Hierarchical Propagation Networks for Fake News Detection: Investigation and Exploitation

Kai Shu[1], **Deepak Mahudeswaran**[1], **Suhang Wang**[2] **and Huan Liu**[1]

[1]Computer Science and Engineering, Arizona State University, Tempe, 85281, USA
[2]College of Information Sciences and Technology, Penn State University, University Park, PA, 16802, USA
{kai.shu, dmahudes, huan.liu}@asu.edu, szw494@psu.edu

Abstract

Consuming news from social media is becoming increasingly popular. However, social media also enables the wide dissemination of *fake news*. Because of the detrimental effects of fake news, fake news detection has attracted increasing attention. However, the performance of detecting fake news only from news content is generally limited as fake news pieces are written to mimic true news. In the real world, news pieces spread through *propagation networks* on social media. The news propagation networks usually involve multi-levels. In this paper, we study the challenging problem of *investigating* and *exploiting* news hierarchical propagation network on social media for fake news detection.

In an attempt to understand the correlations between news propagation networks and fake news, first, we build hierarchical propagation networks for fake news and true news pieces; second, we perform a comparative analysis of the propagation network features from structural, temporal, and linguistic perspectives between fake and real news, which demonstrates the potential of utilizing these features to detect fake news; third, we show the effectiveness of these propagation network features for fake news detection. We further validate the effectiveness of these features from feature importance analysis. We conduct extensive experiments on real-world datasets and demonstrate the proposed features can significantly outperform state-of-the-art fake news detection methods by at least 1.7% with an average F1>0.84. Altogether, this work presents a data-driven view of hierarchical propagation network and fake news and paves the way towards a healthier online news ecosystem.

1 Introduction

Social media platforms are easy to access, support fast dissemination of posts, and allow users to comment and share, which are attracting more and more users to seek out and receive timely news information online. For example, the Pew Research Center announced that approximately 68% of US adults get news from social media in 2018, while in 2012, only 49% reported seeing news on social media[1]. However, social media also enables the wide dissemination of large amounts of fake news, i.e., news stories with intentionally false information (Allcott and Gentzkow 2017; Shu et al. 2017). For example, a report estimated that over 1 million tweets were related to the fake news story "Pizzagate" [2] by the end of 2016 presidential election.

The widespread of fake news has detrimental societal effects. First, it weakens public trust in governments and journalism. For example, the social engagements of fake news (e.g. share, like) during the 2016 U.S. presidential election campaign for top twenty fake news pieces was, ironically, larger than the top twenty most-discussed true stories [3]. Second, fake news may change the way people respond to legitimate news. Study showed that 45% of people who do not trust media are because of fake news[4]. Third, rampant fake news can lead to real-life societal events. For example, fake news claiming that Barack Obama was injured in an explosion wiped out $130 billion in stock value [5].

However, detecting fake news on social media presents unique challenges. First, fake news is intentionally written to mislead readers, which makes it nontrivial to detect simply based on content; Second, social media data is large-scale, multi-modal, mostly user-generated, sometimes anonymous and noisy. In the real world, news pieces spread in networks on social media. These propagation networks have a hierarchical structure, including macro-level and micro-level propagation networks (see Figure 1). On one hand, macro-level propagation networks demonstrate the spreading path from news publishers to the social media posts sharing the news, and those reposts of these posts. Macro-level networks for fake news are shown to be deeper, wider, and includes more social bots than real news (Shao et al. 2017; Vosoughi, Roy, and Aral 2018), which provides clues for detecting fake news. On the other hand, micro-level propagation networks illustrate the user conversations under the posts or reposts, such as replies/comments. Micro-level networks contain user discussions towards news pieces, which brings auxiliary cues such as sentiment polarities (Gilbert 2014), stance

[1]http://www.journalism.org/2018/09/10/news-use-across-social-media-platforms-2018/

[2]https://en.wikipedia.org/wiki/Pizzagate_conspiracy_theory

[3]https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook

[4]https://www.cjr.org/the_media_today/trust-in-media-down.php

[5]https://www.forbes.com/sites/kenrapoza/2017/02/26/can-fake-news-impact-the-stock-market/#4986a6772fac

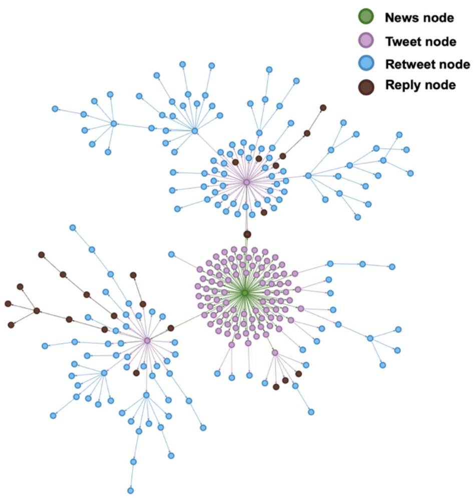


Figure 1: An example of the hierarchical propagation network of a fake news piece fact-checked by Politifact [7]. It consists of two types: **macro-level** and **micro-level**. The macro-level propagation network includes the news nodes, tweet nodes, and retweet nodes. The micro-level propagation network indicates the conversation tree represented by cascade of reply nodes.

signals (Jin et al. 2016a), to differentiate fake news. Studying macro-level and micro-level propagation network provides fine-grained social signals to understand fake news and can facilitate fake news detection. Despite the seminal work in analyzing the macro-level propagation network from temporal or structural perspectives (Vosoughi, Roy, and Aral 2018), no principled study is conducted on characterizing the propagation network from a hierarchical perspective on social media, let alone exploring whether/how these features can help fake news detection. In addition, there is no research that actually provides a deep understanding of (i) how fake news and true news propagate differently from micro-level and macro-level; (ii) whether features extracted from hierarchical propagate networks are useful for fake news detection; and (iii) how discriminative these features are. To give a comprehensive understanding, we investigate the following two research questions:

- **RQ1:** *What are the characteristics of the structure, temporal and linguistic of hierarchical propagation networks of fake and real news?*
- **RQ2:** *How well do the extracted features serve the task of detecting fake news?*

By investigating **RQ1**, we aim to assess whether the propagation network features of fake and real news are different or not from micro-level and macro-level, and to what extent and in what aspects they are different. In addition, by studying **RQ2**, we explore different ways to model propagation network features, analyze the importance of each feature,

[7]https://bit.ly/2H8FnR5

Table 1: The statistics of FakeNewsNet

Platform	PolitiFact	GossipCop
# True news	624	16,817
# Fake news	432	5,323
# True news with propagation network	277	6,945
# Fake news with propagation network	351	3,684
# Users	384,813	739,166
# Tweets	275,058	1,058,330
# Retweets	293,438	530,833
# Replies	125,654	232,923

and show the feature robustness to various learning algorithms. By answering these research questions, we made the following contributions:

- We study a novel problem of understanding the relationships between hierarchical propagation network and fake news, which lays the foundation of exploiting them for fake news detection;
- We propose a principled way to characterize and understand hierarchical propagation network features. We perform a statistical comparative analysis over these features, including micro-level and macro-level, of fake news and true news; and
- We demonstrate the usefulness of the extracted hierarchical network features to classify fake news, whose performance consistently outperforms the existing state-of-the-art methods. We also show that the extracted propagation network features are robust to different learning algorithms, with an average $F1 > 0.84$. We further validate the effectiveness of these features through feature importance analysis and found that temporal and structure features perform better than linguistic features.

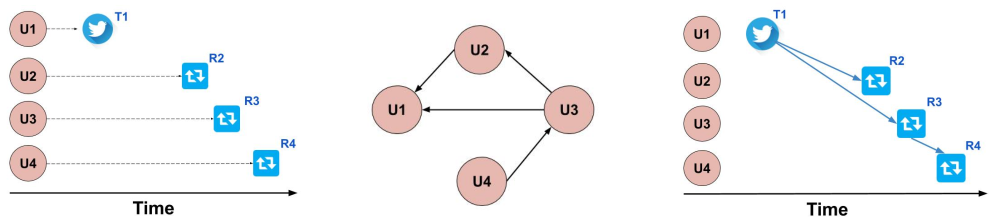
2 Constructing Propagation Networks

In this section, we investigate how to construct the hierarchical propagation networks of news. We explore how we can capture the news spreading process in a propagation network with different granularity such as micro-level and macro-level, which can be further utilized to extract discriminative features from different perspectives for fake news detection.

2.1 Datasets

We utilize public fake news detection data repository FakeNewsNet (Shu et al. 2017). The repository consists of news data related to different fact-checking websites and the correspondent information of news content, social context, and dynamic information.

We use the data from following fact-checking websites: *GossipCop* and *PolitiFact*, both containing news content with labels annotated by professional journalists, social context, and temporal information. News content includes the



(a) Twitter timeline (b) Friendship network (c) Macro-level propagation network

Figure 2: Illustration of macro-level network construction from twitter timeline and friendship network.

meta attributes of the news (e.g., body text), the social context includes the related user social engagements of news items (e.g., user posting/sharing/commenting news on Twitter), and dynamic information includes the timestamps of users' engagements. The detailed statistics of the datasets are shown in Table 1. Next, we introduce how to build hierarchical propagation networks from FakeNewsNet.

2.2 Hierarchical Propagation Networks

The hierarchical propagation network is constructed at different levels of granularity including **micro-level** and **macro-level**. Micro-level networks represent the network of replies where information is shared on the local level. Macro-level networks represent global propagation of information through a cascade of retweets. Through hierarchical propagation network, both local and global pattern of information diffusion related to fake and real news can be studied.

For the macro-level propagation network, nodes represent the tweets and the edges represent the retweet relationship among them. In a macro network, an edge exists from node u to v when a tweet u is retweeted by some user x and node v is created as a result of it. In Twitter, a tweet or a retweet can be retweeted. However, in the retweet data collected from official Twitter API, the retweet of a retweet points to the original tweet. So the retweet network cannot be explicitly constructed from the data available from official Twitter API data. Hence a different strategy using social network (Goel et al. 2015) of the users is used to construct a macro propagation network. For inferring the source of the retweet, we can identify the potential user's friends who retweeted the tweet. For example in Figure 2, $U4$ has a retweet $R4$, and $U3$ has a retweet $R3$; and $U3$ is the only friend that has retweeted earlier than $U4$, so we connect $R3$ and $R4$. In the case when two or more friends have retweeted the tweet, we consider the earliest retweet as the source. For example, $U3$ has both $U1$ and $U2$ as friends and $U1$ has tweeted $T1$ earlier than $U2$'s retweet $R2$, so we connect $T1$ to $R3$. If the timestamp of the user's retweets is greater than the timestamp of the one of the user friend's retweet time stamp, then the user has most likely seen the tweet from one of his/her friends and retweeted it. In a case where immediate retweet from a user's friend is not found, we can consider the retweet is from the original tweet rather than retweet of another retweet.

For the micro-level propagation network, the nodes represent the replies to the tweets posting news and edges represent the relationship among them. In Twitter, a user can reply to actual tweet or reply of another user. In cases where the user replies to the original tweet, then an edge is between tweet posting news and the current node. In case where users reply to the reply of another user, a conversation thread is formed and it is represented as the chain of replies in the propagation path. Users can only reply to the original tweet and not to the retweets and hence micro-level network exists only at post level and not at the retweet nodes of macro-level network.

3 Characterizing Propagation Networks

In this section, we address **RQ1** by performing a analysis of the constructed hierarchical propagation networks for fake news and real news.

3.1 Macro-Level Propagation Network

Macro-level propagation network encompasses information on tweets posting pattern and information sharing pattern. We analyze the macro-level propagation network in terms of structure and temporal aspects. Since the same textual information related to a news article is shared across the macro-level network, linguistic analysis is not applicable.

Structural analysis Structural analysis of macro-level networks helps to understand the global spreading pattern of the news pieces. Existing work has shown that learning latent features from the macro-level propagation paths can help to improve fake news detection while lacking an in-depth understanding of why and how it is helpful (Wu and Liu 2018; Liu and Wu 2018). Thus, we characterize and compare the macro-level propagation networks by looking at various network features as follows.

- $(\mathbf{S_1})$ *Tree depth*: The depth of the macro propagation network, capturing how far the information is spread-/retweeted by users in social media.
- $(\mathbf{S_2})$ *Number of nodes in macro-network*: This indicates the number of users who share the new article and can be a signal for understanding the spreading pattern.
- $(\mathbf{S_3})$ *Maximum Outdegree*: Maximum outdegree in macro network could reveal the tweet/retweet with the most influence in the propagation process.
- $(\mathbf{S_4})$ *Number of cascades*: The number of original tweets posting the original news article.

Table 2: Statistics of structural features for macro propagation network. Stars denotes statistically significant under t-test.

Features	PolitiFact						GossipCop					
	Fake			Real			Fake			Real		
	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg
$\mathbf{S_1}$	2	14	5.93 *	2	13	5.49*	2	12	3.89 *	2	10	3.43*
$\mathbf{S_2}$	2	35,189	774.65*	2	23,494	1,205.46*	2	5339	272.14 *	2	2,497	108.76 *
S_3	0	145	27.37	0	95	31.35	0	198	14.42 *	0	98	12.44 *
S_4	1	17,548	415.59	1	9,577	537.0	1	2568	158.67 *	1	1625	80.19 *
$\mathbf{S_5}$	0	7	1.17*	0	5	1.03*	0	4	1.15 *	0	6	0.94 *
S_6	0	3207	56.79	0	1640	84.55	0	421	18.57*	0	214	3.58*
$\mathbf{S_7}$	0	1	0.16*	0	1	0.08*	0	1	0.15*	0	1	0.05*
S_8	0	2462	63.26	0	1735	77.05	0	461	12.86*	0	125	3.37*
S_9	0.01	0.68	0.23	0.03	0.8	0.21	0.01	0.89	0.29 *	0.01	0.97	0.32*

- ($\mathbf{S_5}$) *Depth of node with maximum outdegree*: The depth at which node with maximum outdegree occurs. This indicates steps of propagation it takes for a news piece to be spread by an influential node whose post is retweeted by more users than any other user's repost.
- ($\mathbf{S_6}$) *Number of cascades with retweets*: It indicate number of tweet cascade those were retweeted at least once.
- ($\mathbf{S_7}$) *Fraction of cascades with retweets*: It indicates the fraction of tweets with retweets among all the cascades.
- ($\mathbf{S_8}$) *Number of bot users retweeting*: It captures the number of bot users among all users retweeting a piece of news.
- ($\mathbf{S_9}$) *Fraction of bot users retweeting*: It is the ratio of bot users among all the users who tweeting and retweeting a news piece. This feature can show whether news pieces are more likely to be disseminated by bots or real humans.

We obtain the aforementioned structural features for macro-level propagation networks of fake news and real news in both Politifact and Gossipcop datasets. As shown in Table 2, we analyze the distribution of structural features and have the following observations:

- The features S_1, S_2, S_5 and S_7 are consistently different from fake and real news in both datasets, under the statistical t-test and bootstrap test. In addition, feature S_3 is statistically different under bootstrap test on both datasets.
- The average depth of the macro-level propagation network (S_1) of fake news is larger than that of real news in both PolitiFact and GossipCop significantly. This shows fake news has a longer chain of retweets than real news, which is consistent with the observation in (Vosoughi, Roy, and Aral 2018).
- Further, the depth of the node with the maximum outdegree (S_5) of fake news is greater than that of real news on both datasets, which indicates fake news takes longer steps to be reposted by an influential user.
- We can see that the fraction of cascades with retweets is larger for macro-level propagation network for fake news than that for real news. It shows that there are more number of tweets posting fake news are retweeted on average.

Temporal analysis The temporal user engagements in macro-level network reveal the frequency and intensity of news dissemination process. The temporal features extracted are interpretable and can provide explainable abilities over existing deep temporal modeling approaches to learn features (Ruchansky, Seo, and Liu 2017; Shu, Mahudeswaran, and Liu) for fake news detection. Following are the features we extracted from the macro propagation network,

- ($\mathbf{T_1}$) *Average time difference between the adjacent retweet nodes*: It indicates how fast the tweets are retweeted in the news dissemination process.
- ($\mathbf{T_2}$) *Time difference between the first tweet and the last retweets*: It captures the life span of the macro-network.
- ($\mathbf{T_3}$) *Time difference between the first tweet and the tweet with maximum outdegree*: Tweets with maximum outdegree in propagation network represent the most influential node. This feature demonstrates how long it took for a news article to be retweeted by the most influential node.
- ($\mathbf{T_4}$) *Time difference between the first and last tweet posting news*: This indicates how long the tweets related to a news article are posted on Twitter.
- ($\mathbf{T_5}$) *Time difference between the tweet posting news and last retweet node in deepest cascade*: Deepest cascade represents the most propagated network in the entire propagation network. This time difference indicates the lifespan of the news in the deepest cascade and can show whether news grows in a bursty or slow manner.
- ($\mathbf{T_6}$) *Average time difference between the adjacent retweet nodes in the deepest cascade*: It indicates how frequent a news article is retweeted in the deepest cascade.
- ($\mathbf{T_7}$) *Average time between tweets posting news*: This time indicates whether tweets are posted in short interval related to a news article.
- ($\mathbf{T_8}$) *Average time difference between the tweet post time and the first retweet time*: The average time difference between the first tweets and the first retweet node in each cascade can indicate how soon the tweets are retweeted.

We compare the temporal features of the macro-level propagation network of fake and real news in Figure 3 (from T_1 to T_8) and have the following observations:

- The temporal features T_2, T_3, T_4, T_7 and T_8 from macro-level are statistically significant between fake and real news, under t-test and bootstrap test.

Table 3: Statistics of structural features for micro propagation network. Stars denote statistically significant under t-test.

Features	PolitiFact						GossipCop					
	Fake			Real			Fake			Real		
	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg
$\mathbf{S_{10}}$	2	6	4.57*	2	6	4.25*	2	6	3.40*	2	6	2.51*
$\mathbf{S_{11}}$	2	21,923	544.91*	2	18,522	853.89*	2	3,453	213.97*	2	1696	90.75*
S_{12}	0	204	24.29	0	210	28.45	0	234	9.82	0	191	4.54
S_{13}	0	1089	26.29*	0	1185	45.33 *	0	401	12.36*	0	145	1.64*
$\mathbf{S_{14}}$	0	1	0.09*	0	1	0.06*	0	1	0.06*	0	1	0.02*

- The time difference between the first tweet and the last retweets (T_2) is smaller for fake news than real news. This indicates that fake news lives shorter than real news on social media on average in our datasets.
- Time difference between the first tweet and tweet with maximum outdegree (T_3) is smaller for fake news than real news in both datasets. It shows that fake news pieces are more likely to be shared earlier by an influential user than real news.
- Further, the time difference between the first and last tweet posting news (T_4) is shorter for fake news. This shows tweets related to fake news are posted in a shorter interval of time and spread faster than real news, which aligns with findings in (Vosoughi, Roy, and Aral 2018).

3.2 Micro-Level Propagation Network

Micro-level propagation networks involve users conversations towards news pieces on social media over time. It contains rich information of user opinions towards news pieces. Next, we introduce how to extract features from micro-level propagation networks from structural, temporal and linguistic perspectives.

Structure analysis : Structural analysis in the micro-level network involves identifying structural patterns in conversation threads of users who express their viewpoints on tweets posted related to news articles.

- ($\mathbf{S_{10}}$) *Tree depth* : Depth of the micro propagation network captures how far is the conversation tree for the tweets spreading a news piece.
- ($\mathbf{S_{11}}$) *Number of nodes*: The number of nodes in the micro-level propagation network indicates the number of comments that are involved. It can measure the popularity of a tweet.
- ($\mathbf{S_{12}}$) *Maximum Outdegree*: This indicates the maximum number of new comments in the chain starting from a particular reply node.
- ($\mathbf{S_{13}}$) *Number of cascade with with micro-level networks*: This features indicates the number of cascades having at least one reply.
- ($\mathbf{S_{14}}$) *Fraction of cascades with micro-level networks*: This feature indicates the fraction of the cascades that have at least one replies among all cascades.

The comparison of structural features for micro-level propagation networks of fake news and real news is demonstrated in Table 3.1. We can see that:

- Structural features S_{10}, S_{11}, and S_{14} are statistically different between fake news and real news in both datasets.
- The micro-level propagation networks of fake news is deeper (S_{10}) than real news significantly under t-test and bootstrap test in both datasets, which is consistent with the observations in macro-level propagation networks revealed previously and in (Vosoughi, Roy, and Aral 2018).
- In addition, the fraction of cascades with micro-level networks (S_{14}) of fake news is greater than that of real news significantly under t-test in both datasets. The reason may be that fake news articles are more likely to be related to controversial and trending topics, which drives more engagements in terms of comments than real news articles.

Temporal analysis Micro-level propagation network depicts users' opinions and emotions through a chain of replies over time. The temporal features extracted from the micro network can help understand exchange of opinions in terms of time. Following are some of the features extracted from the micro-level propagation network,

- ($\mathbf{T_9}$) *Average time difference between adjacent replies in cascade*: It indicates how frequent users reply to one another.
- ($\mathbf{T_{10}}$) *Time difference between the first tweet posting news and first reply node*: It indicates how soon the first reply is posted in response to a tweet posting news.
- ($\mathbf{T_{11}}$) *Time difference between the first tweet posting news and last reply node in micro network*: It indicates how long a conversation tree lasts starting from the tweet posting a new piece.
- ($\mathbf{T_{12}}$) *Average time difference between replies in the deepest cascade*: It indicates how frequent users reply to one another in the deepest cascade.
- ($\mathbf{T_{13}}$) *Time difference between first tweet posting news and last reply node in the deepest cascade*: Indicates the life span of the conversation thread in the deepest cascade of the micro-level network.

The differences in the distribution of temporal features from micro-level networks of fake and real news are visualized in Figure 3 and we make the following observations:

- The temporal features T_9, T_{10} and T_{11} for fake news and real news are statistically significant under t-test and bootstrap test for both datasets.
- The average time difference between adjacent replies T_9 is longer for fake news than real news, and it shows users

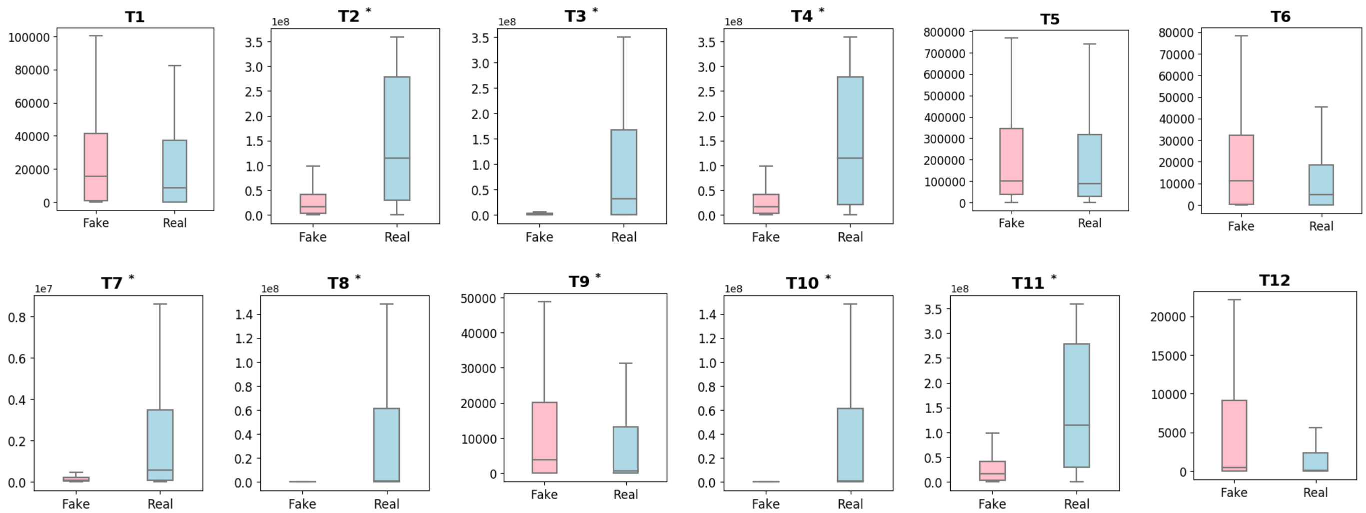


Figure 3: The box plots demonstrating the differences in the distribution of temporal features of fake and real news pieces from PolitiFact dataset. Statistically significant features are represented by asterisk in the feature title. We observe similar patterns in Gossipcop dataset, and we omit the results due to the space limitation.

Table 4: Statistics of linguistic features from micro propagation network. Stars denote statistically significant under t-test.

Features	PolitiFact						GossipCop					
	Fake			Real			Fake			Real		
	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg
L_1	0.0	6.0	1.12	0.0	15.0	1.305	0.0	23.5	1.094*	0.0	20.0	1.010*
$\mathbf{L_2}$	-0.772	0.855	0.007*	-0.585	0.894	0.045*	-0.961	0.9607	0.051*	-0.922	0.969	0.077*
$\mathbf{L_3}$	-0.693	0.855	-0.001*	-0.885	0.894	0.0428 *	-0.961	0.961	0.046*	-0.921	0.969	0.074*
L_4	-0.811	0.879	0.0156	-0.934	0.664	0.027	-0.960	0.938	0.044*	-0.893	0.969	0.066*
L_5	-0.811	0.879	0.011	-0.934	0.851	0.028	-0.961	0.993	0.039*	-0.896	0.969	0.062*

take a longer time to respond to each other. The time difference between the tweet and the first reply T_{10} is shorter for fake news, which may indicate that users take less time to reply to tweets related fake and it takes more time to reply to another users comments.

Linguistic analysis People express their emotions or opinions towards fake news through social media posts, such as skeptical opinions, sensational reactions, etc. These textual information has shown to be related to the content of original news pieces. Thus, it is necessary to extract linguistic-based features to help find potential fake news via reactions from the general public as expressed in comments from micro-level propagation network. We demonstrate the sentiment features extracted from the comments, as the representative of linguistic features. We utilize the widely-used pre-trained model VADER (Gilbert 2014) to predict the sentiment score for each user reply, and extract a set of features related to sentiment as follows,

- ($\mathbf{L_1}$) *Sentiment ratio*: We consider a ratio of the number of replies with positive sentiment to the number of replies with negative sentiment as a feature for each news articles because it helps to understand whether fake news gets more number of positive or negative comments.
- ($\mathbf{L_2}$) *Average sentiment*: Average sentiment scores of the nodes in the micro-level propagation network. Sentiment ratio does not capture the relative difference in the scores of the sentiment and hence average sentiment is used.
- ($\mathbf{L_3}$) *Average sentiment of first level replies*: This indicates whether people post positive or negative comments on the immediate tweets posts sharing fake and real news.
- ($\mathbf{L_4}$) *Average sentiment of replies in deepest cascade*: Deepest cascade generally indicate the nodes that are most propagated cascade in the entire propagation network. The average sentiment of the replies in the deepest cascade capture the emotion of user comments in the most influential information cascade.
- ($\mathbf{L_5}$) *Sentiment of first level reply in the deepest cascade*: The sentiment of the first level reply indicates the user emotions to most influential information cascade.

We obtain the aforementioned linguistic features for micro-level propagation networks of fake news and real news in both Politifact and Gossipcop datasets. As shown in Table 4, we analyze the distribution of linguistic features and have the following observations:

- The linguistic features L_2, and L_3 are statistically significantly different for fake news and real news in both

Table 5: Best Performance comparison for fake news detection with different feature representations

Datasets	Metric	RST	LIWC	STFN	GCNFN	HPFN	RST_HPFN	LIWC_HPFN	STFN_HPFN
PolitiFact	Accuracy	0.796	0.830	0.649	0.837	0.843	0.875	0.872	0.856
	Precision	0.821	0.855	0.605	0.880	0.835	0.873	0.869	0.809
	Recall	0.752	0.792	0.836	0.785	0.851	0.876	0.872	0.927
	F1	0.785	0.822	0.702	0.830	0.843	0.875	0.871	0.864
GossipCop	Accuracy	0.600	0.725	0.796	0.798	0.861	0.861	0.869	0.863
	Precision	0.623	0.773	0.812	0.784	0.854	0.850	0.856	0.857
	Recall	0.596	0.637	0.770	0.823	0.869	0.876	0.887	0.871
	F1	0.614	0.698	0.791	0.803	0.862	0.863	0.871	0.864

datasets under t-test and bootstrap test.

- The average sentiment of replies (L_2) is lower for fake news than real news in both the datasets. It shows that tweets related to fake news receive more negative sentiment comments over real news. A similar result is observed in the sentiment of comments posted directly to tweets captured by feature L_3.

4 Evaluating Propagation Features

In this section, we address **RQ2**. We explore whether the hierarchical propagation network features can help improve fake news detection, and how we can build effective models based on them. Moreover, we perform feature importance and model robustness analysis. We first introduce how to represent the hierarchical propagation network features $\mathbf{f}_i$ for a news item a_i. Let $\mathcal{G}_i$ denote the temporal propagation network of news piece a_i. For $\mathcal{G}_i$, we extract all types of propagation features and concatenate them into one feature vector $\mathbf{f}_i$. We also denote the proposed *H*ierarchical *P*ropagation *N*etwork *F*eature vector $\mathbf{f}_i$ as HPFN.

4.1 Experimental Settings

To evaluate the performance of fake news detection, we use the following metrics, which are commonly used to evaluate classifiers in related areas: Accuracy (Acc), Precision (Prec), Recall (Rec), and F1. We randomly choose 80% of news pieces for training and the remaining 20% for testing, and the process is performed for 5 times and the average performance is reported. The details of baseline feature representations with the feature dimensions are given as below[8]:

- **RST** (37 features): RST (Ji and Eisenstein 2014) can capture the writing style of a document by extracting the relations from rhetorical structure theory systematically. It learns a transformation from a bag-of-words surface representation into a latent feature representation [9].
- **LIWC** (93 features): LIWC (Pennebaker et al. 2015) extracts lexicons that fall into different psycholinguistic categories, and learn a feature vector through multiple measures for each document [10].
- **STFN** (8 features): STFN includes the structural and temporal features proposed in (Vosoughi, Roy, and Aral 2018) for macro-level propagation network, i.e., tree height, number of nodes, max breadth of the tree, fraction of unique users, time taken to reach depth 1 in propagation, time taken to reach depth of 2 in propagation, number of unique users within level 1 and number of unique users within level 3 of propagation.
- **GCNFN** (200 node-level features): GCNFN (Monti et al. 2019) utilizes deep geometric learning such as graph convolutional neural networks to model propagation network along with textual node embedding features for fake news detection. It includes two graph convolution layers, two fully connected layers, and a softmax layer for prediction.
- **RST_HPFN** (69 features): RST_HPFN represents the concatenated features of RST and HPFN, which includes features extracted from both news content and hierarchical propagation network.
- **LIWC_HPFN** (125 features): LIWC_HPFN represents the concatenated features of LIWC and HPFN, which includes features extracted from both news content and hierarchical propagation network.
- **STFN_HPFN** (40 features): STFN_HPFN represents the concatenated features of STFN and HPFN, which includes features structural and temporal features discussed in STFN and hierarchical propagation network features.

Note that for a fair and comprehensive comparison, we choose the above feature extraction methods from following aspects: (1) **news content**, such as RST and LIWC; and (2) **propagation network**, such as Structure and Temporal features for Fake News Detection (STFN) and GCNFN. We also combine RST, LIWC and STFN feature with HPFN to further explore if HPFN provides complementary information. For a fair comparison, we use the classifier that performs best on each feature set and compare the effectiveness of these different feature representations.

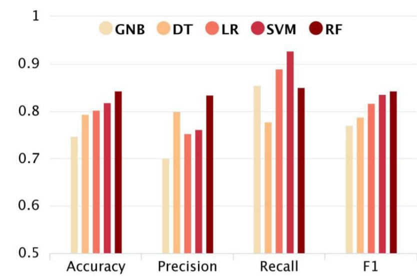
4.2 Fake News Detection Performance Comparison

We test the baseline features on different learning algorithms and choose the one that achieves the best performance (see Table 5). The algorithms include Gaussian Naive Bayes (GNB), Decision Tree (DT), Logistic Regression (LR), and Random Forest (RF). To ensure a fair comparison of the proposed features and baseline features, we ran all the algorithms using default parameter settings of scikit-learn. The

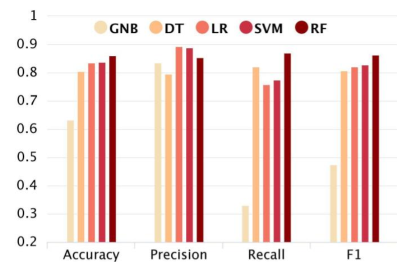
[8]all code and data are available at: http://tiny.cc/ixor6y

[9]The code is available at: https://github.com/jiyfeng/DPLP

[10]The software and description of measures are available at: http://liwc.wpengine.com/



(a) PolitiFact dataset



(b) GossipCop dataset

Figure 4: Detection Performance for HPFN with Different Learning Algorithms

Table 6: Best Detection Performance with Different Group of Features from HPFN

Datasets	Level	Acc	Prec	Rec	F1
PolitiFact	Micro	0.834	0.823	0.847	0.835
	Macro	0.807	0.816	0.789	0.802
	All	**0.843**	**0.835**	**0.851**	**0.843**
GossipCop	Micro	0.843	0.841	0.845	0.843
	Macro	0.852	0.841	0.868	0.854
	All	**0.861**	**0.854**	**0.869**	**0.862**

Table 7: Best Detection Performance with Different Group of Features from HPFN

Datasets	Type	Acc	Prec	Rec	F1
PolitiFact	Structural	0.681	0.681	0.672	0.676
	Temporal	0.793	0.716	**0.963**	0.821
	Linguistic	0.659	0.648	0.683	0.665
	All	**0.843**	**0.835**	0.851	**0.843**
GossipCop	Structural	0.826	0.828	0.823	0.826
	Temporal	0.826	0.827	0.825	0.826
	Linguistic	0.578	0.594	0.491	0.538
	All	**0.861**	**0.854**	**0.869**	**0.862**

experimental results are shown in Table 5. We have the following observations:

- For news content-based methods, we see that LIWC performs better than RST. This indicates that the LIWC vocabulary can better capture the deceptiveness in news content, which reveals that fake news pieces are very different from real news in terms of word choice from psychometrics perspectives.
- Our proposed HPFN can achieve the best performance in both datasets on most of the metrics compared with all other baseline methods. This shows that the extracted features from macro-level and micro-level propagation networks can help improve fake news detection significantly.
- For propagation network-based methods, we can see that HPFN performs better than STFN and GCNFN consistently in both datasets. This is because HPFN includes features from both macro and micro-level networks from structural, temporal and linguistic perspectives that are useful for fake news detection. STFN and GCNFN only encodes some features from the macro-level network.
- In addition, we observe that by combining HPFN features with existing features can further improve the detection performances. For example, RST_HPFN performs better than either RST or HPFN, which reveals that they are extracted from orthogonal information spaces, i.e., RST features are extracted from news content and HPFN features from hierarchical propagation network on social media, and have complementary information to help fake news detection. We have similar observations for other features, i.e., (LIWC_HPFN >LIWC, LIWC_HPFN>HPFN) and (STFN_HPFN>HPFN, STFN_HPFN>STFN).

We further evaluate the robustness of the extracted features HPFN. We demonstrate the fake news detection performances using different classifiers (see Figure 4). These algorithms have different learning biases, and thus their performance is often different for the same task. While we observe that: (1) RF achieves the best overall performance on both datasets; and (2) while the performance of RF is slightly better than other learning algorithms, the results are not significantly different across algorithms. This demonstrates that when sufficient information is available in the hierarchical propagation network features and so the performance is not very sensitive to the choice of learning algorithms.

4.3 Feature Importance Analysis

In this subsection, we analyze the importance of the features in different granular levels to understand how each type of features contributes to the detection performance. We analyze feature importance in the Random Forest by computing a feature importance score based on the Gini impurity [11].

First, we evaluate the fake news detection performance on different levels of hierarchical propagation network including a) Micro-level; b) Macro-level; and c) both micro-level and macro-level (All) and compare their contributions to fake news detection in table 6. We have the following observations: (i) The combination of micro-level and macro-level features can achieve better performance than either micro-level or macro-level features in both datasets consistently. This shows that features from different levels provide complementary information in feature dimension and thus help fake news detection; (ii) In general, we observe that micro-level features can achieve good performance, which demonstrates the necessitates of exploring micro-level features; (iii) Compared with macro-level features, micro-level

[11] https://bit.ly/2T1j29K

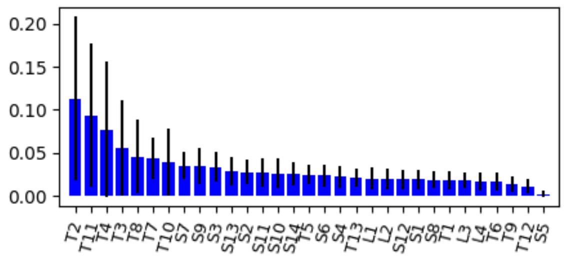


Figure 5: Feature importance in Politifact dataset

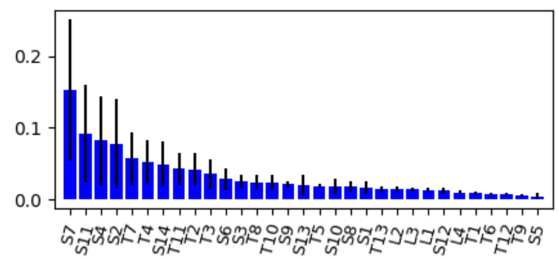


Figure 6: Feature importance in GossipCop dataset

features may not always perform better. For example, features from micro-level networks perform better than that from macro-level networks in PolitiFact dataset, and vice versa for GossipCop dataset.

Next, we evaluate the performance of different types of features from hierarchical propagation network including a) Structural; b) Temporal; c) Linguistic and d) combination of structural, temporal and linguistic (All), and compare their classification performance in Table 7. We have the following observations i) Temporal features perform better than both structural and linguistic features in both the datasets and this shows that temporal features have more importance in the classification task; ii) Structural features performs better than linguistic features in both the datasets as the micro-level network have limited linguistic contents; and iii) When the features from all perspectives are considered, the classification performance is better than considering either of the three features and this shows the features have complementary information to differentiate fake news from real news.

From the Figure 5, we can observe that: (i) the temporal features of PolitiFact dataset have higher importance scores over the structural and linguistic features; (ii) The feature T_{11} shows that lifespan of the engagements in the micro-level network is the most important feature in fake news classification. Similarly, the life span of news in the macro network captured by T_2 shows the second-highest importance score. This indicates that the longevity of fake and real news in social media is different; and (iii) Among structural features extracted from Politifact, the maximum out-degree of the macro network S_3 has more importance than other structural features. Figure 6 demonstrates the feature importance results on Gossipcop dataset. We make the following observations: (i) the fraction of cascades with retweets S_7 in the macro network has the highest importance score. This shows difference in the scale of spreading scope of fake and real news; (ii) In addition, the number of cascades in macro network S_4 has the second-highest importance score; and (iii) the time difference between the first and last tweet posting news T_4 has higher importance score among temporal features. This confirms our findings that fake news tends to spread shortly on social media than real news.

4.4 Early Fake News Detection

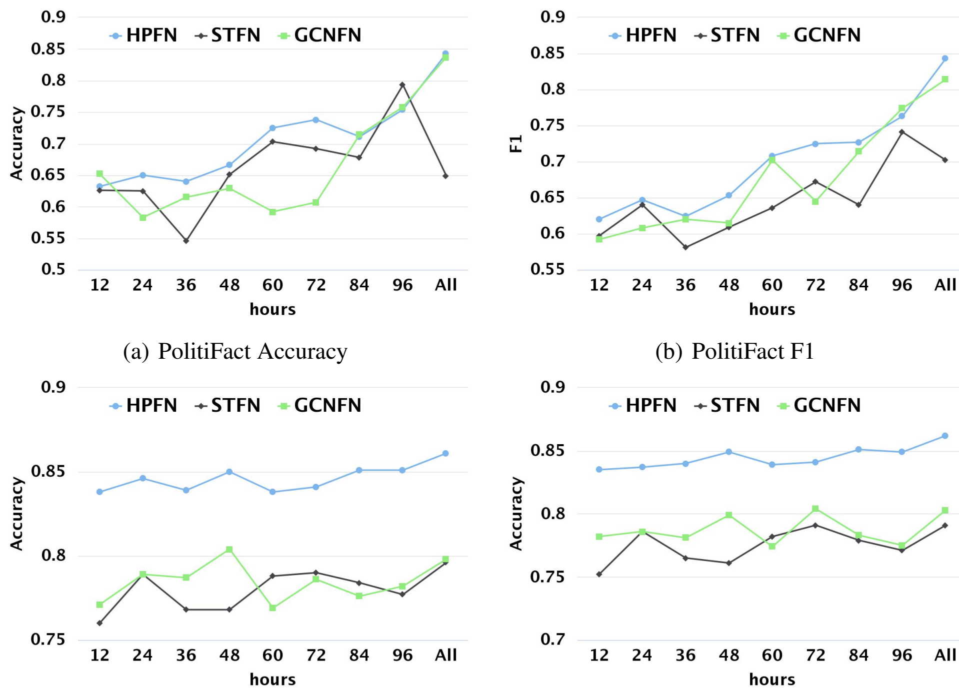
Early detection of fake news is very desirable to restrict the dissemination scope of fake news and prevent its further propagation. It aims to give an early alert of fake news, by only considering the limited social context within a specific range of time delay of original news posted. To evaluate the effectiveness of the framework for early fake news detection, we vary the delay time T as $[12, 24, 36, 48, 60, 72, 84, 96]$ hours and only use the social context within T. From the figure, we have the following observations: (i) the proposed HPFN features can achieve good performance of fake news detection even in the early stage, in terms of both Accuracy and F1, compared with GCNFN and STFN. This indicates the effectiveness of utilizing features from structural, temporal, and linguistic perspectives of hierarchical propagation networks; (ii) For methods that model macro-level networks, we can see that GCNFN>STFN holds on both datasets. It may be because GCNFN can better capture the non-linearity of network structure through the graph convolution operations, while STFN only utilizes several raw features from the propagation networks; and (iii) We can see that in the very early stage (e.g., 12, and 24 hours), the performances of HPFN and GCNFN are similar. This may because the propagation networks are mostly macro-level networks, and GCNFN is powerful on modeling the textual and structure information in macro-level networks.

5 Related Work

In this section, we introduce the related from two-folds: fake news detection and fake news propagation.

5.1 Fake News Detection

Fake news detection approaches generally fall into two categories: (1) using *news content*; and (2) using *social contexts* (Shu et al. 2017; Zafarani et al. 2019). For news content based approaches, features are extracted as linguistic-based such as writing styles (Potthast et al. 2017), and visual-based such as fake images (Gupta et al. 2013). Linguistic-based features capture specific writing styles and sensational headlines that commonly occur in fake news content (Potthast et al. 2017), such as lexical and syntactic features. Visual features are extracted from visual elements (e.g. images and videos) to capture the different characteristics for fake news (Jin et al. 2016b). News content based models include i) knowledge-based: using external sources to fact-checking claims in news content (Magdy and Wanas 2010; Wu et al. 2014), and 2) style-based: capturing the manipulators in writing style, such as deception (Rubin and Lukoianova 2015) and non-objectivity (Potthast et al. 2017).



(a) PolitiFact Accuracy

(b) PolitiFact F1

(c) GossipCop Accuracy

(d) GossipCop F1

Figure 7: The Performance on early detection of fake news.

In addition, latent textual representations are modeled using tensor factorization (Hosseinimotlagh and Papalexakis 2018), deep neural networks (Wang 2017; Karimi and Tang 2019), which achieve good performance to detect fake news with news contents.

Different from content-based approaches, social context-based approaches incorporate features from social media user profiles, post contents, and social networks (Shu and Liu 2019). User features can measure users' characteristics and credibilities (Castillo, Mendoza, and Poblete 2011). Post features represent users' social responses, such as stances (Jin et al. 2016a). Network features are extracted by constructing specific social networks, such as diffusion networks (Kwon et al. 2013) or co-occurrence networks (Ruchansky, Seo, and Liu 2017). Most of these social context models can basically be grouped as either stance-based or propagation-based. Stance-based models utilize users' comments, sentiments or opinions towards the news to infer news veracity (Jin et al. 2016a; Shu et al. 2019). Propagation-based models apply propagation methods to model unique patterns of information spread such as the interactions among publishers, news pieces, and consumers (Shu, Wang, and Liu 2019). Recently, research also focuses on challenging problems of fake news detection, such as fake news early detection by adversarial learning (Wang et al. 2018) and user response generating (Qian et al. 2018), semi-supervised detection (Guacho et al. 2018) and unsupervised detection (Yang et al. 2019).

Existing approaches that exploit user social engagements simply extract features to train classifiers without a deep understanding of these features, which makes it a black-box that is difficult to interpret. Thus, we perform, to our best knowledge, the first in-depth investigation of various aspects of hierarchical propagation network for their usefulness for fake news detection. Early detection is necessary for practical setting and some studies on early detection (Varol et al. 2017; Gupta et al. 2014) show some promising early results. Thus, we also explore the capacity of the proposed HPFN features for early fake news detection.

5.2 Fake News Propagation

Diffusion-based models typically focus on modeling how fake news spreads/diffuses on social media. One recent work (Vosoughi, Roy, and Aral 2018) analyzes the diffusion of falsehoods (rumors) and truth on Twitter, which reveals that falsehoods tend to diffuse faster than truthful claims on social networks. Consequently, several works follow up on this line of work to comprehensively characterize the nature of fake news diffusion and dissemination (Bovet and Makse 2018; Wang, Pang, and Pavlou 2018). Babcock *et al.* (Babcock, Cox, and Kumar) study the user community structures of different types of fake news during the propagation process. (Shao et al. 2017) analyze the role of social bots in the diffusion of low credibility content and suggest that such automated accounts are particularly active in disseminating low-credibility content before the content becomes viral. Similarly (Babcock, Cox, and Kumar) note that not all fake news is the same and its effect on campaigns and communities differ. Consequently, they explore the reactions of different communities to fake news conversations on Twitter.

Moreover, rumor propagation has shown different characteristics from fake news. For example, (Vosoughi, Roy, and Aral 2018) reveals that fake news spread faster and deeper than true news, while in (Friggeri et al. 2014) that for rumors covered by Snopes, it was the true ones that ran deeper.

Existing fake news propagation work mainly focuses on analyzing the macro-level propagation and does on perform an in-depth study on utilizing various propagation network features for fake news detection. To fill this gap, we construct a hierarchical propagation network from both macro and micro levels and exploit the features from structural, temporal and linguistic perspectives for fake news detection.

6 Conclusion and Future Work

In this paper, we aim to answer questions regarding the correlation between hierarchical propagation networks and fake news and provide a solution to utilize features from different perspectives from hierarchical propagation networks for fake news detection. Now we, summarize our findings of each research question and discuss the future work.

RQ1 *What are the characteristics of the structure, temporal and linguistic of hierarchical propagation networks of fake and real news?* To perform this study, we first construct the hierarchical propagation networks from macro-level and micro-level. For each type of network, we extract various features from structural, temporal and linguistic perspectives for fake news and real news. We compare these features to see if they are different or not for fake and real news with statistical analysis.

RQ2 *How well do the extracted features serve the task of detecting fake news?* With the quantitative analysis of news hierarchical propagation network features, we build different learning algorithms to detect fake news. We evaluate the effectiveness of the extracted features by comparing with several existing baselines. The experiments show that: (1) these features can make significant contributions to help detect fake news; (2) these features are overall robust to different learning algorithms; and (3) temporal features are more discriminative than linguistic and structural features and macro and micro-level features are complimentary.

This work opens up the doors for many areas of research. First, we can learn to predict whether a user will spread a fake news piece or not by studying the structures of hierarchical propagation networks, which is a precursor for mitigating fake news dissemination. Second, we can exploit the hierarchical structure of propagation networks to perform unsupervised fake news detection. Third, we may combine the extracted explicit propagation network features with deep learning models to further boost the performance of fake news detection.

Ackowledgments

This material is in part supported by the NSF awards #1909555 and #1614576.

References

Allcott, H., and Gentzkow, M. 2017. Social media and fake news in the 2016 election. Technical report, National Bureau of Economic Research.

Babcock, M.; Cox, R. A. V.; and Kumar, S. Which fake fits? how the diffusion of disinformation varies by campaign type and community in.

Bovet, A., and Makse, H. A. 2018. Influence of fake news in twitter during the 2016 us presidential election. *arXiv preprint arXiv:1803.08491*.

Castillo, C.; Mendoza, M.; and Poblete, B. 2011. Information credibility on twitter. In *WWW*.

Friggeri, A.; Adamic, L.; Eckles, D.; and Cheng, J. 2014. Rumor cascades. In *ICWSM*.

Gilbert, C. H. E. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text.

Goel, S.; Anderson, A.; Hofman, J.; and Watts, D. J. 2015. The structural virality of online diffusion. *Management Science* 62(1):180–196.

Guacho, G. B.; Abdali, S.; Shah, N.; and Papalexakis, E. E. 2018. Semi-supervised content-based detection of misinformation via tensor embeddings. In *ASONAM*, 322–325. IEEE.

Gupta, A.; Lamba, H.; Kumaraguru, P.; and Joshi, A. 2013. Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy. In *WWW*.

Gupta, A.; Kumaraguru, P.; Castillo, C.; and Meier, P. 2014. Tweetcred: Real-time credibility assessment of content on twitter. In *International Conference on Social Informatics*, 228–243. Springer.

Hosseinimotlagh, S., and Papalexakis, E. E. 2018. Unsupervised content-based identification of fake news articles with tensor decomposition ensembles. In *Proceedings of the Workshop on Misinformation and Misbehavior Mining on the Web (MIS2)*.

Ji, Y., and Eisenstein, J. 2014. Representation learning for text-level discourse parsing. In *ACL'2014*, volume 1, 13–24.

Jin, Z.; Cao, J.; Zhang, Y.; and Luo, J. 2016a. News verification by exploiting conflicting social viewpoints in microblogs. In *AAAI*, 2972–2978.

Jin, Z.; Cao, J.; Zhang, Y.; Zhou, J.; and Tian, Q. 2016b. Novel visual and statistical image features for microblogs news verification. *IEEE transactions on multimedia* 19(3):598–608.

Karimi, H., and Tang, J. 2019. Learning hierarchical discourse-level structure for fake news detection. *arXiv preprint arXiv:1903.07389*.

Kwon, S.; Cha, M.; Jung, K.; Chen, W.; and Wang, Y. 2013. Prominent features of rumor propagation in online social media. In *ICDM*, 1103–1108. IEEE.

Liu, Y., and Wu, Y.-F. B. 2018. Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks. In *AAAI*.

Magdy, A., and Wanas, N. 2010. Web-based statistical fact checking of textual documents. In *Proceedings of the 2nd international workshop on Search and mining user-generated contents*, 103–110. ACM.

Monti, F.; Frasca, F.; Eynard, D.; Mannion, D.; and Bronstein, M. M. 2019. Fake news detection on social media using geometric deep learning. *arXiv preprint arXiv:1902.06673*.

Pennebaker, J. W.; Boyd, R. L.; Jordan, K.; and Blackburn, K. 2015. The development and psychometric properties of liwc2015. Technical report.

Potthast, M.; Kiesel, J.; Reinartz, K.; Bevendorff, J.; and Stein, B. 2017. A stylometric inquiry into hyperpartisan and fake news. *arXiv preprint arXiv:1702.05638*.

Qian, F.; Gong, C.; Sharma, K.; and Liu, Y. 2018. Neural user response generator: Fake news detection with collective user intelligence. In *IJCAI*, 3834–3840.

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

Ruchansky, N.; Seo, S.; and Liu, Y. 2017. Csi: A hybrid deep model for fake news detection. In *CIKM*.

Shao, C.; Ciampaglia, G. L.; Varol, O.; Yang, K.; Flammini, A.; and Menczer, F. 2017. The spread of low-credibility content by social bots. *arXiv preprint arXiv:1707.07592*.

Shu, K., and Liu, H. 2019. Detecting fake news on social media. *Synthesis Lectures on Data Mining and Knowledge Discovery* 11(3):1–129.

Shu, K.; Sliva, A.; Wang, S.; Tang, J.; and Liu, H. 2017. Fake news detection on social media: A data mining perspective. *KDD exploration newsletter*.

Shu, K.; Cui, L.; Wang, S.; Lee, D.; and Liu, H. 2019. defend: Explainable fake news detection. In *KDD*, 395–405.

Shu, K.; Mahudeswaran, D.; and Liu, H. Fakenewstracker: a tool for fake news collection, detection, and visualization. *Computational and Mathematical Organization Theory* 1–12.

Shu, K.; Wang, S.; and Liu, H. 2019. Beyond news contents: The role of social context for fake news detection. In *WSDM*, 312–320.

Varol, O.; Ferrara, E.; Menczer, F.; and Flammini, A. 2017. Early detection of promoted campaigns on social media. *EPJ Data Science* 6(1):13.

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

Wang, Y.; Ma, F.; Jin, Z.; Yuan, Y.; Xun, G.; Jha, K.; Su, L.; and Gao, J. 2018. Eann: Event adversarial neural networks for multi-modal fake news detection. In *KDD*, 849–857.

Wang, S.; Pang, M.-S.; and Pavlou, P. A. 2018. 'cure or poison?'identity verification and the spread of fake news on social media. *Identity Verification and the Spread of Fake News on Social Media (September 14, 2018)*.

Wang, W. Y. 2017. "liar, liar pants on fire": A new benchmark dataset for fake news detection. *arXiv preprint arXiv:1705.00648*.

Wu, L., and Liu, H. 2018. Tracing fake-news footprints: Characterizing social media messages by how they propagate. In *WSDM*, 637–645. ACM.

Wu, Y.; Agarwal, P. K.; Li, C.; Yang, J.; and Yu, C. 2014. Toward computational fact-checking. *VLDB* 7(7):589–600.

Yang, S.; Shu, K.; Wang, S.; Gu, R.; Wu, F.; and Liu, H. 2019. Unsupervised fake news detection on social media: A generative approach. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, 5644–5651.

Zafarani, R.; Zhou, X.; Shu, K.; and Liu, H. 2019. Fake news research: Theories, detection strategies, and open problems. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 3207–3208. ACM.