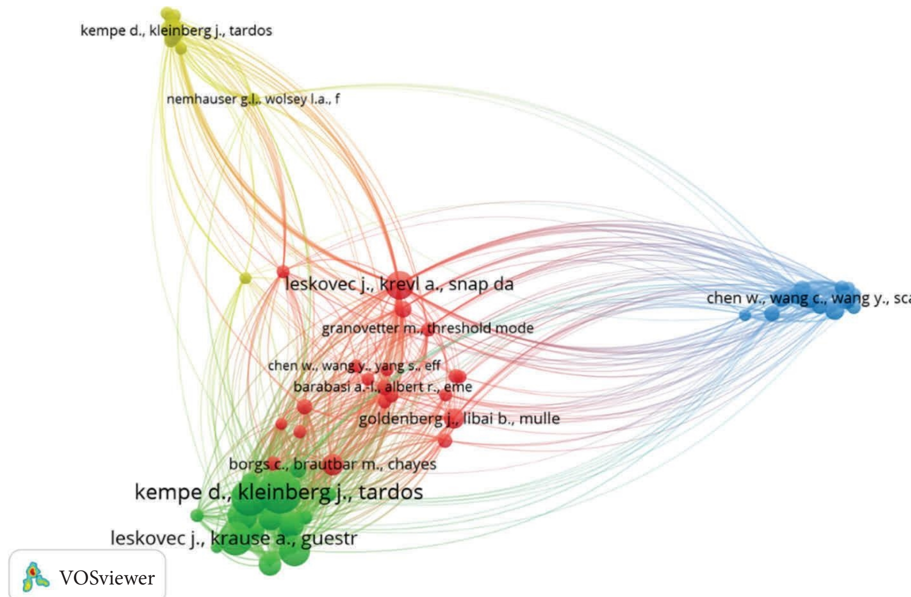

Figure 5: Reference co-citation network, illustrating the relationships between references that are frequently cited together within the analyzed corpus.

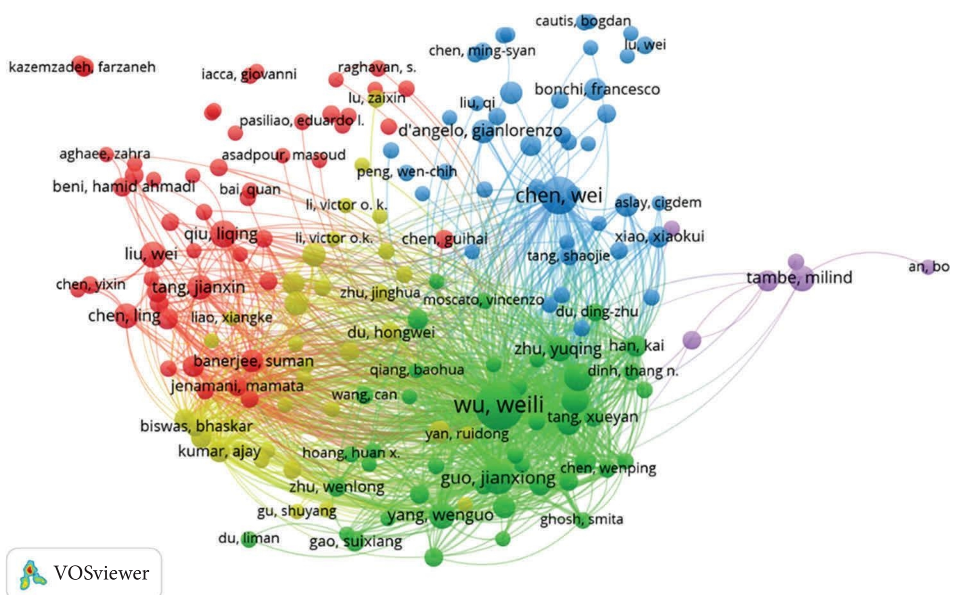

Figure 6: Author bibliographic coupling network, illustrating the strength of connections between authors based on shared references. Thicker links represent stronger coupling, indicating a higher degree of overlap in cited literature.

that approach advances in scalability are explicitly grounded in the earlier submodular framework.

- Figure 6 depicts the author's map. It reveals several collaboration communities with dense intragroup ties and sparser cross-group bridges. A large blue community led by Chen Wei exhibits high centrality and tight internal collaboration around scalable IM; a large green community cantered on Wu Weili and Guo Jianxiong indicates another active hub; a red community around Qiu Liying and Tang Jianxin points to learning-based, social computing lines; and a smaller purple cluster around Tambe, Milind proposes a focused stream (e.g., public interest/fairness), moderately connected to the main hubs. Overall, node size (productivity and impact) and proximity (collaboration strength) illustrated a mature field clustered around algorithmic scalability and diffusion modeling, with developing more specialized subcommunities that align with our review's thematic trends (learning, robustness, fairness, and application-driven IM).
- The outcomes of this bibliometric analysis have practical implications for both industry practitioners

and academic researchers. In particular, the increasing demand for developing scalable IM algorithms underlines the growing need for approaches that can be applied to large-scale, real-world social networks. Scalable algorithms are important in platforms such as LinkedIn, Facebook, and Twitter where millions of nodes (users) and edges (connections) exist. Effectively detecting key influences in these networks can significantly improve public health messaging, viral marketing, emergency response communication, and opinion shaping in elections or awareness campaigns. For example, adopting heuristic-based algorithms or scalable greedy models assist decision makers to simulate and optimize knowledge spread within limited computational resources and time. This made IM not just a theoretical problem but also a critical solution of high impact in the digital age. Moreover, our findings showed there are attempts toward deep learning-based graph neural network approaches, indicating future IM models may suggest both scalability and adaptiveness, further broadening their real-world applicability.

- While this study offers valuable insights into research trends, productivity, and influential contributions, several limitations should be acknowledged. First, citation bias can influence the observed impact of publications, as older published papers or those published in high impact journals are more likely to collect citations regardless of quality or significance. Second, depending on specific data sources, although comprehensive, may exclude relevant publications indexed exclusively in other databases such as IEEE Xplore or Google Scholar. This could lead to imperfect coverage, particularly for interdisciplinary or region-specific research. Furthermore, bibliometric approaches primarily extract quantitative patterns and may inspect methodological contributions that need qualitative assessment. These limitations emphasize the need to interpret the results with caution and, where possible, complement bibliometric findings with contents analysis.
- For further analysis, we conducted temporal keyword trend investigation. Figure 4 represents a heatmap to illustrate keyword trends. The temporal keyword analysis showed how research interests have changed over time, offering valuable insights into developing and declining topics within the field. By employing yearly co-occurrence heatmaps, it becomes possible to visualize not only the frequency of individual keywords but also their interrelationships across different years. The obtained results showed that there are shifts in research emphasis, for example, the terms like "deep reinforcement learning" and "competitive influence maximization" have been used noticeably in recent years, indicating emerging subfields. To improve readability and highlight the most relevant topics, we focused on the top 30 most frequent keywords across the entire period. This allowed us to clearly visualize the growth of research motivations over time and to identify emerging subtopics in IM.
- We conducted thematic evolution across three time slices (2007–2012, 2013–2018, and 2019–2024). Themes were gathered at each slice and synthetic link weights were allocated to simulate merging/splitting dynamics (e.g., IM $\longrightarrow$ information diffusion $\longrightarrow$ deep learning for influence; social network analysis $\longrightarrow$ community detection $\longrightarrow$ graph neural networks). Figure 7 reports a thematic evolution map based on top keywords from 2007 to 2024. For each period, the top eight keywords were selected as representative themes. Continuity between periods was established by matching identical keywords across time slices, with ribbon thickness proportional to the shared frequency. This mapping enables the visualization of stable core topics—such as IM and social networks and highlights how emerging areas (e.g., information diffusion and viral marketing) evolved or diminished over time. Figure 8 shows the heatmap of the temporal evolution of top research keywords.

4.1. Research Themes. The problem of maximizing influence in social networks is to identify a small group of users whose interaction leads to the maximum spread of information or products within the network [39]. Most studies rely on the graph structure, where nodes represent users and edges reflect their interaction, making influence related to the individual's position within the network. However, the actual spread of information depends not only on the geometric structure but also is affected by individuals' behavior, interaction patterns, and influence transfer probabilities. Therefore, the challenge is to design models that can select influencers effectively, taking into account factors such as dynamic interaction, shared interests, and different levels of influence to ensure maximum spread at the lowest possible cost.

4.2. Project Management. The first research paper that used network analysis in complex networks was published in 2001 by Domingos, the concept of seed set selection to spread influence in a network. This approach is based on selecting a small group of customers (or nodes in the network) that have the maximum ability to spread information or influence others within the social network. The research aims to analyze the value of customers based on their influence within the network, not just based on their characteristics. In addition, Chen (2009) has received the highest citations in the field of IM in social networks. Subsequently, the publication of IM in the social networks literature has increased significantly.

The study [40] provides a standard scientific analysis of the development of IM research in social and information networks, focusing on research trends, methods used, and the most influential institutions. 1026 research papers were collected and analyzed from the Scopus database for the period 2009–2023 using VOSviewer. The results show that research in this field has seen significant growth since 2015,

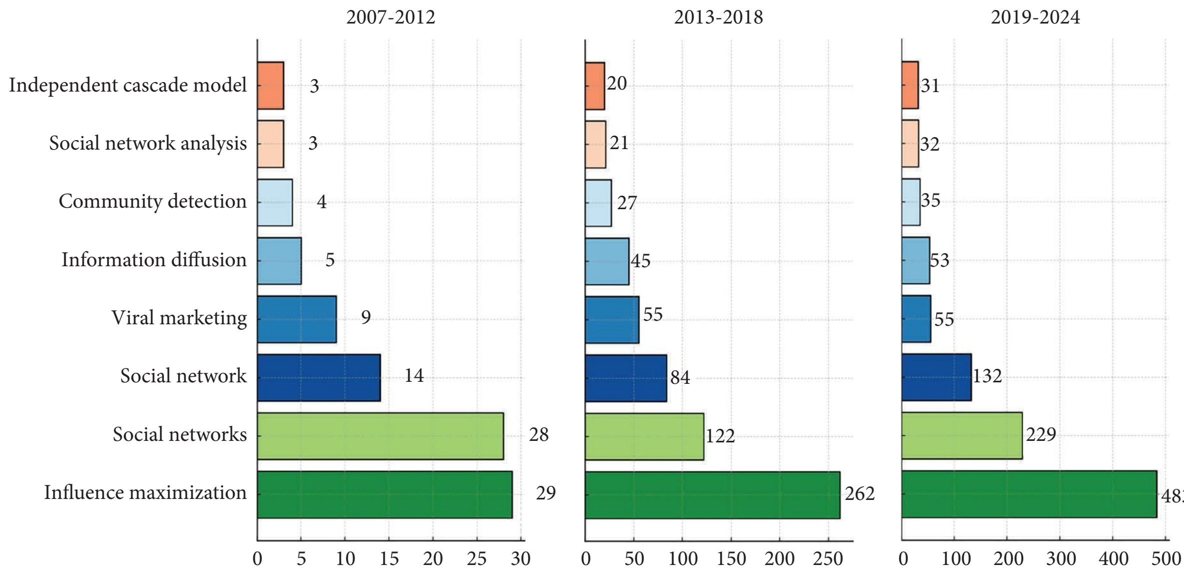

Figure 7: Thematic evolution map based on top keywords from 2007 to 2024.

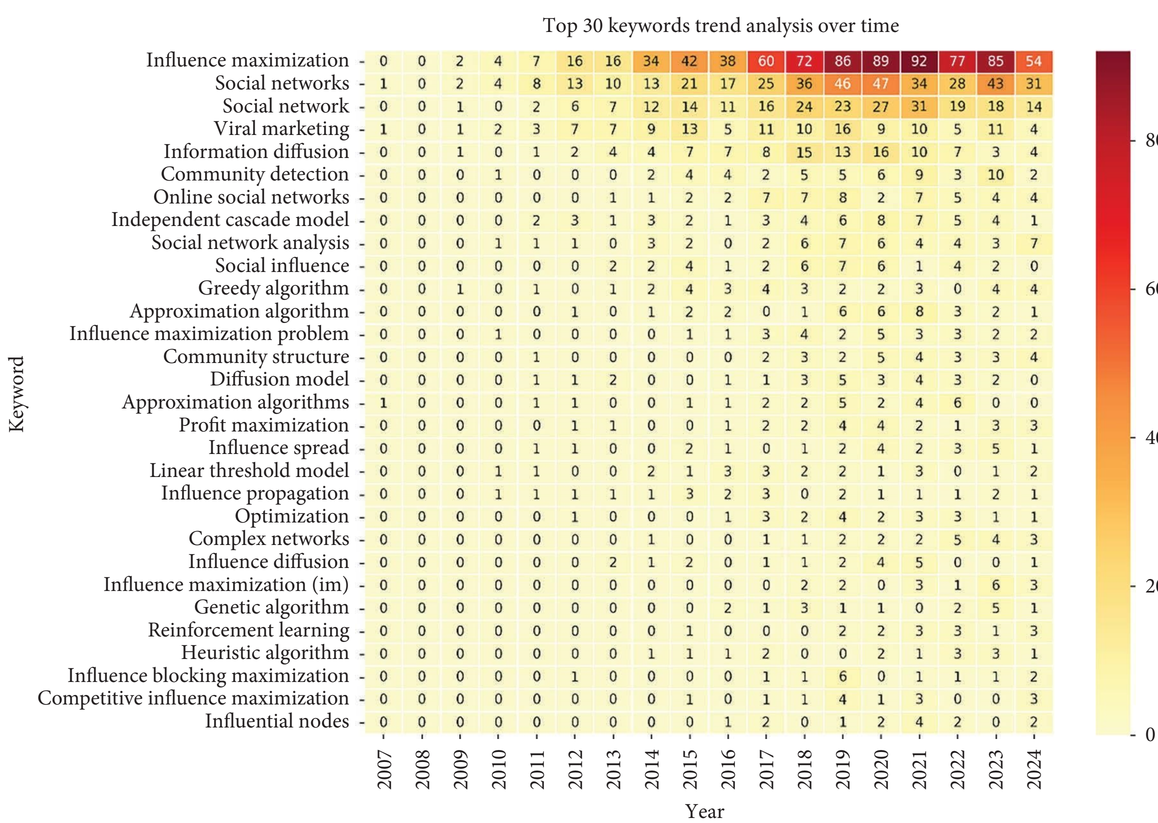

Figure 8: Heatmap showing the temporal evolution of top research keywords. Darker cells represent higher co-occurrence frequencies, indicating periods of strong thematic association.

with China, the United States, and India dominating scientific production. Recent developments are moving toward integrating artificial intelligence and deep learning to improve the efficiency of IM. The study recommends expanding research in developing countries and developing hybrid algorithms to enhance applications in various fields.

4.3. Social Networks Online. The IM approach in social network subjects was widespread, with "social network online" among the most frequent and top 30 keywords (see Table 2). Wu, Weili was the most productive author who used IM in social network research. The exponential increase in the use of social networks has drawn the attention of researchers to link prediction, community detection, and information diffusion. There are several potential uses for these, including social recommendation, rumor control, viral marketing, and income maximization [41–43].

Besides, the study discovered three significant journals with the most publications in social network research: Lecture Notes in Computer Science, Information Sciences, and IEEE Access, similar to [44, 45], whereby Lecture Notes in Computer Science was the most productive journal. Recently, the Information Sciences surpassed IEEE Access.

4.4. Social Network Analysis and Mining. The IM usage in social network analysis and mining varies; For example, Chin and others in [46] are used in viral marketing, including how a marketing campaign can achieve maximum reach at the lowest cost, while taking into account the important issue of time. IM [47] has also been applied in social network analysis to understand the mechanism of information diffusion within a network, in addition to controlling crisis management through information diffusion [48] and more. The top productive journals are the Social Network Analysis and Mining, Information Sciences, and IEEE Access.

4.5. Recommendation System Management. The IM is used for a recommendation system. Keywords under this subject include "viral marketing," "marketing," "information diffusion," and others. The study presented a group recommendation model that considers maximizing the impact of diffusion.

It is considered optimizing recommendations provided to the group, not just to include group members but also to have the potential to spread beyond the group. Meanwhile, the problem of maximizing influence also contributed to data poisoning, increasing attacks on recommendation systems, whereas the attack focused on influential users rather than regular users, increasing the impact of the attack at the lowest cost, as the targeted users are fewer in number due to their influence. Specifically, the top three journals were ACM SIGKDD International Conference, IEEE, and Information Sciences.

4.6. Future Work Direction. In this section, we highlight the future direction of our current study; we will (1) design seed selection and spread estimators for heterogeneous/multiplex networks and report advantages over single-layer baselines on real cross-platform data (Magnani and Rossi, 2013; Li et al., 2019); (2) move beyond static IC/LT assumptions to temporal and continuous-time diffusion, enabling staged seeding and deadline-aware objectives evaluated with latency/arrival metrics (Gomez-Rodriguez et al., 2011; Du et al., 2013); (3) develop robust, distributionally robust IM that optimizes worst-case spread under uncertainty in edges/probabilities and reports robustness–utility frontiers (He and Kempe, 2015); (4) apply graph learning and bandit/RL models for influence prediction and adaptive seeding in repeated campaigns (Qiu et al., 2018; Wang and Chen, 2017); (5) incorporate fairness constraints to guarantee minimum exposure for protected groups and publish Pareto curves for spread versus disparity (Tsang et al., 2019; Rahmattalabi et al., 2020); (6) enforce privacy and security via differentially private IM with formal utility bounds and resilience to poisoning/rumor attacks (Lyu et al., 2017; Fang et al., 2018); and (7) strengthen reproducibility through open, temporally annotated benchmarks and transparent seed/budget protocols grounded in submodular IM foundations (Kempe et al., 2003; Leskovec et al., 2007).

5. Conclusion

The development, state, and prospects of IM research in social networks are clarified by this bibliometric analysis. With important contributions from organizations and writers like Wu Weili and Chen Weili, the study emphasizes the dominance of China and the US in information management research. Publications have increased dramatically since 2006, indicating the expanding significance of information management in crisis management, recommendation systems, and viral marketing. Key findings highlight the necessity of behavior-aware models to handle dynamic user interactions and the significance of scalable algorithms, as demonstrated by highly cited works. Even with advancements, there are still gaps, especially when it comes to integrating deep learning and artificial intelligence to improve the effectiveness of information management. The study emphasizes how hybrid algorithms and underrepresented fields should be explored in future research. This study's mapping of the information management environment lays the groundwork for future research into specialized fields and innovation, ultimately leading to more efficient methods of disseminating information and influencing social networks.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

No funding was received for this manuscript.

References

[1] M. A. I. Gazi, M. A. Rahaman, M. F. Rabbi, M. Masum, M. N. Nabi, and M. R. B. S. Senathirajah, "The Role of Social Media in Enhancing Communication Among Individuals: Prospects and Problems," *Environment and Social Psychology* 9, no. 11 (2024): https://doi.org/10.59429/esp.v9i11.2979.

[2] Z. A. Abutiheen, E. A. Mohammed, and M. H. Hussein, "Behavior Analysis in Arabic Social Media," *International Journal of Speech Technology* 25, no. 3 (2022): 659–666, https://doi.org/10.1007/s10772-021-09856-6.

[3] E. A. Abbas, N. A. Hussein, and R. Alubady, "Machine Learning for Link Prediction Between Nodes in Complex Networks," *Fusion: Practice and Applications* 18, no. 1 (2025): 41–55, https://doi.org/10.54216/FPA.180104.

[4] D. Weber and F. Neumann, "Amplifying Influence Through Coordinated Behaviour in Social Networks," *Social Network Analysis and Mining* 11, no. 1 (2021): 111, https://doi.org/10.1007/s13278-021-00815-2.

[5] E. A. Abbas and H. N. Nawaf, "Influence Maximization Based on a Non-dominated Sorting Genetic Algorithm," *Karbala International Journal of Modern Science* 7, no. 2 (2021): 139–150, https://doi.org/10.33640/2405-609X.2891.

[6] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the Spread of Influence Through a Social Network," in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2003), 137–146, https://doi.org/10.1145/956750.956769.

[7] P. Domingos and M. Richardson, "Mining the Network Value of Customers," *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining* (2001): 57–66, https://doi.org/10.1145/502512.502525.

[8] *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM Digital Library, 2013).

[9] M. Li, Z. Wang, Q. L. Han, et al., "Influence Maximization in Multiagent Systems by a Graph Embedding Method: Dealing With Probabilistically Unstable Links," *IEEE Transactions on Cybernetics* 53, no. 9 (2023): 6004–6016.

[10] Y. Zhang, J. Guo, W. Yang, and W. Wu, "Supplementary Influence Maximisation Problem in Social Networks," *IEEE Trans Comput Soc Syst* 11, no. 1 (2023): 986–996.

[11] Q. Gong, Y. Chen, X. He, et al., "Cross-Site Prediction on Social Influence for cold-start Users in Online Social Networks," *ACM Transactions on the Web* 15, no. 2 (2021): 1–23, https://doi.org/10.1145/3409108.

[12] A. De Salve, P. Mori, B. Guidi, L. Ricci, and R. D. Pietro, "Predicting Influential Users in Online Social Network Groups," *ACM Transactions on Knowledge Discovery from Data* 15, no. 3 (2021): 1–50, https://doi.org/10.1145/3441447.

[13] M. N. Zareer and R. R. Selmic, "Modeling Interactions in Social Media Networks Using an Asynchronous and Synchronous Opinion Dynamics," *Social Network Analysis and Mining* 14, no. 1 (2024): 1–28, https://doi.org/10.1007/s13278-024-01402-x.

[14] F. Zhang, J. Tang, X. Liu, et al., "Understanding WeChat User Preferences and 'Wow' Diffusion," *IEEE Transactions on Knowledge and Data Engineering* 34, no. 12 (2022): 6033–6046, https://doi.org/10.1109/TKDE.2021.3064233.

[15] A. Zareie and R. Sakellariou, *Influence Maximisation in Social Networks: A Survey of behaviour-aware Methods* (Springer, 2023).

[16] Y. Ye, Y. Chen, and W. Han, *Influence Maximisation in Social Networks: Theories, Methods and Challenges* (Elsevier B.V, 2022).

[17] M. Jaouadi and L. Ben Romdhane, "Influence Maximisation Problem in Social Networks: an Overview," in *Proceedings of IEEE/ACS International Conference on Computer Systems and Applications, AICCSA* (IEEE Computer Society, 2019).

[18] A. E. Ezugwu, A. K. Shukla, R. Nath, et al., "Metaheuristics: a Comprehensive Overview and Classification Along with Bibliometric Analysis," *Artificial Intelligence Review* 54, no. 6 (2021): 4237–4316, https://doi.org/10.1007/s10462-020-09952-0.

[19] N. Carmona-serrano, J. López-belmonte, J. L. Cuesta-gómez, and A. J. Moreno-guerrero, *Documentary Analysis of the Scientific Literature on Autism and Technology in Web of Science* (MDPI AG, 2020).

[20] R. Pranckutė, *Web of Science (WoS) and Scopus: The Titans of Bibliographic Information in Today's Academic World* (Multidisciplinary Digital Publishing Institute (MDPI), 2021).

[21] J. E. Hirsch, "An Index to Quantify an Individual's Scientific Research Output," *Proceedings of the National Academy of Sciences of the United States of America*, 102, no. 46 (2005), 16569–16572, https://doi.org/10.1073/pnas.0507655102.

[22] J. E. Hirsch and M. Cardona, "Does the H Index Have Predictive Power?" *Proceedings of the National Academy of Sciences* 104, no. 49 (2007): 19193–19198.

[23] P. Dhamija and S. Bag, "Role of Artificial Intelligence in Operations Environment: a Review and Bibliometric Analysis," *The TQM Journal*, 32, no. 4 (Emerald Group Holdings Ltd, 2020), 869–896, https://doi.org/10.1108/TQM-10-2019-0243.

[24] B. Pesta, J. Fuerst, and E. O. W. Kirkegaard, "Bibliometric Keyword Analysis Across Seventeen Years (2000–2016) of Intelligence Articles," *Journal of Intelligence* 6, no. 4 (2018): 1–12, https://doi.org/10.3390/jintelligence6040046.

[25] *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM Digital Library, 2013).

[26] W. Chen, C. Wang, and Y. Wang, "Scalable Influence Maximisation for Prevalent Viral Marketing in large-scale Social Networks," in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2010), 1029–1038.

[27] "*Proceedings of the 20th International Conference Companion on World Wide Web*. ACM Digital Library" (2013).

[28] W. Chen, Y. Yuan, and L. Zhang, "Scalable Influence Maximisation in Social Networks Under the Linear Threshold Model".

[29] C. Borgs, M. Brautbar, J. Chayes, and B. Lucier, "Maximizing Social Influence in Nearly Optimal Time," (2014), https://epubs.siam.org/terms-privacy.

[30] Y. Tang, Y. Shi, and X. Xiao, "Influence Maximisation in near-linear Time: a Martingale Approach," in *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data* (2015), 1539–1554.

[31] *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM Digital Library, 2013).

[32] A. Goyal, W. Lu, and V. S. Lakshmanan, "SIMPATH: an Efficient Algorithm for Influence Maximisation Under the Linear Threshold Model".

[33] Y. Li, J. Fan, Y. Wang, and K. L. Tan, "Influence Maximization on Social Graphs: a Survey," *IEEE Transactions on Knowledge and Data Engineering* 30, no. 10 (2018): 1852–1872, https://doi.org/10.1109/TKDE.2018.2807843.

[34] S. Bharathi, D. Kempe, and M. Salek, "Competitive Influence Maximisation in Social Networks".

[35] K. Jung, W. Heo, and W. Chen, "IRIE: Scalable and Robust Influence Maximisation in Social Networks," (2011), http://arxiv.org/abs/1111.4795.

[36] X. He, G. Song, W. Chen, and Q. Jiang, "Influence Blocking Maximisation in Social Networks Under the Competitive Linear Threshold Model," https://epubs.siam.org/terms-privacy.

[37] A. Goyal, F. Bonchi, and L. V. S. Lakshmanan, "A Data-based Approach to Social Influence Maximisation," (2011), http://arxiv.org/abs/1109.6886.

[38] H. T. Nguyen, M. T. Thai, and T. N. Dinh, "Stop-and-Stare: Optimal Sampling Algorithms for Viral Marketing in billion-scale Networks," in *Proceedings of the ACM SIGMOD International Conference on Management of Data* (Association for Computing Machinery, 2016).

[39] G. Schoenebeck, B. Tao, and F.-Y. Yu, "Limitations of Greed: Influence Maximisation in Undirected Networks Re-visited" (2020), https://doi.org/10.5555/3398761.3398903.

[40] M. Venunath, P. Sujatha, P. Koti, and S. Dharavath, "The Evolution of Influence Maximisation Studies: a Scientometric Analysis," in *XVIII International Conference on Data Science and Intelligent Analysis of Information* (Springer, 2023).

[41] S. S. Singh, D. Srivastva, M. Verma, and J. Singh, *Influence Maximisation Frameworks, Performance, Challenges and Directions on Social Network: A Theoretical Study* (King Saud bin Abdulaziz University, 2022).

[42] Y. R. Al-Saadi, M. S. Tapou, A. A. Badi, S. Abdulla, and M. Diykh, "Developing Smart Self Orienting Solar Tracker for Mobile PV Power Generation Systems," *IEEE Access* 10 (2022): 79090–79099, https://doi.org/10.1109/access.2022.3194026.

[43] M. Diykh, S. Abdulla, A. Y. Oudah, H. A. Marhoon, and S. Siuly, "A Novel Alcoholic EEG Signals Classification Approach Based on Adaboost k-means Coupled with Statistical Model," in *International Conference on Health Information Science* (Cham: Springer International Publishing, 2021).

[44] A. Sahi, M. Albdair, M. Diykh, et al., "SGDM-GRU: Spectral Graph Deep Learning Based Gated Recurrent Unit Model for Accurate Fake News Detection," *Expert Systems with Applications* 281 (2025): 127572, https://doi.org/10.1016/j.eswa.2025.127572.

[45] A. Sahi, D. Lai, and Y. Li, "Three-Party Password-based Authenticated Key Exchange Protocol Based on the Computational Diffie-Hellman Assumption," *International Journal of Communication Networks and Distributed Systems* 21, no. 4 (2018): 560–581, https://doi.org/10.1504/ijcnds.2018.095373.

[46] A. Goyal, F. Bonchi, L. V. S. Lakshmanan, and S. Venkatasubramanian, "On Minimizing Budget and Time in Influence Propagation over Social Networks," *Social Network Analysis and Mining* 3, no. 2 (2013): 179–192, https://doi.org/10.1007/s13278-012-0062-z.

[47] J. Tang, X. Tang, and J. Yuan, "An Efficient and Effective Hop-based Approach for Influence Maximization in Social Networks," *Social Network Analysis and Mining* 8, no. 1 (2018): 10, https://doi.org/10.1007/s13278-018-0489-y.

[48] C.-H. Tsai, J.-K. Lou, W.-C. Lu, and S.-D. Lin, "Exploiting rank-learning Models to Predict the Diffusion of Preferences on Social Networks," in *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)* (IEEE, 2014).