


## Review

# Information Diffusion Model in Twitter: A Systematic Literature Review

Firdaniza Firdaniza <sup>1,\*</sup> , Budi Nurani Ruchjana <sup>1</sup> , Diah Chaerani <sup>1</sup>  and Jaziar Radianti <sup>2</sup>

<sup>1</sup> Department of Mathematics, Universitas Padjadjaran, Sumedang 45363, Indonesia; budi.nurani@unpad.ac.id (B.N.R.); d.chaerani@unpad.ac.id (D.C.)

<sup>2</sup> Department of Information Systems, University of Agder, 4630 Kristiansand, Norway; jaziar.radianti@uia.no

\* Correspondence: firdaniza@unpad.ac.id

**Abstract:** Information diffusion, information spread, and influencers are important concepts in many studies on social media, especially Twitter analytics. However, literature overviews on the information diffusion of Twitter analytics are sparse, especially on the use of continuous time Markov chain (CTMC). This paper examines the following topics: (1) the purposes of studies about information diffusion on Twitter, (2) the methods adopted to model information diffusion on Twitter, (3) the metrics applied, and (4) measures used to determine influencer rankings. We employed a systematic literature review (SLR) to explore the studies related to information diffusion on Twitter extracted from four digital libraries. In this paper, a two-stage analysis was conducted. First, we implemented a bibliometric analysis using VOSviewer and *R-bibliometrix* software. This approach was applied to select 204 papers after conducting a duplication check and assessing the inclusion–exclusion criteria. At this stage, we mapped the authors’ collaborative networks/collaborators and the evolution of research themes. Second, we analyzed the gap in research themes on the application of CTMC information diffusion on Twitter. Further filtering criteria were applied, and 34 papers were analyzed to identify the research objectives, methods, metrics, and measures used by each researcher. Nonhomogeneous CTMC has never been used in Twitter information diffusion modeling. This finding motivates us to further study nonhomogeneous CTMC as a modeling approach for Twitter information diffusion.

**Keywords:** information diffusion; social media; Twitter; systematic literature review; bibliometric; continuous time Markov chain



**Citation:** Firdaniza, F.; Ruchjana, B.N.; Chaerani, D.; Radianti, J. Information Diffusion Model in Twitter: A Systematic Literature Review. *Information* **2022**, *13*, 13. <https://doi.org/10.3390/info13010013>

Academic Editor: Arkaitz Zubiaga

Received: 4 December 2021

Accepted: 24 December 2021

Published: 28 December 2021

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Social media such as blogs, forums, chat applications, and social networking are platforms for online interactions regardless of a user’s physical location [1] and have become integrated into daily life. Twitter is a growing microblog that allows users to send text messages composed of up to 280 characters [2]. User activities on social media have been subjected to data collection and monitoring and are considered meaningful data sources for various public and private organizations, including in the industry and academia. The largest use of social media data comes from microblogs, as much as 46% [1]. Twitter data can be used in a remarkably diverse number of research studies, such as sentiment analyses [3,4], text analyses [5–8], opinion analyses [9,10], as well as analyses of influence or information diffusion [11–14].

Studies of information diffusion on Twitter are important as it is a topic that continues to attract researchers’ attention and is a subject useful for scrutinization and various advanced analyses. Information diffusion is defined as the process of information travelling from a sender to a set of receivers through a carrier. In the case of Twitter, the sender is the user who posted the tweet, the carrier is the tweet that was posted, and the recipients are followers of the user who posted the tweet [15].

A user is influential regarding delivering tweets if their messages can spread to many other Twitter users. In this case, these people with a strong influence in the Twitter network are called influencers [16].

Various themes and methods related to the study of information diffusion on Twitter have been investigated, such as comparing influential measures [17], finding the most influential users [11,18–21], maximizing influence [22–24], measuring the influencer index [25,26], and modeling the diffusion of information [11,27].

Although some studies on information diffusion in social networks exist, such as [28–35], the following have not been studied in previous literature reviews: article selection reviewed according to the SLR procedure, a bibliometric analysis and theme evolution approach, the *determination of* research objectives related to the information diffusion model on Twitter, and the metrics and measures used by researchers.

This study applied a systematic literature review (SLR) as an approach to obtain an overview of existing studies and trends related to information diffusion and social media, especially on Twitter. Our work proposes a bibliometric analysis that allows us to know popular places and networks from which authors conducting research on information diffusion and social influencers gather their data. We also conduct an evolution analysis, allowing us to detect the changes in topics over time.

In other words, this paper gives a survey report and systematically presents existing studies on information diffusion and Twitter analysis from the perspective of the methods, metrics, and measures used. We added a bibliometric analysis to assess the evolution of themes year after year, which has not been covered in previous review articles.

To be precise, the SLR that is used in this paper is a method of identifying, evaluating, and interpreting all existing research relevant to a phenomenon of interest [36]. Referring to [36], the determination of a research question (RQ) is based on the research objectives. We studied the topic of information diffusion on Twitter in our research, so we read several articles related to information diffusion, such as [12,14,21]. Varshney et al. [12] aimed to predict the probability of information diffusion and used a Bayesian network method based on tweet/retweet metrics. Kumar et al. [14] used the susceptible–exposed–infected (SEI) model to model the diffusion of information on Twitter. Meanwhile, Oo and Lwin [21] used PageRank to measure the influence of users on Twitter. Through a discussion of the study results from all of these authors, we determined the RQs. Thus, in this paper, we focus on information diffusion on Twitter and answer the following research questions (RQ):

1. What are the purposes of information diffusion-related research on Twitter?
2. What methods have researchers used regarding the information diffusion model on Twitter?
3. What metrics from Twitter data do researchers use?
4. What measures do researchers use to determine influencer rankings?

Information on Twitter spreading from one user to another fulfills the nature of a Markov chain, that is, the information redistributed by the next user depends only on the current user's information and does not depend on the history of the previous dissemination of information [11]. This means that information diffusion on Twitter can be modeled with the continuous time Markov chain (CTMC). In this paper, we aim to find gaps in the research theme of CTMC in modeling information diffusion on Twitter.

To conduct this research, we applied bibliometric analysis software, namely VOSviewer and *R-bibliometrix*. VOSviewer was used to implement co-authorship–author, co-authorship–country, and co-occurrence–word analyses. Meanwhile, *R-bibliometrix* with the biblioshiny web interface was employed to assess the evolution of themes.

This paper is structured as follows: Section 2 provides a related literature review. Section 3 presents the research method, especially the method used to collect the articles for analysis. In Section 4, we describe the bibliometric analysis of the papers obtained using two software and the results of a study on the selected articles covering research about information diffusion on Twitter, the methods used in information diffusion modeling on Twitter, the metrics used in Twitter data, and the measures used to determine the rank

of influencers. Section 5 comprises the discussion and further research agenda. Section 6 concludes this article.

## 2. Related Literature Review

Several surveys or literature reviews have been conducted regarding the study of information diffusion on social networks (SN) in general and not specifically on Twitter. No previous review articles performed a complete systematic literature review. Kakar and Mehrotra [28] conducted a review of 90 filtered papers from six databases: namely, Scopus, Science Direct, ACM Digital Library, Springer, IEEE, and Google Scholar. This work focuses on three research areas under the umbrella of information diffusion in social networks, namely influence modeling, influence maximization, and retweet prediction. However, it did not discuss the metrics and measures used by researchers.

Razaque et al. [29] conducted a review focused on the classification of information diffusion models and the vulnerability of each model. However, this work did not discuss network influencers and how to maximize influence within the network. Alamsyah and Rahardjo [30] investigated social networks (SNs) from the perspective of graphical representations to find the SN taxonomy. It included the SN topology, structural modeling, community detection, tie strength, community detection, as well as metrics. Graphical representation techniques were the focus of this study, rather than discussing information diffusion models. Hamzah [31] reviewed 49 articles filtered from six databases, namely the ACM Digital Library, Google Scholar, IEEE Xplore, Science Direct, Springer Link, and Taylor & Francis Online. This article discussed machine learning and visualization techniques used for Twitter analytics. In this work, neither network influencers nor information diffusion analysis techniques based on the perspective of user behaviors were discussed, and the focus was more toward identifying social vulnerabilities.

Meanwhile, a survey conducted by [32] divided the diffusion model into two categories, namely explanatory models and predictive models. This explanatory model included epidemic models and influence models, while the predictive model included the independent cascade model (ICM), the linear threshold model (LTM), and the game theory model (GTM).

Singh [33] analyzed how information dissemination is carried out through networks and how people influence each other on social networks. This survey focused on maximizing influence with LTM, ICM, and epidemic models but did not discuss the metrics and measures used by researchers.

Firdaus et al. [34] conducted a survey on the information diffusion mechanism on Twitter. The authors focused on predicting tweets: starting from how to retrieve Twitter data, users who post tweets, tweet content, and predicting whether a tweet will be retweeted. Machine learning techniques were used for the tweet prediction. Additionally, some of the measures used to evaluate the performance of the model were discussed. This work did not discuss influence maximization or influential users on Twitter.

Riquelme and González-Cantergiani [35] conducted a survey on the size of a user's influence on Twitter. This work collected and classified various measures of influence on Twitter. Some were based on simple metrics, and some were based on complex mathematical models. Various criteria were given to determine the most influential users on Twitter. However, an information diffusion model was not discussed in this work.

What is different about our work compared with that of previous researchers is the systematic literature search conducted on studies of information diffusion models, especially on Twitter. We also added a bibliometric analysis to describe the distribution of research topics regarding the diffusion of information and changes in themes that occurred during each time period. The aspects covered in our article and in existing reviews are summarized in Table 1.

The SLR procedure that we carried out was referenced from [36]. SLR procedures include the search strategy, quality assessment, data extraction and monitoring, and data synthesis. Starting with database selection, we all agreed to choose four databases, i.e., Sco-

pus, Science Direct, Dimensions, and Google Scholar. These four databases provide many articles from various domains such as science, engineering, computer science, medicine, social science, and others. We hope that our research can be considered a contribution to the field of research. The data from these four databases are quite demanding on our efforts to conduct a bibliometric analysis because the metadata of each database are different. Referring to [37], our keyword selection was performed through RQ analysis and by looking at related papers. We all agreed on the keywords applied to the four databases. The keywords included “information diffusion”, “user influence”, “influence maximization”, “social network”, and “Twitter”. Capturing all articles about information diffusion studies, especially on Twitter, was assumed to be sufficient. In the inclusion process, we decided together that the selected articles had to be articles that had been published in an English-language journal from 2000 to February 2021 (when the data were collected). In the manual article-exclusion stage, two authors read the abstract and the article content to mark the article as being relevant or not. If opinions differed, another author participated in reading the abstract and the contents of the article and made the final decision. The complete research procedure carried out can be seen in Section 3.

**Table 1.** Summary of the aspects covered in our article and in existing review articles.

Author	SLR?	Search Strategy?	Data Extraction?	Information Diffusion Model?	Influence Analysis?	Metrics and Measures?	Twitter/SN?	Bibliometric Analysis?	Article Time Span
[28]	×	✓	✓	✓	✓	×	SN	×	2000–2019
[29]	×	×	×	✓	×	×	SN	×	2002–2017
[30]	×	×	×	×	×	✓	SN	×	1998–2012
[31]	×	✓	×	×	×	×	Twitter	×	2011–2016
[32]	×	×	×	✓	✓	×	SN	×	2001–2017
[33]	×	×	×	✓	✓	✓	SN	×	1978–2012
[34]	×	×	×	✓	×	✓	Twitter	×	2001–2015
[35]	×	×	×	×	✓	✓	Twitter	×	1972–2015
<b>Our article</b>	✓	✓	✓	✓	✓	✓	Twitter	✓	2000–2021

### 3. Methods

In this section, we describe how this research process was carried out: namely, the collection of article data and the selection method.

This study began with a systematic search for publications indexed in four selected databases: Scopus, Science Direct, Dimensions, and Google Scholar. The keywords used in this first search were (“information diffusion”) OR (“influence analysis”) OR (“influence maximization”) OR (“user influence”). We limited publication time to 2000 as Twitter was founded in 2006, and we ended collection time in February 2021. We limited our search by only looking for journal article publications and excluded conference proceedings or books. We only included articles written in English and published (final) in peer-reviewed in international journals.

Data retrieval in the Scopus, Science Direct, and Dimensions digital libraries was carried out with a few keywords applied to the “title, abstract, and keywords”. Meanwhile, in the Google Scholar database, the keywords were only applied to the title, since the Google Scholar search engine does not provide a search process using the abstract. The first step was to find all papers related to information diffusion. From these search activities, 2675 papers were obtained from Scopus, 850 papers were obtained from Science Direct, 2950 papers were obtained from Dimensions, and 5950 papers were obtained from Google Scholar. Then, we applied the inclusion filter using two new keywords, namely (“social networks”) OR (“social media”), aiming to acquire papers that included social media data or social networks when scrutinizing information diffusion. In this second round, several papers from the Scopus database were removed, reducing the number to 1211 papers. The number

of papers from Science Direct decreased to 367 papers, while the number of papers from Dimensions and Google Scholars were brought down to 1090 and 110 papers, respectively.

The filtering process continued by applying the keyword “Twitter” to capture papers examining information diffusion studies on Twitter. From this search, we obtained 199 papers from Scopus, 50 papers from Science Direct, 172 papers from Dimensions, and 3 papers from Google Scholar.

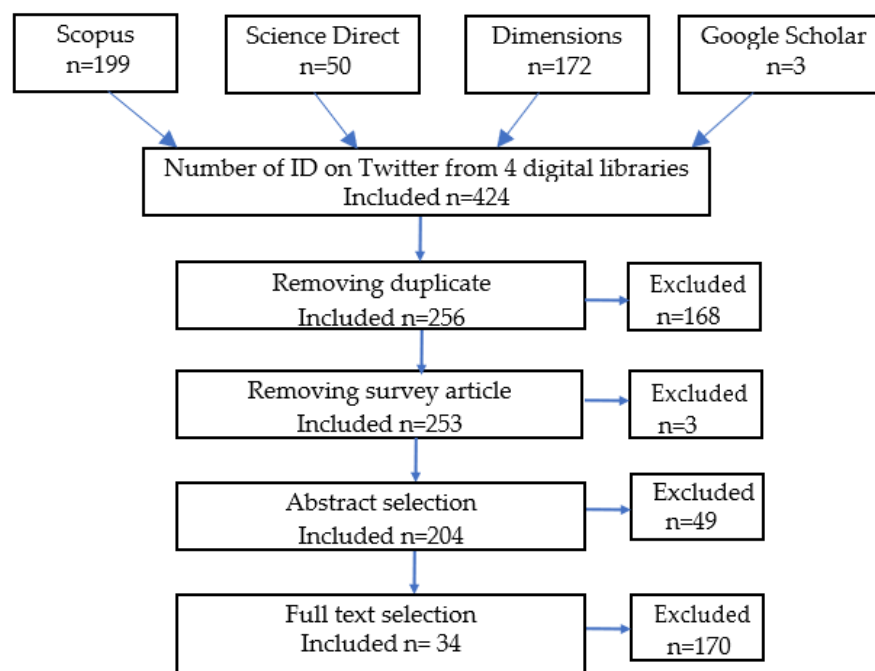
A summary of the search results from the three filtering processes on the four databases can be seen in Table 2. Note that the “Type” column in Table 2 represents the use of the following keywords:

- A. (“information diffusion”) OR (“influence analysis”) OR (“influence maximization”) OR (“user influence”);
- B. (“social network”) OR (“social media”); and
- C. “Twitter”.

**Table 2.** Number of publications from four databases with three types of keywords.

Keywords	Type	Scopus	Science Direct	Dimensions	Google Scholar
Keyword 1	A	2675	850	2950	5950
Keyword 2	A AND B	1211	367	1090	110
Keyword 3	A AND B AND C	199	50	172	3

Furthermore, after semi-automatic selection of all articles with the three keywords in the four digital libraries, we removed 168 duplicate articles and 3 survey articles. Then, the selection of articles was carried out through the abstracts, obtaining 204 relevant articles. Next, we performed manual filtering by reading the full text and obtained 34 articles. Our general selection process is shown in Figure 1.



**Figure 1.** Diagram of the data selection process.

### 3.1. Semi-Automatic Selection

We developed a simple script using Python to select duplicate documents. We used Scopus articles as a reference for viewing duplicates in Dimensions, Science Direct, and

Google Scholar. From this process, 122 Dimensions articles, 45 Science Direct articles, and 1 Google Scholar article were found to be redundant. After removing the duplicate articles, we obtained a total of 256 unique articles.

### 3.2. Manual Selection

The manual selection process was carried out in three stages:

- First: We examined the title, abstract, and full text from the filtered articles to find articles *conducting a survey or literature review*. We removed three articles in the form of surveys, namely two articles from Scopus and one article from Dimensions. Thus, in total, from this stage, we obtained 253 articles.
- Second: We examined the abstract to assess *the relevance of the article* to our research focus. Based on the abstracts, we discarded a total of 49 out of 253 articles, so we obtained 204 selected articles (hereinafter referred to as “Dataset 1”). Note that the original raw data returned from each digital library came in different formats. The selection results of this article originally had a different data format. Hence, we adjusted the article data for Dimensions, Science Direct, and Google Scholar in such a way that their formats were uniform to the raw file from Scopus. After restructuring all datasets into a homogeneous structure, bibliometric analysis was carried out for Dataset 1 (see Section 4).
- Third: We thoroughly read the full text and the content and discussion of the articles to further evaluate their relevance. At this point, we obtained 34 articles (henceforth referred to as “Dataset 2”), which were used further for our systematic literature review analysis.

To sum up, we used Dataset 1 to conduct the bibliometric analysis as presented in Section 4.1 and Dataset 2 to discuss the results from the systematic literature review as presented in Section 4.2.

The results of this semi-automatic and manual selection process are shown in Table 3.

**Table 3.** Results of the semi-automatic and manual selection.

Database	Data Keyword 3	Semi-Automatic		Manual Selection					
		Duplicate		Survey		Abstract		Full Text	
		Excluded	Included	Excluded	Included	Excluded	Included	Excluded	Included
Scopus	199	0	199	2	197	29	168	142	26
Science Direct	50	45	5	0	5	2	3	3	0
Dimensions	172	122	50	1	49	17	32	24	8
Google Scholar	3	1	2	0	2	1	1	1	0
Total	424	168	256	3	253	49	204	170	34

### 3.3. Bibliometric Analysis

We performed a bibliometric analysis for Dataset 1. This analysis technique is often used for literature analyses intent on obtain bibliographic overviews of scientific selections of highly cited publications. It can recover a list of author productions, national or subject bibliographies, or other specialized subject patterns [38]. We performed the bibliometric analysis using VOSviewer and *R-bibliometrix*. VOSviewer is a computer program used for bibliometric mapping [39], while *R-bibliometrix* is a package from the open source R software with a shiny web interface capable of conducting comprehensive analyses and scientific mapping of data with complete bibliographic information [40]. Both software have their respective advantages in bibliometric analysis. For example, VOSviewer has better visualization and clear links among different nodes in the network images compared with *R-bibliometrix*. In contrast, *R-bibliometrix* has a Sankey diagram feature that is particularly useful in conducting thematic evolution analyses.



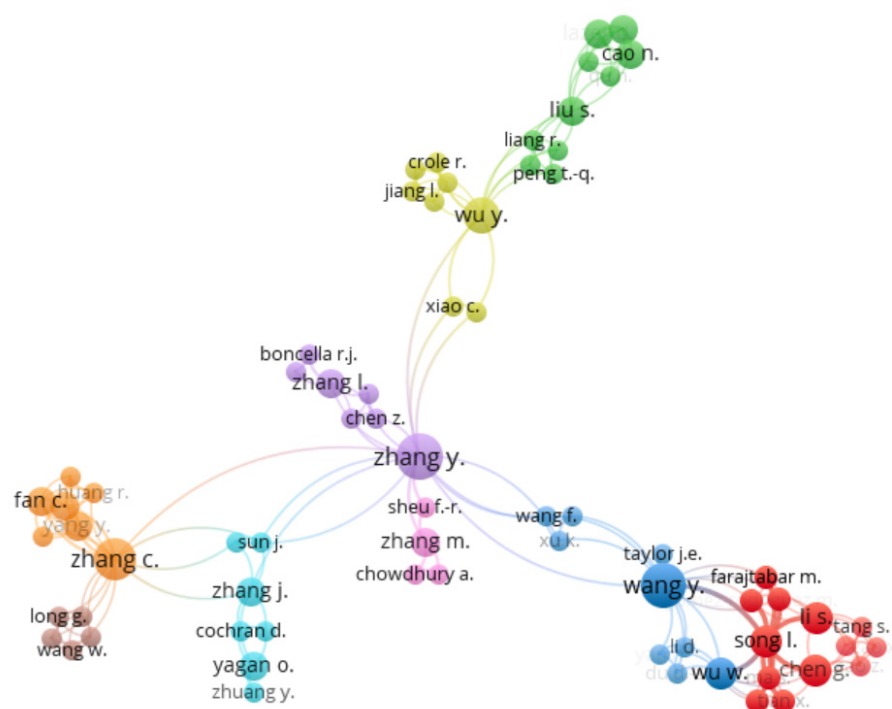
## 4. Results

### 4.1. Results from Bibliometric Analysis

In this section, we present the results of the analysis using the network visualization, grid matrix, and Sankey diagram techniques. This analysis was divided into three parts: (1) co-authorship–author and co-authorship–country, (2) co-occurrence–words, and (3) thematic evolution.

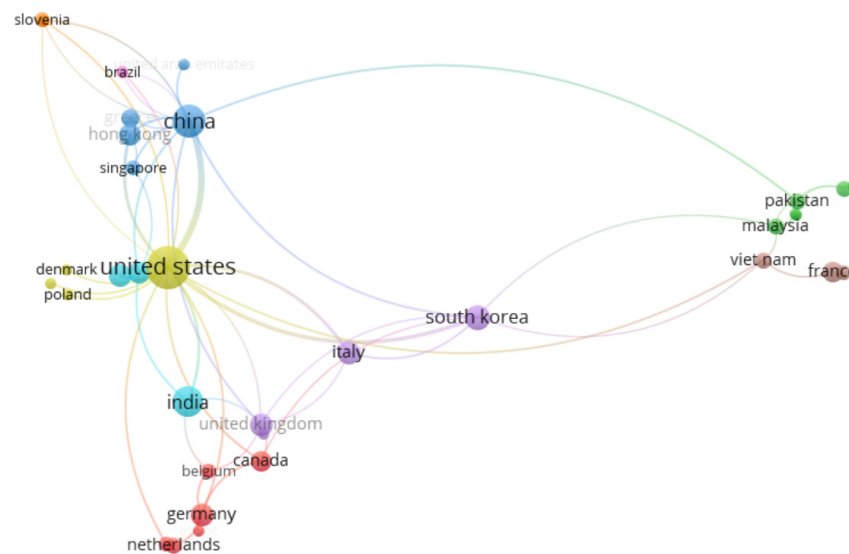
#### 4.1.1. Visualization of the Co-Authorship–Author and Co-Authorship–Country Relations

In this section, the co-authorship analysis was conducted by examining the relationship between authors and their countries of origin. In VOSviewer, the co-authorship–author menu was selected by limiting each author to a minimum of one article. This means that all articles were analyzed. Based on this provision, VOSviewer obtained 559 authors, but only 70 authors were connected with other authors. The co-authorship–author relation was divided into nine clusters, namely red, yellow, green, blue, orange, pink, aqua, brown, and purple, as shown in Figure 2. In this case, the most productive author on the topic under study was Zhang, Y with five articles, followed by Zhang, C and Wang, Y with four articles each.



**Figure 2.** Network visualization of the co-authorship–author relation for Dataset 1.

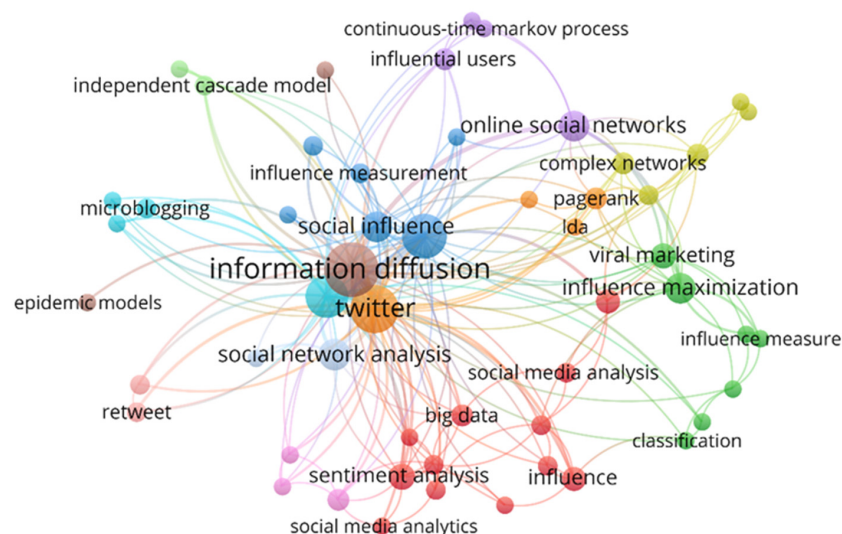
Furthermore, a bibliometric analysis was also carried out to assess the countries of origin of the authors involved in the network. The type of analytic used was the co-authorship–country relation, with the minimum number of documents from a country for a co-authorship being 1. VOSviewer detected 44 countries in our Dataset 1; however, only 36 countries had connections with other countries in the context of the co-authorship–country relation. The 36 countries were divided into nine clusters, as shown in the network visualization in Figure 3. The clusters are indicated with different colors. From these results, the US had the most, with 68 articles (27%); followed by China, with 29 articles (12%); and then, India, with 23 articles (9%). As an example, the visualization also tells us that the authors in the US cooperated with authors in various countries such as Denmark, Poland, Slovenia, China, Hongkong, Brazil, Vietnam, South Korea, Italy, India, the Netherlands, Germany, Canada, and the United Kingdom.



**Figure 3.** Network visualization of the co-authorship–countries relation for Dataset 1.

#### 4.1.2. Visualization of Co-Occurrence–Word Relation

To conduct a co-occurrence analysis in Dataset 1, we searched for the most frequent words that appeared in all documents. Dataset 1 contains data taken based on the title, keywords, and abstract only. VOSviewer has a support feature allowing us to conduct an assessment of the co-occurrence–author keyword relation from the menu on VOSviewer. We set up the minimum number of occurrences of a word in a document at two. From this, VOSviewer returned 467 words, and only 54 passed the threshold. The words that appeared at least two times in each document were divided into 12 clusters. The results show that the most frequent words appearing in Dataset 1 are “information diffusion”, with 66 events; followed by “Twitter”, with 44 events; and “social networks”, with 37 events. This co-occurrence–word network visualization is shown in Figure 4. Note that the co-occurrence network has extensively been used in social media analyses and text analyses for discovering the relationships among people, organizations, concepts, and other areas of interests. Here, we observe that, for example, the information diffusion concept is often linked to various concepts, especially Twitter, social influence, social networks, user influence, contagion, and popularity prediction, to name a few.

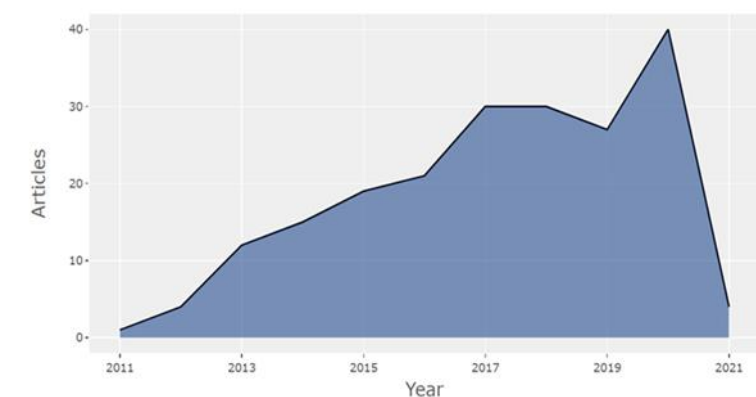


**Figure 4.** Network visualization of the co-occurrence–word relation for Dataset 1.



#### 4.1.3. Thematic Evolution

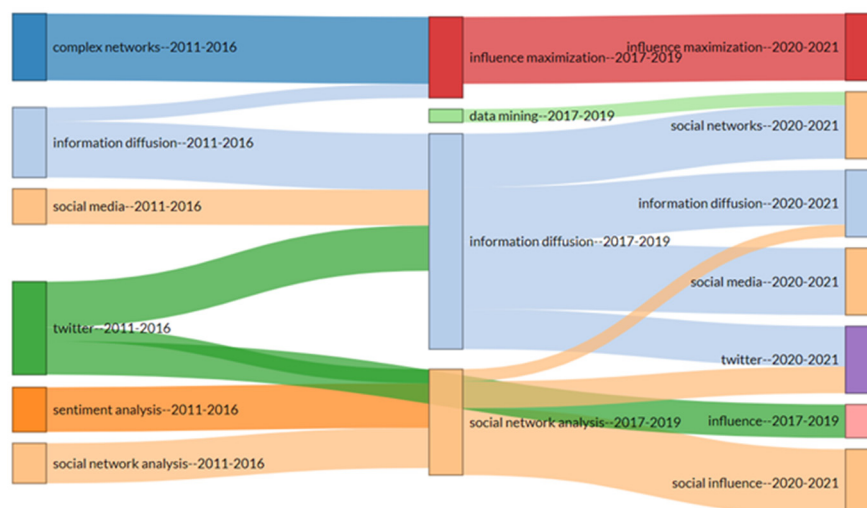
Using *R-bibliometrix*, we also acquired an overview of the evolution of themes. Topics that were in a certain quadrant in the previous period could be shifted to another quadrant in the next period. This evolution was presented as a Sankey diagram. To determine the distribution of time periods or time slices used for thematic evolution analysis, the overall number of published articles in Dataset 1 was analyzed. The number of issues per year from Dataset 1 for the four databases can be seen in Figure 5.



**Figure 5.** Number of publications in Dataset 1 per year.

As seen in Figure 5, the publication of information diffusion on Twitter (Dataset 1) began in 2011 because Twitter was only established in 2006, and no scientific publications related to this platform existed before 2011. The number of information diffusion publications on Twitter increased and slightly flattened between 2015 until the end of 2016. Afterwards, the publications again sharply increased; then, they flattened before they decreased by the end of 2018 and then increased again at the beginning of 2019. The decline in 2021 was due to our data collection only covering publications until 6 March 2021. To sum up, from this observation, we can conclude that two points can be considered cutting points for our thematic evolution analysis, namely the years 2016 and 2019.

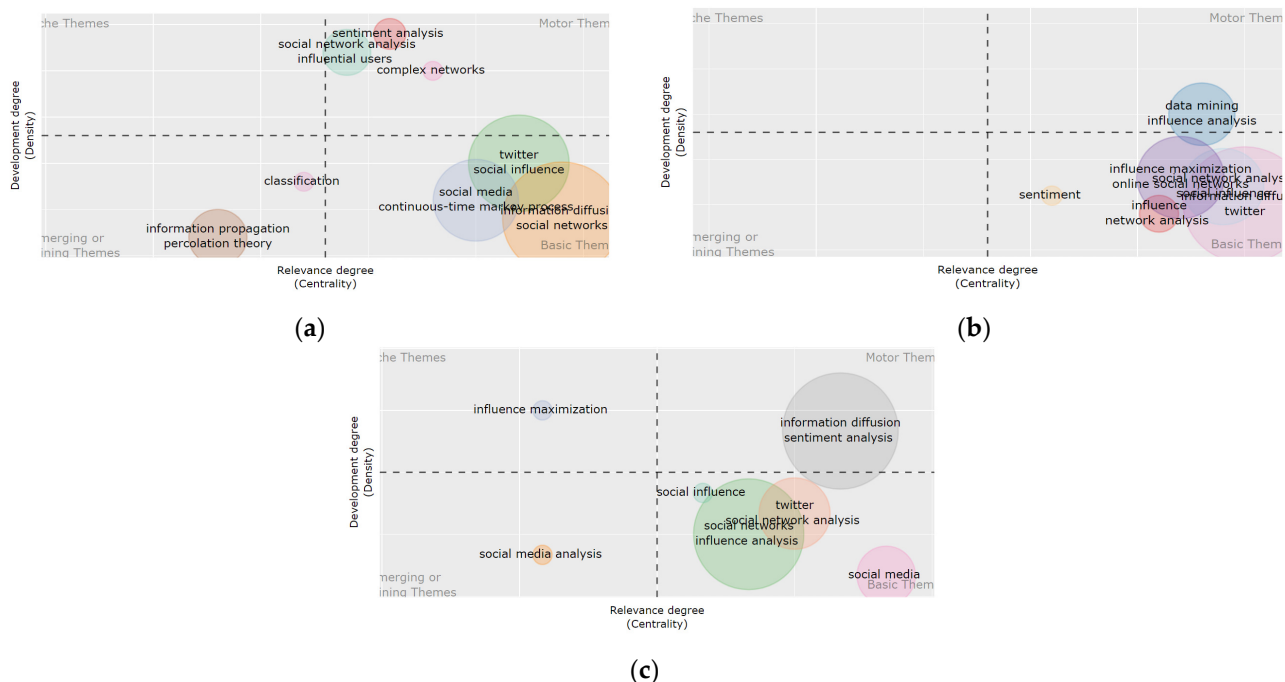
The selected thematic evolution parameters were author keywords, the number of words was 450, the minimum cluster frequency was 5, the number of labels was 2, and the number of cutting points was 2. A visualization of the evolution of these is presented based on three time slices, namely time slice 1 (2011–2016), time slice 2 (2017–2019), and time slice 3 (2020–2021), in Figure 6.



**Figure 6.** Thematic evolution for Dataset 1.

From Figure 6, notice that a small part of the topics of “information diffusion” in the 2011–2016 period joined the topic of “influence maximization” in the 2017–2019 period, but some information diffusion remained popular until the 2020–2021 period. In the 2017–2019 period, the topic of information diffusion became the most studied topic. The topic of information diffusion partly remained until the 2020–2021 period, and some of it spread to become the topics social networks, social media, and Twitter. Twitter in the 2011–2016 period joined the topics information diffusion, influence, and social network analysis in the 2017–2019 period. In contrast, the topic influence maximization recently emerged in the 2017–2019 period, remaining stable until the 2020–2021 period. The topic social network analysis started in the 2011–2016 period; remained until the 2017–2019 period; and joined the topics information diffusion, Twitter, and social influence in the 2020–2021 period.

For each time slice, this evolution in themes can be described more fully with the Callon centrality method [41]. The most frequently discussed themes in the literature are portrayed and mapped as clusters plotted in the grid diagram consisting of the four quadrants. The clusters are depicted in the form of circles of diverse sizes and colors. The size of the cluster represents the frequency that the word appears in the documents. The first quadrant includes *the motor themes*. In this quadrant, the cluster has a large centrality and density. This means that clusters have links with other clusters and strong internal links. The second quadrant includes *the niche themes*. In this quadrant, the links with other clusters are weak, but internally, the links are strong. Quadrant 3 includes *the emerging or declining themes*. In this quadrant, the centrality and density are small, which describes a new topic developing or having decreased. Quadrant 4 includes *the basic themes*. In this quadrant, the cluster is strongly connected to other clusters, but the internal link intensity is low. Using the *R-bibliometrix* tool for visualization, the thematic evolution in each time slice can be seen in Figure 7.



**Figure 7.** (a) Thematic map Dataset 1 for time slice 1 (2011–2016), (b) thematic map Dataset 1 for time slice 2 (2017–2019), and (c) thematic map Dataset 1 for time slice 3 (2020–2021).

Figure 7 shows that the topics discussed in Dataset 1 are presented in certain quadrants and clusters that experience changes in each period. In the 2011–2016 period, all topics were spread out into three quadrants and eight different clusters. The three biggest clusters, namely information diffusion and social networks, Twitter and social influence, social media and continuous time Markov chain (CTMC), are in the *basic themes*, meaning that

links with other clusters are strong, but internal link intensity is weak. Note that in some literature, CTMC is often called CTMP. In this paper, we sometimes use these two terms interchangeably to refer to the same concept, especially if the literature explicitly uses the term CTMP instead of CTMC.

In the 2017–2019 period, the topic of CTMP did not appear anymore, while information diffusion joined Twitter, and a new topic emerged, namely influence maximization. Furthermore, in the 2020–2021 period, the topic of information diffusion remained the most studied topic and occupied the motor themes quadrant, meaning that the links with other clusters and internal links between clusters are strong. Twitter, which is in the same cluster as the social network analysis, is in the basic themes quadrant, meaning that links with other clusters are strong even though the internal cluster is weak. The topic influence maximization, which was originally in the basic themes quadrant but moved to the *niche themes*, meaning that links with other clusters were weak.

The topic CTMP is in the *basic themes* quadrant, and it is different from the clusters with “information diffusion” and “Twitter”. This indicates that the topic CTMP has a very strong link with “information diffusion” and “Twitter”. In brief, the topic of CTMP has not been frequently studied and is open to further research in connection with information diffusion on Twitter. This is our contribution in our next study.

#### 4.2. Results from Systematic Literature Review

In this section, we present the results of a study on Dataset 2, namely 34 selected articles discussing information diffusion on Twitter. Articles in Dataset 2 were published within the 2012–2020 timeframe (with 41% published in 2020).

##### 4.2.1. The Purpose of Research on Information Diffusion on Twitter

We conducted an analysis of information diffusion on Twitter to answer RQ1: *What are the purposes behind information diffusion-related research on Twitter?*

After examining all of the selected articles, we sorted them into three categories of study areas based on the purpose of each article. The percentages of these three categories can be seen in Figure 8. The three categories are as follows:

1. Information Difference Model on Twitter—articles that focus on modeling how information diffuses or spreads on Twitter;
2. Influential User on Twitter—articles that discuss how to find the most influential users on Twitter or to rank Twitter users; and
3. Influence Maximization on Twitter—articles that discuss how to maximize the influence of users who share information on Twitter.

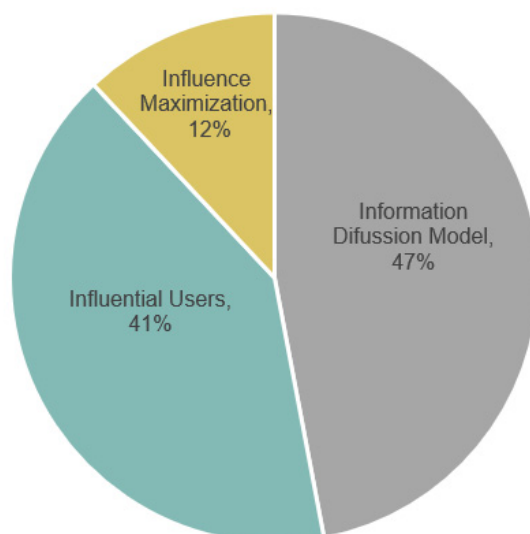


Figure 8. Categories of studies from Dataset 2.

From Figure 8, we can see that the information diffusion model (47%) and then influential users (41%) are the two most important purposes behind why scholars study Twitter analytics.

#### 4.2.2. Methods Used in Information Diffusion Modeling on Twitter

This section intends to answer RQ2: *What methods have researchers used regarding information diffusion models on Twitter?* Based on our review of the information diffusion model on Twitter, various methods have been used by researchers, such as epidemic models [14,42–44], the stochastic model [11,45,46], machine learning [47–49], and the independent cascade model (ICM) [50].

In the stochastic model, Foroozani [45] used discrete time-random walk (DTRW) and continuous time-random walk (CTRW); meanwhile, Li et al. [11] used the continuous time Markov process (CTMP). In their study, the authors of [11] used homogeneous CTMP, meaning that the rate of information dissemination was assumed to be constant. The methods used by researchers to conduct studies related to the complete information diffusion Model on Twitter can be seen in Tables 4–6.

**Table 4.** Studies of the information diffusion model on Twitter.

#	Author	Methods	Metrics
1	[42]	Ignorant–spreader–counterspreader–recovered (ISCR), Ignorant–spreader–spreader–recovered–recovered (ISSRR), Stochastic ISI Information Model, Stochastic ISR Information Model	-
2	[50]	Time decay features cascade model (TDF-C), Time decay features cascade threshold model (TDF-CT)	retweet, quote, reply, like
3	[43]	Expectation Maximation-Independent Propagation Model with Susceptible–Infected State (EM-IPSI)	retweet
4	[45]	Discrete time random walk and Continuous time random walk (DTRW-CTRW)	mention, retweet
5	[44]	SIR model	retweet
6	[12]	Bayesian network	tweet, retweet
7	[47]	Regression analysis	retweet
8	[11]	Homogeneous continuous time Markov process (CTMP)	retweet
9	[51]	Evolutionary Game Theory (EGT)	tweet with a specific hashtag
10	[46]	A stochastic model	retweet
11	[48]	Regression	retweet
12	[52]	Quantum q-Attention Model	retweet
13	[27]	Modified forest-fire model based on mentioned, similarity score, user activity, topic significance	retweet
14	[14]	Susceptible–exposed–infected (SEI)	tweet, retweet, reply, mention
15	[53]	Textual-Homo-IC, Textual-Homo-PCM	follow
16	[49]	Poisson regression model	retweet, quote

**Table 5.** Studies of influential users on Twitter.

#	Author	Methods	Metrics	Measures
1	[20]	Analytic hierarchy process (AHP)	tweet–retweet	analytic hierarchy process (AHP) algorithm
2	[54]	Average Consensus Ranking Aggregation (ACRA) and Weighted Average Consensus Ranking Aggregation (WACRA)	retweet, mention, follow	degree centrality, closeness, betweenness, eigenvector, PageRank
3	[55]	Weighted multimodal ensemble average influence (WMMEAI) for rank top-k	tweet, follower	degree centrality, closeness, betweenness, eigenvector
4	[56]	PageRank and betweenness	retweet, quote, favorite	PageRank, betweenness
5	[21]	PageRank	retweet, mention, reply	PageRank
6	[57]	Smart Inf (stands for Smart Influencer) algorithm; Classical Metrics to Compute Node Influence PageRank centrality (PR), eigenvector centrality (EV), and betweenness centrality (BW)	Retweet	PageRank, eigen vector, betweenness centrality
7	[58]	Influence Factorization	retweet	PageRank
8	[59]	Personalized PageRank (PPR)	retweet	PPR
9	[60]	T and HT measure	retweet	T and HT
10	[61]	Topic-based Social Influence Measure (TSIM)	retweet, mention	PageRank, network centrality
11	[62]	Hidden Markov Model (HMM)	retweet, reply, mention	buzz rank, reach buzz rank
12	[13]	The Clauset–Newman–Moore community detection algorithm, The Latent Dirichlet Allocation (LDA) method	retweet	reciprocity degree, closeness, betweenness, diameter, modularity, transitivity
13	[63]	Aggregation Consensus Rank Algorithm (ACRA)	retweet, mention	indegree, closeness centrality, betweenness centrality, eigenvector centrality
14	[64]	Cluster-based fusion techniques	retweet, mention	retweet impact, mention impact, signal strength

**Table 6.** Studies of influence maximization on Twitter.

#	Author	Methods	Metrics
1	[23]	Evidential influence maximization models	tweet, follow, retweet, mention
2	[65]	Artificial Bee Colony (ABC) algorithm	retweet
3	[66]	Independent Cascade (IC)	retweet
4	[24]	Potential-Driven Model	reply, retweet

#### 4.2.3. Twitter Metrics

Our third research question (RQ3) was as follows: *What metrics do researchers use from Twitter data?* Our study reveals that the most common metric used by researchers to process Twitter data is the number of “retweets”; however, some studies included replies, mentions, and follows.

The use of metrics from Twitter data can be seen in full in Tables 4–6.

#### 4.2.4. Measures for Determining Influencers

This subsection tries to answer the fourth research question (RQ4): *What measures do researchers use to rank influencers?* For articles with a focus on the study of influential users on Twitter, the measures used to rank influential users are traditional measures such as closeness, betweenness, and PageRank [54–57,61,63], analytic hierarchy process [20], and buzz rank [62]. The types of measure used to assess influential users can be seen in Table 5.



## 5. Discussion

In this section, we discuss the results of the analysis that we obtained from the literature review (Dataset 2), totaling 34 articles.

### 5.1. The State-of-the-Art of Information Diffusion Application on Twitter

A review of Dataset 2, which includes the state-of-the-art of our research, is presented in a table, which consists of information diffusion model studies on Twitter in Table 4, influential user studies in Table 5, and influence maximization studies in Table 6. The tables were completed with the research objectives, the methods used, the metrics used, and the measurements used by the researcher.

### 5.2. Research Gaps

From the modeling perspective of Twitter data, our analysis from the previous literature selected showed three research gaps.

First, *research in homogeneous CTMP for the information diffusion model on Twitter*. Referring to Table 4, we observe that one of the methods used in the information diffusion model is the stochastic model. We only found one study, Li et al. [11], that applied homogeneous CTMP for the information diffusion model on Twitter. In such an approach, the rate of transition of information dissemination from Twitter users to other users is considered constant. We notice in Figure 7 that CTMP only appeared in the 2011–2016 period, and we did not observe such an approach used afterwards. On the contrary, information dissemination and Twitter are continuously the most studied topics in each period analyzed. Our analysis on time slice 1 of the thematic maps shows that CTMP lies in the basic theme quadrant, likely linking to the topic Twitter information diffusion. This can be observed from the strong link between this quadrant and other clusters. We also notice in the analysis in Section 4 that only a few studies are related to the application of homogeneous CTMP in Twitter analytics.

Second, *research in non-homogeneous CTMP for the information diffusion model on Twitter*. Looking at the phenomenon of information sharing among Twitter users, the transition rate from spreading information on Twitter is not completely constant but depends on the timing of information dissemination. In this case, the *nonhomogeneous* CTMP method can be considered an alternative to modeling the dissemination of information on Twitter.

Third, *research on maximizing influence*. As seen in Table 6, the number of publications about maximizing influence is still small (only four studies). Likewise, as seen in Figure 7b,c, in 2017–2019, the topic influence maximization is in the basic theme quadrant, which means that the connections with other clusters are strong but the links to internal clusters are weak. Then, in the 2020–2021 period, it moved to the niche theme quadrant, meaning that links with other clusters are weak but internal links are strong. This means that more opportunities are available to study influence maximization in the future.

## 6. Conclusions

In this paper, we presented a systematic literature review on information diffusion on Twitter. We screened 424 papers from four digital libraries, namely Scopus, Science Direct, Dimensions, and Google Scholar. After going through the selection of duplicates, titles, and abstracts, 204 articles were obtained (Dataset 1). We performed a bibliometric analysis for Dataset 1. We showed how the usage of the bibliographic mapping technique can reveal an overview of the existing themes as well as the changes over time. This study demonstrates that publications on the diffusion of information are continuously increasing every year. This description of themes can serve as a basis for deciding further studies.

Moreover, we conducted a manual selection of full texts and obtained 34 articles (Dataset 2). From the results of the SLR in Dataset 1, we found that 47% of the publications studied the information diffusion model, 41% studied the influence of users, and 12% studied influence maximization. We answered our research questions raised in the Introduction. To sum up, we found that publications on information diffusion models on

Twitter have used various methods, such as epidemic models (e.g., SIR, EM-IPSI, and SEI), stochastic models (CTMC), machine learning, Regression, Bayesian, EGT, and those based on independent cascade models and threshold models. Additionally, the metrics used by researchers in general are retweets, mentions, and replies.

Publications about the influence of researchers use the methods of AHP, ACRA, WACRA, WMMEAI, PageRank, influence factorization, T-HT, LDA, and the cluster-based fusion technique. The measures used are degree centrality, closeness, betweenness, eigenvector, PageRank, buzz rank, and the T and HT measures.

Our study shows three gaps that could be future directions of influence analysis and information diffusion models on Twitter. First, very limited studies examine influence maximization, which is openly available as a future research direction. Second, we also noticed from our endeavor in this article that studies on the information diffusion model using homogeneous continuous time Markov chain are limited, where we only found one study on a homogeneous CTMC variant. Research in homogeneous CTMC for information diffusion model on Twitter can still be conducted as a future study. Third, the study of information diffusion models with nonhomogeneous CTMC is very open to future research, considering that this model is very realistic and that the transition rate of information dissemination on Twitter is not constant but depends on the time of information dissemination.

Our systematic literature review is not without limitations. First, we used four databases to mine data: Scopus, Science Direct, Dimension, and Google Scholar. We hope that most of the articles from other databases are already contained in the database we used. Second, we chose keywords related to our specific topic. Third, to minimize the subjectivity in articles selection, we performed a standard procedure regarding the title and abstract.

**Author Contributions:** Writing—original draft preparation, F.F.; writing—review and editing, F.F. and J.R.; conceptualization and methodology, F.F.; supervision, J.R., B.N.R. and D.C.; literature review and analysis, all authors; funding acquisition, B.N.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by *Riset Disertasi Doktor Unpad (RDDU)* or Doctoral Research Dissertation 2019 (contract number 5879/UN6.D/LT/2019).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We thank the Rector of the Universitas Padjadjaran, Indonesia, which financed this research. We are grateful for the discussion on social media analytics study with the collaborators in the RISE-SMA H2020 Project.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ghani, N.A.; Hamid, S.; Targio Hashem, I.A.; Ahmed, E. Social media big data analytics: A survey. *Comput. Hum. Behav.* **2019**, *101*, 417–428. [\[CrossRef\]](#)
2. Fernando, S.; Amador Díaz López, J.; Șerban, O.; Gómez-Romero, J.; Molina-Solana, M.; Guo, Y. Towards a large-scale twitter observatory for political events. *Futur. Gener. Comput. Syst.* **2019**, *110*, 976–983. [\[CrossRef\]](#)
3. Ruz, G.A.; Henríquez, P.A.; Mascareño, A. Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers. *Futur. Gener. Comput. Syst.* **2020**, *106*, 92–104. [\[CrossRef\]](#)
4. Hoeber, O.; Hoeber, L.; El Meseery, M.; Odoh, K.; Gopi, R. Visual Twitter Analytics (Vista): Temporally changing sentiment and the discovery of emergent themes within sport event tweets. *Online Inf. Rev.* **2016**, *40*, 25–41. [\[CrossRef\]](#)
5. Ibrahim, N.F.; Wang, X. A text analytics approach for online retailing service improvement: Evidence from Twitter. *Decis. Support Syst.* **2019**, *121*, 37–50. [\[CrossRef\]](#)
6. Radianti, J.; Hiltz, S.R.; Labaka, L. An overview of public concerns during the recovery period after a major earthquake: Nepal twitter analysis. In Proceedings of the 49th Hawaii International Conference on System Sciences (HICSS), Koloa, HI, USA, 5–8 January 2016; pp. 136–145.

7. Zahra, K.; Imran, M.; Ostermann, F.O. Automatic identification of eyewitness messages on twitter during disasters. *Inf. Processing Manag.* **2020**, *57*, 102107. [\[CrossRef\]](#)
8. Sakaki, T.; Okazaki, M.; Matsuo, Y. Earthquake shakes Twitter users: Real-time event detection by social sensors. In Proceedings of the 19th International World Wide Web Conference, Raleigh, NC, USA, 26–30 April 2010; pp. 851–860.
9. Bolzern, P.; Colaneri, P.; De Nicolao, G. Opinion influence and evolution in social networks: A Markovian agents model. *Automatica* **2019**, *100*, 219–230. [\[CrossRef\]](#)
10. Barnaghi, P.; Ghaffari, P.; Breslin, J.G. Opinion Mining and Sentiment Polarity on Twitter and Correlation between Events and Sentiment. In Proceedings of the 2016 IEEE Second International Conference on Big Data Computing Service and Applications (BigDataService), Oxford, UK, 29 March–1 April 2016; pp. 52–57.
11. Li, J.; Peng, W.; Li, T.; Sun, T.; Li, Q.; Xu, J. Social network user influence sense-making and dynamics prediction. *Expert Syst. Appl.* **2014**, *41*, 5115–5124. [\[CrossRef\]](#)
12. Varshney, D.; Kumar, S.; Gupta, V. Predicting information diffusion probabilities in social networks: A Bayesian networks based approach. *Knowl. Based Syst.* **2017**, *133*, 66–76. [\[CrossRef\]](#)
13. Mostafa, M.M. Information Diffusion in Halal Food Social Media: A Social Network Approach. *J. Int. Consum. Mark.* **2021**, *33*, 471–491. [\[CrossRef\]](#)
14. Kumar, S.; Saini, M.; Goel, M.; Aggarwal, N. Modeling Information Diffusion in Online Social Networks Using SEI Epidemic Model. *Procedia Comput. Sci.* **2020**, *171*, 672–678. [\[CrossRef\]](#)
15. Kushwaha, A.K.; Kar, A.K.; Ilavarasan, P.V. Predicting information diffusion on twitter a deep learning neural network model using custom weighted word features. *Responsib. Des. Implement. Use Inf. Commun. Technol.* **2020**, *12066*, 456.
16. Peng, S.; Zhou, Y.; Cao, L.; Yu, S.; Niu, J.; Jia, W. Influence analysis in social networks: A survey. *J. Netw. Comput. Appl.* **2018**, *106*, 17–32. [\[CrossRef\]](#)
17. Cha, M.; Haddai, H.; Benevenuto, F.; Gummadi, K.P. Measuring user influence in Twitter: The million follower fallacy. In Proceedings of the International AAAI conference on web and social media, Washington, DC, USA, 23–26 May 2010; Volume 4, pp. 10–17.
18. Salavati, C.; Abdollahpouri, A. Identifying influential nodes based on ant colony optimization to maximize profit in social networks. *Swarm Evol. Comput.* **2019**, *51*, 100614. [\[CrossRef\]](#)
19. Li, D.; Wang, W.; Jin, C.; Ma, J.; Sun, X.; Xu, Z.; Li, S.; Liu, J. User recommendation for promoting information diffusion in social networks. *Phys. Stat. Mech. Its Appl.* **2019**, *534*, 121536. [\[CrossRef\]](#)
20. Rezaie, B.; Zahedi, M.; Mashayekhi, H. Measuring time-sensitive user influence in Twitter. *Knowl. Inf. Syst.* **2020**, *62*, 3481–3508. [\[CrossRef\]](#)
21. Oo, M.M.; Lwin, M.T. Detecting Influential Users in a Trending Topic Community Using Link Analysis Approach. *Int. J. Intell. Eng. Syst.* **2020**, *13*, 178–188. [\[CrossRef\]](#)
22. Gong, M.; Yan, J.; Shen, B.; Ma, L.; Cai, Q. Influence maximization in social networks based on discrete particle swarm optimization. *Inf. Sci.* **2016**, *367*, 600–614. [\[CrossRef\]](#)
23. Jendoubi, S.; Martin, A.; Liétard, L.; Ben Hadji, H.; Ben Yaghlane, B. Two evidential data based models for influence maximization in Twitter. *Knowl. Based Syst.* **2017**, *121*, 58–70. [\[CrossRef\]](#)
24. Felfli, Z.; George, R.; Shujaee, K.; Kerwat, M. Potential-driven model for influence maximization in social networks. *IEEE Access* **2020**, *8*, 189786–189795. [\[CrossRef\]](#)
25. Arora, A.; Bansal, S.; Kandpal, C.; Aswani, R.; Dwivedi, Y. Measuring social media influencer index-insights from facebook, Twitter and Instagram. *J. Retail. Consum. Serv.* **2019**, *49*, 86–101. [\[CrossRef\]](#)
26. Deborah, A.; Michela, A.; Anna, C. How to quantify social media in fl uencers: An empirical application at the Teatro alla Scala. *Heliyon* **2019**, *5*, e01677. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Kumar, S.; Saini, M.; Goel, M.; Panda, B.S. Modeling information diffusion in online social networks using a modified forest-fire model. *J. Intell. Inf. Syst.* **2020**, *56*, 355–377. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Kakar, S.; Mehrotra, M. A review of critical research areas under information diffusion in social networks. *Int. J. Adv. Comput. Sci. Appl.* **2020**, *11*, 383–396. [\[CrossRef\]](#)
29. Razaque, A.; Rizvi, S.; Almiani, M.; Al Rahayfeh, A. State-of-art review of information diffusion models and their impact on social network vulnerabilities. *J. King Saud Univ. Comput. Inf. Sci.* **2019**, *34*, 1275–1294. [\[CrossRef\]](#)
30. Alamsyah, A.; Rahardjo, B. Social Network Analysis Taxonomy Based on Graph Representation. *arXiv* **2021**, arXiv:2102.08888.
31. Hamzah, M. A Taxonomy of Twitter Data Analytics Techniques. In Proceedings of the 32nd IBIMA Conference, Seville, Spain, 15–16 November 2018.
32. Li, M.; Wang, X.; Gao, K.; Zhang, S. A survey on information diffusion in online social networks: Models and methods. *Information* **2017**, *8*, 118. [\[CrossRef\]](#)
33. Singh, S.S.; Singh, K.; Kumar, A.; Shakya, H.K.; Biswas, B. A Survey on Information Diffusion Models in Social Networks. In *International Conference on Advanced Informatics for Computing Research*; Springer: Singapore, 2018; Volume 956, pp. 426–439.
34. Firdaus, S.N.; Ding, C.; Sadeghian, A. Retweet: A popular information diffusion mechanism—A survey paper. *Online Soc. Netw. Media* **2018**, *6*, 26–40. [\[CrossRef\]](#)
35. Riquelme, F.; González-Cantergiani, P. Measuring user influence on Twitter: A survey. *Inf. Processing Manag.* **2016**, *52*, 949–975. [\[CrossRef\]](#)
36. Kitchenham, B. *Procedures for Performing Systematic Reviews*; Joint Technical Report; Keele University: Staffordshire, UK, 2004.

37. Silva, R.L.S.; Neiva, F.W. *Systematic Literature Review in Computer Science—A Practical Guide*; Technical Report; Federal University of Juiz de Fora: Juiz de Fora, Brazil, 2016.
38. Ellegaard, O.; Wallin, J.A. The bibliometric analysis of scholarly production: How great is the impact? *Scientometrics* **2015**, *105*, 1809–1831. [[CrossRef](#)] [[PubMed](#)]
39. van Eck, N.J.; Waltman, L. *VOSviewer Manual*; Univeristeit Leiden: Leiden, The Netherlands, 2013; Volume 1, pp. 1–53.
40. Aria, M.; Cuccurullo, C. Bibliometrix: An R-tool for comprehensive science mapping analysis. *J. Informetr.* **2017**, *11*, 959–975. [[CrossRef](#)]
41. Callon, M.; Courtial, J.P.; Laville, F. Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemsitry. *Scientometrics* **1991**, *22*, 155–205. [[CrossRef](#)]
42. Muhlmeyer, M.; Agarwal, S.; Huang, J. Modeling Social Contagion and Information Diffusion in Complex Socio-Technical Systems. *IEEE Syst. J.* **2020**, *14*, 5187–5198. [[CrossRef](#)]
43. Kim, Y.; Seo, J. Detection of Rapidly Spreading Hashtags via Social Networks. *IEEE Access* **2020**, *8*, 39847–39860. [[CrossRef](#)]
44. Zheng, Z.; Yang, H.; Fu, Y.; Fu, D.; Podobnik, B.; Stanley, H.E. Factors influencing message dissemination through social media. *Phys. Rev. E* **2018**, *97*, 062306. [[CrossRef](#)] [[PubMed](#)]
45. Foroozani, A.; Ebrahimi, M. Anomalous information diffusion in social networks: Twitter and Digg. *Expert Syst. Appl.* **2019**, *134*, 249–266. [[CrossRef](#)]
46. Kawamoto, T. A stochastic model of tweet diffusion on the Twitter network. *Phys. Stat. Mech. Its Appl.* **2013**, *392*, 3470–3475. [[CrossRef](#)]
47. Yoo, E.; Rand, W.; Eftekhar, M.; Rabinovich, E. Evaluating information diffusion speed and its determinants in social media networks during humanitarian crises. *J. Oper.* **2016**, *45*, 123–133. [[CrossRef](#)]
48. Stieglitz, S.; Dang-Xuan, L. Emotions and information diffusion in social media—Sentiment of microblogs and sharing behavior. *J. Manag. Inf.* **2013**, *29*, 217–247. [[CrossRef](#)]
49. Kwon, J.; Han, I.; Kim, B. Effects of source influence and peer referrals on information diffusion in Twitter. *Ind. Manag. Data Syst.* **2017**, *117*, 896–909. [[CrossRef](#)]
50. Agarwal, S.; Mehta, S. Effective influence estimation in twitter using temporal, profile, structural and interaction characteristics. *Inf. Process. Manag.* **2020**, *57*, 102321. [[CrossRef](#)]
51. Jiang, C.; Chen, Y.; Liu, K.J.R. Evolutionary dynamics of information diffusion over social networks. *IEEE Trans. Signal Process.* **2014**, *62*, 4573–4586. [[CrossRef](#)]
52. Shuai, X.; Ding, Y.; Busemeyer, J.; Chen, S.; Sun, Y.; Tang, J. Modeling indirect influence on Twitter. *Int. J. Semant. Web Inf. Syst.* **2012**, *8*, 20–36. [[CrossRef](#)]
53. Ho, T.K.T.; Bui, Q.V.; Bui, M. Information Diffusion on Complex Networks: A Novel Approach Based on Topic Modeling and Pretopology Theory. *Vietnam J. Comput. Sci.* **2019**, *6*, 285–309. [[CrossRef](#)]
54. Tidke, B.; Mehta, R.; Dhanani, J. Consensus-based aggregation for identification and ranking of top-k influential nodes. *Neural Comput. Appl.* **2020**, *32*, 10275–10301. [[CrossRef](#)]
55. Tidke, B.; Mehta, R.; Dhanani, J. Multimodal ensemble approach to identify and rank top-k influential nodes of scholarly literature using Twitter network. *J. Inf. Sci.* **2020**, *46*, 437–458. [[CrossRef](#)]
56. Zhang, L. Product information diffusion in a social network. *Electron. Commer. Res.* **2020**, *20*, 3–19. [[CrossRef](#)]
57. Bhowmick, A. Temporal Sequence of Retweets Help to Detect Influential Nodes in Social Networks. *IEEE Trans. Comput. Soc. Syst.* **2019**, *6*, 441–455. [[CrossRef](#)]
58. Alp, Z.Z.; Ögüdücü, Ş.G. Influence Factorization for identifying authorities in Twitter. *Knowl. Based Syst.* **2019**, *163*, 944–954. [[CrossRef](#)]
59. Alp, Z.Z. Identifying topical influencers on twitter based on user behavior and network topology. *Knowl. Based Syst.* **2018**, *141*, 211–221.
60. Qasem, Z.; Jansen, M.; Hecking, T.; Hoppe, H.U. Using attractiveness model for actors ranking in social media networks. *Comput. Soc. Networks* **2017**, *4*, 3. [[CrossRef](#)]
61. Hamzehei, A.; Jiang, S.; Koutra, D.; Wong, R.; Chen, F. Topic-based social influence measurement for social networks. *Australas. J. Inf. Syst.* **2017**, *21*. [[CrossRef](#)]
62. Simmie, D.; Vigliotti, M.G.; Hankin, C. Ranking twitter influence by combining network centrality and influence observables in an evolutionary model. *J. Complex Netw.* **2014**, *2*, 495–517. [[CrossRef](#)]
63. Mittal, D.; Suthar, P.; Patil, M.; Pranaya, P.G.S.; Rana, D.P.; Tidke, B. Social Network Influencer Rank Recommender Using Diverse Features from Topical Graph. *Procedia Comput. Sci.* **2020**, *167*, 1861–1871. [[CrossRef](#)]
64. Kanavos, A.; Georgiou, A.; Makris, C. Estimating Twitter Influential Users by Using Cluster-Based Fusion Methods. *Int. J. Artif. Intell. Tools* **2019**, *28*, 1960010. [[CrossRef](#)]
65. Sankar, C.P.; Asharaf, S.; Kumar, K.S. Learning from bees: An approach for influence maximization on viral campaigns. *PLoS ONE* **2016**, *11*, e0168125. [[CrossRef](#)]
66. Kim, H.; Beznosov, K.; Yoneki, E. A study on the influential neighbors to maximize information diffusion in online social networks. *Comput. Soc. Netw.* **2015**, *2*, 3. [[CrossRef](#)]