Network-based Fake News Detection: A Pattern-driven Approach

Xinyi Zhou and Reza Zafarani
Data Lab, EECS Department, Syracuse University
{zhouxinyi,reza}@data.syr.edu

ABSTRACT

Fake news gains has gained significant momentum, strongly motivating the need for fake news research. Many fake news detection approaches have thus been proposed, where most of them heavily rely on news content. However, network-based clues revealed when analyzing news propagation on social networks is an information that has hardly been comprehensively explored or used for fake news detection. We bridge this gap by proposing a network-based pattern-driven fake news detection approach. We aim to study the patterns of fake news in social networks, which refer to the news being spread, spreaders of the news and relationships among the spreaders. Empirical evidence and interpretations on the existence of such patterns are provided based on social psychological theories. These patterns are then represented at various network levels (i.e., node-level, ego-level, triad-level, community-level and the overall network) for being further utilized to detect fake news. The proposed approach enhances the explainability in fake news feature engineering. Experiments conducted on real-world data demonstrate that the proposed approach can outperform the state of the arts.

1. INTRODUCTION

With "post-truth" named as the Word of the Year in 2016 by the Oxford Dictionary, discussion around fake news has sparked, especially in the period around the 2016 U.S. presidential election and the U.K. Brexit referendum [44]. The rise of social media and its popularity play an indispensable role in this surge of interest. Social media breaks the physical distance barrier among individuals, and provides rich platforms for users to participate and discuss online news, where the most popular story during the critical months of the 2016 U.S. presidential election campaign ("*Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement*", which was fake news) can generate 960,000 shares, reactions, and comments on Facebook [34].

The situation becomes worse with the existence of an echo chamber effect on social media, where the biased information can be amplified and reinforced [10]. Meanwhile, studies have shown that humans can be irrational and vulnerable differentiating between truth and falsehood when overloaded with deceptive information; studies in social psychology and communications have demonstrated that human ability to detect deception is only slightly better than chance - with a mean accuracy of 54% over 1,000 participants in over 100 experiments [27]. Various manual fact-checking websites and platforms (e.g., PolitiFact[1]and Snopes[2]) have emerged to serve the public on this matter. Nevertheless, manual fact-checking does not scale well with the volume of newly created information, especially on social media, hence motivating the need for automatic fake news detection.

Current research on automatic fake news detection heavily relies on news content [22]. These studies have significantly contributed to fake news detection (see "Related Work" in Section 2) while often face multiple challenges.

First, the traditional approach to detect fake news is to use a *knowledge-based* fact-checking system [5; 32]. The system compares relational knowledge extracted from to-be-verified news content with that stored in a knowledge graph, often a ground truth dataset collected from the Web [6; 25]. However, the most serious issue by using such system is that it can only detect false news instead of fake news (i.e., intentionally false news) [41]. Second, another common approach is to use a *style-based* fake news detection system by assuming that fake news exhibits a distinguishable writing style from that of the truth [43], where malicious entities can disguise the writing style to bypass these linguistic models. Recently, neural networks and deep learning techniques have been well developed to detect fake news by incorporating multi-modal or social-network data, e.g., images within news content [37; 40] and users (news spreaders) [15; 29; 42]; nevertheless, these models often face the problems with computational efficiency or interpretability [44].

Present Work: Considering that social-network data related to news propagation and spreaders has hardly been comprehensively explored (across network levels) and used in an explainable way for fake news detection, we propose a network-based pattern-driven fake news detection model, robust against manipulations by malicious entities on news content. To that end, our work aims to utilize patterns in fake news dissemination on social networks, which reveal that compared to the truth, fake news can (i) spread farther and (ii) attract more spreaders, where these spreaders are often (iii) more strongly engaged with the news and (iv) more densely connected within the network. Machine learning features representing these patterns are designed at different levels of a network (i.e., node-, ego-, triad-, community-, and network-level), which will be further used within a supervised learning framework to detect fake news. Overall, the specific contributions of this paper are as follows:

[1]https://www.politifact.com/
[2]https://www.snopes.com/

1. A network-based pattern-driven approach is proposed, which can detect fake news in an explainable way. Experiments conducted on real-world data demonstrate that the proposed approach can perform comparatively well compared to the state of the art.
2. Fake news patterns in social networks are investigated and summarized, which relate to the news being spread, spreaders of the news, and relationships among the news spreaders. Empirical studies and social psychological theories are provided to validate and interpret the existence of these patterns;
3. Fake news patterns are represented and quantified across multiple network levels, i.e., node, ego, triad, community, and the overall network level. Experimental results indicate that the proposed approach can perform stably with limited available network information, which makes it suitable for fake news early detection.

The rest of this paper is organized as follows. Section 2 reviews current fake news detection research. Fake news patterns in social networks are summarized and represented in Section 3. Experiments are conducted and presented in Section 4. Section 5 concludes the paper.

2. RELATED WORK

As an emerging topic, the development of fake news detection is in its early stages, where the existing research can be generally grouped into content-based and network-based fake news detection.

Content-based Fake News Detection. Content-based fake news detection investigates news content. One traditional way of detection is based on *knowledge*, often represented as a set of (Subject, Predicate, Object) triples [6; 21]. Knowledge-based approaches aim to assess news authenticity by comparing the knowledge extracted from to-be-verified news content with true knowledge (i.e., ground truth) [5; 32]. Such ground truth is generally provided in a knowledge graph such as Knowledge Vault [6], which contains massive manually processed relational knowledge from the open Web. However, the timeliness and completeness of knowledge graphs are still open issues, and importantly, such approaches developed can only detect false news rather than fake news (intentionally false news) [44].

Another common way is based on writing *style*, a set of self-defined [non-latent] features well representing news writing style. Style features can be those capturing content structure at various language levels such as discourse level by employing rhetorical structure theory [28; 12]; or those capturing specific attributes in the content such as sentiment and readability [24; 23; 43], which can be supported by forensic psychological theories such as Undeutsch hypothesis [35]. Such fundamental theories are a double-edged sword for content-based fake news detection: features inspired can help achieve explainable fake news detection, while some linguistic cues that they reveal might not be applicable for news articles (e.g, non-immediacy) [44].

In addition to non-latent features, fake news detection based on latent representation of news content has been well developed recently, where neural networks such as Convolutional Neural Network (CNN) [37] have been utilized to automatically select content features. Nevertheless, these features are often difficult to be comprehended.

While content-based approaches can detect fake news by analyzing news content from various perspectives, auxiliary information revealed in news propagation, e.g., news spreaders, is not considered. In addition, approaches can be sensitive to news content when heavily relying on it, where malicious entities might manipulate the results of detection by disguising their writing styles. Hence, network-based fake news detection has been emerged recently.

Network-based Fake News Detection. Network-based fake news detection utilizes social context information revealed in news propagation. In general, it investigates two types of networks: *homogeneous* and *heterogeneous* networks.

Homogeneous networks contain single type of nodes and edges. A typical example is the stance network, which represents the stance (e.g., *for* or *against*) similarity among news or posts of news. Based on such network, Jin et al. evaluate news credibility by mining the stance correlations within a graph optimization framework [11]. Another typical example of homogeneous networks is the propagation graph (tree), which presents post-repost relationships for each news article on social media, e.g., tweets and retweets on Twitter [38; 17]. Using propagation trees, for instance, Vosoughi et al. discover that fake news spreads faster, farther and more broadly than the truth [36].

Heterogeneous networks have multiple types of nodes or edges. By exploring relationships among entities such as news articles, publishers, users (spreaders) and user posts, PageRank-like algorithm [7], matrix/tensor factorization [8; 33], and Recurrent Neural Networks (RNN) [29; 42] have been developed for fake news detection.

In general, our work is a complement of the current [network-based] studies. Compared to current studies, our work investigates a homogeneous network, the friendship network. To our best knowledge, studying fake news with respect to the friendship network is yet to be explored, which allows one to better understand news spreaders and their social relationships on various network levels. Additionally, we aim to detect fake news in an explainable way - by utilizing its propagation characteristics on social networks, which will be detailed in the next section.

3. FAKE NEWS PATTERNS AND REPRESENTATION IN NETWORKS

Fake news dissemination in networks exhibits distinguishable patterns from the diffusion of true news. In this section, we summarize these patterns and discuss social psychological theories that can explain the existence of these patterns. In terms of fake news patterns, we demonstrate ways to represent news articles as a set of features across network levels (i.e., node-, triad-, community- and network-level), which can be further utilized to detect fake news within a supervised machine learning framework.

Broadly speaking, fake news patterns involved in this study relate to (1) the news being spread (Section 3.1 and Section 3.2), (2) spreaders of the news (Section 3.3), and (3) relationships among the news spreaders (Section 3.4). Before further elaboration, we first define Fake News Network (FNN) in Definition 1.

Definition 1 (Fake News Network, FNN). *Fake News Network (FNN) is a subgraph* $G_{\mathcal{F}} = (V_{\mathcal{F}}, E_{\mathcal{F}})$ *of the social network* $G = (V, E)$*, where* $V_{\mathcal{F}} \in V$ *are the users that*

Table 1: Key Notations

Notation	Description
$\mathcal{F}; \mathcal{T}$	Fake news events; True news events
$\mathrm{G} = (\mathrm{V}, \mathrm{E})$	Social (friendship) network
$\mathrm{G}_X = (\mathrm{V}_X, \mathrm{E}_X)$	$X = \mathcal{F}$: Fake news network; $X = \mathcal{T}$: True news network
E_{NS}	Relationships from a normal user to a susceptible user
$\mathrm{E}_{\triangle>0}; \mathrm{E}_{\triangle=0}; \mathrm{E}_{\triangle<0}$	Relationships satisfying $\mathbf{S}(v_i) - \mathbf{S}(v_j) > 0; \mathbf{S}(v_i) - \mathbf{S}(v_j) = 0; \mathbf{S}(v_i) - \mathbf{S}(v_j) < 0$
V_X; Tr_X; M_X	Nodes (Spreaders); Triads; Communities within G_X
$\mathbf{B}(*)$	$\mathbf{B} = 1$ if $*$ is true; otherwise, $\mathbf{B} = 0$
$\mathbf{C}(v)$	Influence (centrality) of user v
$\mathbf{S}(v)$	Susceptibility of user v
θ	Threshold of user susceptibility, $\mathbf{S}(v) < \theta$ ($\mathbf{S}(v) > \theta$) indicates a normal (susceptible) user
$\mathbf{T}(v, X)$	Spreading frequency of user v for news event X

have engaged with fake news $\mathcal{F}$, and $\mathrm{E}_{\mathcal{F}} \in \mathrm{E}$ represents the relationships among these users.

True News Network (TNN) is similarly defined, which is denoted as $\mathrm{G}_{\mathcal{T}} = (\mathrm{V}_{\mathcal{T}}, \mathrm{E}_{\mathcal{T}})$ for a true news event $\mathcal{T}$. The key notations in this section are presented in Table 1.

3.1 More-Spreader Pattern

Evidence has been provided that fake news is in general more "popular" than true news within the same population of users. For instance, during the critical months of the 2016 U.S. presidential election campaign, top twenty frequently-discussed fake election stories generated 8,711,000 shares, reactions, and comments on Facebook, ironically, greater than the total of 7,367,000 for the top twenty most-discussed election stories posted by nineteen major news medium [34]. Fake news popularity can be attributed to two reasons. First, as stated by information gap theory [16], rather than telling the truth, fake news creators make great efforts to produce an information gap between the news content and individuals' knowledge. Such information gap produces the feeling of deprivation labeled curiosity, which motivates individuals to obtain the missing information to reduce such feeling. Secondly, to greatly influence online users, those who can benefit from fake news often create or recruit malicious accounts (e.g., bots [30]) to spread or discuss the fake content. For example, millions of malicious accounts have participated in 2016 U.S. presidential election online discussions.[3]

News popularity can be characterized in terms of the number of users that spread such news, where Vosoughi et al. [36] have empirically validated that:

Pattern 1 (More-Spreader Pattern). *More users spread fake news than true news.*

To capture the number of news spreaders, we investigate the number and proportion of (**I**) general (i.e., *non-attributed*) spreaders and (**II**) specific (i.e., *attributed*) spreaders in news propagation.

I. General (Non-Attributed) Spreaders. In general, the More-Spreaders Pattern can be quantified by the number of users involved in spreading each fake or true news story. This number is basically the number of nodes within each FNN and TNN, which we use as a feature.

[3] https://comprop.oii.ox.ac.uk/research/public-scholarship/resource-for-understanding-political-bots/

II. Specific (Attributed) Spreaders. Principles like homophily [18] and social validation theory [4] suggest that in a social network, users with similar characteristics tend to become connected or form groups and exhibit similar behavior. These observations imply that spreaders of fake (true) news stories may also share some similar attributes; hence, allowing one to distinguish fake news from true news by studying specific users (i.e., with specific attributes) participated in news dissemination. Here we consider (a) user susceptibility [to fake news] and (b) user influence, both of which are attributes that can be computed with information provided by FNNs and TNNs.

a. User Susceptibility. We investigate user susceptibility to fake news based on (i) the number of involvements in the propagation of different fake news and (ii) the frequency of such involvements.

i. Number of Involvements. Susceptibility in terms of involvements is defined as the proportion of fake news among all news that user v_i has participated in spreading, which is denoted as $\mathbf{S}(v_i)$:

$$\mathbf{S}(v_i) = \frac{\sum_j \mathbf{B}(v_i \in \mathrm{V}_{\mathcal{F}_j})}{\sum_k \mathbf{B}(v_i \in \mathrm{V}_{\mathcal{T}_k}) + \sum_j \mathbf{B}(v_i \in \mathrm{V}_{\mathcal{F}_j})}, \quad (1)$$

where $\mathbf{B}(v_i \in \mathrm{V}_X) = 1$ if $v_i \in \mathrm{V}_X$, otherwise $\mathbf{B}(v_i \in \mathrm{V}_X) = 0$. $\mathbf{S}(v_i) = 1$ ($\mathbf{S}(v_i) = 0$) indicates that all news stories spread through v_i are fake (true), i.e., v_i is completely susceptible (immune) to fake news.

ii. Frequency of Involvements. Consider the special case where a user spreads a true news story once and a fake news story multiple times, this user may need to be considered more susceptible than a user who posts each story once. Hence, as an alternative way, we define user susceptibility as the ratio between the spreading frequency of fake news stories and that of all news stories a user has spread. Mathematically,

$$\mathbf{S}(v_i) = \frac{\sum_j \mathbf{B}(v_i \in \mathrm{V}_{\mathcal{F}_j})\mathbf{T}(v_i, \mathcal{F}_j)}{\sum_k \mathbf{B}(v_i \in \mathrm{V}_{\mathcal{T}_k})\mathbf{T}(v_i, \mathcal{T}_k) + \sum_j \mathbf{B}(v_i \in \mathrm{V}_{\mathcal{F}_j})\mathbf{T}(v_i, \mathcal{F}_j)}, \quad (2)$$

where $\mathbf{T}(v_i, X)$ is the number of times that v_i has spread news story X.

Being assigned with a susceptibility score $\mathbf{S}(v_i)$, users can be further labeled as susceptible ($\mathbf{S}(v_i) > \theta$) or normal ($\mathbf{S}(v_i) < \theta$) based on a fixed threshold value $\theta \in [0, 1]$. Such labeling allows us to represent More-Spreaders Pattern

by recording the (i) number and (ii) proportion of susceptible spreaders (nodes) in each FNN or TNN, as well as the (iii) number and (iv) proportion of normal spreaders within each FNN and TNN. We include (i-iv) as features representing the pattern. Without such labeling one can represent spreaders involved in each FNN or TNN in terms of their mean and median of susceptibility scores, which are also considered into our feature set.

b. User Influence. An approximation of a node (user) influence can be computed via a centrality score within the network. One can consider the following well-established criteria for computing centrality: (i) [in-, out-] degrees, (ii) [in-, out-] closeness, (iii) betweenness, (iv) PageRank score, (v) hub and authority score, all of which use the connections among nodes to identify their positions within the network. We avoid grouping users into influential and non-influential users as many parameters will be introduced (each centrality measure requires a threshold value), which in turn can affect the performance of fake news detection. Therefore, based on each centrality measure, we directly calculate the mean and median user influence within each FNN and TNN, and include both as features.

3.2 Farther-Distance Pattern

In addition to the number of users that spread news articles, news popularity can be also characterized by how far the news can spread, which leads to the corresponding pattern:

PATTERN 2 (FARTHER-DISTANCE PATTERN). *Fake news spreads farther than true news.*

This pattern has been observed and validated by Vosoughi et al. [36]; they found that the propagation trees of fake news are generally deeper than that of truth, i.e., an original post referring to a fake news event is often more iteratively forwarded than a true news event. On the other hand, given a news story, how far it spreads can be approximated by computing the shortest "distance" between the two most distant spreaders (nodes) within the corresponding FNN or TNN (i.e., network diameter). To represent FARTHER-DISTANCE PATTERN and calculate such "distance", we investigate (**I**) shortest (geodesic) distance which refers to the paths existing between two nodes, and (**II**) effective distance which considers the information flow between two nodes [2].

I. Geodesic Distance. Based on geodesic distance, the diameter of each FNN and TNN is equivalent to the shortest path length between the two most distant spreaders within the network.

II. Effective Distance. Besides conventional shortest distance, we introduce effective distance to help assess the network diameter, which was initially proposed by Brockmann and Helbing [2]. The initial binary (unweighted) FNNs and TNNs is hence transformed into weighted networks, where the weights are determined by the volume of information flow among nodes. Given a network, the effective distance among nodes is defined as follows.

DEFINITION 2 (EFFECTIVE DISTANCE). *Given a network* G*, we assume* F *denotes the flow matrix whose entities* F_{ij} *represent the volume of information flow from node* i *to node* j*. Based on the flow matrix, the effective distance* $d_{\mathsf{Eff}}(i,j)$ *from node* i *to node* j *is defined as*

$$d_{\mathsf{Eff}}(i,j) = 1 - \log \frac{\mathrm{F}_{ij}}{\sum_l \mathrm{F}_{lj}}, \tag{3}$$

where $d_{\mathsf{Eff}}(i,j)$ *satisfies* $d_{\mathsf{Eff}(i,j)} \geq 1$.

Information flow has been defined differently in various networks. For instance, it can be the passenger flux in global mobility networks or the transport flow in transportation networks [2]. In FNNs and TNNs it is the news flow among nodes (users) in the network which can be defined as (i) the total number of news stories both users have spread, i.e., $\mathrm{F}_{ij} = \sum_X \mathbf{B}(e_{ij} \in \mathrm{E}_X)$, or (ii) the overall number of times both users have at least spread the same news stories, i.e., $\mathrm{F}_{ij} = \sum_X \mathbf{B}(e_{ij} \in \mathrm{E}_X) \times \min\{\mathbf{T}(u_i, X), \mathbf{T}(u_j, X)\}$. The diameter of each FNN and TNN based on effective distance is then equivalent to the minimum [sum of] effective distance between the two most distant spreaders within the network. We include diameters computed using geodesic and effective distances as features representing FARTHER-DISTANCE PATTERN.

3.3 Stronger-Engagement Pattern

The statistics in [34] have revealed that fake news stories can engage users more compared to true news stories. Note that a user may decide to engage with a fake news story (e.g., post it) more than one time, such "more engagements" can be attributed to the number of users engaging with fake news, which has been summarized as MORE-SPREADER PATTERN investigated in Section 3.1, and/or to the number of times each user engages with a fake news story, leading to the following pattern:

PATTERN 3 (STRONGER-ENGAGEMENT PATTERN). *Spreaders engage more strongly with fake news than with true news.*

To quantify the "engagements" of users for each news story, one can concentrate on (**I**) group level engagements, i.e., the engagements of all spreaders, and (**II**) individual level engagements, i.e., the engagements of a single spreader.

I. Group Engagements. On a group level, quantifying spreader engagements for a certain news story can be equivalent to counting the total number of times that the news story has been spread. With specific user attributes (susceptible or normal), such engagements can be further quantified as the (i) number or (ii) proportion of times that the news story has been spread by susceptible users, as well as (iii) number or (iv) proportion of times that the news story has been spread by normal users.

II. Individual Engagements. Individual engagements of a news story can be evaluated by the average spreading frequencies of (susceptible, normal, all) users who have participated in the news propagation. In this case, the impact of the number of such news spreaders (i.e., MORE-SPREADERS PATTERN) is divided and removed.

All above ways of representing fake news patterns are on the level of nodes, e.g., individual engagement, and the whole network, e.g., network diameter. Next we will specify how to represent DENSER-NETWORKS PATTERN for fake news detection, which will be represented at different network levels: ego, triad and community.

3.4 Denser-Network Pattern

Research in social psychology such as homophily [18] and social validation theory [4] suggests that connected users in social networks share similar attributes, interests and behaviors, e.g., sharing the same news article. On the other hand, malicious users often form cohesive groups, taking collective action that are more purposeful than normal users [20; 39]. These fundamental theories suggest the possibility to assume that fake and true news articles can be distinguished by the relationships among their corresponding spreaders, which can be summarized as:

PATTERN 4 (DENSER-NETWORK PATTERN). *Fake news spreaders form denser networks compared to truth spreaders.*

To capture the "density" of connections among news spreaders, we analyze news networks at different levels: (**I**) ego, (**II**) triad and (**III**) community levels.

I. Ego Level. At the ego level, to compute density of networks formed by users that have engaged with a certain news story, we look at the numbers and proportions of connections that these users (spreaders) have (*i*) generally formed with other spreaders, and (*ii*) specifically with other normal or susceptible spreaders.

i. General Ego Relations. We include as a feature the total number of ego relationships among spreaders for each news story, i.e., the number of edges within each FNN and TNN ($|\mathrm{E}_X|$). To eliminate the impact of the number of news spreaders (i.e., MORE-SPREADERS PATTERN), for each FNN or TNN G_X we also record $|\mathrm{E}_X|/|\mathrm{V}_X|$ and $|\mathrm{E}_X|/\binom{|\mathrm{V}_X|}{2}$, which calculate the average number of ego relationships per spreader and network density, respectively. Here, $|\mathrm{V}_X|$ is the number of spreaders (nodes) in G_X and $\binom{|\mathrm{V}_X|}{2}$ is the number of edges within a fully connected version of G_X.

ii. Specific Ego Relations. Labeling users as susceptible or normal allows one to group all directed ego relationships into four subsets: (1) E_{NN} containing relationships from a normal user to a normal user, (2) E_{NS} containing relationships from a normal user to a susceptible one, (3) E_{SN} containing relationships from a susceptible user to a normal one, (4) E_{SS} containing relationships from a susceptible user to a susceptible one. We include the number and proportion of each type of edges within a FNN or TNN as features being used for fake news detection. In addition, each edge e_{ij} can be also classified into one of the following set: (1) $\mathrm{E}_{\triangle>0}$ if $\triangle = \mathbf{S}(v_i) - \mathbf{S}(v_j) > 0$, (2) $\mathrm{E}_{\triangle=0}$ if $\mathbf{S}(v_i) - \mathbf{S}(v_j) = 0$, (3) $\mathrm{E}_{\triangle<0}$ if $\mathbf{S}(v_i) - \mathbf{S}(v_j) < 0$ which does not require partitioning users into susceptible or normal ones. We also include as features the number and proportion of each above type of edges within a FNN or TNN.

II. Triad Level. Triads (a set of three connected users) are the most basic subgraphs of networks. Similar to our study at the ego level, we investigate (i) general triads and (ii) specific triads formed between [susceptible and normal] users within networks.

i. General Triads. One simple way to represent the DENSER-NETWORK PATTERN is to directly count the total number of triads $|\mathrm{Tr}_X|$ within a G_X. Similarly, to control for MORE-SPREADERS PATTERN, we also include as features the value of $|\mathrm{Tr}_X|/|\mathrm{V}_X|$ and $|\mathrm{Tr}_X|/\binom{|\mathrm{V}_X|}{3}$ where $\binom{|\mathrm{V}_X|}{3}$ is the number of triads within a fully connected version of G_X.

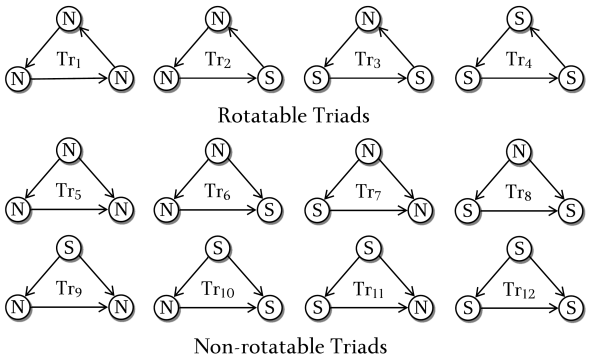


Figure 1: Specific Triads. N indicates normal users and S indicates susceptible users. $A \rightarrow B$ denotes A follows B.

ii. Specific Triads. Regarding each user as either a susceptible or normal user, we can have twelve different triads to be further explored (shown in Figure 1). We include as features the number and proportion of every type of triads within each FNN and TNN.

III. Community Level. In networks, a community structure refers to the occurrence of groups of nodes in a network that are more densely connected internally than with the rest of the network. Therefore, the number and proportion of communities within each FNN and TNN can be used to represent DENSER-NETWORKS PATTERN and, broadly speaking, should be negatively correlated to the network density. As features, we include the number of communities $|\mathrm{M}_X|$ within each FNN and TNN, and the proportion of communities (assuming in the worst case each node is its own community) which removes the impact of the number of news spreaders, i.e., the value of $|\mathrm{M}_X|/|\mathrm{V}_X|$. Note that $|\mathrm{M}_X|$ can be obtained either from (i) global or (ii) local perspective. From a global perspective, communities that nodes (spreaders) belong to within a FNN or TNN, as a subgraph of the social network, are based on the structure of the overall social network. From a local perspective, communities can be detected within a FNN or TNN. We include counts and proportion features for both types of communities.

Integrated Representation of Patterns. To represent each fake news pattern, we have used network information such as network diameters, the number of news spreaders (size), and the number of relationships among the spreaders (density). Networks with various diameters, sizes and densities exhibit various overall structures. Hence, the overall network structure can be regarded as the integrated representation of all related patterns. On the other hand, including such "structure" features to detect fake news helps to evaluate if the fake news patterns and their representations defined in this section have well captured the difference of dissemination between fake news and the truth. To quantify such "structure", one can compare the similarities among FNNs and TNNs, where graph kernel and graph embedding [31; 38] methods can be useful. Here, we consider FNNs and TNNs as *labeled graphs* for further comparison, where node labels can be either (i) user identities or (ii) user attributes (susceptible or normal).

Overall, Table 2 presents all features defined and involved in our work to detect fake news, and their corresponding formulations for reproducibility.

Table 2: Network-based Pattern-driven Feature Set for Fake News Detection

Pattern	No.	Feature(s)	Formulation(s)
More-Spreaders Pattern	1	# News Spreaders	$\lvert \mathrm{V}_X \rvert$
	2-9	# Normal Spreaders, where user susceptibility, a.k.a., $\mathbf{S}(v)$, is based on Equation (1) (#news) or Equation (2) (frequency)	$\lvert \sum_{v_j \in \mathrm{V}_X} \mathbf{B}(\mathbf{S}(v_j) \leq \theta) \rvert$
		# Susceptible Spreaders, where $\mathbf{S}(v)$ is based on #news or frequency	$\lvert \sum_{v_j \in \mathrm{V}_X} \mathbf{B}(\mathbf{S}(v_j) \geq \theta) \rvert$
		% Normal Spreaders, where $\mathbf{S}(v)$ is based on #news or frequency	$\frac{\lvert \sum_{v_j \in \mathrm{V}_X} \mathbf{B}(\mathbf{S}(v_j) \leq \theta) \rvert}{\lvert \mathrm{V}_X \rvert}$
		% Susceptible Spreaders, where $\mathbf{S}(v)$ is based on #news or frequency	$\frac{\lvert \sum_{v_j \in \mathrm{V}_X} \mathbf{B}(\mathbf{S}(v_j) \geq \theta) \rvert}{\lvert \mathrm{V}_X \rvert}$
	10-13	Average Spreader Susceptibility, where $\mathbf{S}(v)$ is based on #news or frequency	$\frac{\sum_{v_j \in \mathrm{V}_X} \mathbf{S}(v_j)}{\lvert \mathrm{V}_X \rvert}$
		Median Spreader Susceptibility, where $\mathbf{S}(v)$ is based on #news or frequency	$\mathbf{P}(\mathbf{S}(v_j) \leq \mathrm{MSS}) = 0.5$ for $v_j \in \mathrm{V}_X$
	14-29	Average Spreader Influence, where influence is based on (in-, out-) degree, (in-, out-) closeness, betweenness, PageRank score, hub and authority score	$\frac{\sum_{v_j \in \mathrm{V}_X} \mathbf{C}(v_j)}{\lvert \mathrm{V}_X \rvert}$
		Median Spreader Influence, where influence is based on (in-, out-) degree, (in-, out-) closeness, betweenness, PageRank score, hub and authority score	$\mathbf{P}(\mathbf{C}(v_j) \leq \mathrm{MSI}) = 0.5$ for $v_j \in \mathrm{V}_X$
Farther-Distance Pattern	30-32	Maximum, Average, and Median Geodesic Distance	-
	33-38	Maximum, Average, and Median Effective Distance, information flow is based on #news and frequency	See Definition 2
Stronger-Engagement Pattern	39	# User Engagements	$\sum_{v_j \in \mathrm{V}_X} \mathbf{T}(v_j, X)$
	40-47	# Normal User Engagements, where $\mathbf{S}(v)$ is based on #news or frequency	$\sum_{v_j \in \mathrm{V}_X; \mathbf{S}(v_j) \leq \theta} \mathbf{T}(v_j, X)$
		# Susceptible User Engagements, where $\mathbf{S}(v)$ is based on #news or frequency	$\sum_{v_j \in \mathrm{V}_X; \mathbf{S}(v_j) \geq \theta} \mathbf{T}(v_j, X)$
		% Normal User Engagements, where $\mathbf{S}(v)$ is based on #news or frequency	$\frac{\sum_{v_j \in \mathrm{V}_X; \mathbf{S}(v_j) \leq \theta} \mathbf{T}(v_j, X)}{\sum_{v_j \in \mathrm{V}_X} \mathbf{T}(v_j, X)}$
		% Susceptible User Engagements, where $\mathbf{S}(v)$ is based on #news or frequency	$\frac{\sum_{v_j \in \mathrm{V}_X; \mathbf{S}(v_j) \geq \theta} \mathbf{T}(v_j, X)}{\sum_{v_j \in \mathrm{V}_X} \mathbf{T}(v_j, X)}$
	48	Average User Engagements	$\frac{\sum_{v_j \in \mathrm{V}_X} \mathbf{T}(v_j, X)}{\lvert \mathrm{V}_X \rvert}$
	49-52	Avg. Normal User Engagements, where $\mathbf{S}(v)$ is based on #news or frequency	$\frac{\sum_{v_j \in \mathrm{V}_X; \mathbf{S}(v_j) \leq \theta} \mathbf{T}(v_j, X)}{\lvert \sum_{v_j \in \mathrm{V}_X} \mathbf{B}(\mathbf{S}(v_j) \leq \theta) \rvert}$
		Avg. Susceptible User Engagements, $\mathbf{S}(v)$ is based on #news or frequency	$\frac{\sum_{v_j \in \mathrm{V}_X; \mathbf{S}(v_j) \geq \theta} \mathbf{T}(v_j, X)}{\lvert \sum_{v_j \in \mathrm{V}_X} \mathbf{B}(\mathbf{S}(v_j) \geq \theta) \rvert}$
Denser-Networks Pattern	53	# Relationships among Spreaders	$\lvert \mathrm{E}_X \rvert$
	54	Average # Relationships of Spreaders	$\lvert \mathrm{E}_X \rvert / \lvert \mathrm{V}_X \rvert$
	55	Ego Density	$\lvert \mathrm{E}_X \rvert / \binom{\lvert \mathrm{V}_X \rvert}{2}$
	56-71	#/% $N \to N$, where $\mathbf{S}(v)$ is based on #news and frequency	$\lvert \mathrm{E}_{NN} \cap \mathrm{E}_X \rvert$; $\frac{\lvert \mathrm{E}_{NN} \cap \mathrm{E}_X \rvert}{\lvert \mathrm{E}_X \rvert}$
		#/% $N \to S$, where $\mathbf{S}(v)$ is based on #news and frequency	$\lvert \mathrm{E}_{NS} \cap \mathrm{E}_X \rvert$; $\frac{\lvert \mathrm{E}_{NS} \cap \mathrm{E}_X \rvert}{\lvert \mathrm{E}_X \rvert}$
		#/% $S \to N$, where $\mathbf{S}(v)$ is based on #news and frequency	$\lvert \mathrm{E}_{SN} \cap \mathrm{E}_X \rvert$; $\frac{\lvert \mathrm{E}_{SN} \cap \mathrm{E}_X \rvert}{\lvert \mathrm{E}_X \rvert}$
		#/% $S \to S$, where $\mathbf{S}(v)$ is based on #news and frequency	$\lvert \mathrm{E}_{SS} \cap \mathrm{E}_X \rvert$; $\frac{\lvert \mathrm{E}_{SS} \cap \mathrm{E}_X \rvert}{\lvert \mathrm{E}_X \rvert}$
	72-83	#/% $\mathbf{S}(v_i) > \mathbf{S}(v_j)$, where $\mathbf{S}(v)$ is based on #news and frequency	$\lvert \mathrm{E}_{\triangle > 0} \cap \mathrm{E}_X \rvert$; $\frac{\lvert \mathrm{E}_{\triangle > 0} \cap \mathrm{E}_X \rvert}{\lvert \mathrm{E}_X \rvert}$
		#/% $\mathbf{S}(v_i) = \mathbf{S}(v_j)$, where $\mathbf{S}(v)$ is based on #news and frequency	$\lvert \mathrm{E}_{\triangle = 0} \cap \mathrm{E}_X \rvert$; $\frac{\lvert \mathrm{E}_{\triangle = 0} \cap \mathrm{E}_X \rvert}{\lvert \mathrm{E}_X \rvert}$
		#/% $\mathbf{S}(v_i) < \mathbf{S}(v_j)$, where $\mathbf{S}(v)$ is based on #news and frequency	$\lvert \mathrm{E}_{\triangle < 0} \cap \mathrm{E}_X \rvert$; $\frac{\lvert \mathrm{E}_{\triangle < 0} \cap \mathrm{E}_X \rvert}{\lvert \mathrm{E}_X \rvert}$
	84	# Triads	$\lvert \mathrm{Tr}_X \rvert$
	85	Average # Triads of Spreaders	$\lvert \mathrm{Tr}_X \rvert / \lvert \mathrm{V}_X \rvert$
	86	Triad Density	$\lvert \mathrm{Tr}_X \rvert / \binom{\lvert \mathrm{V}_X \rvert}{3}$
	87-110	# Tr_1 to Tr_{12} (see Figure 1), where $\mathbf{S}(v)$ is based on #news and frequency	$\lvert \mathrm{Tr}_k \cap \mathrm{Tr}_X \rvert$; $\frac{\lvert \mathrm{Tr}_k \cap \mathrm{Tr}_X \rvert}{\lvert \mathrm{Tr}_X \rvert}$; $k = 1, 2, \cdots, 12$
	111-134	% Tr_1 to Tr_{12}, where user susceptibility is based on # news and frequency	
	135-136	# Communities (from global and local perspective, see Section 3.4 for details)	$\lvert \mathrm{M}_X \rvert$
	137-138	Community Density (from global and local perspective)	$\lvert \mathrm{M}_X \rvert / \lvert \mathrm{V}_X \rvert$

Table 3: Data Statistics

Data	PolitiFact	BuzzFeed
# Users	23,865	15,257
# News–Users	32,791	22,779
# Users–Users	574,744	634,750
# News Stories	240	182
# True News	120	91
# Fake News	120	91
# Triads	6,972,189	6,885,951
# Communities	163	46

4. EXPERIMENTS

Fake news patterns in networks have been specified as well as how they can be represented as a set of quantifiable and meaningful features. In this section, various experiments are conducted to verify the effectiveness of the proposed approach in detecting fake news. We first present the experimental setup in Section 4.1, followed by the evaluations of experimental results in Section 4.2.

4.1 Experimental Setup

We detail data used in experiments in Section *4.1.1*, followed by how data is prepared for experiments in Section *4.1.2*, and the baselines which the proposed approach is compared with in Section *4.1.3*.

4.1.1 Datasets

Our experiments are conducted on two public benchmark datasets of fake news detection [33]. News articles in these datasets are collected from PolitiFact and BuzzFeed, respectively. Ground truth labels (*true* or *fake*) of news articles in both datasets are provided by fact-checking experts. In addition to (i) news content and labels, both datasets also provide information on (ii) social network of Twitter which contains Twitter users and their following relationships, i.e., user-user relationships, and (iii) how the news has propagated (tweeted/re-tweeted) by users, i.e., news-user relationships. Based on the original datasets, we further identify triads and communities in the social network. Communities are detected using Louvain algorithm, a fast and widely-accepted modularity-based community detection algorithm [1]. Statistics of two datasets are shown in Table 3.

4.1.2 Data Preparation

Following dataset collection, feature values are computed for both datasets, which will be utilized in a supervised learning framework for fake news detection. However, an extra step is necessary to take when computing user susceptibility scores. In Section 3, two ways are defined for determining user susceptibility [to fake news] (see Equation (1) and (2), respectively). Both ways rely on the historical information of users on how they previously engaged in fake news dissemination, where the news labels (*true* or *fake*) are necessary in the calculation. To avoid *information leakage* (i.e., features having an unfair prior knowledge of labels), when dividing a dataset into the training and testing one, we dynamically calculate user susceptibility by using the historical information provided in training dataset, rather than the whole dataset. For users with no historical information in training dataset, we treat their susceptibility as the threshold value, indicating that their susceptibility to fake news is unknown.

4.1.3 Baselines

The performance of the proposed method is compared with several benchmark fake news detection methods on the same datasets. These methods include (1) content-based (linguistic) models, which rely on non-latent ([23; 43]) or latent representation ([19; 14]) of news content, (2) network-based models ([3]), which investigate information revealed in news propagation, and hybrid models ([33]), which utilize both content and network information to detect fake news.

I. Pérez-Rosas et al. [23] propose a comprehensive linguistic model for fake news detection, involving the following features: (i) n-grams (i.e., uni-grams and bi-grams) and (ii) CFGs based on TF-IDF encoding; (iii) word and phrase proportions referring to all categories provided by LIWC; and (iv) readability. Features are computed and used to predict fake news within a supervised machine learning framework.

II. Zhou et al. [43]. In our previous study, forensic psychological theories are studied and used to detect fake news in a supervised learning framework, which provide the evidence of distinguishing fake news in content style from the truth. Such content style is captured by the frequency of (i) lexicons relying on Bag-Of-Words (BOW) model, (ii) Part-Of-Speech (POS) tags and Context Free Grammers (CFGs) at syntax-level, (iii) Rhetorical Relationships (RRs) at discourse-level, and by assessing a set of theory-driven (iv) DisInformation-related Attributes (DIAs) and (v) ClickBait-related Attributes (CBAs) at semantic-level.

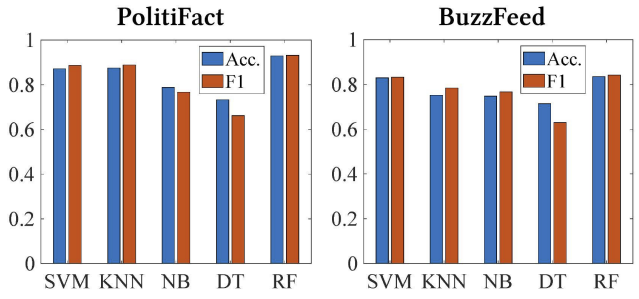
III. Castillo et al. [3] design features that exploit information from user profiles, tweets and propagation trees to evaluate news credibility within a supervised learning framework. Specifically, these features are based on (i) quantity, sentiment, hash-tag and URL information from user tweets, (ii) user profiles such as registration age, (iii) news topics through mining tweets of users, and (iv) propagation trees (e.g., the number of propagation trees for each news topic).

IV. Shu et al. [33] detect fake news by exploring and embedding the relationships among news articles, publishers and spreaders on social media. Such embedding involves (i) news content by using non-negative matrix factorization, (ii) users on social media, (iii) news-user relationships (i.e., user engagements in spreading news articles), and (iv) news-publisher relationships (i.e., publisher engagements in publishing news articles). Fake news detection is then conducted within a semi-supervised machine learning framework.

Additionally, fake news detection based on latent representation of news articles is also investigated in comparative studies, where we consider as baselines supervised classifiers with features being **(V) Word2Vec** [19] and **(VI) Doc2Vec** [14] embedding of news articles.

4.2 Performance Evaluation

Various supervised learners with 5-fold cross-validation were used in our experiments. The performance is evaluated using accuracy and F_1 score. In the following, we will first present the general performance of the proposed approach in Section *4.2.1*. Based on that, the importance of patterns (see Section *4.2.2*) and features (see Section *4.2.3*) in fake news detection is further assessed. The sensitivity of the proposed approach is evaluated to the threshold and calculation of user susceptibility in Section *4.2.4*, as well as its sensitivity to how much labeled news articles are available



Classifier	PolitiFact		BuzzFeed	
	Acc.	F_1	Acc.	F_1
Support Vector Machine (SVM)	.871	.886	.830	.833
k-Nearest Neighbors (kNN)	.875	.888	.752	.784
Naïve Bayes (NB)	.788	.767	.748	.767
Decision Tree (DT)	.733	.662	.714	.630
Random Forest (RF)	.929	.932	.835	.842

Figure 2: General Performance of Fake News Detection by Using Different Classifiers, where random forests perform best among all on both datasets.

and what proportion between two labels (true vs. fake) in Section *4.2.5*. The performance of our approach on fake news early detection is finally examined in Section *4.2.6*.

4.2.1 General Performance Evaluation

We experimented with various classifiers to detect fake news using our features, including Support Vector Machine (SVM), k-Nearest Neighbors (k-NN), Naïve Bayes (NB), Decision Trees (DT) and Random Forests (RF). The results obtained are all provided in Figure 2. It can be observed from the Figure 2 that RF performs best on both datasets, achieving an accuracy and F_1 score of around 0.93 on PolitiFact and around 0.84 on BuzzFeed.

Such performance is further compared with that of baselines, where the results are presented in Table 4. Compared to the content-based ([23; 43; 19; 14]) and network-based models ([3]) among baselines, the proposed approach can perform relatively well on both datasets. Compared to the hybrid one ([33]), the proposed approach can be comparable with it and can outperform it when introducing the linguistic features in the proposed approach ("Our Approach + [43]").

4.2.2 Performance of Fake News Patterns

We further analyze the performance of each fake news pattern and their combinations on fake news detection. The results are presented in Table 5, which supports the following observations. First, More-Spreader Pattern and Stronger-Engagement Pattern perform best compared to the others when being separately utilized to detect fake news, achieving around 89% (81%) accuracy and F_1 score using PolitiFact (BuzzFeed) data. The performance of Denser-Network Pattern follows. Second, when combining different patterns, their performance is in general better than when separately using them, which can achieve an accuracy and F_1 score of around 93% (82%) on PolitiFact (BuzzFeed). Third, when using all patterns to detect fake news, a significantly better performance is achieved compared to using network similarity, which provides a mix of patterns from a higher network view, which is a positive sign for our summarized fake news patterns and defined representations of patterns in networks. Fourth, network similarity features can slightly improve the performance of the combination of four fake news propagation patterns, which finally achieves an accuracy and F_1 score of around 93% (84%) on PolitiFact (BuzzFeed).

Table 4: General Performance of Fake News Detection Methods. The proposed network-based approach can perform relatively well compared to the content-based ([23; 43; 19; 14]) and network-based approaches ([3]) among baselines. Compared to the hybrid one ([33]), the proposed approach can be comparable with it and can outperform it when introducing the linguistic features in the proposed approach ("Our Approach + [43]").

Method	PolitiFact		BuzzFeed	
	Acc.	F_1	Acc.	F_1
Pérez-Rosas et al. [23]	.811	.811	.755	.757
Zhou et al. [43]	.865	.865	.855	.856
• BOWs	.856	.858	.823	.823
• POS Tags	.755	.755	.745	.745
• CFGs	.877	.877	.778	.778
• DIAs	.729	.735	.667	.647
• CBAs	.604	.612	.638	.628
• RRs	.621	.621	.658	.658
Word2Vec-based [19]	.688	.667	.703	.718
Doc2Vec-based [14]	.698	.698	.615	.615
Castillo et al. [3]	.794	.822	.789	.794
Shu et al. [33]	.878	.880	.864	.870
Our Approach	.929	.932	.835	.842
Our Approach + [43]	.933	.939	.865	.884

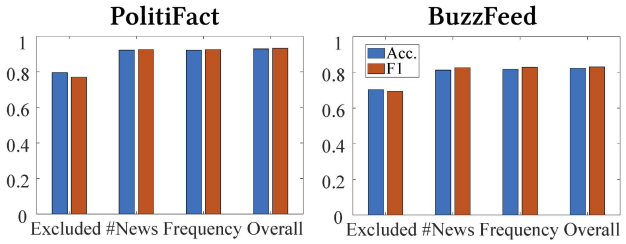
4.2.3 Feature Importance Analysis

Features are ranked by their importance in fake news detection. Results are shown in Table 6, which are obtained by Relief algorithm, a widely-accept feature selection algorithm [13]. Consistent with the performance of patterns, features representing More-Spreader Pattern, Stronger-Engagement Pattern, and Denser-Network Pattern are relatively more discriminative in predicting fake news compared to the other features. In addition, it can be observed from Table 6 and Figure 7 that for More-Spreader Pattern, features contributing most to fake news detection are the mean or median of (i) spreader susceptibility and (ii) spreader influence, where fake news spreaders often share a higher susceptibility and centrality score compared to true news spreaders. For Stronger-Engagement Pattern, such features relate to (iii) susceptible and normal user engagements, where susceptible (normal) users engage more strongly in fake (true) news compared to true (fake) news. For Denser-Network Pattern, such features are generally at (iv) ego and (v) community level. Specifically, FNNs are characterized with a higher proportion of connections between susceptible users ($S \to S$) while TNNs are characterized with a higher proportion of other ego relations ($S \to N$, $N \to S$ and $N \to N$). With a same network size, a

[4]The results are based on Weisfeiler-Lehman graph kernel [31]. As a widely-accept graph kernel, Weisfeiler-Lehman graph kernel can measure the similarities among labeled graphs, which we treat TNNs and FNNs as.

Table 5: Pattern Performance in Fake News Detection. More-Spreader Pattern and Stronger-Engagement Pattern perform best compared to the others when being separately utilized to detect fake news. When combining different patterns, their performance is in general better than when separately using them, and than when using network similarity, as a mix of patterns from a higher view.

Pattern(s)	PolitiFact		BuzzFeed	
	Accuracy	F_1 Score	Accuracy	F_1 Score
More-Spreaders	.891	.901	.808	.817
Farther-Distance	.639	.587	.678	.698
Stronger-Engagement	.898	.898	.807	.808
Denser-Networks	.746	.718	.687	.704
More-Spreaders + Farther-Distance	.846	.803	.824	.824
More-Spreaders + Stronger-Engagement	.879	.864	.830	.847
More-Spreaders + Denser-Networks	.919	.919	.770	.796
Farther-Distance + Stronger-Engagement	.917	.923	.814	.824
Farther-Distance + Denser-Networks	.742	.710	.786	.798
Stronger-Engagement + Denser-Networks	.921	.926	.829	.840
All Patterns - Denser-Networks	.908	.916	.814	.819
All Patterns - Stronger-Engagement	.929	.928	.819	.815
All Patterns - Farther-Distance	.913	.914	.780	.759
All Patterns - More-Spreaders	.879	.871	.802	.803
All Patterns	.929	.928	.828	.823
Network Similarity (Mix of Patterns)[4]	.808	.770	.671	.689
All Patterns + Network Similarity	.929	.932	.835	.842



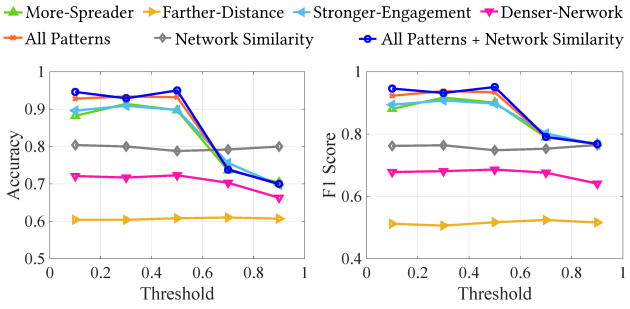
User Susceptability	PolitiFact		BuzzFeed	
	Acc.	F_1	Acc.	F_1
Excluded	.796	.771	.704	.694
Based on #News	.922	.925	.813	.827
Based on Frequency	.921	.925	.818	.830
Both Included	.929	.932	.824	.831

Figure 3: Impact of the Utilizing and Means to Calculate User Susceptibility on Fake News Detection. Considering user susceptibility can improve fake news prediction, while two methods of computing user susceptibility perform similarly.

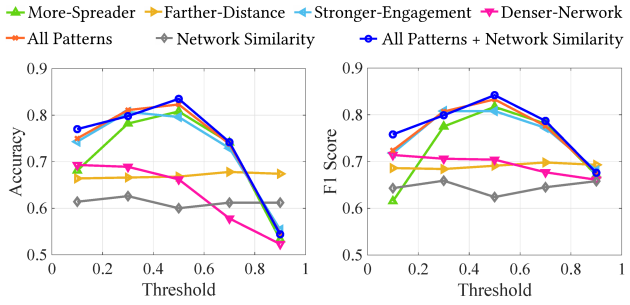
FNN often has less communities compared to a TNN, which indicate that a denser network structure is often within a FNN compared to a TNN.

4.2.4 *User Susceptibility Analysis*

Two methods have been defined to compute user susceptibility [to fake news] which plays an important role in representing patterns - one is based on the number of fake news that a user has spread (Equation (1)) and the other is based on the frequency of a user on spreading fake news (Equation (2)). Once such susceptibility is computed, whether a user is susceptible or normal relies on the selection of a threshold. Thus, here we assess the impact of user susceptibility on fake news prediction by (I) how the susceptibility



(a) PolitiFact



(b) BuzzFeed

Figure 4: Impact of User Susceptibility Threshold on Fake News Detection. When the threshold changes, (i) the performance of Farther-Distance Pattern and network similarity is hardly impacted due to no relevance to user susceptibility; (ii) More-Spreader Pattern or Stronger-Engagement Pattern can outperform Denser-Network Pattern though they are less stable; (iii) using all patterns (with/without network similarity) can always perform comparatively well compared to the others and achieve the highest performance when the threshold value is 0.5.

Table 6: Top 20 Important Features

Rank	PolitiFact	BuzzFeed
1	Avg. Spreader susceptibility (#News)	Median Spreader susceptibility (Frequncy)
2	Avg. Spreader susceptibility (Frequncy)	Median Spreader susceptibility (Frequncy)
3	Median Spreader susceptibility (Frequncy)	Avg. Spreader susceptibility (#News)
4	Median Spreader susceptibility (Frequncy)	Avg. Spreader susceptibility (Frequncy)
5	Avg. Normal User Engagement (Frequency)	Global Community Density
6	% Normal User Engagement (Frequency)	Median Spreader Influence (Authority)
7	% Susceptible User Engagement (Frequency)	% Normal User Engagement (#News)
8	% Normal User Engagement (#News)	% Susceptible User Engagement (#News)
9	% Susceptible User Engagement (#News)	Median Spreader Influence (In-degrees)
10	% Normal Spreaders (Frequency)	% Normal User Engagement (Frequency)
11	% Susceptible Spreaders (Frequency)	% Susceptible User Engagement (Frequency)
12	% Normal Spreaders (#News)	% Normal Spreaders (#News)
13	% Susceptible Spreaders (#News)	% Susceptible Spreaders (#News)
14	Global Community Density	Median Spreader Influence (In-closeness)
15	% Egos (S → S, #News)	% Egos (S → S, Frequency)
16	% Egos (S → S, Frequency)	% Normal Spreaders (Frequency)
17	Avg. Normal User Engagement (#News)	% Susceptible Spreaders (Frequency)
18	% Egos (S → N, Frequency)	% Egos (S → S, #News)
19	% Egos (N → S, #News)	Median Spreader Influence (Hub)
20	% Egos (N → S Frequency)	% Triads (N → S, S → S, S → N, #News)

Blue: More-Spreader Pattern; Green: Stronger-Engagement Pattern; Yellow: Denser-Network Pattern

is computed and (II) how the threshold is determined.

I. Computation of User Susceptibility. To evaluate the impact of the [non-] existence of user susceptibility and how it is computed, we conduct fake news detection (i) without user susceptibility features, and by using features that compute user susceptibility based on the (ii) number of news spread, (iii) frequency of the spreading, and (iv) both number and frequency. The results are provided in Figure 3. It can be observed that utilizing user susceptibility can enhance the performance by ∼10% when predicting fake news based on both PolitiFact and BuzzFeed datasets. However, no significant performance difference exists between the two ways that user susceptibility can be calculated.

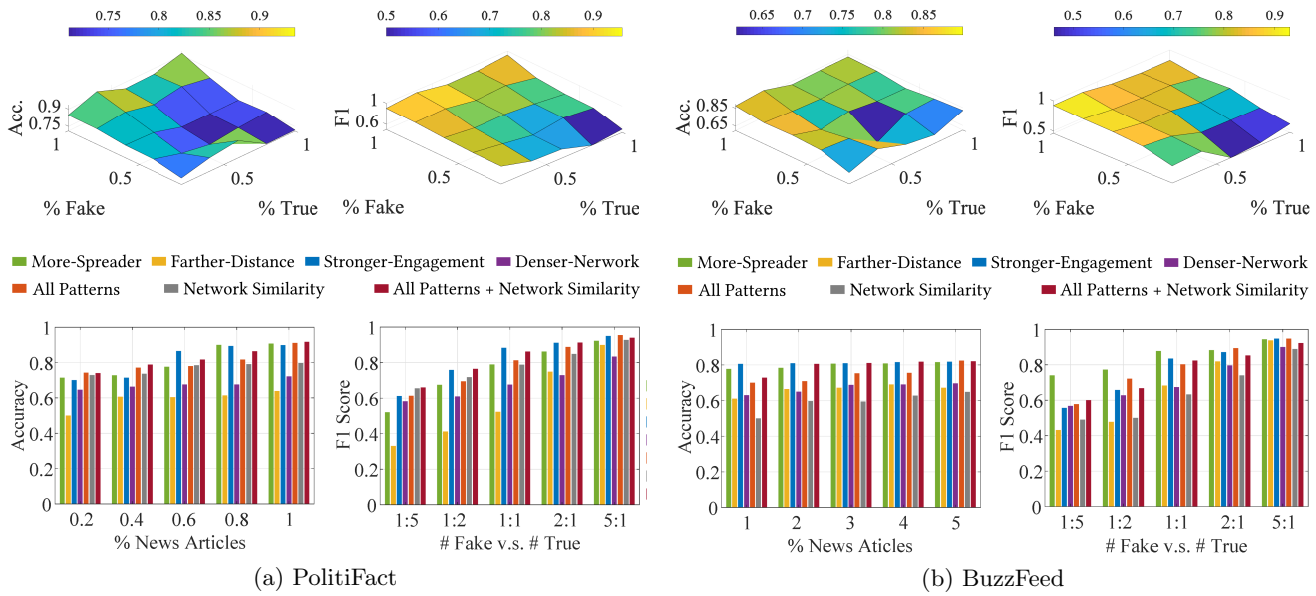
II. Evaluating User Susceptibility Threshold. To evaluate the impact of susceptibility threshold on fake news prediction, we set the threshold value from 0 (i.e., all users are susceptible) to 1 (i.e., all users are normal) and use the proposed approach to predict fake news based on different threshold values. The results are plotted in Figure 4. It can be observed that when the threshold changes, (i) as the features representing Farther-Distance Pattern and network similarity do not need to compute susceptibility scores of users, the performance is rarely impacted; (ii) More-Spreader Pattern or Stronger-Engagement Pattern can outperform Denser-Network Pattern while are less stable; (iii) the combination of all patterns (with/without network similarity) can always perform comparatively well compared to the others and achieves the highest performance when the threshold value is 0.5.

4.2.5 *Impact of News Number and Distribution*

As in the practice, the number and distribution (the proportion between fake and true news) of news articles on social networks can be dynamic and change, here we evaluate the impact of the (i) number and (ii) distribution of news articles available on the performance of the proposed method. To that end, a certain proportion ($\in [0, 1]$) of samples is randomly selected from the population of true news stories in a dataset and that of fake news stories in that dataset, respectively. The performance of the proposed approach with each proportion of true and fake news is plotted in Figure 5 (the upper row). Results in Figure 5 indicate that, in general, the proposed approach can perform an accuracy rate of ∼0.7 to ∼0.85 and an F_1 score of ∼0.65 to ∼0.9 in most cases on both datasets.

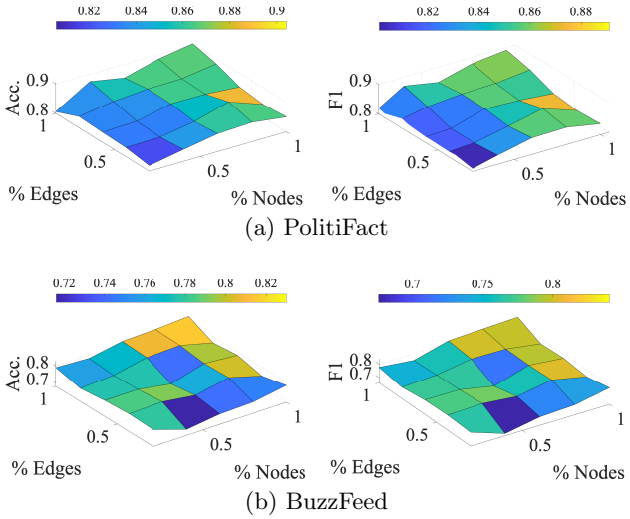
Note that two variables (i.e., the number and distribution of news articles) both exist and change in this process. For a clear observation, we first control the sampled news distribution to be the same as that in original datasets, and record the performance of the proposed method with various number of overall news articles available for training and predicting fake news. On the other hand, we keep the a fixed number of news articles while vary the proportion between fake and true news in it.

Results are all provided in Figure 5 (the lower row). It can be observed that (i) the impact of the number of news articles is less significant compared to the news distribution when predicting fake news based on the proposed method; (ii) when varying the number of news articles, an accuracy rate (we only present the accuracy performance in Figure 5 as the datasets are balanced at this time) between ∼0.73 (∼0.8) to ∼0.9 (∼0.82) can be achieved by using PolitiFact (BuzzFeed) data and all patterns plus network similarity (Stronger-Engagement Pattern); and (iii) when varying the news distribution, a F_1 score (we evaluate the performance only based on F_1 score here as the datasets can be unbalanced) ranging from ∼0.65 (∼0.75) to ∼0.93 (∼0.92) can be achieved by using PolitiFact (BuzzFeed) data and all patterns plus network similarity (More-Spreader Pattern).



(a) PolitiFact (b) BuzzFeed

Figure 5: Impact of News Number and Distribution on Fake News Prediction. In general, (i) the proposed approach can achieve an accuracy rate (a F_1 score) ∼0.7 (∼0.65) to ∼0.85 (∼0.9) in most cases on both datasets (see the upper four figures). When only the number of news articles varies, an accuracy rate (here the datasets are class-balanced) between ∼0.73 (∼0.8) to ∼0.9 (∼0.82) can be achieved on PolitiFact (BuzzFeed) data and overall features (STRONGER-ENGAGEMENT PATTERN). When the news distribution varies, a F_1 score (here the dataset can be unbalanced) ranging from ∼0.65 (∼0.75) to ∼0.93 (∼0.92) can be achieved on PolitiFact (BuzzFeed) data and overall features (MORE-SPREADER PATTERN).



(a) PolitiFact

(b) BuzzFeed

Figure 6: Impact of Available Network Information on Fake News Prediction. The proposed approach is generally stable with an accuracy and F_1 score between ∼0.8 (∼0.7) and ∼0.9 (∼0.82) on PolitiFact (BuzzFeed) data.

4.2.6 *Early Detection Analysis*

Fake news *early* detection is an arduous but important task. It aims to detect fake news at an early stage before it has widely spread on social networks, when only limited information is available. Early detection is crucial for fake news, especially due to *validity effect*, which indicates that the more individuals get exposed to certain fake news, the more they may trust it. Meanwhile, it is difficult to correct one's cognition after fake news has gained their trust [26]. Effective early detection of fake news helps take early actions on fake news intervention. As few temporal information (e.g., the time that users spread the news articles or form relationships) is available in the datasets, the experiment to verify the early detection ability of the proposed approach is designed based on the following intuition. In our study, each FNN or TNN provides all network information for the corresponding [fake or true] news story. If the dissemination of a news story is at its early stages, the number of spreaders (i.e., nodes) and the involved relationships among spreaders (i.e., edges) should be relatively small compared to when it has been widely spread. Hence, we randomly select a certain proportion ($\in [0, 1]$) of nodes or edges for each FNN or TNN and detect fake news on these [sub-] FNNs and [sub-] TNNs. The results are presented in Figure 6. It can be observed from Figure 6 that the proposed approach is generally stable with an accuracy and F_1 score between ∼0.8 (∼0.7) and ∼0.9 (∼0.82) by using PolitiFact (BuzzFeed) data, which is friendly to fake news early detection.

5. CONCLUSION

With the rampancy of fake news and the damage it has inflicted on societies, there is a demand for a deep understanding of fake news and effective approaches to detect it. Integrating empirical studies and social psychological theories, our work can deepen the understanding of fake news by investigating its patterns in social networks. These patterns are further exploited and represented at multiple network levels (i.e., node-, ego-, triad-, community- and network-level) to detect fake news in an explainable way. Experiments on two real-world datasets validate the effectiveness of the proposed approach, which can perform relatively well

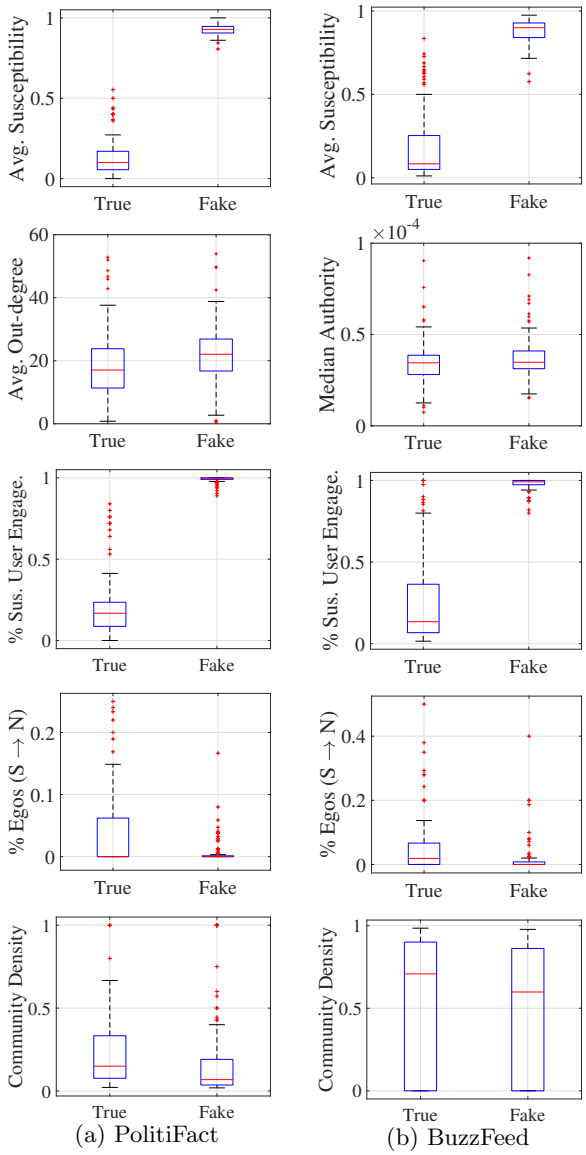


Figure 7: Statistics of Illustrated Important Features for Fake and True News

compared to the state-of-the-art. It should be pointed out that compared to content-based models, the proposed approach can hardly detect fake news before it has been propagated on social media, while it can detect fake news with a stable performance by using limited amount of network (propagation) information and a very small number of training news articles. Additionally, by rarely relying on news content, it provides the other perspective to detect fake news which is being robust to the possible manipulation writing styles by malicious entities. Clearly, the proposed approach can be enhanced by introducing more patterns and user attributes that are defined using network information such as network roles [9], and validated on cross-domain and language fake news data to assess its generalization power. Both will be part of our future work.

6. REFERENCES

[1] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, 2008.

[2] D. Brockmann and D. Helbing. The hidden geometry of complex, network-driven contagion phenomena. *Science*, 342(6164):1337–1342, 2013.

[3] C. Castillo, M. Mendoza, and B. Poblete. Information credibility on twitter. In *Proceedings of the 20th International Conference on World Wide Web*, pages 675–684. ACM, 2011.

[4] R. B. Cialdini. *Influence: Science and Practice*, volume 4. Pearson education Boston, MA, 2009.

[5] G. L. Ciampaglia, P. Shiralkar, L. M. Rocha, J. Bollen, F. Menczer, and A. Flammini. Computational fact checking from knowledge networks. *PloS one*, 10(6):e0128193, 2015.

[6] X. Dong, E. Gabrilovich, G. Heitz, W. Horn, N. Lao, K. Murphy, T. Strohmann, S. Sun, and W. Zhang. Knowledge vault: A web-scale approach to probabilistic knowledge fusion. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 601–610. ACM, 2014.

[7] M. Gupta, P. Zhao, and J. Han. Evaluating event credibility on twitter. In *Proceedings of the 2012 SIAM International Conference on Data Mining*, pages 153–164. SIAM, 2012.

[8] S. Gupta, R. Thirukovalluru, M. Sinha, and S. Mannarswamy. CIMTDetect: A Community Infused Matrix-Tensor Coupled Factorization Based Method for Fake News Detection. *arXiv preprint arXiv:1809.05252*, 2018.

[9] K. Henderson, B. Gallagher, T. Eliassi-Rad, H. Tong, S. Basu, L. Akoglu, D. Koutra, C. Faloutsos, and L. Li. RolX: Structural Role Extraction and Mining in Large Graphs. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1231–1239. ACM, 2012.

[10] K. H. Jamieson and J. N. Cappella. *Echo Chamber: Rush Limbaugh and the Conservative Media Establishment*. Oxford University Press, 2008.

[11] Z. Jin, J. Cao, Y. Zhang, and J. Luo. News Verification by Exploiting Conflicting Social Viewpoints in Microblogs. In *AAAI*, pages 2972–2978, 2016.

[12] H. Karimi and J. Tang. Learning Hierarchical Discourse-level Structure for Fake News Detection. *arXiv preprint arXiv:1903.07389*, 2019.

[13] K. Kira and L. A. Rendell. A practical approach to feature selection. In *Machine Learning Proceedings 1992*, pages 249–256. Elsevier, 1992.

[14] Q. Le and T. Mikolov. Distributed Representations of Sentences and Documents. In *International Conference on Machine Learning*, pages 1188–1196, 2014.

[15] Y. Liu and Y.-f. B. Wu. Early Detection of Fake News on Social Media Through Propagation Path Classification with Recurrent and Convolutional Networks. In *AAAI*, 2018.

[16] G. Loewenstein. The psychology of curiosity: A review and reinterpretation. *Psychological bulletin*, 116(1):75, 1994.

[17] J. Ma, W. Gao, and K.-F. Wong. Rumor Detection on Twitter with Tree-structured Recursive Neural Networks. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, pages 1980–1989, 2018.

[18] M. McPherson, L. Smith-Lovin, and J. M. Cook. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1):415–444, 2001.

[19] T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.

[20] A. Mukherjee, A. Kumar, B. Liu, J. Wang, M. Hsu, M. Castellanos, and R. Ghosh. Spotting Opinion Spammers Using Behavioral Footprints. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 632–640. ACM, 2013.

[21] M. Nickel, K. Murphy, V. Tresp, and E. Gabrilovich. A review of relational machine learning for knowledge graphs. *Proceedings of the IEEE*, 104(1):11–33, 2016.

[22] R. Oshikawa, J. Qian, and W. Y. Wang. A Survey on Natural Language Processing for Fake News Detection. *arXiv preprint arXiv:1811.00770*, 2018.

[23] V. Pérez-Rosas, B. Kleinberg, A. Lefevre, and R. Mihalcea. Automatic Detection of Fake News. *arXiv preprint arXiv:1708.07104*, 2017.

[24] M. Potthast, J. Kiesel, K. Reinartz, J. Bevendorff, and B. Stein. A Stylometric Inquiry into Hyperpartisan and Fake News. *arXiv preprint arXiv:1702.05638*, 2017.

[25] X. Ren, N. Peng, and W. Y. Wang. Scalable Construction and Reasoning of Massive Knowledge Bases. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Tutorial Abstracts*, pages 10–16, 2018.

[26] A. Roets et al. 'fake news': Incorrect, but hard to correct. The role of cognitive ability on the impact of false information on social impressions. *Intelligence*, 65:107–110, 2017.

[27] V. L. Rubin. On deception and deception detection: Content analysis of computer-mediated stated beliefs. *Proceedings of the Association for Information Science and Technology*, 47(1):1–10, 2010.

[28] V. L. Rubin and T. Lukoianova. Truth and deception at the rhetorical structure level. *Journal of the Association for Information Science and Technology*, 66(5):905–917, 2015.

[29] N. Ruchansky, S. Seo, and Y. Liu. CSI: A Hybrid Deep Model for Fake News Detection. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pages 797–806. ACM, 2017.

[30] C. Shao, G. L. Ciampaglia, O. Varol, K.-C. Yang, A. Flammini, and F. Menczer. The spread of low-credibility content by social bots. *Nature Communications*, 9(1):4787, 2018.

[31] N. Shervashidze, P. Schweitzer, E. J. v. Leeuwen, K. Mehlhorn, and K. M. Borgwardt. Weisfeiler-Lehman Graph Kernels. *Journal of Machine Learning Research*, 12(Sep):2539–2561, 2011.

[32] B. Shi and T. Weninger. Discriminative predicate path mining for fact checking in knowledge graphs. *Knowledge-Based Systems*, 104:123–133, 2016.

[33] K. Shu, S. Wang, and H. Liu. Beyond News Contents: The Role of Social Context for Fake News Detection. In *WSDM*, 2019.

[34] C. Silverman. This analysis shows how viral fake election news stories outperformed real news on Facebook. *BuzzFeed News*, 16, 2016.

[35] U. Undeutsch. Beurteilung der glaubhaftigkeit von aussagen. *Handbuch der psychologie*, 11:26–181, 1967.

[36] S. Vosoughi, D. Roy, and S. Aral. The spread of true and false news online. *Science*, 359(6380):1146–1151, 2018.

[37] Y. Wang, F. Ma, Z. Jin, Y. Yuan, G. Xun, K. Jha, L. Su, and J. Gao. EANN: Event Adversarial Neural Networks for Multi-Modal Fake News Detection. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 849–857. ACM, 2018.

[38] K. Wu, S. Yang, and K. Q. Zhu. False rumors detection on sina weibo by propagation structures. In *Data Engineering (ICDE), 2015 IEEE 31st International Conference on*, pages 651–662. IEEE, 2015.

[39] S. Xie, G. Wang, S. Lin, and P. S. Yu. Review spam detection via temporal pattern discovery. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 823–831. ACM, 2012.

[40] Y. Yang, L. Zheng, J. Zhang, Q. Cui, Z. Li, and P. S. Yu. TI-CNN: Convolutional Neural Networks for Fake News Detection. *arXiv preprint arXiv:1806.00749*, 2018.

[41] R. Zafarani, X. Zhou, K. Shu, and H. Liu. Fake News Research: Theories, Detection Strategies, and Open Problems. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. ACM, 2019.

[42] J. Zhang, L. Cui, Y. Fu, and F. B. Gouza. Fake News Detection with Deep Diffusive Network Model. *arXiv preprint arXiv:1805.08751*, 2018.

[43] X. Zhou, A. Jain, V. V. Phoha, and R. Zafarani. Fake News Early Detection: A Theory-driven Model. *arXiv preprint arXiv:1904.11679*, 2019.

[44] X. Zhou and R. Zafarani. Fake News: A Survey of Research, Detection Methods, and Opportunities. *arXiv preprint arXiv:1812.00315*, 2018.