

Isyaku HASSAN*
Mohd Nazri Latiff AZMI**
Akibu Mahmoud ABDULLAHI***

Evaluating the Spread of Fake News and its Detection. Techniques on Social Networking Sites

Abstract

The phenomenon of fake news has become a much contentious issue recently. The controversy regarding this issue has further been intensified by the openness of social media platforms. Via a systematic review, this paper offers a discussion on the spread and detection techniques of fake news on Social Networking Sites (SNSs). A total of 47 articles eventually fulfilled the inclusion criteria and were coded for the literature synthesis. The overall findings from the literature on fake news and social media have been extracted and synthesized to explore the creation, influence and popular techniques and dimensions used for fake news detection on SNSs. The results showed that various entities are involved in the creation and spread of fake news on SNSs, including malicious social and software agents. It was also found that early registered users, old people, female users, delusion-prone persons, dogmatic persons, and religious fundamentalists are more likely to believe in fake news than other groups of individuals. One of the major problems of the existing techniques is their deficiency in datasets. Therefore, future studies on fake news detection should focus on developing an all-inclusive model with comprehensive datasets. Social media users require fake news detection skills especially using linguistic approach. This study provides the public with valuable information about the spread and detection of fake news on SNSs. This is because SNSs are an important avenue for fake news providers.

Keywords: Fake news, social media, detection techniques, news content, social network

Introduction

In the last few years, researchers have designed and developed a variety of models and techniques for fake news detection purposes (e.g., Wang, 2017; Granik & Mesyura, 2017; Ahmed, Traore, & Saad, 2017; Aldwairi & Alwahedi, 2018; Zhang, Cui, Fu, & Gouza, 2018; Papanastasiou, 2018; Yang, Zheng, Zhang, Cui, Li, & Yu, 2018; Reis, Correia, Murai, Veloso, Benevenuto, & Cambria, 2019). Their contributions focused on different dimensions of news characteristics including social context (Shu, Bernard, & Liu, 2019) using various linguistic

* Universiti Sultan Zainal Abidin, Kuala Terengganu, Malaysia. Email: isyaku87@gmail.com

** Universiti Sultan Zainal Abidin, Kuala Terengganu, Malaysia. Email: mnazrix@gmail.com

*** Taylor's University, Selangor, Malaysia. Email: akibu1989@hotmail.com

and network-based approaches (Conroy, Rubin, Chen, 2015; Reis et al., 2019). Fake news detection on Social Networking Sites (SNSs) is an emerging research area with various challenges (Ahmed et al., 2017) particularly related to techniques and datasets.

Fake news in general is defined as misinformation intentionally fabricated and spread to mislead the audience for political and/or financial gains (Pennycook & Rand, 2017; Wu & Liu, 2018; Tandoc, Lim, & Ling, 2018). Types of fake news include hoax, propaganda, trolling, and satire (Verstraete, Bambauer, & Bambauer, 2017). Fake news on social media is a trans-disciplinary area of research interest adopted by science and social science researchers (Zhou, Zafarani, Shu, & Liu, 2019). The major challenge is how to develop a fake news detection technique that incorporates every news dimension such as news content, news sources, and social dimensions. Nevertheless, empirical research focusing on the design and development of models for fake news detection on SNSs is still limited and sporadic (Ahmed et al., 2017; Niklewicz, 2017).

The advent of SNSs has changed the manner in which people consume news and increased their exposure to a variety of news (Shu, Wang, & Liu, 2017; Flintham, Karner, Bachour, Creswick, Gupta, & Moran, 2018). There are currently over 3.2 billion social media users in the world, constituting about 42% of the earth's population (Oberlo, 2019). This situation has further intensified the spread of news articles containing false content intentionally created for a variety of financial and political purposes with potential detrimental effects on individuals and the society (Shu et al., 2017). This is because news readers often lack the knowledge required to interpret news critically. Detecting fake news is imperative to ensure that the audience receive credible information and help maintain a reliable news ecosystem (Shu et al., 2019).

It was found that studies on fake news generally focus on its effects other than causes, which suggest the need for further research incorporating the origin and causes of fake news (Dumitrache, 2019). Previous reviews on the current trends of fake news detection on social media focused on effects, solutions and process of manifestation (Campan, Cuzzocrea, & Truta, 2017) other than origin, formation and category of people affected by fake news. None of the previous reviews focused on systematic review of fake news studies on SNSs. Therefore, via a systematic review of studies on fake news and social media, this paper aims to explore the formation of fake news on SNSs, identify the category of news readers who are likely to fall for fake news on SNSs, and determine the popular techniques and dimensions of fake news detection on SNSs. The current study offers a comprehensive evaluation of the studies on fake news on SNSs, in view of its formation, potential effects, and various detection techniques. This is because detection of fake news on SNSs is an important phenomenon that requires scholarly attention (Ahmed, Hinkelmann, & Corradini, 2019).

Approaches to Fake News Detection

There are two main approaches used by researchers to tackle fake news detection problem (Yang et al., 2018): linguistic approach and network approach. Linguistic approaches, such as natural language processing, focus on news content with particular attention to news patterns through analysis of underlying semantics. Currently, this method is supplemented by network approaches in the process of fake news detection (Zhang, Gupta, Kauten, Deokar, & Qin, 2019). First, various linguistic approaches were used, such as natural language processing techniques

(Mihalcea & Strapparva, 2009), sentiment analysis, lexical and content similarity, as well as style similarity (Liu, 2012). Some researchers combined this method alongside other linguistic or network approaches (Feng, Banerjee & Choi, 2012). Other researchers paid attention to semi-supervised learning methods (Hai, Zhao, Cheng, Yang, & Li, 2016). Subsequently, some deep language structures, such as the syntax tree, were utilized (Yang et al., 2018). It was further discovered that even deep syntax analysis is often insufficient for fake news prediction.

According to Feng and Hirst (2013), most fake news providers use language to mislead the audience. Though they try to control their communication, language “leakage” occurs. In this regard, linguistic approach examines such leakage occurrences within the news content. In data representation, texts are represented using the “bag of words” approach. In the bag of words approach, n-grams (individual words) frequencies are collected and evaluated to reveal deception signals (Markowitz & Hancock, 2014). In addition, semantic analysis extends the n-gram by integrating the features of profile compatibility to enhance classification performance (Feng & Hirst, 2013). Another way of detecting fake news is discourse analysis in which deception cues are found both in CMC communication and in news content, by identifying instances of rhetoric relations between linguistic elements (Potthast, Kiesel, Reinartz, Bevendorff, & Stein, 2017; Rubin & Lukoianova, 2015).

On the other hand, network-based approaches focus on the analyses of network structure and social network behaviors. With the advancement of knowledge graph, methods based on the knowledge graph analysis can achieve high accuracy. In addition, deep learning models are generally used for fake news detection. Recently, various deep learning models have been proposed based on the word vectors for QA and summarization (Yang et al., 2018). Network approaches, which are innovative and diverse, are meant to supplement linguistic approaches. This is because for researchers to investigate the value of “unstructured data”, there is a need for robust fake news detection techniques (Zhang et al., 2019).

Generally, the features used in fake detection are obtained from two main sources: news contents and social contexts. News content techniques utilize features obtained from linguistic and visual information. Linguistic features aim to capture specific writing styles and sensational headlines while visual features identify fake images that are intentionally created to mislead the audience. On the other hand, social context techniques make use of features obtained from aspects such as user profiles, posts and shares as well as social networks. Features obtained from user profiles are used to measure characteristics and reliability of the users. Posts features characterize the user’s social responses while social network features are obtained through certain networks, for instance co-occurrence network and diffusion network (Shu, Wang, and Liu, 2018a; Shu et al., 2019). This is one of the aspects taken into consideration in this article to identify the most popular techniques and dimensions of fake news detection on SNSs.

In summary, the aforesaid techniques and features are used for fake news detection on SNSs (Yang, Shu, Wang, Gu, Wu, & Liu, 2019). They are largely adopted in natural language processing, data mining, and machine learning with various tools and datasets. Fake news detection techniques adopt various news-related features, such as body text, headline, publisher, news sources as well as and social-related features including feedback, propagation paths and propagators (Zhou et al., 2019). They focus on differences in writing styles between fake and true news, information provided in news dissemination, and credibility headlines, publishers, source websites, comments, and users to detect fake news. The current advances in

news verification approaches attempt to address the rising demand for automated means of fake news detection (Rubin, Chen, & Conroy, 2015).

Method

Search procedure and inclusion criteria

A systematic review was carried out with multidisciplinary literature on the fake news and SNSs, using “fake news” and “social media” as keywords. An internet-based search was conducted to locate recent and relevant articles published between 2014 and 2019. Only articles that focus on detection of fake news in social media were considered. The articles were collected from online sources bibliographic databases. In the selection criteria, content relevance (empirical and conceptual papers focusing on fake news detection), year of publication between 2014–2019, and language (English) were considered. The search for literature was inclusive within the data pool consisting of empirical and conceptual papers in official science and social science journals, as well as refereed conference proceedings, refereed studies, and research reports that investigated the phenomenon of fake news on SNSs. This is because relevant and up-to-date information could be sourced through synthesis of articles from various sources. The following figure depicts information on the article selection process.

Figure 1. Article selection process.

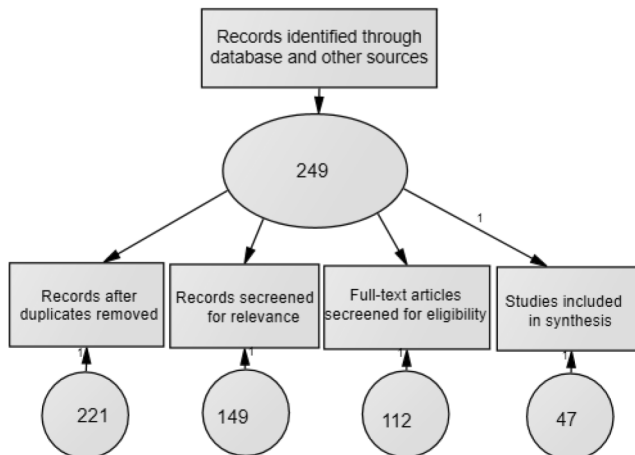


Figure 1 shows the article selection process used in this study. The preliminary online search of the above-mentioned data pool identified 249 papers. The abstracts and references were separately read and collected for decision. After eliminating the duplicates, 221 articles were considered for full text reading, after which 149 articles were screened for relevance. Subsequently, 112 articles fulfilled the initial selection criteria. Finally, 47 articles were selected and coded for literature synthesis. Studies on fake news fall into three categories, namely: occurrence of fake news within public discourse, effects of fake news, and detection of fake news (Egelhofer & Lecheler, 2019). This review specifically considered studies within the aforesaid categories.

Coding procedure

During the literature search and preliminary coding, special attention was paid to articles expected to form initial components of a systematic framework. Then, an initial coding framework was created to classify the articles based on the study purpose, approach, and major findings. The coding framework was revised as the process of analysis progresses. Considering the constant comparative method (Strauss & Corbin 1990), the data obtained from each article were continuously compared for reanalysis and revision of the framework. Three coders were used to achieve consistency, accuracy and validity of the findings and analysis during the process of coding (Creswell, 2012). The coding was conducted with the aim of identifying themes and patterns on the spread of fake news, users who are likely to fall for fake news, and the most popular techniques and dimensions used for fake news detection on SNSs. The coders reached 100 percent agreement on the final results.

Results

To mitigate the spread of fake news on SNSs, it is imperative to ascertain how the fake news is created and identify an effective technique to tackle the problem. The first objective of this study was to investigate the spread of fake news on SNSs. It was found that fake news is more likely to spread faster than the truth. Identifying key constructors of fake news is highly necessary in developing an effective technique for its detection. It is evident in the previous studies on fake news detection on SNSs that fake news is co-authored and co-constructed by malicious users of social media, journalists, malicious software agents, and organizations. These co-constructors create fake news purposely for social, political and financial incentives (n=31). Fake news can be artificially or automatically generated, but human beings bear great responsibility for fake news creation. For example, software agents can mass-send fake news content to manipulate the opinion of a group of agents and users in the network (Niklewicz, 2017). Also, a social bot could generate and use a personal account by simulating the use of a Web browser programmatically to penetrate a group of users who are already engaged in conversations about a particular subject, thereby using them as spreaders of fake news to their extended subnetworks (n=6). Most fake news stories are intentionally created and are likely to spread through sequential sharing by social media users. According to Papanastasiou (2018), agents are logically more likely to share a story the more it has been previously shared by their peers, resulting in fake news content going viral.

The above findings indicate that journalism is in danger of being repressed by satirists who speed up the spread of falsehoods. It should be noted that fake-news related issues, such as the open circulation of malicious falsehoods and lack of effective fact-checking seem to pose a great challenge to the fundamental ethics of journalism (Egelhofer & Lecheler, 2019). In this case, the mainstream media may use the internet to expose, at least, the most prominent fake news stories (Vargo, Guo, & Amazeen, 2018), since individual fact-checking is only possible by educated audience. Some SNSs, such as Facebook, have promised action to limit the spread of fake news, but other issues remain, including confirmation bias – the use of “likes” and information sharing with friends (White, 2017). This situation corresponds to the communication hypotheses, which presumes that audience are more likely to form their beliefs based on how the information is presented other than its sources. A such, providing an all-inclusive method for fake news detection is highly imperative.

The second objective of this study was to identify the category of people who are likely to fall for fake news on SNSs. Emerging themes from the previous studies suggest that early registered users, old people, female users, delusion-prone persons, dogmatic persons, religious fundamentalists and those who overly accept weak claims are likely to fall for fake news. In addition, identifying the potential groups of individuals who are likely to be affected by fake news on SNSs is useful for decision-makers to combat the spread of fake news. However, due to the large number of social media users, countering the spread of fake news could be quite challenging, especially because not every user has the ability to personally fact-check news. Individual fact-checking requires skills and time, and if possible, both the news content and the news source need to be verified. According to Egelhofer and Lecheler (2019), various factors could influence the decision of social media users on the authenticity of information, and therefore extensive verification is needed to avoid falling for fake news.

The third objective of this study was to identify the popular techniques and dimensions used for fake news detection on SNSs. This is because providing a well-developed web-based technique for fake news detection could supplement the individual fact-checking process. This study revealed that most of the previous studies ($n=23$) on fake news detection on SNSs utilized machine learning methods, focusing on techniques such as supervised learning (Reis et al., 2019), semi-supervised learning (Shu et al., 2017), reinforcement learning (Aymanns, Foerster, Georg, 2017), feature extraction (Zhang et al., 2018; Karimi, Roy, Saba-Sadiya, & Tang, 2018), stacked ensemble of classifiers (Thorne, Chen, Myriantous, Pu, Wang, & Vlachos, 2017), nonnegative matrix factorization (Shu, et al., 2019), flat classification and hierarchical classification (Rajdev & Lee, 2015). Some studies utilized Bayesian approaches such as Bayesian network (Yang, et al., 2019), Bayesian inference (Kim, Tabibian, Oh, Schölkopf, & Gomez-Rodriguez, 2018; Tschitschek, Singla, Gomez Rodriguez, Merchant, & Krause, 2018), Naive Bayes Classifier (Granik, & Mesyura, 2017), and a combination of Support Vector Machines and Naive Bayes (Chen, Conroy, & Rubin, 2015).

Some studies ($n=6$) combined machine learning with natural language processing techniques such as sentiment analysis and N-gram analysis (Atodiresei, Tănăselea, & Iftene, 2018; Rubin, Conroy, Chen, & Cornwell, 2016; Shu, Mahudeswaran, Wang, Lee, & Liu, 2018b; Ahmed et al., 2017). Natural language processing approaches in combination with machine learning could manage content directly through detection of language patterns (Rubin et al., 2016). Other studies focused on crowd sourcing as a knowledge-based approach (Saez-Trumper, 2014; Tacchini, Ballarin, Della Vedova, Moret, & de Alfaro, 2017; Sethi, 2017). Very few studies utilized data mining (Jin, Cao, Zhang, & Luo, 2016) and deep learning approaches such as convolutional networks (Azzimonti, & Fernandes, 2018; Yang et al., 2018). Despite the existing efforts toward fake news detection though various computational methods, lack of comprehensive datasets with various features needed for the detection of fake news becomes one of major challenges (Shu et al., 2018b). Various datasets, such as news content, news sources, social and temporal dimensions must be considered in mitigating the spread of fake news. The following table depicts the dataset information used for fake news detection techniques utilized in the previous literature.

Table 1. Popular dimensions used for fake news detection.

<i>Themes</i>	<i>Feature</i>	
Theme 1	Fake pages	1
Theme 2	News content	14
Theme 3	News content and user profile	3
Theme 4	News content and number of story shares	4
Theme 5	News content and readers	6
Theme 6	News content, social context, and spatiotemporal information	1
Theme 7	Users (believers and (fact-checkers)	1
Theme 8	News sources	2
Theme 9	News content, user social engagements, publishers' partisan	1
Theme 10	Speaker profiles, topic information of the news articles	1
Theme 11	Users, social engagements	1
Theme 12	Networks and news articles	4
Theme 13	User beliefs and network connectivity	1
Theme 14	Headline and article body	5
Theme 15	Users, responders, moderators	1
Theme 16	News contents, user, user-news interaction, publisher-news relation	1

Table 1 shows the features used for fake news detection in the previous studies. Most of the studies relied on news content alone ($n=14$), followed by studies focusing on news content and readers, headline and article body. Other features used for fake news detection are: fake pages, user profile, story shares, fact-checkers, user engagement, publishers' partisan, speaker profile, topic information, networks, moderators, and news sources. These features generally fall under two major categories: news contents and social context. According to Shu et al. (2018a), fake news detection methods generally focus on using news contents and social contexts. But in most cases, other important features are ignored, such as news sources. By detecting malicious information created by non-credible sources, fake news researchers can develop algorithms to mitigate the prominence of such information. Each feature used in fake news detection has limitation on its own (Shu et al., 2017), and therefore, every feature is essential.

Of the 47 articles coded, only three utilized some aspects of social context in addition to news content for fake news detection on SNSs (i.e., Shu et al., 2017; Shu et al., 2018a; Papanastasiou, 2018; Shu et al., 2019). News contents was exclusively utilized in 14 studies (e.g., Chen et al., 2015; Rubin et al., 2016; Zhang et al., 2018; Atodiresei et al., 2018; Wu & Liu, 2018; Vosoughi, Roy, & Aral, 2018; Reis et al., 2019). Other articles relied to a large extent on features such as news headlines or social media pages alone. Only one study (Shu et al., 2018) included features such as content, social and temporal dimensions but did not take much consideration of the news sources, though to some extent, they tried to detect social bots

using Botometer. To achieve high accuracy in automatic fake news detection, the problem of dataset must be addressed first, though research on fake news is still developing (Aldwairi & Alwahedi, 2018).

Researchers have various options for fake news analysis. This study found that most of the current studies on fake news detection on SNSs paid particular attention to Twitter as an SNS (eg Rajdev & Lee, 2015; Thorne, et al., 2017; Atodiresei et al., 2018; Wu & Liu, 2018; Yang et al., 2018; Jang, Geng, Li, Xia, Huang, Kim, & Tang, 2018). Some researchers focus on Facebook as an SNS (eg Granik & Mesyura, 2017; Tacchini et al., 2017; Flintham et al., 2018; Tschatschek et al., 2018). Nevertheless, in an attempt to widen the scope and context of fake news analysis on SNSs, some studies focused on more than one SNS, for example, Twitter and Facebook (eg Shu et al., 2017; Yang et al., 2019), Weibo and Twitter (eg Liu & Wu, 2018; Kim et al., 2018), Facebook, Forex and Reddit (eg Aldwairi & Alwahedi, 2018), as well as Twitter, Sina Weibo, Facebook (Zubiaga, Aker, Bontcheva, Liakata, & Procter, 2018).

Some studies focused specifically on particular fact-checking websites such as PolitiFact (eg Long, Lu, Xiang, Li, & Huang, 2017; Wang, 2017; Zhang et al., 2018). Some of them considered PolitiFact and GossipCop (eg Shu et al., 2018b) while others combined PolitiFact and BuzzFeed (eg Shu et al., 2018a; Shu, et al., 2019). In addition, there are some studies that focus on particular geographical locations especially US and Canada (eg Conroy et al., 2015; Rubin et al., 2016; Allcott, & Gentzkow, 2017), focusing on political news per se. Some of the studies (eg. Liu & Wu, 2018) considered the situation of fake news in China, focusing on its early detection. Some of the researchers gave reasons for limiting their data to particular social media, stating that the field of fake news detection is a relatively new area of research interest.

Discussion

The problem of fake news is not a new phenomenon, despite its complexity especially in the digital age when the detection of fake news becomes more challenging than ever before (Ahmed et al., 2017; Yang et al., 2018). Burkhardt (2017) asserted that increase in the means of spreading fake news is correlated with increase in the seriousness of its consequences. This study found that various parties are involved in the creation and spread of fake news on SNSs, such as journalists, users, and social bots. Research has shown that fake news is either deliberately created as pseudo-journalistic disinformation or used as a political instrument to delegitimize the news media (Egelhofer & Lecheler, 2019). The creation of fake news can be socially, politically or financially motivated. Influential individuals and organizations, such as politicians and advertisers provide incentives for writers to create fake news. According to Verstraete et al. (2017), fake news is created and circulated by platforms motivated more by financial incentives than by journalistic norms. Generally, as pointed out by Egelhofer and Lecheler (2019), fake news is intentionally created to deceive its audience.

Understanding how fake news is created is an important step in tackling the fake news phenomenon on SNSs (Louni & Subbalakshmi, 2014). According to Guille, Hacid, Favre, and Zighed (2013), analysis of information diffusion consists of modeling of information, detection of popular topics, and identification of influential users. To improve the flow of quality information and combat fake news on social media, it is essential to identify its origin and diffusion pattern (Campan et al., 2017; Jang et al., 2018). According to Mustafaraj and Metaxas (2017), steps in the spread of fake news on SNSs include creating websites to accommodate

fake news, creating anonymous accounts, identifying and following influential users, and targeting the users with fake messages. It is evident that internet bots are also involved in the creation and spread of fake news. They are social media accounts controlled by a computer algorithm and can be used specifically to spread fake news on SNSs (Shu et al., 2019).

The social media has made information disseminated faster, but they serve as a platform for the spread of fake news (Tambuscio et al., 2015; Tacchini et al., 2017; Shu et al., 2018a; Yang et al., 2018). According to Koohikamali and Sidorova (2017), perceived relevance and reliability of information must be considered when sharing the information. Nevertheless, few people seem to verify the reliability of news before sharing them with their friends. This might be due to decentralization of information provided by social media to a large extent. All users of social media are potential news sources (Tambuscio et al., 2015). The spread of fake news in social media affects the accuracy of online news content (Jin et al., 2016). According to Rubin et al. (2016), fake news in social media may mislead inattentive readers.

The effect of fake news on social media users could significantly extend to the offline society (Zhang et al., 2018; Shu et al., 2018a). Tackling the problem of fake news on SNSs could improve the quality of information, which will in turn decrease the rate of potential socio-political crisis in the society at large. This study found that the victims of fake news include early registered users, old people, female users, delusion-prone persons, dogmatic persons, and religious fundamentalists. This shows that fake news has significant influence in social media community and the society at large. It should be noted that fake news is consumed enthusiastically by readers for whom it reinforces existing beliefs (Verstraete et al., 2017). Users can be believers or fact-checkers (Tambuscio et al., 2015). The user's ability to fact-check news on social media is determined by their news verification skills. Readers who fact-check news might be skeptical about the news credibility. Chen et al. (2015) asserted that the ability to critically evaluate the quality of information is an essential skill, which is lacking in a large portion of the population. Rubin et al. (2016) argued that news readers often lack the literacy skills necessary for critical interpretation of news.

In addition, this study found that previous studies on fake news detection focused on using news content as dataset, which is considered ineffective (Shu et al., 2019). To curtail the spread of fake news, there is a need to incorporate various datasets beyond news content per se. According to Shu et al. (2019), various dimensions must be considered in fake news detection, including news content, temporal user engagement, credibility. There are various weaknesses in the existing models in relation to incorporating the information needed for accurate fake news detection, such as news content and user profile (Yang et al., 2019). In addition, very few studies incorporated semantic content (Sethi, 2017) while some analyses were performed on headlines only, arguing that users often share links they have not opened (Sommariva, Vamos, Mantzarlis, Dào, & Tyson, 2018). For example, one can verify news posted by bots by looking at the sentence construction, because automatic fake news creation could result in poorly constructed sentences which could raise suspicion (Wang, Angarita, & Renna, 2018).

It is shown that majority of the studies on fake news focused on Twitter while some studies focus, to a great extent, on Facebook as an SNS. Nevertheless, readers are possibly exposed to fake news in various SNSs other than Twitter and Facebook (Jin et al., 2016). Another issue is that in terms of geographical focus, most of the previous studies were conducted in the US and China. The phenomenon of fake news, being prevalent in several other parts of the world, needs to be addressed globally. Researchers, in different parts of the world, must

direct their efforts toward mitigating the phenomenon of fake news. This is because fake is likely to be understood within its specific context of production, because the journalism practices vary across the globe (Egelhofer & Lecheler, 2019). In addition, approaches to the ethics of journalism among reporters differ across contexts. These instances signify the need to curtail the fake news phenomenon both in specific and general contexts.

Surprisingly, research found that fake news on SNSs is more popular than factual news based on their number of shares and likes (Howard, Bolsover, Bradshaw, Kollanyi, & Neudert, 2017). This shows how prevalent the phenomenon of fake news becomes in the digital age of communication. Therefore, fact-checking might not be much effective to counteract the problem of fake news on SNSs. Tracing the root of fake news is much significant in regulating its effects, but very few studies focused on news sources, particularly websites as potential source of fake news (Saez-Trumper; 2014; Aldwairi & Alwahedi, 2018). The original source of fake news on SNSs is difficult to track down especially due to the anonymity of its authors and lack of journalism ethical standard (Jang et al. (2018). There is a need for computer-assisted techniques to detect fake news, but research on automatic detection of fake news is still in its infancy stage.

In summary, fake news is co-constructed by various parties, such as social media users, journalists, malicious software agents, and organizations. It is created for social, political and financial reasons. Nevertheless, the datasets used in most studies on fake news detection on social media are deficient. Though news content is the fundamental means through which fake news can be detected, additional datasets such as user profile and social behavior of users on social media are needed (Shu et al., 2018b; Karimi et al., 2018). For example, comments by users, such as “this is fake news”, could help in analyzing the news credibility. Therefore, it is necessary to utilize comprehensive datasets in fake news detection on SNSs. Shu et al. (2018a) noted that the news ecosystem on social media involves news publisher, social media users and the news itself. In this regard, Mustafaraj and Metaxas (2017) asked a very important question “in the current context of the ubiquitous, web-based, socio-technical systems, what decisions should be made by humans and algorithms?” In order to tackle the problem fake news detection, both behavioral and social entities must be incorporated while knowledge and computer-assisted techniques need to be combined (Ahmed et al., 2019).

Conclusions

This study has evaluated the spread of fake news and its detection techniques on social networking sites (SNSs). The spread of fake news and its various detection techniques for effective and improved quality of information on SNSs have been highlighted. Several fake news detection techniques were proposed in the previous studies. Nevertheless, the manner in which fake news spreads and the purpose for which it is created makes its detection challenging, especially by utilizing news content as dataset per se. One of the major problems of the existing techniques is their deficiency in datasets. This is because majority of the existing detection techniques focus on identifying clues from news contents, which are largely unreliable since fake news is often intentionally created by imitating true news. As such, misperceptions are simply formed through fake news. Researchers across the world must pay attention to the fake news phenomenon both in specific and general contexts.

This study proved that most of the current efforts to detect fake news on SNSs utilized features that control information from specific datasets. Therefore, future research on fake news detection should focus on developing an all-inclusive model with comprehensive datasets. The multifaceted nature of information on social media, with its potential effects, must be considered. The most important step is to ensure that malicious information is controlled, because this kind of content, if accessible, is sometimes used as a reference especially among the uneducated audiences (Egelhofer & Lecheler, 2019). This could be achieved by developing an all-inclusive fake news detection technique that incorporates the news content and sources – such as malicious websites, alternative and partisan media – as well as social context. This recommendation requires a comprehensive computer-assisted network approach that is capable of capturing malicious information, supported by research on media literacy in general. This proposed technique, if developed with high level of precision, is expected to detect both intentional and unintentional fake news.

In addition, there is a need to safeguard the public from the potential threat of fake news. Hence, for social media users, knowledge and skills on fake news detection especially using natural language processing or linguistic approach are necessary. In addition, creating awareness amongst the public on the existence of fake news on SNSs is crucial in mitigating this phenomenon. This study provides the public with valuable information about the spread and detection of fake news on SNSs. The current findings are expected to broaden our understanding of the origin of fake news and its detection techniques on SNSs, which is necessary to enrich the quality of information the digital age of communication. This is because SNSs are an important avenue for fake news providers. To tackle fake news on SNSs, there is a need for collective efforts of researchers, journalists and professional bodies.

References

- Ahmed, H., Traore, I., & Saad, S. (2017). Detection of online fake news using n-gram analysis and machine learning techniques. In Traore I., Woungang I., Awad A. (Eds.), *Proceeding of International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments* (pp. 127-138). Springer, Cham.
- Ahmed, S., Hinkelmann, K., & Corradini, F. (2019). Combining Machine Learning with Knowledge Engineering to detect Fake News in Social Networks-a survey. In A. Martin, K. Hinkelmann, A. Gerber, D. Lenat, F. van Harmelen, P. Clark (Eds.), *Proceedings of the AAAI 2019 Spring Symposium on Combining Machine Learning with Knowledge Engineering* (AAAI-MAKE). Palo Alto, California, USA: Stanford University,
- Aldwairi, M., & Alwahedi, A. (2018). Detecting Fake News in Social Media Networks. *Procedia Computer Science*, 141, 215-222. doi:10.1016/j.procs.2018.10.171
- Atodiresei, C. S., Tănăseselea, A., & Iftene, A. (2018). Identifying Fake News and Fake Users on Twitter. *Procedia Computer Science*, 126, 451-461. doi:10.1016/j.procs.2018.07.279
- Aymanns, C., Foerster, J., Georg, C. (2017). Fake News in Social Networks. Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3023320
- Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130-1132. doi:10.1126/science.aaa1160
- Burkhardt, J. M. (2017). History of Fake News. *Library Technology Reports*, 53(8), 5-9. Retrieved from <https://journals.ala.org/index.php/ltr/article/view/6497/8636>

- Campan, A., Cuzzocrea, A., & Truta, T. M. (2017). Fighting fake news spread in online social networks: Actual trends and future research directions. *Proceedings of IEEE International Conference on Big Data (Big Data)* (pp. 4453-4457). Boston, MA, USA.
- Chen, Y., Conroy, N. J., & Rubin, V. L. (2015). Misleading online content: Recognizing clickbait as false news. *Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection* (pp. 15-19). Seattle, Washington, USA.
- Conroy, N. J., Rubin, V. L., & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology*, 52(1), 1-4. <https://doi.org/10.1002/pr2.2015.145052010082>
- Creswell, J. W. (2012). *Research design: Qualitative, quantitative and mixed methods approaches* (2nded.). London, UK: SAGE publications.
- Dumitrache, A. C. (2019). Shades of fake news: Manifestation, effects and ways to combat false information. *Romanian Journal of Communication and Public Relations*, 21 (2): 41-57. [doi: 10.21018/rjcrp.2019.2.276](https://doi.org/10.21018/rjcrp.2019.2.276)
- Egelhofer, J. L., & Lecheler, S. (2019). Fake news as a two-dimensional phenomenon: A framework and research agenda. *Annals of the International Communication Association*, 43(2), 97-116.[doi:10.1080/23808985.2019.1602782](https://doi.org/10.1080/23808985.2019.1602782)
- Feng, S., Banerjee, R., & Choi, Y. (2012). Syntactic stylometry for deception detection. *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2* (pp. 171-175). Jeju Island, Korea.
- Feng, V. W., & Hirst, G. (2013). Detecting deceptive opinions with profile compatibility. *Proceedings of the Sixth International Joint Conference on Natural Language Processing* (pp. 338-346). Nagoya, Japan
- Flintham, M., Karner, C., Bachour, K., Creswick, H., Gupta, N., & Moran, S. (2018). Falling for fake news: investigating the consumption of news via social media. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (p. 376). Montreal QC, Canada
- Granik, M., & Mesyura, V. (2017). Fake news detection using naive Bayes classifier. *Proceedings of IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)* (pp. 900-903). Kiev, Ukraine.
- Guille, A., Hacid, H., Favre, C., & Zighed, D. A. (2013). Information diffusion in online social networks: A survey. *ACM Sigmod Record*, 42(2), 17-28. [doi:10.1145/2503792.2503797](https://doi.org/10.1145/2503792.2503797)
- Hai, Z., Zhao, P., Cheng, P., Yang, P., Li, X. L., & Li, G. (2016). Deceptive review spam detection via exploiting task relatedness and unlabeled data. *Proceedings of the 2016 conference on empirical methods in natural language processing* (pp. 1817-1826). Austin, Texas, USA.
- Markowitz DM, Hancock JT (2014) Linguistic Traces of a Scientific Fraud: The Case of Diederik Stapel. *PLoS ONE*, 9(8): e105937.[doi:10.1371/journal.pone.0105937](https://doi.org/10.1371/journal.pone.0105937)
- Howard, P. N., Bolsover, G., Bradshaw, S., Kollanyi, B., & Neudert, L.-M. (2017). Junk news and bots during the U.S. Election: What were michigan voters sharing over twitter? Retrieved from <http://275rzy1ul4252pt1hv2dqyuf.wpengine.netdnacloud.com/wpcontent/uploads/2017/07/2206.pdf>
- Jang, S. M., Geng, T., Li, J. Y. Q., Xia, R., Huang, C. T., Kim, H., & Tang, J. (2018). A computational approach for examining the roots and spreading patterns of fake news: Evolution tree analysis. *Computers in Human Behavior*, 1 (84): 103-113. [doi: 10.1016/j.chb.2018.02.032](https://doi.org/10.1016/j.chb.2018.02.032)
- Jin, Z., Cao, J., Zhang, Y., & Luo, J. (2016). News verification by exploiting conflicting social viewpoints in microblogs. *Proceedings of Thirtieth AAAI Conference on Artificial Intelligence*. Retrieved from: <https://dl.acm.org/citation.cfm?id=3016318>
- Karimi, H., Roy, P., Saba-Sadiya, S., & Tang, J. (2018). Multi-source multi-class fake news detection. *Proceedings of the 27th International Conference on Computational Linguistics* (pp. 1546-1557). Santa Fe, New Mexico, USA.
- Kim, J., Tabibian, B., Oh, A., Schölkopf, B., & Gomez-Rodriguez, M. (2018). Leveraging the crowd to detect and reduce the spread of fake news and misinformation. *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining* (pp. 324-332). Marina Del Rey, CA, USA

- Koohikamali, M., & Sidorova, A. (2017). Information re-sharing on social network sites in the age of fake news. *The International Journal of an Emerging Transdiscipline*, 1 (20): 215-235. Retrieved from <http://www.informingscience.org/Publications/3871>
- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 5(1), 1-167. doi: 10.2200/S00416ED1V01Y201204HLT016
- Liu, Y., & Wu, Y. F. B. (2018). Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks. *Proceeding of the Thirty-Second AAAI Conference on Artificial Intelligence*. Retrieved from: <https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view-Paper/16826>
- Long, Y., Lu, Q., Xiang, R., Li, M., & Huang, C. R. (2017). Fake news detection through multi-perspective speaker profiles. *Proceedings of the Eighth International Joint Conference on Natural Language Processing* (pp. 252-256). Retrieved from <https://www.aclweb.org/anthology/I17-2043>
- Louni, A., & Subbalakshmi, K. P. (2014). A two-stage algorithm to estimate the source of information diffusion in social media networks. In *Computer Communications Workshops (INFOCOM WKSHPS), 2014 IEEE Conference on* (pp. 329-333). doi:10.1109/INFOCOMW.2014.6849253
- Mihalcea, R., Strapparava, C. (2009). The lie detector: Explorations in the automatic recognition of deceptive language. *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers* (pp. 309-312). Suntec, Singapore.
- Mustafaraj, E., & Metaxas, P. T. (2017, June). The fake news spreading plague: was it preventable?. *Proceedings of the 2017 ACM on Web Science Conference* (pp. 235-239). Troy, New York, USA.
- Mustafaraj, E., & Metaxas, P. T. (2017, June). The fake news spreading plague: was it preventable?. In *Proceedings of the 2017 ACM on Web Science Conference* (pp. 235-239). Troy, New York, USA
- Niklewicz, K. (2017). Weeding out fake news: an approach to social media regulation. *European View*, 16(2), 335-335. doi:10.1007/s12290-017-0468-0
- Oberlo (2019). 10 Social Media Statistics You Need to Know in 2019 [Infographic]. Retrieved from: <https://my.oberlo.com/blog/social-media-marketing-statistics>
- Papanastasiou, Y. (2018). Fake news propagation and detection: A sequential model. Retrieved from: <https://ssrn.com/abstract=3028354>
- Pennycook, G., & Rand, D. G. (2017). Who falls for fake news? The roles of analytic thinking, motivated reasoning, political ideology, and bullshit receptivity. *SSRN Electronic Journal*. Retrieved from: <https://www.researchgate.net/>
- Pennycook, G., Cannon, T., & Rand, D. G. (2017). Prior exposure increases perceived accuracy of fake news but has no effect on entirely implausible statements. *Journal of Experimental Psychology*, 147(12):1865-1880. doi: 10.1037/xge0000465
- Pothast, M., Kiesel, J., Reinartz, K., Bevendorff, J., & Stein, B. (2017). A Stylometric Inquiry into Hyper-partisan and Fake News. Retrieved from: <https://arxiv.org/abs/1702.05638>
- Rajdev, M., & Lee, K. (2015). Fake and spam messages: Detecting misinformation during natural disasters on social media. *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)* (pp. 17-20). Singapore, Singapore.
- Reis, J. C., Correia, A., Murai, F., Veloso, A., Benevenuto, F., & Cambria, E. (2019). Supervised Learning for Fake News Detection. *IEEE Intelligent Systems*, 34(2), 76-81. doi: 10.1109/MIS.2019.2899143
- Rubin, V. L., & Lukoianova, T. (2015). Truth and deception at the rhetorical structure level. *Journal of the Association for Information Science and Technology*, 66(5), 905-917. doi: 10.1002/asi.23216
- Rubin, V. L., Chen, Y., & Conroy, N. J. (2015). Deception detection for news: three types of fakes. *Proceedings of the Association for Information Science and Technology* (pp. 1-4). St. Louis, Missouri.
- Rubin, V., Conroy, N., Chen, Y., & Cornwell, S. (2016). Fake news or truth? using satirical cues to detect potentially misleading news. In *Proceedings of the second workshop on computational approaches to deception detection* (pp. 7-17). doi:10.18653/v1/w16-0802
- Saez-Trumper, D. (2014). Fake tweet buster: a webtool to identify users promoting fake news on twitter. *Proceedings of the 25th ACM conference on Hypertext and social media* (pp. 316-317). doi:10.1145/2631775.2631786

- Sethi, R. J. (2017). Crowdsourcing the verification of fake news and alternative facts. *Proceedings of the 28th ACM Conference on Hypertext and Social Media* (pp. 315-316). Prague, Czech Republic.
- Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 19(1), 22-36. doi:10.1145/3137597.3137600
- Shu, K., Wang, S., & Liu, H. (2018a). Understanding user profiles on social media for fake news detection. *Proceedings of the IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)* (pp. 430-435). Miami, FL, USA
- Shu, K., Mahudeswaran, D., Wang, S., Lee, D., & Liu, H. (2018b). Fakenewsnet: A data repository with news content, social context and dynamic information for studying fake news on social media. Retrieved from: <https://arxiv.org/abs/1809.01286>
- Shu, K., Bernard, H. R., & Liu, H. (2019). Studying fake news via network analysis: detection and mitigation. Retrieved from: https://link.springer.com/chapter/10.1007/978-3-319-94105-9_3
- Sommariva, S., Vamos, C., Mantzaris, A., Dào, L. U. L., & Martinez Tyson, D. (2018). Spreading the (fake) news: exploring health messages on social media and the implications for health professionals using a case study. *American Journal of Health Education*, 49(4), 246-255. <https://eric.ed.gov/?id=EJ1183667>
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and Techniques*. Newbury Park: Sage Publications.
- Tacchini, E., Ballarin, G., Della Vedova, M. L., Moret, S., & de Alfaro, L. (2017). Some like it hoax: Automated fake news detection in social networks. Retrieved from <https://arxiv.org/abs/1704.07506>.
- Tambuscio, M., Ruffo, G., Flammini, A., & Menczer, F. (2015). Fact-checking effect on viral hoaxes: A model of misinformation spread in social networks. *Proceedings of the 24th international conference on World Wide Web* (pp. 977-982). Florence, Italy.
- Tandoc Jr, E. C., Lim, Z. W., & Ling, R. (2018). Defining “fake news” A typology of scholarly definitions. *Digital Journalism*, 6(2), 137-153. doi: 10.1080/21670811.2017.1360143
- Thorne, J., Chen, M., Myriantous, G., Pu, J., Wang, X., & Vlachos, A. (2017). Fake news stance detection using stacked ensemble of classifiers. *Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism* (pp. 80-83). Copenhagen, Denmark.
- Tschiatschek, S., Singla, A., Gomez Rodriguez, M., Merchant, A., & Krause, A. (2018). Fake news detection in social networks via crowd signals. *Proceedings of the The Web Conference 2018* (pp. 517-524). Lyon, France.
- Vargo, C. J., Guo, L., & Amazeen, M. A. (2018). The agenda-setting power of fake news: A big data analysis of the online media landscape from 2014 to 2016. *New Media & Society*, 20(5), 2028–2049. doi: 10.1177/1461444817712086
- Verstraete, M., Bambauer, D. E., & Bambauer, J. R. (2017). Identifying and countering fake news. *Arizona Legal Studies Discussion*, 17 (15). doi:10.2139/ssrn.3007971
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 3(6), 1146-1151. doi:10.1126/science.aap9559
- Wang, P., Angarita, R., & Renna, I. (2018). Is this the Era of Misinformation yet: Combining Social Bots and Fake News to Deceive the Masses. *Proceedings of the The Web Conference 2018 on The Web Conference 2018* (pp. 1557-1561). Lyon, France.
- Wang, W. Y. (2017). Liar, liar pants on fire: A new benchmark dataset for fake news detection. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics* (pp. 422–426). Vancouver, Canada.
- White, A. (2017). Facebook and matters of fact in the post-truth era. Retrieved from: <https://ethicaljournalismnetwork.org/resources/publications/ethics-in-the-news/fake-news>
- Wu, L., & Liu, H. (2018). Tracing fake-news footprints: Characterizing social media messages by how they propagate. *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining* (pp. 637-645). Marina Del Rey, CA, USA.
- Yang, S., Shu, K., Wang, S., Gu, R., Wu, F., & Liu (2019). Unsupervised fake news detection on social media: A generative approach. *Proceedings of 33rd AAAI Conference on Artificial Intelligence*. Retrieved from: <https://aaai.org/ojs/index.php/AAAI/article/view/4508>

- Yang, Y., Zheng, L., Zhang, J., Cui, Q., Li, Z., & Yu, P. S. (2018). TI-CNN: Convolutional neural networks for fake news detection. Retrieved from: <https://arxiv.org/abs/1806.00749>
- Zhang, J., Cui, L., Fu, Y., & Gouza, F. B. (2018). Fake news detection with deep diffusive network model. Retrieved from: arXiv preprint arXiv:1805.08751.
- Zhang, C., Gupta, A., Kauten, C., Deokar, A. V., & Qin, X. (2019). Detecting Fake News for Reducing Misinformation Risks Using Analytics Approaches. *European Journal of Operational Research*, 1 (1): 1036-1052. doi: 10.1016/j.ejor.2019.06.022
- Zhou, X., Zafarani, R., Shu, K., & Liu, H. (2019). Fake News: Fundamental Theories, Detection Strategies and Challenges. *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining* (pp. 836-837). Melbourne, Australia.
- Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M., & Procter, R. (2018). Detection and resolution of rumours in social media: A survey. *ACM Computing Surveys (CSUR)*, 51(2), 32. doi:10.1145/3161603