Studies in Economics and Business Relations
https://sebr.sabapub.com
ISSN:2709-670X

2022Volume 1, Issue 1:17 – 30
DOI: 10.48185/sebr.v3i1.401

The spread of Misinformation on social media: An insightful countermeasure to restrict

Jamal Abdul Nasir Ansari[1], Mohd Azhar[2], Mohd Junaid Akhtar [3]

Department of Commerce, Aligarh Muslim University, Aligarh, 202002 India
Mobile No. +91-7618057190, jnasir4@gmail.com
Department of Commerce, Aligarh Muslim University, Aligarh, 202002 India
Mobile No. +91- 8630373750, azharqamar@gmail.com
Department of Commerce, Aligarh Muslim University, Aligarh, 202002 India
Mobile No. +91- 8273065877, junaidakhtar10@gmail.com

Received: 26.11.2021 • Accepted: 25.01.2022 • Published: 27.01.2022 • Final Version: 27.01.2022

Abstract: The term misinformation on social media has got significant attention in public sermons over the last few decades. This research article explores the growing tendency of misinformation on social media, how it influences people and prescribes insightful measures to counter the spreading of misinformation on social media. Systematic Literature Review (SLR) was employed on the three databases; Google Scholar; Scopus, Web of Science, following keywords; "misinformation", "disinformation", and "social media". A total of 34 articles were finally found suitable for the study. This study confirmed that self-motive and election campaigns are the major causes of misinformation on social media. This study manifested that machines can detect fake news to some extent but cannot be relied upon solely. Human intervention is equally important in identifying misinformation. Moreover, an efficient conceptual model has been proposed to counter the misinformation spread on social media.

Keywords: Misinformation, Disinformation, Social Media, Fake News, Countermeasure

1. Introduction

The emersion of social media as a strong medium of information has torn down all the barriers of physical boundaries and squeezed the whole world into a tiny space (Koo *et al.*, 2011; Chung *et al.*, 2012; Choi, 2013). Creating, spreading and getting the information is no longer a herculean task as it used to be in the past (Koo *et al.*, 2011).This phenomenon has made social media one of the most popular platforms for information diffusion (Shu *et al.,* 2020)that has got widespread recognition and lot of scholarly attention in the recent past (Collins *et al.*, 2020). In the last few years, the spread and pervasive use of social media has grown simultaneously worldwide and has become the modus operandi of the 21st century (Lange-Faria& Elliot, 2012).As technological advancement brings comfort and ease with it, at the same time; it causes distrust and distress too. The excess information on social media is becoming one of the leading causes of chaos and pandemonium among the masses(Collins *et al.*, 2020).Most of the information on social media is dubious and intended to mislead (Zhang &Ghorbani, 2020).Social media is flooded with fake news, misleading information, doctored videos, tempered facts and rumours. It was found in a survey conducted by statista that around 52% of users encounter fake news on a daily basis (Ahuja & Kumar, 2020). The rapid diffusion of misinformation is shaking human cognizance and making the decision-making process more cumbersome.

Rumours are one of the consistent features of social life. As a sense-modality progression, it could work as an anti-stress oxidant and help people grip with the unfamiliar situation (Wood, 2018).Social media has billions of active users globally. As per Statista, a German Company specializing in market and consumer data, Facebook has 2701 million, Instagram has 1158 million, and Twitter has 353 million active users worldwide as of October 2020 (Statista, 2020). The flagrancy and high discrepancy of information promulgated through large user communities influence the public sermons in the democratic society (Bagheri *et al.,* 2020). During global health crises such as the COVID-19 pandemic, common people tend to rely on (mis)information spread over social media to cope with their anxiety and make sense of baffling state.

The political scenario has a long history of misinformation on social media instigated pervasive alarm in recent years (Allcott*et al.,* 2019). The fact-checking contents usually lag behind themis/disinformation by 10-20 hours. Fake news is predominantly circulated by active users (Shao *et al.,* 2016). There have been growing tendencies noted regarding Online Social Media (OSM) services as the best and cheapest way to share the information pertaining to the events (Gupta *et al.,* 2013).

Over the recent past, the mis/disinformation spread on social media created venomous over it has been acknowledged (Shu *et al.,* 2020). Circulation of misinformation on social media specially fabricated news has raised concerned about an "infodemic," which exacerbates people muddle and daunts preventive measures. The results confirmed that a higher level of social media uses increased worry and vaccine misinformation in context with Covid-19 (Su, 2021). Social media-based misinformation rapidly increasing in this technological era (Eysenbach*et al.,* 2002). During the Covid-19 pandemic, fake news propagation has been increased specifically vaccine-related (Apuke& Omar, 2021). Circulation of health-related misinformation on social media constitutes a probable threat to public health (Waszak*et al.,* 2018). The widespread misinformation on social media seemed a menace to national integrity, security and democratic society particularly onward of a national election in Israel, Mexico Sweden and India. Levush (2019) reported that "as per the Facebook announcement, 783 unauthenticated pages have been removed which tied to Afghanistan, Albania, Algeria, Bahrain, Egypt, France, Germany, India, Indonesia, Iran, Iraq, Israel, Libya, Mexico, Morocco, Pakistan, Qatar, Saudi Arabia, Serbia, South Africa, Spain, Sudan, US, and Yemen. Approximately 21600 tweets were self-suspected to originated from Russia, Iran and Venezuela targeting Canadians with messages. Although, despite strick regulation on the Internet in China, 6.7 million illegal and false information were allegedly disseminated in a single month of July 2018" (Levush, 2019). Therefore, handling this global problem, addressing these concerns means identifying and intervening disinformation on social media is a matter of great concerns.

Subsequently, researches related to mis/disinformation detection are captivating the attention of industry practitioners, marketers and academicians to carry out more studies (Tucker *et al.,* 2018). However, many studies have been undertaken on mis/disinformation in social media setting, but the majority of them have talked upon health-related misinformation spread on social media and that too in natural calamity or health crisis like the recent COVID-19 pandemic. Moreover, traditional researches mainly focus on descriptive analysis of the early detection of fake news, its causes and ways to prevent the fabrication, which is intentionally spreading on social media for self-motive. Literature is lacking in this domain (Collins *et al.,* 2020). Hence an attempt has been made to address the issues pertaining to misinformation created causes and insightful countermeasure to restrict the spread and alleviating its negative outcome. Therefore, this paper explores and investigates the very nature of misinformation spread on social media and countermeasures to limit it. In other words, the purpose of this paper is more conceptual than practical. The aim is to get the fullest and most profound conceptual understanding of misinformation spread on social media and countermeasures to restrict it.

The present study is to systematically reviewed the existing literature that deals with tools, techniques and models of combating mis/disinformation spread over social media and to bring forth misinformation detection methods out of the existing literature in a more comprehensive and lucid way. Furthermore, it also proposes a model to counter misinformation.

In a nutshell, the present study delineates new insights in the existing literature and makes the following contribution to the subject matter: It represents the most recent studies on countering misinformation on social media and highlights various tools, techniques, methods, models and frameworks to mitigate the dispersion of misinformation in the social media context. Moreover, it gives a concrete review of the most recent trends and proposes a counter measurement model to extenuate the spread of misinformation on social media.

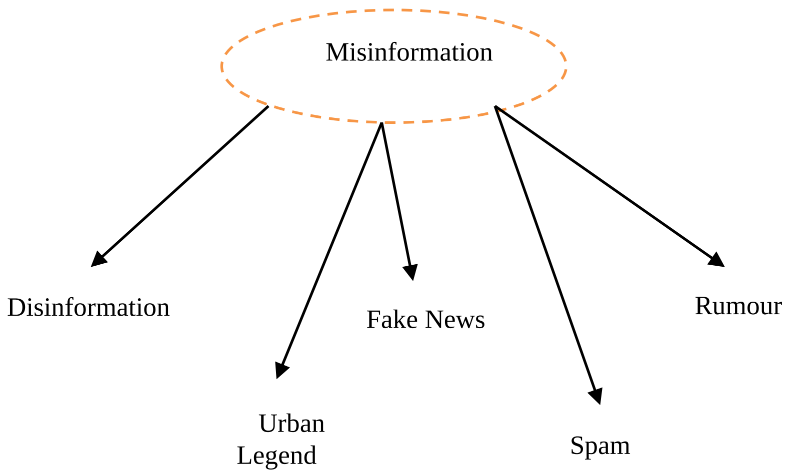
This paper is organized as follows: The first section presents the theoretical background that inquiries into different domains or fields related to mis/disinformation on social media, tracked by the methodology adopted for the study. The subsequent section deals with data extraction and analysis, succeeding by insightful measure to counter the misinformation spread on social media. This section summarises various methods of detecting misinformation on social media. The last section concludes the paper with a discussion and humble suggestions on how to mitigate the effect of misinformation on social media.

2. Theoretical background

Elucidation of Misinformation

The Oxford dictionary defines the term information as "facts provided or learned about something or someone". Oxford University described other forms of information as "misinformation, false or inaccurate, especially that is deliberately intended to deceive". Misinformation denotes false and inaccurate information circulated intentionally or unintentionally on social media (Alina *et al.*, 2017; Søe, 2018; Wu *et al.*, 2019) deliberately intended to betray. Similarly, disinformation is a faux statement that is aimed to mislead, specifically government propaganda, which is intentionally created by a rival power or media (Kumar and Geethakumari, 2014). There are many terms, other than misinformation and disinformation, which are used interchangeably but possess different connotation (Wu *et al.*, 2019). Figure 1 depicts the key terms which are often used interchangeably with misinformation.

Figure 1. Different terms related to misinformation



Source: Designed by Authors

According to Wu *et al*. (2019), these key terms are defined as "disinformation refers to inaccurate information which is usually distinguished from misinformation by the intention of deception, fake news refers to false information in the form of news (which is not necessarily disinformation since innocent users may unintentionally share it), rumour refers to unverified information that can be either true or false,spam refers to irrelevant information that is sent to a large number of users and urban legend is intentionally-spread misinformation that is related to fictional stories about local events" (Wu *et al.,* 2019).

Misinformation on Social Media

In a democratic society, spreading misinformation could cause harm to the majority of the people. Misinformation is strongly opposed to a democratic society. Empirical studies on mis/disinformation on social media can be additionally fragmented into various domains that explore a particular incident and those which consider the explicit type of content. The majority of the used data come from social media platforms such as Facebook and, more frequently, Twitter. During the political election campaign, misinformation through social media, specifically Facebook, remains high, and Fact checker websites play a dominant role in diffusion (Allcott*et al.,* 2019). It can be diffused on social media by political figures and aggravated by trolls and bots (Andi *et al.,* 2020). A USA based fact-checking organization reported that 70% of claims made during the 2016 US presidential campaign were delusive and false (Lewandowsky *et al.*, 2017).

A few years back, purposefully fabricated misinformation spread on social media has been subject to huge controversy and potential threats to a democratic society. Some social media events tended to be political, such as the 2012/2016 United State Presidential Election (Badawy*et al.,* 2018; Shao *et al.*, 2018; Maryam & Ali, 2019) and the 2013 Italian electoral campaign (Mocanu*et al.*, 2015). At the same time, some of them were crises-based, such as the 2020 Corona virus pandemic (Mian & Khan, 2020), the 2010 Chile earthquake (Mendoza *et al.*, 2010), the 2014 Ebola virus crisis (Jin*et al.,* 2014), the Boston Marathon Bombings (Gupta *et al.*, 2013; (Butler, 2014), and Hurricane Sandy (Gupta *et al.*, 2013). Based on the conspiracy theories, online social media (Facebook) user, which are blatantly interacting with confederacy information, seemed more susceptible to the deliberate false claim (Mocanu*et al.,* 2015). It was highlighted in the report of *The World Economic Forum* that "massive digital misinformation" as one of the big threats to a modern democratic society (Howell, 2013). Common people perceptions, belief, knowledge about the world are tempered through the information they access from electronic media and print media (McCombs & Shaw, 1972). Extraordinary practices to cop-up misinformation-both on social media and other online digital platforms have emerged, cheering sharp and repeated corrections of misinformation to reduce associated wrong perception (Lewandowsky *et al.,* 2017).

During the 2016 U.S. Presidential Election Campaign, it was highlighted by an independent fact-checker agency '*Politifact'* that nearly 70% of all statements posted by the then-presidential candidate Donald Trump to be false, mostly false, similarly opponent to this Hillary Clinton, this rate was much lower at 26 % (Puyosa, 2017). Mostly, the literature on misinformation on social media comes from the medical domain (Chen *et al.*, 2018). The study confirmed that gynaecologic cancer-related tweets on social media medically precise, and 30 % contain misinformation. Additionally, cancer-related misinformation spread needs to be controlled and reduced by the service providers' efforts (Chen *et al.,* 2018).According to Tucker *et al*. (2018),"we need better estimates of the effects of exposure to information and disinformation online as well as more research to see its influence"(Tucker *et al.,* 2018, p. 7). Therefore, this study discusses the effects of misinformation on social media across different domains and derives important tools, techniques, methods and models to mitigate thespread through an in-depth survey of the literature.

3. Methodology of the Study

Literature Search and Evaluation

Inclusion Criterion

Studies related to only tools, technique, methods and models of combating mis/disinformation on social media are included in this study. Therefore, other studies that have considered different aspects, other than the purview of the present study, were excluded (Alalwan*et al*., 2017). Researchers came across n number of studies related to misinformation in various domains, predominantly health crisis and emergency, COVID-19, politics, marketing etc. but focusing on the aim of the present study; all those studies kept excluded from the present study and only studies that discuss techniques, methods and models of combating misinformation were considered for evaluation. These inclusion and exclusion criteria have been adapted from Xiao and Watson (2019) research paper entitled "Guidance on Conducting a Systematic Literature Review".

Literature Identification

Literature was searched by using the keywords "social media", "misinformation", "disinformation". The preliminary relevance of the manuscripts was determined by the title. If the title was found pertinent, it was obtained with full reference, including title, author(s), year, journal name and abstract for further evaluation.

Three frequently and widely used databases, Google Scholar, Scopus and Web of Science, were searched to get the apposite literature using the keywords mentioned earlier. Researchers limit the period of search from June 2020 to November 2020 so that the most recent literature could be found on the current tools, techniques, models and advanced methods to counter mis/disinformation on social media and to mitigate its effect. Combining all the three databases (Google Scholar, Scopus and Web of Science), researchers identified a total of60studies for further assessment.

Screening for Inclusion

Researchers read the abstract of all the60studies to decide their relevance for inclusion in the study. They performed the task of assessing manuscripts parallelly and independently and selected full-text articles of 42 studies for quality assessment.

Quality and Eligibility Assessment

Researchers went through the full-text of the articles to measure the quality of the studies. Articles published in reputed journals with high impact factors written in the English language (Alalwan*et al*., 2017) were deemed high-quality and included in the review. Most of the online presentations and other reports were excluded because of the lack of peer-review process and unscientific citation and referencing. Hence, out of the full-text of 42 studies, 38 studies were carried forward to the next step. The concept-driven systematic review approach, proposed by Webster and Watson (2002), was carried out for the present study. This approach analyses literature from the stance of the concept given by all the authors instead of the author-driven approach (Alalwan*et al*., 2017).Therefore, this method was found appropriate for the present study as the concept of mis/disinformation on social media is one of the most challenging, debatable and emerging phenomena in the modern era. This approach served the purpose of the study in capturing all the related studies carried out during the stipulated time (Alalwan*et al*., 2017).

Iterations

Researchers identified related studies through the backward and forward search. This technique was adapted from Xiao and Watson (2019). Preference was given only to those studies that dealt with tools, techniques, models and methods of countering misinformation on social media. Following the iteration process, four studies were found beyond the present study's objective and were dropped for further inclusion. Overall, 34 studies were finally selected for inclusion in this study.

Data Extraction and Analysis

Researchers extracted information from the selected studies pertinent to mis/disinformation and countermeasures to restrict it. They deduced the concept of misinformation in the context of social media, inferred the measures to counter it and derived some significant suggestions to mitigate its effect.

Related Studies

In the middle of 2020, Ahuja and Kumar created the S-HAN model to identify fake news on social media. It was an improved version of the Hierarchical Attention Networks to identify misinformation by pointing out the news's critical words and sentences. This model works on deep-level text structures and uses the Hierarchical Attention Network to extract essential features from the given text. This model converts the shared documents in a hierarchical structure and works on stacked RNN (Recurrent Neural Network). The researchers found 93.63% accuracy of the model in the real-world dataset.

Jarouche*het al.* (2020) carried out their research to combat fake news by presenting a high-level overview of TRUSTD, blockchain, and collective signature-based ecosystem to help the creator get their content backed by the community and facilitating users in assessing the correctness and credibility of these contents. In this paper, researchers focused on the human element while addressing the issues of fake content. They emphasized that machines can detect fake content to some extent, but there is no substitute for human intervention. They further said that these automated tools to detect fake content could not be relied upon solely. They can be used only as a part of the user policy.

In another research, Dordevic, Pourghomi and Safieddine (2020) identified twenty-seven variables through an extensive review of the literature. These variables control and affirm the dispersion of mis/disinformation. They categorized these variables based on three key players involved in the process-user, content and social networks. The variables related to social media users are *contact rate, susceptible, exposed, infected, sceptic, recovered, total population, edges, vertices, influential node, counter influence node and threshold.* Dordevic*et al.* (2020) reported variables related to social media content. They are"*time-sensitive, reference source, fake news dataset, dynamic timestamp, enquiry phrases and propagation path analysis",* and the variable linked to social media platforms are "*sharing, passing on information, authentication, crosswire, same level (cluster) communication, reverse validation, newsgroup, filter bubble and platform policy".* By identifying and collecting the variables, the researchers generated a greater and holistic view of the environment in which fake news blooms.

Kirchner and Reuter (2020) researched countering fake news when someone knows the inappropriateness of social media claims. They adopted a user-centric approach to counter identified mis/disinformation. Using a three-step design, they analyzed the effectiveness of measures to counter fake news in social media and user perspective about them. They found that warning based approaches lessen the effect of misinformation and that too combined with an explanation. Warning based approaches were found to be effective in bringing down the estimated accuracy of mis/disinformation headlines.

Shao *et al.* (2020) developed a cluster system to detect fake accounts on social media called FADE, based on group behaviour to point out suspicious groups. On the basis of similarity in the flow of sent messages and timing of account creation, they found these accounts to be fake. They further added that "a wider array of cluster-level features can help classify groups into fake or real using supervised learning algorithms. These algorithms are shown to work better when applied to cluster-level features than when applied to individual account features. This is because individual accounts have larger variability along these feature dimensions".

Torusdağ*et al*. (2020) analyzed the vulnerabilities of Botometer. They found that Botometer was not able to determine the presence of social bots on Twitter. Thus, more advance bot detections models are needed to mitigate the effects of social bots in countering mis/disinformation on social media. In a similar way, Shu *et al*. (2020) provided a comprehensive fake news data repository, FakeNewsNet. This repository comprises two comprehensive data sets with diverse characteristics in news content, social context and spatiotemporal information. They suggested a principled strategy to get pertinent data from various resources. Ksieniewicz*et al.* (2020) used machine learning methods to detect fake news. Unlike many other studies, they treated incoming messages as stream data.

Pasi*et al.* (2020) adopted a Multi-Criteria Decision Making (MCDM) approach for assessing the news credibility in the micro blogging sites, particularly on Twitter. They suggested this approach by exploiting the aggregation operator and prior domain knowledge. This approach is not entirely dataset dependent. Malhotra and Vishwakarma (2020) proposed a method to detect rumour and fake news. They prepared their own dataset by extracting user profile features using Twitter API. They said that "Our proposed method leverages textual features from source tweets to get the linguistic cue of news using LSTM networks and RoBERTa based vector". Zhou and Zafarani (2020) have explained news detection strategies based on four criteria: *1. Knowledge-based, 2. Style based, 3. Propagation based, and 4. Source-based.* They suggested detecting fake news from multiple perspectives.

Due to the rapidly growing usage of social media platforms such as Facebook and Twitter, rumours and distorted news are able to propagate widely in a very short span of time (Hunt *et al.*, 2020).On the social media platforms, Facebook and Twitter, information travel across the globe at a very high speed, allowing active users to obtain information related to politics, sports and educations related (Hunt *et al.*, 2020).

Table 1. Latest studies on fake news counter

Author(s)	*Formulation*	*Findings*
Varshney & Vishwakarma (2020)	Model	Researchers developed a model named "Hoax News Inspector" to detect fake news with an accuracy of 95%.
Hunt et al. (2020)	Framework	Developed a machine learning framework that detects the accuracy of tweets that are spread during the crisis.
Wang et al. (2020)	Model	Developed automatic detectors that can instantiate antivaccine messages on Twitter. The model is specially designed for both visual and textual information.
Ahmad et al. (2020)	Approach	Researchers proposed a machine learning ensemble approach to distinguish fake contents from multiple domains.
Pinnaparaju et al. (2020)	Model	Researchers put forward a method to detect the false news spreaders and achieve an accuracy of 71.5% and 70% in both English and Spanish test set respectively.
Aphiwongsophon & Chongstitvatana (2020)	Model	Proposed a machine learning framework to identify misinformation on Twitter through the methods, naïve Bayes, neural network and a support vector machine with an accuracy of 95.55%, 97.09% and 98.15% respectively.
Yesugade et al. (2020)	Model	Researchers suggested methods and models detect deep fake videos by using various algorithms from machine learning, NLP, CNN, RNN and LSTM etc.
Cueva et al. (2020)	Approach	Researchers identify a method to detect the fake news on Twitter using artificial intelligence through study LSTM, NLP and GRU networks.
Ahuja & Kumar (2020)	Model	Researchers created an improved version of Hierarchical Attention Network named S-HAN model to detect fake news with an accuracy of 93.63% in a real-world dataset.
Jaroucheh et al. (2020)	Approach	Researchers presented a high-level overview of TRUSTD, blockchain, and collective signature-based ecosystem to assess the correctness and credibility of contents.
Dordevic et al. (2020)	Variables	Identified twenty-seven variables that control the dispersion of mis/disinformation and differentiate these variableson the basis of three key players involved in the process-user, content and social networks.

Kirchner & Reuter (2020)	Approach	Researchers adopted a user-centric approach to counter identified mis/disinformation using a three-step design and found that warning-based approaches lessen the effect of misinformation.
Shao et al. (2020)	Framework	Researchers developed a cluster system to detect fake accounts on social media called FADE, based on group behaviour to point out suspicious groups.
Shu et al. (2020)	Fake News Data Repository	Provided a comprehensive fake news data repository, FakeNewsNet.

Source: Prepared by the Authors

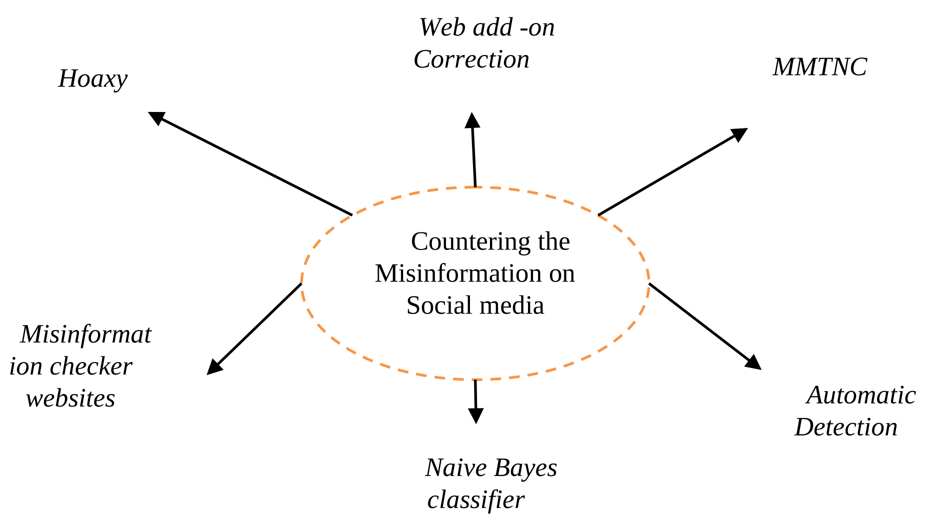
Analyzing the existing literature on countering the spread of misinformation on social media, researchers came to know the paucity of available literature on the given topic. However, there are many studies that have discussed *mis/disinformation on social media* focusing on a particular domain, either politics, marketing, emergency, health or the recent emerged pandemic COVID-19; still, there are insufficient studies related to counter measurement models to restrict the spread of misinformation and a very few that have carried out a systematic literature review process to analyze the previous studies. This study is novel in the sense that it focuses only on tools, techniques, methods, models and frameworks to restrict the spread of misinformation and on the basis of that, proposes a counter measurement model.

4. Insightful measure to counter the misinformation spread on social media

- *Web add-on Correction*

The spread of misinformation on social media platforms such as Facebook and Twitter are long-standing issues and matter significant concerns over the last few decades, and formulating effective strategies remains a defy (Lee, 2020).Web add-on proved to be a useful countermeasure to reduce the belief in misinformation. Internet browsers have developed an automatic detection extension that immediate siren users about the falsification of news. Recently, Google Chrome has introduced an extension that enables active social media users that posts they were reading are real or fake (Darbyshire, 2018). To assess the credibility of fake news on social media and convey its falsity more efficiently and transparently, an academician has attempted to develop a web plug-in based on the white box approach that enables suspicious tweet (Bhuiyan *et al.*, 2018).The researchers further developed a web browser extension *Feed Reflect* (Bhuiyan *et al.*, 2018). Web add-on correction working based on an algorithm (Lee, 2020).

Figure 2. Counter Measurement Model



Source: Designed by Authors

5. Mental model theory and Narrative Correction

This theory provides a psychological understanding of narrative correction. According to this theory, "people tend to retain the information that already built in their mind, irrespective of the alternative available" (Lewandowsky *et al*., 2012). The alternative option provides to active users, alter their pre-tense perceptions. This theory enables the allied scholars to set up effective correction by providing substitute explanation of the misinformation. The researches have demonstrated that alternative option successfully replaces misinformation instead of simple statements pertaining to the falsehood of the disinformation (Goldvarg & Johnson-Laird, 2001; Tenney*et al.*, 2009; Lewandowsky *et al*., 2012).

The narrative-based approached enhances active web-based users' engagement with fact-based information. The narrative format provides tools in reducing health-related risks (Murphy *et al*., 2013). The narrative correction stated as the 'heightened rectification' and further advocated engaging emotional response, narrative norms and causal reasoning (Murphy *et al.*, 2013; Cappella *et al.*, 2015).

- *Hoaxy*

Hoaxy deals with misinformation as "false or inaccurate information" likely that rumours, false news and elaborate the conspiracy theories (Shao *et al.*, 2016). An advance algorithm has been developed to detect deliberate the spread of misinformation on social media. The computationally advance mechanism enables active users to measure the credibility of the news's information and quality (Kumar &Geethakumari, 2014).

Massive amounts of misinformation circulate on social media such as on Facebook and Twitter. The spread of misinformation observed to be a fashion in the forms of fake news, rumours and conspiracy theory. At this point, several journalist organizations tried to address the fact-checking and resultant figures in cascades instances of both credible and incredible. All these pretences challenge studying the social dynamics of news sharing (Shao *et al.*, 2016). To counter such issues, *Hoaxy*, a platform, was developed to collect, detect and analyze the online misinformation on social media and its fact-checking. The *Hoaxy* was developed by the researchers of the Indiana University Network Science Institute (IUNI) and School of Informatics and Computing Centre for Complex Network and System Research (CnetS). This platform enables researchers, academician and common people to identify the factors that influence the triumph and vindication of the misinformation (CnetS, 2016). The main goal of the hoaxy is to track misinformation spread on social media and diagnosis fact-checking. It uses the "POST statues/filter" API endpoints to bring together all public tweets, including links to fact-checking and misinformation articles (Hui *et al*., 2018). Hoaxy deals with the misinformation and false information with examples such as rumours, false news and elaborate conspiracy theory (Shao *et al.*, 2016, p.745). Hoaxy searches for claims and fact-checking. It is working of links sharing basis from low-credibility sources. Hoaxy platform was also recognized with the National Science Foundation award (Hoaxy, Observatory on social media, 2016)

- *Misinformation checker websites*

With the tidal waves of misinformation and tsunami of fake news on social media, I.T. specialist and software developers have developed many fact-checking websites to counter the spread of misinformation such as *BuzzSumo.com, Politifact, FactCheck.org and Snopes.com*etc. BuzzSumo.com is an influential tool to find out the popular content by the relevancy of topic and provides the content what people want to get. In such a way, it gets to know the most viral stories on social media and protects the user from other irrelevant content by being focused on the desired content. Politifact is another fact-checking website which checks political newsworthy and significant statements. It rates these statements "Mostly True" to "Pants on Fire." Similarly, Factcheck.org monitors the factual accuracy present in American politics. The prime focus of this website is to concentrate on the statements made by U.S. Politicians. Snopes.com is a fact-checking website that sorted out myths and rumours on social media by conducting extensive fact-checking research.

- *Naive Bayes classifier*

Naïve Bayes Classifier is a collection of algorithms based on Bayes 'Theorem. It is a family of algorithms which works on a shared principle. It classifies data based on probability. It determines any news as fake by counting the occurrence of a word in the headline. Further, it changes the occurrence to a probability and calculates the probability of the headline to be fake or real.

- *Automatic Detection*

False news detection is an intriguing task for humans as there are bombardments of fake news on social media in every second. So automatic detection models are required to detect fake news with little or no human intervention. Zubiaga*et al.* (2018) started a project coined PHEME to detect rumours automatically, and Zubiaga*et al.* (2016) provided a dataset of tweets consisting of 1972 rumours, 3830 non-rumours and 5 breaking news. That model is named on the Greek Goddess of rumour, report and gossip PHEME. The model checks the veracity of information that is apparently credible but hard to verify and create sufficient doubts and anxiety. PHEME works on the 4-way typology of support, deny, query and comment.

6. Outcomes

The proposed counter measurement model suggests that web add-on is a valuable technique to reduce misinformation. It is an extension that immediate siren users about the falsification of news (Darbyshire, 2018). Thus, it can be one of the effective methods to mitigate the spread of misinformation. The Mental Model Theory and Narrative Correction (MMTNC) is another very effective method to counter misinformation by setting up effective correction and providing substitute explanation of the misinformation. Because the alternative option successfully replaces misinformation (Goldvarg & Johnson-Laird, 2001; Tenney*et al.,* 2009; Lewandowsky *et al*., 2012). Hoaxy is a platform that collect, detect and analyze the online misinformation on social media and its fact-checking. It uses the "POST statues/filter" API endpoints to bring together all public tweets, including links to fact-checking and misinformation articles (Hui *et al*., 2018). It yields productive results and can be used to detect misinformation. There are some other websites that check misinformation such as *BuzzSumo.com, Politifact, and FactCheck.org and Snopes.com* etc. Based on the requirement and news content, any of them can be used accordingly (Allcott*et al.,* 2019).Naive Bayes classifier and automatic detection models are also efficient methods in countering misinformation (Zubiaga*et al.,* 2018).

7. Discussion and Conclusion

Literature reviews build a foundation for academic enquiries (Xiao & Watson, 2019) and educate the researchers to fill the gap left behind in prior studies. After going through the literature survey, researchers found that automated tools to detect mis/disinformation are significant to some extent, but they cannot be relied upon solely. There is no substitute for human intervention. A blend of the two, automated tools and techniques to detect fake news and human intervention, is needed to mitigate misinformation more efficiently and effectively. Jindal & Anand (2020) also proposed a combination of non-cognitive text analytics with cognitive machine learning to explore, analyze and filter the information effectively. As of now, the majority of the researchers have worked upon two approaches of fake news detection: *1. Fact and news checker, 2. Artificial intelligence algorithms for news analysis and manipulation detection.* There is a dire need to develop more models to counter the spread of fake news as misinformation on social media is blooming like the wild thorns and need to be trimmed promptly. It was found that most of the approaches to detecting fake news are based on machine learning. More Multi-Criteria Decision-Making approaches are required to assess the credibility of information on social media. Misinformation has different types with varied motives, so one method or model cannot detect all kinds of misinformation. Detecting fake news is not effective if only one perspective is taken into consideration. Multiple perspectives need to be addressed while detecting misinformation. Warning based approaches can also be used to lessen the effect of misinformation, and that too combined with an explanation. They are effective in bringing down the estimated accuracy of mis/disinformation headlines. This discourse unwraps the fact that detecting misinformation will remain a mighty challenge and a complex issue in the recent future. Thus, developing more sophisticated methods and models is the need of the hour. The potential harm caused by misinformation can never be underestimated. Thus, it requires serious deliberations and graved concerns of researchers, academicians, and government agencies as it is considered a threat to national security and the democratic system (Levush, 2019).

References

[1] Ahmad, I., Yousaf, M., Yousaf, S., & Ahmad, M. O. (2020). Fake News Detection Using Machine Learning Ensemble Methods. Complexity, 2020, 1-11.

[2] Ahuja, N., & Kumar, S. (2020, June). S-HAN: Hierarchical Attention Networks with Stacked Gated Recurrent Ujnit for Fake News Detection. In 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) (pp. 873-877). IEEE.

[3] Alalwan, A. A., Rana, N. P., Dwivedi, Y. K., &Algharabat, R. (2017). Social media in marketing: A review and analysis of the existing literature. Telematics and Informatics, 34(7), 1177-1190.

[4] Alina, C., Cuzzocrea, A., &Truta, T. M. (2017). Fighting fake news spread in online social networks: Actual trends and future research directions. 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, 2017 (pp. 4453–4457). https://doi.org/10.1109/BigData.2017.8258484

[5] Andı, S., Aytaç, S. E., &Çarkoğlu, A. (2020). Internet and social media use and political knowledge: Evidence from Turkey. Mediterranean Politics, 25(5), 579-599.

[6] Aphiwongsophon, S., &Chongstitvatana, P. (2020). Identifying misinformation on Twitter with a support vector machine. Engineering and Applied Science Research, 47(3), 306-312.

[7] Badawy, A., Ferrara, E., &Lerman, K. (2018). Analysing the digital traces of political manipulation: The 2016 Russian interference Twitter campaign. 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). https://doi.org/10.1109/asonam.2018.8508646

[8] Bagheri, E., Liu, H., Shu, K., &Zarrinkalam, F. (2020). The 5th international workshop on mining actionable insights from social networks (Maison 2020). Proceedings of the 29th ACM International Conference on Information & Knowledge Management. https://doi.org/10.1145/3340531.3414078

[9] Barsi, J. A., Schott, J. R., Palluconi, F. D., & Hook, S. J. (2005). Validation of a web-based atmospheric correction tool for single thermal band instruments. Earth Observing Systems X. https://doi.org/10.1117/12.619990

[10] Bhuiyan, M. M., Zhang, K., Vick, K., Horning, M. A., & Mitra, T. (2018). FeedReflect. Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing. https://doi.org/10.1145/3272973.3274056

[11] Butler, K. (2014). The Boston marathon bombing: After-action review of the Brigham and women's hospital emergency radiology response. The Journal of Emergency Medicine, 47(6), 753-754. https://doi.org/10.1016/j.jemermed.2014.10.017

[12] Cappella, J. N., Maloney, E., Ophir, Y., & Brennan, E. (2015). Interventions to correct misinformation about tobacco products. Tobacco Regulatory Science, 1(2), 186-197. https://doi.org/10.18001/trs.1.2.8

[13] Choi, S. (2013). An empirical study of social network service (SNS) continuance: incorporating the customer value-satisfaction-loyalty model into the I.S. continuance model. Asia Pac. J. Inf. Syst. 23 (4), 1–28

[14] Chung, N., Han, H., Koo, C. (2012). Mediating roles of attachment for information sharing in social media: social capital theory perspective. Asia Pac. J. Inf. Syst. 22 (4), 101–123

[15] CnetS. (2016, December 21). Hoaxy: A platform for tracking online misinformation. CNetS. https://cnets.indiana.edu/blog/2016/12/21/hoaxy/

[16] Collins, B., Hoang, D. T., Nguyen, N. T., & Hwang, D. (2020). Trends in combating fake news on social media–a survey. Journal of Information and Telecommunication, 1-20.

[17] Cueva, E., Ee, G., Iyer, A., Pereira, A., Roseman, A., & Martinez, D. (2020). Detecting Fake News on Twitter Using Machine Learning Models.

[18] Darbyshire, P. (2018). Fake news. Fake journals. Fake conferences. What we can do. Journal of Clinical Nursing, 27(9-10), 1727-1729. https://doi.org/10.1111/jocn.14214

[19] Dordevic, M., Pourghomi, P., &Safieddine, F. (2020, October). Identifying Fake News from the Variables that Governs the Spread of Fake News. In 2020 15th International Workshop on Semantic and Social Media Adaptation and Personalization (SMA (pp. 1-6). IEEE.

[20] Dordevic, M., Pourghomi, P., &Safieddine, F. (2020, October). Identifying Fake News from the Variables that Governs the Spread of Fake News. In 2020 15th International Workshop on Semantic and Social Media Adaptation and Personalisation (SMA (pp. 1-6). IEEE.

[21] Ebrahim Bagheri,HuanLiu,KaiShu,FattaneZarrinkalam. (2020, August 27). Information processing & management. https://www.journals.elsevier.com/information-processing-and-management/call-for-papers/special-issue-on-dismisinformation-mining-from-social-media

[22] Eysenbach, G., Powell, J., Kuss, O., & Sa, E. (2002). Empirical studies assessing the quality of health information for consumers on the World Wide Web. JAMA, 287(20), 2691. https://doi.org/10.1001/jama.287.20.2691

[23] Goldvarg, E., & Johnson-Laird, P. (2001). Naive causality: A mental model theory of causal meaning and reasoning. Cognitive Science, 25(4), 565-610. https://doi.org/10.1207/s15516709cog2504_3

[24] Gupta, A., Lamba, H., Kumaraguru, P., & Joshi, A. (2013). Faking sandy. Proceedings of the 22nd International Conference on World Wide Web - WWW '13 Companion. https://doi.org/10.1145/2487788.2488033

[25] Hoaxy (Observatory on social media). (2016). RAND Corporation Provides Objective Research Services and Public Policy Analysis | RAND. https://www.rand.org/research/projects/truth-decay/fighting-disinformation/search/items/hoaxy-observatory-on-social-media.html

[26] Hoaxy: A platform for tracking online misinformation. (2017, January 18). CNetS. https://cnets.indiana.edu/blog/2016/12/21/hoaxy/

[27] Hunt, K., Agarwal, P., & Zhuang, J. (2020). Monitoring misinformation on Twitter during crisis events: a machine learning approach. Risk analysis.

[28] Hunt, K., Agarwal, P., & Zhuang, J. (2020). Monitoring misinformation on Twitter during crisis events: a machine learning approach. Risk analysis.

[29] Jaroucheh, Z., Alissa, M., Buchanan, W. J., & Liu, X. (2020, July). TRUSTD: Combat Fake Content using Blockchain and Collective Signature Technologies. In 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC) (pp. 1235-1240). IEEE.

[30] Jindal, S., & Anand, G. (2020). Fighting an Infodemic in the Time of the COVID-19 Pandemic in India: Leveraging Technology and Social Media. Journal of Preventive Medicine and Public Health, 53(5), 311.

[31] Kirchner, J., & Reuter, C. (2020). Countering Fake News: A Comparison of Possible Solutions Regarding User Acceptance and Effectiveness. Proceedings of the ACM on Human-Computer Interaction, 4(CSCW2), 1-27.

[32] Koo, C., Wati, Y., Jung, J.J. (2011). Examination of how social aspects moderate the relationship between task characteristics and usage of social communication technologies (SCTS) in organizations. Int. J. Inf. Manage. 31, 445–459

[33] Ksieniewicz, P., Zyblewski, P., Choraś, M., Kozik, R., Giełczyk, A., &Woźniak, M. (2020, July). Fake News Detection from Data Streams. In 2020 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE.

[34] Kumar, K.P.K. and Geethakumari, G. (2014), "Detecting misinformation in online social networks using cognitive psychology", Human-centric Computing and Information Sciences, Vol. 4 No. 1, 22pp.

[35] Lange-Faria, W., & Elliot, S. (2012). Understanding the role of social media in destination marketing. Tourismos, 7(1).

[36] Levush, R. (2019, September). Government Responses to Disinformation on Social Media Platforms: Comparative Summary. Library of Congress. https://www.loc.gov/law/help/social-media-disinformation/compsum.php

[37] Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction. Psychological Science in the Public Interest, 13(3), 106-131. https://doi.org/10.1177/1529100612451018

[38] Malhotra, B., & Vishwakarma, D. K. (2020, September). Classification of Propagation Path and Tweets for Rumor Detection using Graphical Convolutional Networks and Transformer based Encodings. In 2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM) (pp. 183-190). IEEE.

[39] Maryam, A., & Ali, R. (2019). Misinformation source identification in an online social network. 2019 IEEE 5th International Conference for Convergence in Technology (I2CT). https://doi.org/10.1109/i2ct45611.2019.9033558

[40] Mendoza, M., Poblete, B., & Castillo, C. (2010). Twitter under crisis. Proceedings of the First Workshop on Social Media Analytics - SOMA '10. https://doi.org/10.1145/1964858.1964869

[41] Mian, A., & Khan, S. (2020). Coronavirus: The spread of misinformation. BMC Medicine, 18(1). https://doi.org/10.1186/s12916-020-01556-3

[42] Mocanu, D., Rossi, L., Zhang, Q., Karsai, M., &Quattrociocchi, W. (2015). undefined. Computers in Human Behavior, 51, 1198-1204. https://doi.org/10.1016/j.chb.2015.01.024

[43] Murphy, S. T., Frank, L. B., Chatterjee, J. S., &Baezconde-Garbanati, L. (2013). Narrative versus Nonnarrative: The role of identification, transportation, and emotion in reducing health disparities. Journal of Communication, 63(1), 116-137. https://doi.org/10.1111/jcom.12007

[44] Pasi, G., De Grandis, M., & Viviani, M. (2020, July). Decision making over multiple criteria to assess news credibility in microblogging sites. In 2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) (pp. 1-8). IEEE.

[45] Pinnaparaju, N., Indurthi, V., & Varma, V. (2020). Identifying Fake News Spreaders in Social Media. In CLEF.

[46] Ruth Levush. (2019, September). Government responses to disinformation on social media platforms. Library of Congress. https://www.loc.gov/law/help/social-media-disinformation/compsum.php

[47] Shao, H., Yao, S., Jing, A., Liu, S., Liu, D., Wang, T., ... &Abdelzaher, T. (2020, August). Misinformation Detection and Adversarial Attack Cost Analysis in Directional Social Networks. In 2020 29th International Conference on Computer Communications and Networks (ICCCN) (pp. 1-11). IEEE.

[48] Shao, C., Ciampaglia, G. L., Flammini, A., &Menczer, F. (2016). Hoaxy. Proceedings of the 25th International Conference Companion on World Wide Web - WWW '16 Companion. https://doi.org/10.1145/2872518.2890098

[49] Shao, C., Hui, P., Wang, L., Jiang, X., Flammini, A., Menczer, F., &Ciampaglia, G. L. (2018). Anatomy of an online misinformation network. PLOS ONE, 13(4), e0196087. https://doi.org/10.1371/journal.pone.0196087

[50] Shu, K., Bhattacharjee, A., Alatawi, F., Nazer, T. H., Ding, K., Karami, M., & Liu, H. (2020). Combating disinformation in a social media age. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(6), e1385.

[51] Shu, K., Mahudeswaran, D., Wang, S., Lee, D., & Liu, H. (2020). FakeNewsNet: A Data Repository with News Content, Social Context, and Spatiotemporal Information for Studying Fake News on Social Media. Big Data, 8(3), 171-188.

[52] Statista. (2020). Most used social media platform. https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

[53] Su, Y. (2021). It doesn’t take a village to fall for misinformation: Social media use, discussion heterogeneity preference, worry of the virus, faith in scientists, and COVID-19-related misinformation beliefs. Telematics and Informatics, 58, 101547. https://doi.org/10.1016/j.tele.2020.101547

[54] Tenney, E. R., Cleary, H. M., & Spellman, B. A. (2009). Unpacking the doubt in “Beyond a reasonable doubt”: Plausible alternative stories increase not guilty verdicts. Basic and Applied Social Psychology, 31(1), 1-8. https://doi.org/10.1080/01973530802659687

[55] Torusdağ, M. B., Kutlu, M., &Selçuk, A. A. (2020, September). Are We Secure from Bots? Investigating Vulnerabilities of Botometer. In 2020 5th International Conference on Computer Science and Engineering (UBMK) (pp. 343-348). IEEE.

[56] Tsourvakas, G., &Yfantidou, I. (2018). Corporate social responsibility influences employee engagement. Social Responsibility Journal, 14(1), 123-137. https://doi.org/10.1108/srj-09-2016-0153

[57] Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., . . . Nyhan, B. (2018). Social media, political polarization, and political disinformation: A review of the scientific literature. Available at SSRN: https://ssrn.com/abstract=3144139

[58] Varshney, D., & Vishwakarma, D. K. (2020). Hoax news-inspector: a real-time prediction of fake news using content resemblance over web search results for authenticating the credibility of news articles. Journal of Ambient Intelligence and Humanized Computing, 1-14.

[59] Wang, Z., Yin, Z., & Argyris, Y. Detecting Medical Misinformation on Social Media Using Multimodal Deep Learning. IEEE journal of biomedical and health informatics.

[60] Waszak, P.M., Kasprzycka-Waszak, W., Kubanek, A., 2018. The spread of medical fake news in social media – The pilot quantitative study. Health Policy and Technology 7 (2), 115–118. Apuke, O. D., & Omar, B. (2021). Fake news and COVID-19: Modelling the predictors of fake news sharing among social media users. Telematics and Informatics, 56, 101475. https://doi.org/10.1016/j.tele.2020.101475

[61] Webster, Jane, and Richard T. Watson. (2002). Analyzing the past to prepare for the future: Writing a literature review. MIS Quarterly, 26(2), xiii-xxiii.

[62] Wood, M. J. (2018). Propagating and debunking conspiracy theories on Twitter during the 2015–2016 Zika virus outbreak. Cyberpsychology, Behavior, and Social Networking, 21(8), 485-490. https://doi.org/10.1089/cyber.2017.0669

[63] Wu, L., Morstatter, F., Carley, K. M., & Liu, H. (2019). Misinformation in social media: definition, manipulation, and detection. ACM SIGKDD Explorations Newsletter, 21(2), 80-90.

[64] Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. Journal of Planning Education and Research, 39(1), 93-112.

[65] Yesugade, T., Kokate, S., Patil, S., Varma, R., & Pawar, S. (2020). FAKE NEWS AND DEEP FAKE DETECTION. Journal of Critical Reviews, 7(19), 3876-3884.

[66] Zhang, X., &Ghorbani, A. A. (2020). An overview of online fake news: Characterisation, detection, and discussion. Information Processing & Management, 57(2), 102025.

[67] Zhou, X., &Zafarani, R. (2020). A survey of fake news: Fundamental theories, detection methods, and opportunities. ACM Computing Surveys (CSUR), 53(5), 1-40.

[68] 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), 1-36.

[69] Zubiaga, A., Liakata, M., & Procter, R. (2016). Learning reporting dynamics during breaking news for rumour detection in social media. arXiv preprint arXiv:1610.07363.