Temporal Fake News Detection

Anonymous Author(s)*

ABSTRACT

this is the abstract

CCS CONCEPTS

• **Computer systems organization** → **Embedded systems**; *Redundancy*; Robotics; • **Networks** → Network reliability.

KEYWORDS

datasets, neural networks, gaze detection, text tagging

ACM Reference Format:

Anonymous Author(s). 2018. Temporal Fake News Detection. In *Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/1122445.1122456

As the popularity of social media platforms continuously grows, so does the dissemination of online information, therefore many deep learning systems have been developed in order to detect false or biased news [37, 61, 69]. While fake news detection is a big step to mitigate the impact of misinformation on our society [70], it is not enough, given that after sharing the information, it is challenging to confine its spread on the web and avoid it's catastrophic effects [10, 33, 65]. Research shows that fact corrections frequently fail in reducing peoples misconception of the truth and occasionally they even have a "backfiring" effect where people's misconception is reinforced [6, 41, 48, 62]. Due to the severe impact of misinformation to our society [32], it is becoming essential to explore ways to prevent the rapid spread of misinformation and address this issue on its origin. This means that there is an emerging need to efficiently identify misinformation spreaders and spurious accounts that are likely to propagate posts from handles of unreliable news sources. To this end, we introduce a model that distinguishes authors that have shared news from unreliable sources in the past from those that, to the best of our knowledge, share news only from reliable sources. We use the terms 'misinformation spreaders' and 'fact checkers' for each class of users respectively [15].

Evidence from previous literature [15, 38, 42] show that misinformation spreaders and fact checkers have certain linguistic and personality features that distinguish them from each other. [24] use a graph-based approach to model the linguistic relationships between texts for fake news detection. Moreover, [13, 44, 45, 53] explore the social network of misinformation spreaders and fact checkers and show that interpersonal trust plays a significant role in differentiating them. Recently, significant attention has garnered towards Graph Neural Networks [64] due to the advances of graph representational learning [66] in various natural language processing domains. Such approaches are able to model user-to-user relationships and therefore provide a promising unexplored research direction for identifying misinformation spreaders.

Recent work has shown that various temporal patterns have been utilized as features for fake news detection [14, 54] and fact-checking [2]. More specifically, [28] show that most of the predictive features of rumors change depending on the observation window and find that structural and temporal features distinguish rumors from non-rumors over a long-term window.[40] extract temporal features of the users' posting activity and show that the temporal features are more discriminative than the static features for fake news detection. [22] examined the impact of time on fake news detection models and showed that the content-based differences of news sources change over time because of the highly dynamic nature of the news topics. This means that most of the fake news detection methods that use non-temporal features would need to be continuously updated with new annotated data in order to stay relevant. We argue that this hypothesis can be generalized for profiling misinformation spreaders. Another gap in existing works on temporal modeling is that they are not specifically designed for learning the time-evolving similarities of the users' content and social interactions. While the topics discussed may change over time (and therefore the users' vocabulary usage), social science suggest that the users' connections tend to follow the same patterns [51]. Building on the limitations of existing research and accounting for the temporal dynamics of user-to-user relationships, we introduce a model that extracts features from the users' content similarities and social interactions and models the temporal evolution of these connections.

Our study aims to answer the following research questions:

- **RQ1** Can we identify misinformation spreaders based on their temporal user-to-user relationships? If so, which types of relationships are the most predictive and which have the ability to generalize on unseen users?
- **RQ2** How do the users' temporal semantic similarities and social interactions fluctuate with respect to the political landscape? Can we detect temporal relationship patterns while accounting for the users' credibility and political bias?

To answer these research questions and address the limitations of existing approaches, we formulate the problem as a binary classification task. First, we construct dynamic graphs by utilizing the content similarities and user interactions in the post thread. Each of these dynamic graphs are constructed based on the users' posting behaviour within consecutive windows of time. Following, the generated temporal graph representations are treated as a sequence of features for the final classification. We conduct a series of exploratory analyses in the user-to-user relationships. Though ablative experiments we show the effectiveness of each model's component for profiling misinformation spreaders. The contributions of this paper can be summarised as follows:

Woodstock '18, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00

https://doi.org/10.1145/1122445.1122456

- We create a dataset that contains a set of roughly 3.6M Reddit posts published by two different groups of users; users that share news posts only from verified news sources (fact checkers) and users that have shared news posts form non-credible sources (misinformation spreaders). The dataset also contains fine-grained labels about the users' factuality level and political bias.
- We extract the users' temporal semantic and social features and analyze the different types of temporal connections.
- We present a temporal Graph Attention Network model that generates features to capture the dynamic nature of the users' content similarities and social interactions.

1 BACKGROUND AND RELATED WORK

Conventional methods. In the recent years, misinformation detection has received a lot of research attention [27], and while user modelling approaches have been used for various tasks such as author profiling [47] and bot detection [7, 23, 26]. However, it wasn't until after the PAN 2020 competition that the problem of misinformation spreaders identification gained significant attention from the research community, therefore this research area is ripe for exploration. Most recent studies are focused on analyzing personality and linguistic patterns [15, 38] and emotional signals [16]. These methods rely on the fact that the content, and therefore the features that are extracted, remains somewhat constant over time. While static linguistic patterns have proven to be useful features for misinformation spreader detection, none of the current methods explore temporal features. In addition, these methods ignore other relevant signals such as the linguistic or social relationships of the users which are more reliant feature predictors in the long term [43, 51]. Our model utilizes the users' linguistic patterns as user feature representations and simultaneously leverages their content similarities and social interactions dynamically.

Graph-based approaches. In the context of modelling the users' behavior, graph representational learning approaches [64] have made significant advances in enhancing NLP models for abusive language detection [35], suicide ideation detection [17, 36, 59], hate speech [12] among others. Several graph-based approaches have been explored in the domain of misinformation detection, [9] propose a relational GNN that models the user's community, [20] propose a hybrid approach that leverages both the users' posting history as well as their social connections. In terms of identifying misinformation spreaders [44, 46] utilize features extracted from the network structure which is build based on interpersonal trust metrics [53]. Despite their success, a limitation of the existing approaches is that they do not account for the temporal characteristics of the linguistic and social connections.

We argue that the users' attributes change dynamically over time due to the dynamic nature of the news cycle, therefore spatial-temporal graphs are more suitable to model the evolution of the user-to-user relationships [67]. The concept of spatial-temporal graphs has been around for some years [21, 52, 60] with numerous applications in traffic flow forecasting [18, 30], action recognition [68] and stock trading [55, 56, 58]. [57] propose an architecture that leverages temporal signals from financial data, social media, and inter-stock relationships via a graph neural network in a hierarchical temporal fashion. We draw inspiration from these approaches and propose a dynamic temporal graph model with attention mechanism for misinformation spreader detection.

2 METHODOLOGY

2.1 Problem Formulation

We formulate the author profiling problem as a binary classification task to predict the class of the user y^i, where $y^i \in \{$misinformation spreader, fact checker$\}$. We denote the user to be classified as a misinformation spreader or not as $u^i \in \mathcal{U} = \{u^1, u^2, \ldots, u^N\}$. Each user u^i is associated with a posting history $\mathcal{H}^i = \{(p_1^i, t_1^i), (p_2^i, t_2^i), \ldots, (p_{L^i}^i, t_{L^i}^i)\}$ where p_k^i is a historic text authored by the user u^i, posted at time t_k^i where $t_1^i < t_2^i < \cdots < t_{L^i}^i$ and L^i is the individual posting history length of each user u^i. The complete posting time period of all users is defined as Ω. We partition Ω in T equal discrete time frames, therefore each time frame $\omega^\tau = [t_\mu^\tau, t_\nu^\tau]$ with $\tau = 1, \ldots, T$ contains the users' historical posts that were posted within that time period. We denote the time frame ω^τ as τ for brevity. Thus, each user's complete posting history is sliced in T (not necessarily equal) parts, so that $\mathcal{H}^i = \{H_1^i, \ldots, H_T^i\}$, where $H_\tau^i = \{(p_k^i, t_k^i), (p_{k+1}^i, t_{k+1}^i), \ldots, (p_{k+\ell_\tau^i}^i, t_{k+\ell_\tau^i}^i)\}$ is the posting history of the user u^i within the time frame τ with $t_\mu < t_k^i < \cdots < t_{k+\ell_i}^i < t_\nu$ and ℓ_τ^i being u^i's posting history length within the time frame τ. Note that if a user u^i has no posts within a time period, τ then $H_\tau^i = \emptyset$.

2.2 Encoding Users

Encoding individual posts. We use Sentence-BERT (SBERT) [49] in order to encode each user's individual historical posts p_k^i. SBERT is a modification of BERT that is specifically designed to produce semantically meaningful sentence embeddings and has been evaluated on seven Semantic Textual Similarity (STS) tasks, achieving state-of-the-art performance on various challenging similarity datasets [1, 8, 31]. SBERT is tailored for producing sentence embeddings that can be compared by using cosine-similarity, rendering this encoding method particularly suitable for capturing the content similarities between users. We encode each user's individual historical post p_k^i as $e_k^i = \mathrm{SBERT}(p_k^i)$; $e_k^i \in \mathbb{R}^{d_b}$.

Encoding the historical context. We wish to encode the users' historical context H_τ^i within each time frame τ, by obtaining their user representations $E_\tau^i \in \mathbb{R}^{d_b}$. [29] empirically showed that simple average sentence embeddings compare favourably to more complex methods. To this end, we average each user's historical encodings over her posting history length within a corresponding time frame τ in order to obtain meaningful user representations:

$$E_\tau^i = \frac{1}{\ell_\tau^i} \sum_{k=1}^{\ell_\tau^i} e_k^i \tag{1}$$

2.3 Graph construction

We define the user graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ denotes a set of user nodes and $\mathcal{E}$ edges between these users. Depending on the edge types between user's, we construct different graphs.

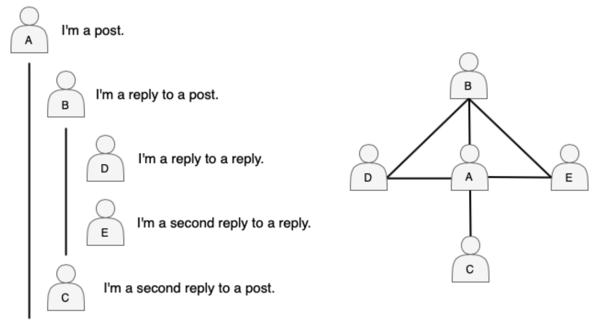


Figure 1: Transforming a post/reply tree in social media into a social graph network.

Semantic graph. To construct the users' semantic graph $\mathcal{G}_{sem} = (\mathcal{V}, \mathcal{E}_{sem})$, we calculate all the pairwise cosine similarities $cos(E^i_\tau, E^j_\tau)\ \forall\ i, j = 1, \dots, N$ between the users' historical context encodings within a time period τ. The E^i_τ capture the users' linguistic patterns within the time-period τ which renders them a suitable method for representing each user's context. Users with semantically similar content are close in the vector space [49] since they have similar context encodings. We form connections between two users only if the cosine similarity of their user embeddings is above a threshold θ, which essentially represents the semantic similarity between two users.

Social graph. On Reddit, users engage in different discussions and topics with their peers. Social science argues that like-minded people tend to interact more with each other [4], therefore we construct the social graph $\mathcal{G}_{soc} = (\mathcal{V}, \mathcal{E}_{soc})$ in a way that captures the users' social interactions with each other. We define as social interaction the replies and mentions in a post thread. For each individual target post of every user, we trace the chain of replies to the root (i.e. the original post) of the conversation and connect the target post with all the intermediate posts in the chain. Following, these post connections are translated to user connections in the social graph. Figure 1 shows how the post/reply tree is transformed into a social graph network.

Fused graph. To compare the effect of different graph construction types, we construct a fused graph by including both social and semantic connections. The new graph is essentially the union of the linguistic and social graphs, i.e. $\mathcal{G}_{fused} = (\mathcal{V}, \{\mathcal{E}_{soc} \cup \mathcal{E}_{sem}\})$.

2.4 Graph Attention Network

To explore the relations between users, we employ graph neural networks. As each user have a different influence on another user, it is important to assign higher weights to more relevant connections with higher influence. To model such dynamic influences of the neighbourhood to a node, we use Graph Attention Networks (GAT) [64]. The input to a GAT layer is a set of users embeddings $E_\tau = \{E^1_\tau, \dots, E^N_\tau\}$ for a time frame τ where $N = |\mathcal{U}|$. A GAT layer produces updated features, $\widetilde{E_\tau} = \{\widetilde{E^1_\tau}, \dots, \widetilde{E^N_\tau}\}$, where $\widetilde{E^i_\tau} \in \mathbb{R}^{d_g}$. First, the GAT layer applies a shared linear transformation by a weight matrix $\mathbf{W} \in \mathbb{R}^{d_g \times d_b}$. Then, we apply a shared self-attention mechanism to each node i, using the neighbourhood $\mathcal{N}(i)$. The normalized attention weight α_{ij} between node i and neighbour node j is computed as follows:

$$\alpha_{ij} = \frac{exp(LeakyReLU(a_w^\intercal\, [\mathbf{W}\, E^i_\tau \parallel \mathbf{W}\, E^j_\tau]))}{\sum_{k \in \mathcal{N}(i)} exp(LeakyReLU(a_w^\intercal\, [\mathbf{W}\, E^i_\tau \parallel \mathbf{W}\, E^k_\tau]))} \tag{2}$$

where $\top$ represents the transpose and $\parallel$ is the concatenation operation. $a_w \in \mathbb{R}^{2d_g}$, is a trainable parameter vector. The attention weights α_{ij} represent the importance of relation from node i to node j. To stabilize the learning process, we employ a multi-head attention [63]. We compute the output representation of a node $\widetilde{E^i_\tau}$ as follows:

$$\widetilde{E^i_\tau} = \sigma\left(\frac{1}{K}\sum_{k=1}^{K}\sum_{j \in \mathcal{N}(i)} \alpha^k_{ij}\mathbf{W}^k E^j_\tau\right) \tag{3}$$

where, σ is the *ReLU* nonlinear function, α^k_{ij} and $\mathbf{W}^k$ denote the normalized attention weight and the linear transformation for k-th head.

2.5 Temporal Representation

Temporal Encoding. To model the sequential dependencies through time for each user, we use a Gated Recurrent Unit (GRU) [11]. The GRU encodes the dynamic user graph representations across time. The output of GRU layer in a time frame of length T is computed as:

$$h_\tau = GRU(\widetilde{E_\tau}, h_{\tau-1}), \qquad \forall\ 1 \le \tau \le T, \tag{4}$$

where $\widetilde{E_\tau}$ is the output features of GAT layer for the time period τ, and $h_\tau \in \mathbb{R}^{d_r}$ is the hidden state of the GRU layer.

Temporal Attention. We employ a temporal attention mechanism [3] to compute an overall representation for the user with adaptive weights. The attention mechanism aims to model the fact that user representations at different time frames have different contributions to its overall representation. We aggregate the GRU hidden states using the attention mechanism as follows:

$$\gamma_\tau = \frac{exp(h_\tau \mathbf{W}\, \overline{h})}{\sum_{t=1}^{T} exp(h_t \mathbf{W}\, \overline{h})} \tag{5}$$

$$\Gamma(\overline{h}) = \sum_\tau \gamma_\tau h_\tau \tag{6}$$

where, $\overline{h} \in \mathbb{R}^{T \times d_r}$ denotes the concatenated hidden states of the GRU layer, γ_τ is the learned attention weight for the time frame τ, W is a learned linear transformation and $\Gamma(\overline{h})$ the overall learned representation of the user.

2.6 Classification Layer

The overall learned representations for each user, are forwarded into a linear layer parameterized by a weight matrix $\mathbf{W}^o \in \mathbb{R}^{d_o \times d_r}$. The final prediction is computed as:

$$\hat{y} = softmax(\mathbf{W}^o\, \Gamma(\overline{h})). \tag{7}$$

Given the true label y for a user, we use cross-entropy loss to calculate the loss L as follows:

$$L = -\sum_{i=1}^{N} y_i \, ln(\hat{y_i}) + (1 - y_i) \, ln(1 - \hat{y_i}). \tag{8}$$

3 DATASET

3.1 Data Collection

Due to its easiness to crawl and higher contextual density, which is beneficial for similarity assessment, we chose the social media platform Reddit[1] as the source for disinformation spreader and fact-checker post histories. The data crawling itself then was performed in a user-centric and iterative fashion. To begin with, a list of 65 subreddits regarding potentially controversial political topics such as the U.S. American presidential race, the SARS-CoV-2 pandemic, vaccines, abortion, feminism, gun control, climate change, 5G or politics in general was manually compiled. Then, for each of those subreddits, the most recent threads were crawled and inserted into a database. This database was designed in a way to keep the dates and the thread structure of all stored posts to later be able to retrieve temporal and interaction information. The maximum number of threads from each subreddit was set to 300.

On this data, we performed the first iteration of the URL domain based disinformation spreader and fact-checker detection to generate a list of Reddit user accounts with equal amounts of users for either class. We then crawled the complete histories of all the users in said list, thus fetched all threads in which they participated in the list of political subreddits. The maximum number of threads before filtering for the political subreddits in this case was set to 1500. To conclude one crawling iteration all of those threads were inserted into the database from which, again, a now larger list of spreaders and checkers can be extracted. This process was then iterated.

3.2 Media Domain Lists

Likewise to the work of [5], the website *mediabiasfactcheck.com*[2] was used as the main source for annotated news outlet domains. It was deemed a suitable resource for the study at hand as it offers annotations for the two dimensions; the *factuality level* and the *political bias* of a large proportion of high frequented online news media.

The range of discrete labels that is assigned to the media outlets by the site curators can be seen in table 1. Since we opted for a binary label for the disinformation spreader detection part of this body of work, a mapping for those labels was created. To be considered a disinformation domain, the *mediabiasfactcheck* label has to be below or at "mixed" factuality or labelled as satire, while the real news domains have to be at least "mostly factual" and between "Right-Center" and "Left-Center" political bias.

As for the credibility of the assigned annotations, the maintainers of *mediabiasfactcheck.com* state that they "are looking at political bias, how factual the information is, and links to credible, verifiable sources" [34]. In the description of their methodology, they also describe that they base the labels on reviews of at least 10 headlines and 5 news stories [34].

[1] https://www.reddit.com
[2] https://mediabiasfactcheck.com

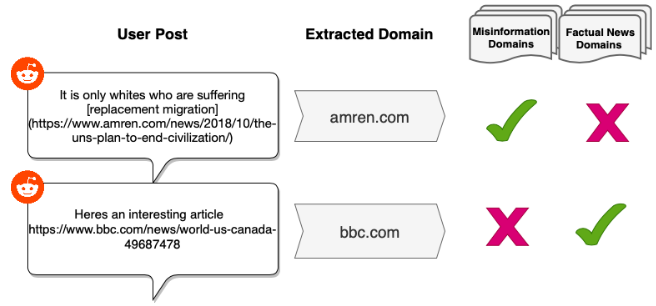


Figure 2: Caption

As a further resource to extend the list of disinformation media sources, an "index of fake-news, clickbait, and hate sites" [50] by the *Columbia Journalism Review*[3] was consulted. It's curators state that it was created by merging pre-existing fake news domain lists from various sources and then checking their actual invalidity with the fact checking platforms "PolitiFact and Snopes" [50].

In total, we on this way performed the post annotations with X disinformation and Y real news domains.

3.3 User-level Fake News Detection

Similar to [39], the detection of misinformation spreaders and fact-checkers to form a ground truth for our studies was done based on the posted web-links in a users' history. More precisely, those links were first extracted from a users' posts using regular expression matching. To decide whether the extracted link was counted as "disinformation" or, "real news" its' domain was matched with two lists of domains of online news outlets, each corresponding to one class of factuality. This process is visualised in **??**.

A user was then labelled as "Misinformation spreader" if he had at least one detected misinformation link in his post history, while for being a fact-checker they had to have no shared links from the fake news list and at least one link posted from the real news list.

3.4 User-level Annotation

In addition to the general separation of users into misinformation spreaders and fact-checkers, each user was furthermore annotated with the two following factors by averaging over a float mapping of the labels from *mediabiasfactcheck.com*, for a more fine-grained annotation:

- Political bias (pb) which represents the level of partisanship, which is in the range of [-3, +3] where each of the labels correspond to the following scales (s_ℓ): extreme left $\rightarrow s_{el} = -3$, left $\rightarrow s_l = -2$, center left $\rightarrow s_{cl} = -1$, least biased $\rightarrow s_{lb} = 0$, center right $\rightarrow s_{cr} = +1$, right $\rightarrow s_r = +2$, extreme right$\rightarrow s_{er} = +3$ and is computed for each author as follows: $pb = \frac{\sum_{\ell \in pbl} s_\ell \cdot N_\ell}{\sum_{\ell \in pbl} N_\ell}$ where N_ℓ in the number of posts labelled as ℓ where $\ell \in [el, l, cl, lb, cr, r, er]$
- Factuality factor (ff) in the range of [-3, +3] where each of the labels correspond to the following scales: very low $\rightarrow s_{vl} = -3$, low $\rightarrow s_{lf} = -2$, mixed $\rightarrow s_m = -1$, mostly factual

[3] https://www.cjr.org

Factuality	Very Low, Low, Mixed, Mostly Factual, High, Very High
Political Bias	Extreme Right, Right, Right-Center, Least Biased, Left-Center, Left, Extreme Left

Table 1: This is a table template

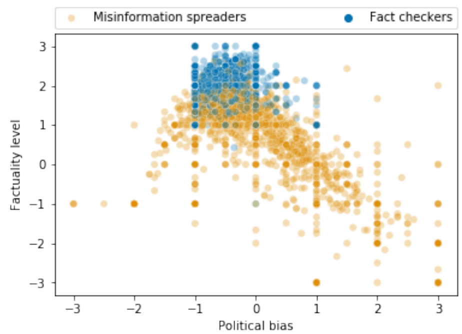


Figure 3: Caption

$\rightarrow s_{mf} = +1$, high $\rightarrow s_h = +2$, very high $\rightarrow s_{vh} = +3$ and is computed for each author as follows: $ff = \frac{\sum_{\ell \in fl} s_\ell \cdot N_\ell}{\sum_{\ell \in fl} N_\ell}$ where N_ℓ in the number of posts labelled as ℓ

3.5 Dataset Stats

In the timeframe from January 2020 till the end of April 2021 that was used for our experiments the dataset contained a total of X posts from Y users. Misinformation spreaders had an average of X posts with this count being at Y for the fact-checkers. In total y percent of the posts contained links to real news media while z percent pointed to domains from the misinformation list.

Using the post-level annotations described in section X the political biases of the users can be looked at: 41.17% of the users that have left wing political bias are fake news spreaders while 58.82% of them are fact-checkers 91.58% of the users that have right wing political bias are fake news spreaders while only 8.41% of them are fact checkers The distribution of factual factors and political biases is also presented in figure 3. It has to be noted here that since we chose a center political bias as a requirement for real news domains the apparent correlation (Pearson correlation of -0.45) between political bias and factuality of the posters could just have arisen by design of our annotation system.

To get an intuition for the actual linguistic differences between the two user groups of misinformation spreaders and fact-checkers we train a linear SVC model to infer on this binary label using a range (1,3) word n-gram feature vector and looked at the learned token weights. [19] showed that this can be done in the linear case to study the predictiveness of the tokens for either class. The most predictive tokens are shown in table 2. It can be seen that there's a tendency for misinformation spreaders to reference politically left-leaning groups as "liber", "dem", "left" or "blm", while fact-checkers use the terms "fascist" and "republican" with a higher frequency.

Dataset issues

Label	Tokens
Misinformation Spreaders	china, video, come, offici, blm, corrupt, media, away, liber, order, new, trump's, seem, wrong, kill, left, dem, riot
Fact Checkers	public, first, week, understand, trial, fascist, republican, war, one, forced-birth, health, please, power, let, shock, view, service

Table 2: Top ranked tokens for each label.

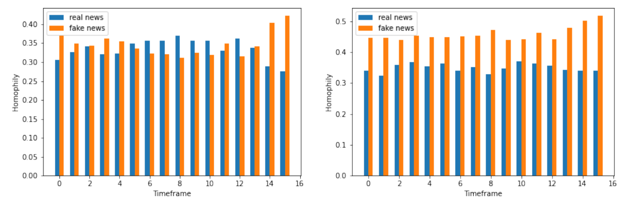


Figure 4: Amount of homophily observed through the time span for both semantic and social graph.

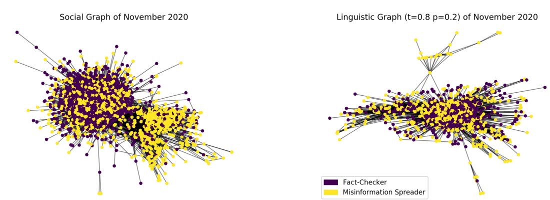


Figure 5: Caption

3.6 Graph Stats

Graph metrics. Like both of our research questions suggest, leveraging the temporal user-to-user connections in the semantic and social graphs is the main goal of this study. Figure 5 shows an example of the semantic and social graphs of November 2020, the month of the U.S. American presidential election. Moreover, we computed several graph metrics and their development over the time span from January 2020 till the end of April 2021. To derive those, the documents of all users were split into time-frames of 30 days. Then, the semantic and social graphs were constructed only using these time-framed documents. In **??**, we show the amount of homophily observed for both semantic and social graphs, which we define as the percentage of edges which connects users who have the same label.

Temporal analysis of edges. We wish to monitor the evolution of the user content similarities between different groups of users over time. We group the users by their label (misinformation spreaders, fact-checkers). Based on this categorization, we can define three different types of edges: (1) edges between misinformation spreaders, which will be denoted as 'm2m' edges for brevity, (2) edges between checkers, denoted as 'c2c' and (3) edges between misinformation spreaders and checkers, denoted as 'm2c'.

We define the connection's percentage of a certain edge type as $\rho_{\text{edge type}} = r^{(\tau)}_{\text{edge type}} / R^{(\tau)}_{\text{edge type}}$, where $r^{(\tau)}_{\text{edge type}}$ is the number of edges (of that edge type) that exist between two users during the time period τ and $R^{(\tau)}_{\text{edge type}}$ is the number of all possible connections (of that edge type) at time period τ which can be computed for each edge type as follows:

$$R^{(\tau)}_{m2m} = N^{(\tau)}_m (N^{(\tau)}_m - 1)/2, \qquad R^{(\tau)}_{c2c} = N^{(\tau)}_c (N^{(\tau)}_c - 1)/2$$

$$R^{(\tau)}_{m2c} = (N^{(\tau)}_m + N^{(\tau)}_c)(N^{(\tau)}_m + N^{(\tau)}_c - 1)/2$$

where $N^{(\tau)}_m$ is the number of misinformation spreaders and $N^{(\tau)}_c$ is the number of fact checkers that have posted at least one post at time period τ. Essentially, this metric shows how similar is the language usage between different groups of users. We partition the users' total posting period (from the start of January until the end of April) to 16 time periods that span 30 days each and compute the connection's percentage within each time period for all edge types. As shown by Figure 6, the 'm2c' connections percentage is consistently the lowest for all time periods, indicating that on an aggregate level misinformation spreaders and fact checkers are not densely connected and therefore do not have as much context similarity to each other. To provide a bit of context as to why there are certain peaks at the connection's percentage during August 2020 (event 1), November 2020 (event 2) and January 2021 (event 3) we provide a list of pivotal political events that happened during these specific months[4]. The events during July are given to provide some context as to why there is a sudden change in the connections percentage prior to (event 1).

July 11 - Early voting expanded and mail-in votes are encouraged.

July 15 - COVID-19 pandemic: The Trump administration politicizes health information by ordering hospitals to send all coronavirus patient information to a central database in Washington rather than to the CDC

July 30 - Donald Trump threatens to postpone the election if it appears mail-in-votes might go against him. (We regard this as if this had happened in August, since the effects of this political event would be still discussed during that month)

August 2020 - Joe Biden chooses Senator Kamala Harris (D-CA) as his running mate (event 1)

November 3 - 2020 United States elections (event 2)

January 2021 - United States Capitol is attacked by supporters of former President Donald Trump (event 3)

We observe that all types of edges follow a similar trend of the connections percentage, where there is an obvious increase of the connections percentage during the months leading to the elections and a couple of months after that major political event the connections percentage decreases again. This means that, while all users talk about the same topics, there are differences in the language usage among these groups

4 EXPERIMENTS

4.1 Dataset Splits

User Split.

[4]https://en.wikipedia.org/wiki/2020_in_United_States_politics_and_government
https://en.wikipedia.org/wiki/2021_in_the_United_States

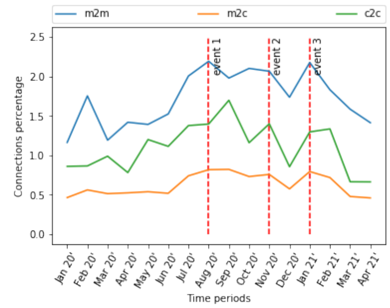


Figure 6: Connections percentage through time

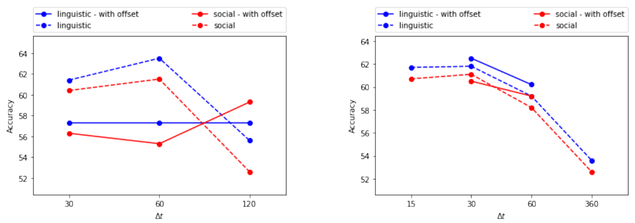
The blue line represents the connections percentage between misinformation spreaders, orange represents the connections percentage between misinformation spreaders and fact checkers and green represents the connections percentage between fact checkers.

Time split.

4.2 Training Setup

We use the paraphrase-mpnet-base-v2 from SBERT pretrained models [5], which has in overall the best performance. This model has max length set to 512, and uses mean pooling and has the output dimension $d_b = 768$. We experimented with different approaches to encode each user's historical representation. However, from our initial experiments, the results with the average approach (AVG), were better than the other two, so we decided to use that approach for the rest of our experiments. For each post in the user history, we masked the links, so that the cosine similarity is not attributed based on the links, but on the linguistic part of the news that the user has posted. We have a time frame of posts that consists from 01.01.2020 until 30.04.2021, a total of 16 months. We evaluate, our model in two different splits of data. First one is user split, where we split the dataset of users train:validation:test in 60:20:20. The second one is the time split, where the data train:validation:test is split in 50:25:25. Specifically for training we have data from 01.01.2020 to 31.08.2020, for validation from 01.09.2020 to 31.12.2020, for testing from 01.01.2021 to 30.04.2021. For training and validation data, we have 1200 and 400 samples. We run experiments with $\delta \in 15, 30, 60, 120$, to partition the time frame. In each sample, we randomly sample $n \in 200, 400, 800, 1200$ users. In each sample, we build a subgraph of users for each discrete time window. In the semantic graph, we connect users with each other based on the hyperparameter $\theta \in [0, 1]$. We perform grid-search to get the best performing hyperparameters. For the time split, we find our model to perform best with the following hyperparameters: $n = 400$, $\delta = 60$, $\theta = 0.8$, $d_b = 768$, $d_g = 256$, $d_r = 128$, and for the user split we have the following hyperparameters: $n = 400, \delta = 30, \theta = 0.8, d_b = 768, d_g = 256, d_r = 128$. The number of attention heads for the GAT layer is set to 8. For both social, and fused graph, we use the same hyperparameters for the same splits. We use Adam optimizer [25], learning rate $5e-6$, weight

[5]https://www.sbert.net/docs/pretrained_models.html



(a) User Split **(b) Time Split**

Figure 7: Plot of the test accuracy vs the time window for the user split and the time split respectively

decay $1e-2$, and train the model for 500 epochs using early stopping with patience 50 on the validation set.

4.3 Proposed Models

We use the following models to evaluate our dataset. As a baseline to compare our proposed graph model, we use a Linear Support Vector Machine classifier [6], optimized on top of features for each user extracted from SBERT and averaged using Equation 1. Moreover, we run ablation studies across different components of our model. Namely, we first remove the graph construction and the GAT layer, which captures the information from such structures. Moreover, we remove the temporal attention layer from our full model, to see the effect of this layer on the overall result. In the last study, we remove the temporal dynamic embeddings. In this setup, we merge all the graphs constructed for every discrete time frame into one, and replace the sequential layer (GRU), with a linear layer. We have a single user representation, which is an average of his representation across all the merged time frames.

4.4 Experimental Results

5 ERROR ANALYSIS

In order to understand and analyze the results of our graph models, we perform a comparative error analysis between them. More specifically, we identify two groups of users in our analysis; users that are consistently correctly classified by all graph models, and users that are consistently assigned to the wrong class by all graph models. 57.98% of the consistently misclassified users are fact checkers and the rest (42.02%) are misinformation spreaders, meaning that the models are able to identify the fake news spreaders more easily. Fig 8a shows a scatter plot of each user's factuality level over their political bias. Analysis of baseline and graph results with regards to the factuality and bias factors here

Analysis of the edge connections with regards to the factuality and bias factors here

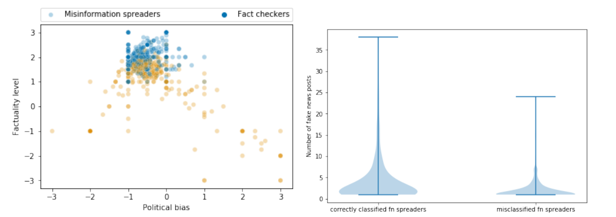
6 CONCLUSION

7 ETHICAL CONSIDERATION

REFERENCES

[1] Eneko Agirre, Mona Diab, Daniel Cer, and Aitor Gonzalez-Agirre. 2012. SemEval-2012 Task 6: A Pilot on Semantic Textual Similarity. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics - Volume 1: Proceedings of the Main Conference and the Shared Task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation* (Montréal, Canada) *(SemEval '12)*. Association for Computational Linguistics, USA, 385–393.

[6] https://scikit-learn.org/stable/modules/generated/sklearn.svm.LinearSVC.html



(a) **(b)**

Figure 8: (a) Scatter plot of each user's factuality level over political bias. The plot contains only the users that were consistently misclassified by all models. Blue corresponds to fact checkers and orange to misinformation spreaders. (b) Violin plot that depicts the distribution of the number of fake news posts for the consistently correctly classified (left) and the consistently misclassified (right) misinformation spreaders. Both plots show the results obtained by the graph models by utilizing the semantic graphs on the time split.

[2] Liesbeth Allein, Isabelle Augenstein, and Marie-Francine Moens. 2020. Time-Aware Evidence Ranking for Fact-Checking. *CoRR* abs/2009.06402 (2020). arXiv:2009.06402 https://arxiv.org/abs/2009.06402

[3] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2014. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473* (2014).

[4] Angela Bahns, Chris Crandall, Omri Gillath, and Kristopher Preacher. 2017. Similarity in Relationships as Niche Construction: Choice, Stability, and Influence Within Dyads in a Free Choice Environment. *Journal of Personality and Social Psychology* 11 (02 2017), 329–355. https://doi.org/10.1037/pspp0000088

[5] Ramy Baly, Georgi Karadzhov, Dimitar Alexandrov, James Glass, and Preslav Nakov. 2018. Predicting factuality of reporting and bias of news media sources. *arXiv preprint arXiv:1810.01765* (2018).

[6] Adam J. Berinsky. 2017. Rumors and Health Care Reform: Experiments in Political Misinformation. *British Journal of Political Science* 47, 2 (2017), 241–262. https://doi.org/10.1017/S0007123415000186

[7] Chiyu Cai, Linjing Li, and Daniel Zengi. 2017. Behavior enhanced deep bot detection in social media. In *2017 IEEE International Conference on Intelligence and Security Informatics (ISI)*. 128–130. https://doi.org/10.1109/ISI.2017.8004887

[8] Daniel Cer, Mona Diab, Eneko Agirre, Inigo Lopez-Gazpio, and Lucia Specia. 2017. SemEval-2017 Task 1: Semantic Textual Similarity Multilingual and Crosslingual Focused Evaluation. *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)* (2017). https://doi.org/10.18653/v1/s17-2001

[9] Shantanu Chandra, Pushkar Mishra, Helen Yannakoudakis, Madhav Nimishakavi, Marzieh Saeidi, and Ekaterina Shutova. 2020. Graph-based Modeling of Online Communities for Fake News Detection. *CoRR* abs/2008.06274 (2020). arXiv:2008.06274 https://arxiv.org/abs/2008.06274

[10] Yang Cheng and Zifei Fay Chen. 2020. The Influence of Presumed Fake News Influence: Examining Public Support for Corporate Corrective Response, Media Literacy Interventions, and Governmental Regulation. *Mass Communication and Society* 23, 5 (2020), 705–729. https://doi.org/10.1080/15205436.2020.1750656 arXiv:https://doi.org/10.1080/15205436.2020.1750656

[11] Kyunghyun Cho, Bart van Merriënboer, Dzmitry Bahdanau, and Yoshua Bengio. 2014. On the Properties of Neural Machine Translation: Encoder–Decoder Approaches. In *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*. Association for Computational Linguistics, Doha, Qatar, 103–111. https://doi.org/10.3115/v1/W14-4012

[12] Shivang Chopra, Ramit Sawhney, Puneet Mathur, and Rajiv Ratn Shah. 2020. Hindi-English Hate Speech Detection: Author Profiling, Debiasing, and Practical Perspectives. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*. AAAI Press, 386–393. https://aaai.org/ojs/index.php/AAAI/article/view/5374

Model	**User Split**			**Time Split**		
BERT + LinSVM	64.2%			56%		
Model	**Semantic**	**Social**	**Fused**	**Semantic**	**Social**	**Fused**
Graph model	63.09%	64.15%	65.3%	**64.93%***	62.03%*	60.7%*
without GAT	65.12%	65.50%	63.5%	62.06%*	62.44%	59.9%*
without temporal attention	64.34%	**66.57%**	65.0%	62.63%*	62.06%*	61.3%
without time frames	62.6%	64.1%	62.2%	64.4%*	62.6%*	61.9%

Table 3: Ablative results. Accuracy performance on the user split and the time split. The results with the asterisk (*) are statistically significant based on the Wilcoxon signed rank test ($p = 0.05$).

[13] E. Fersini, Justin Armanini, and Michael D'Intorni. 2020. Profiling Fake News Spreaders: Stylometry, Personality, Emotions and Embeddings. In *CLEF*.

[14] Adam Fourney, Miklos Z. Racz, Gireeja Ranade, Markus Mobius, and Eric Horvitz. 2017. *Geographic and Temporal Trends in Fake News Consumption During the 2016 US Presidential Election*. Association for Computing Machinery, New York, NY, USA, 2071–2074. https://doi.org/10.1145/3132847.3133147

[15] Anastasia Giachanou, Esteban A. Ríssola, Bilal Ghanem, Fabio Crestani, and Paolo Rosso. 2020. The Role of Personality and Linguistic Patterns in Discriminating Between Fake News Spreaders and Fact Checkers. In *Natural Language Processing and Information Systems*, Elisabeth Métais, Farid Meziane, Helmut Horacek, and Philipp Cimiano (Eds.). Springer International Publishing, Cham, 181–192.

[16] Anastasia Giachanou, Paolo Rosso, and Fabio Crestani. 2021. The impact of emotional signals on credibility assessment. *Journal of the Association for Information Science and Technology* (May 2021). https://doi.org/10.1002/asi.24480

[17] Jeffrey J. Glenn, Alicia L. Nobles, Laura E. Barnes, and Bethany A. Teachman. 2020. Can Text Messages Identify Suicide Risk in Real Time? A Within-Subjects Pilot Examination of Temporally Sensitive Markers of Suicide Risk. *Clinical Psychological Science* 8, 4 (2020), 704–722. https://doi.org/10.1177/2167702620906146 arXiv:https://doi.org/10.1177/2167702620906146

[18] Shengnan Guo, Youfang Lin, Ning Feng, Chao Song, and Huaiyu Wan. 2019. Attention Based Spatial-Temporal Graph Convolutional Networks for Traffic Flow Forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence* 33, 01 (Jul. 2019), 922–929. https://doi.org/10.1609/aaai.v33i01.3301922

[19] Isabelle Guyon, Jason Weston, Stephen Barnhill, and Vladimir Vapnik. 2002. Gene selection for cancer classification using support vector machines. *Machine learning* 46, 1-3 (2002), 389–422.

[20] Tarek Hamdi, Hamda Slimi, Ibrahim Bounhas, and Yahya Slimani. 2020. A Hybrid Approach for Fake News Detection in Twitter Based on User Features and Graph Embedding. In *International Conference on Distributed Computing and Internet Technology, Odisha, India, January 19-12, 2020*, Dang Van Hung and Meenakshi D´Souza (Eds.). Springer International Publishing, Springer International Publishing.

[21] Wentao Han, Youshan Miao, Kaiwei Li, Ming Wu, Fan Yang, Lidong Zhou, Vijayan Prabhakaran, Wenguang Chen, and Enhong Chen. 2014. Chronos: A Graph Engine for Temporal Graph Analysis. In *Proceedings of the Ninth European Conference on Computer Systems* (Amsterdam, The Netherlands) *(EuroSys '14)*. Association for Computing Machinery, New York, NY, USA, Article 1, 14 pages. https://doi.org/10.1145/2592798.2592799

[22] Benjamin D. Horne, Jeppe Nørregaard, and Sibel Adali. 2019. Robust Fake News Detection Over Time and Attack. *ACM Trans. Intell. Syst. Technol.* 11, 1, Article 7 (Dec. 2019), 23 pages. https://doi.org/10.1145/3363818

[23] Sofia Hurtado, Poushali Ray, and Radu Marculescu. 2019. Bot Detection in Reddit Political Discussion. In *Proceedings of the Fourth International Workshop on Social Sensing* (Montreal, QC, Canada) *(SocialSense'19)*. Association for Computing Machinery, New York, NY, USA, 30–35. https://doi.org/10.1145/3313294.3313386

[24] Gihwan Kim and Youngjoong Ko. 2021. Graph-based Fake News Detection using a Summarization Technique. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*. Association for Computational Linguistics, Online, 3276–3280. https://aclanthology.org/2021.eacl-main.287

[25] Diederik P Kingma and Jimmy Ba. 2015. Adam, A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, Vol. 1412.

[26] Dijana Kosmajac and Vlado Keselj. 2020. Twitter User Profiling: Bot and Gender Identification - Notebook for PAN at CLEF 2019. In *Experimental IR Meets Multilinguality, Multimodality, and Interaction - 11th International Conference of the CLEF Association, CLEF 2020, Thessaloniki, Greece, September 22-25, 2020, Proceedings (Lecture Notes in Computer Science, Vol. 12260)*, Avi Arampatzis, Evangelos Kanoulas, Theodora Tsikrika, Stefanos Vrochidis, Hideo Joho, Christina Lioma, Carsten Eickhoff, Aurélie Névéol, Linda Cappellato, and Nicola Ferro (Eds.). Springer, 141–153. https://doi.org/10.1007/978-3-030-58219-7_13

[27] Shailender Kumar, Subodh Kumar, Pooja Yadav, and Meghna Bagri. 2021. A Survey on Analysis of Fake News Detection Techniques. In *2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS)*. 894–899. https://doi.org/10.1109/ICAIS50930.2021.9395978

[28] Sejeong Kwon, Meeyoung Cha, and Kyomin Jung. 2016. Rumor detection over varying time windows. https://doi.org/10.7910/DVN/BFGAVZ

[29] Jae Hee Lee, José Camacho Collados, Luis Espinosa Anke, and S. Schockaert. 2020. Capturing Word Order in Averaging Based Sentence Embeddings. In *ECAI*.

[30] Yaguang Li, Rose Yu, Cyrus Shahabi, and Yan Liu. 2018. Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting. arXiv:1707.01926 [cs.LG]

[31] Marco Marelli, Stefano Menini, Marco Baroni, Luisa Bentivogli, Raffaella Bernardi, and Roberto Zamparelli. 2014. A SICK cure for the evaluation of compositional distributional semantic models. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*. European Language Resources Association (ELRA), Reykjavik, Iceland, 216–223. http://www.lrec-conf.org/proceedings/lrec2014/pdf/363_Paper.pdf

[32] Hannah Marshall and Alena Drieschova. 2018. Post-Truth Politics in the UK's Brexit Referendum. *New Perspectives* 26, 3 (2018), 89–105. https://doi.org/10.1177/2336825X1802600305 arXiv:https://doi.org/10.1177/2336825X1802600305

[33] Spencer McKay and Chris Tenove. 0. Disinformation as a Threat to Deliberative Democracy. *Political Research Quarterly* 0, 0 (0), 1065912920938143. https://doi.org/10.1177/1065912920938143 arXiv:https://doi.org/10.1177/1065912920938143

[34] mediabiasfactcheck.com. 2021. *Methodology*. https://mediabiasfactcheck.com/methodology/

[35] Pushkar Mishra, Marco Del Tredici, Helen Yannakoudakis, and Ekaterina Shutova. 2019. Abusive Language Detection with Graph Convolutional Networks. *CoRR* abs/1904.04073 (2019). arXiv:1904.04073 http://arxiv.org/abs/1904.04073

[36] Rohan Mishra, Pradyumn Prakhar Sinha, Ramit Sawhney, Debanjan Mahata, Puneet Mathur, and Rajiv Ratn Shah. 2019. SNAP-BATNET: Cascading Author Profiling and Social Network Graphs for Suicide Ideation Detection on Social Media. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Student Research Workshop*. Association for Computational Linguistics, Minneapolis, Minnesota, 147–156. https://doi.org/10.18653/v1/N19-3019

[37] Federico Monti, Fabrizio Frasca, Davide Eynard, Damon Mannion, and Michael M. Bronstein. 2019. Fake News Detection on Social Media using Geometric Deep Learning. *CoRR* abs/1902.06673 (2019). arXiv:1902.06673 http://arxiv.org/abs/1902.06673

[38] Yida Mu and Nikolaos Aletras. 2020. Identifying Twitter users who repost unreliable news sources with linguistic information. *PeerJ Comput. Sci.* 6 (2020), e325. https://doi.org/10.7717/peerj-cs.325

[39] Yida Mu and Nikolaos Aletras. 2020. Identifying Twitter users who repost unreliable news sources with linguistic information. *PeerJ Computer Science* 6 (2020), e325.

[40] Taichi Murayama, Shoko Wakamiya, and Eiji Aramaki. 2020. Fake News Detection using Temporal Features Extracted via Point Process. *CoRR* abs/2007.14013 (2020). arXiv:2007.14013 https://arxiv.org/abs/2007.14013

[41] Brendan Nyhan and Jason Reifler. 2010. When Corrections Fail: The Persistence of Political Misperceptions. *Political Behavior* 32, 2 (2010), 303–330. http://www.jstor.org/stable/40587320

[42] Gordon Pennycook and David G. Rand. 2019. Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition* 188 (2019), 39–50. https://doi.org/10.1016/j.cognition.2018.06.011 The Cognitive Science of Political Thought.

[43] Bhavtosh Rath, Wei Gao, and Jaideep Srivastava. 2021. Assessing Individual and Community Vulnerability to Fake News in Social Networks. *CoRR* abs/2102.02434 (2021). arXiv:2102.02434 https://arxiv.org/abs/2102.02434

[44] Bhavtosh Rath, Xavier Morales, and Jaideep Srivastava. 2021. SCARLET: Explainable Attention based Graph Neural Network for Fake News spreader prediction. *CoRR* abs/2102.04627 (2021). arXiv:2102.04627 https://arxiv.org/abs/2102.04627
[45] Bhavtosh Rath, Aadesh Salecha, and Jaideep Srivastava. 2020. Detecting Fake News Spreaders in Social Networks using Inductive Representation Learning. *CoRR* abs/2011.10817 (2020). arXiv:2011.10817 https://arxiv.org/abs/2011.10817
[46] Bhavtosh Rath, Aadesh Salecha, and Jaideep Srivastava. 2020. Detecting Fake News Spreaders in Social Networks using Inductive Representation Learning. arXiv:2011.10817 [cs.SI]
[47] T. R. Reddy. 2016. A Survey on Authorship Profiling Techniques.
[48] David P. Redlawsk, Andrew J. W. Civettini, and Karen M. Emmerson. 2010. The Affective Tipping Point: Do Motivated Reasoners Ever "Get It"? *Political Psychology* 31, 4 (2010), 563–593. https://doi.org/10.1111/j.1467-9221.2010.00772.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9221.2010.00772.x
[49] Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. Association for Computational Linguistics, Hong Kong, China, 3982–3992. https://doi.org/10.18653/v1/D19-1410
[50] Columbia Journalism Review. 2021. *CJR index of fake-news, clickbait, and hate sites*. https://www.cjr.org/fake-beta
[51] Samuel C. Rhodes. 2021. Filter Bubbles, Echo Chambers, and Fake News: How Social Media Conditions Individuals to Be Less Critical of Political Misinformation. *Political Communication* 0, 0 (2021), 1–22. https://doi.org/10.1080/10584609.2021.1910887 arXiv:https://doi.org/10.1080/10584609.2021.1910887
[52] Emanuele Rossi, Ben Chamberlain, Fabrizio Frasca, Davide Eynard, Federico Monti, and Michael Bronstein. 2020. Temporal Graph Networks for Deep Learning on Dynamic Graphs. arXiv:2006.10637 [cs.LG]
[53] Atanu Roy, Chandrima Sarkar, Jaideep Srivastava, and Jisu Huh. 2016. Trustingness amp; trustworthiness: A pair of complementary trust measures in a social network. In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. 549–554. https://doi.org/10.1109/ASONAM.2016.7752289
[54] Natali Ruchansky, Sungyong Seo, and Yan Liu. 2017. CSI: A Hybrid Deep Model for Fake News. *CoRR* abs/1703.06959 (2017). arXiv:1703.06959 http://arxiv.org/abs/1703.06959
[55] Ramit Sawhney, Shivam Agarwal, Arnav Wadhwa, Tyler Derr, and Rajiv Ratn Shah. 2021. Stock Selection via Spatiotemporal Hypergraph Attention Network: A Learning to Rank Approach. *Proceedings of the AAAI Conference on Artificial Intelligence* 35, 1 (May 2021), 497–504. https://ojs.aaai.org/index.php/AAAI/article/view/16127
[56] Ramit Sawhney, Shivam Agarwal, Arnav Wadhwa, and Rajiv Shah. 2021. Exploring the Scale-Free Nature of Stock Markets: Hyperbolic Graph Learning for Algorithmic Trading. In *Proceedings of the Web Conference 2021* (Ljubljana, Slovenia) *(WWW '21)*. Association for Computing Machinery, New York, NY, USA, 11–22. https://doi.org/10.1145/3442381.3450095
[57] Ramit Sawhney, Shivam Agarwal, Arnav Wadhwa, and Rajiv Ratn Shah. 2020. Deep Attentive Learning for Stock Movement Prediction From Social Media Text and Company Correlations. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, Online, 8415–8426. https://doi.org/10.18653/v1/2020.emnlp-main.676
[58] Ramit Sawhney, Shivam Agarwal, Arnav Wadhwa, and Rajiv Ratn Shah. 2020. Spatiotemporal Hypergraph Convolution Network for Stock Movement Forecasting. In *2020 IEEE International Conference on Data Mining (ICDM)*. 482–491. https://doi.org/10.1109/ICDM50108.2020.00057
[59] Ramit Sawhney, Harshit Joshi, Rajiv Ratn Shah, and Lucie Flek. 2021. Suicide Ideation Detection via Social and Temporal User Representations using Hyperbolic Learning. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, Online, 2176–2190. https://doi.org/10.18653/v1/2021.naacl-main.176
[60] Youngjoo Seo, Michaël Defferrard, Pierre Vandergheynst, and Xavier Bresson. 2016. Structured Sequence Modeling with Graph Convolutional Recurrent Networks. *CoRR* abs/1612.07659 (2016). http://dblp.uni-trier.de/db/journals/corr/corr1612.html#SeoDVB16
[61] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake News Detection on Social Media: A Data Mining Perspective. *CoRR* abs/1708.01967 (2017). arXiv:1708.01967 http://arxiv.org/abs/1708.01967
[62] Briony Swire, Ullrich K.H. Ecker, and Stephan Lewandowsky. 2017. The role of familiarity in correcting inaccurate information. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 43, 12 (1 Dec. 2017), 1948–1961. https://doi.org/10.1037/xlm0000422
[63] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is All you Need. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, 4-9 December 2017, Long Beach, CA, USA*, Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett (Eds.). 5998–6008. http://papers.nips.cc/paper/7181-attention-is-all-you-need
[64] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. 2018. Graph Attention Networks. *International Conference on Learning Representations* (2018). https://openreview.net/forum?id=rJXMpikCZ
[65] Marco Visentin, Gabriele Pizzi, and Marco Pichierri. 2019. Fake News, Real Problems for Brands: The Impact of Content Truthfulness and Source Credibility on consumers' Behavioral Intentions toward the Advertised Brands. *Journal of Interactive Marketing* 45 (2019), 99–112. https://doi.org/10.1016/j.intmar.2018.09.001
[66] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and Philip S. Yu. 2019. A Comprehensive Survey on Graph Neural Networks. *CoRR* abs/1901.00596 (2019). arXiv:1901.00596 http://arxiv.org/abs/1901.00596
[67] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and Philip S. Yu. 2021. A Comprehensive Survey on Graph Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems* 32, 1 (Jan 2021), 4–24. https://doi.org/10.1109/tnnls.2020.2978386
[68] Sijie Yan, Yuanjun Xiong, and Dahua Lin. 2018. Spatial Temporal Graph Convolutional Networks for Skeleton-Based Action Recognition. arXiv:1801.07455 [cs.CV]
[69] Rowan Zellers, Ari Holtzman, Hannah Rashkin, Yonatan Bisk, Ali Farhadi, Franziska Roesner, and Yejin Choi. 2019. Defending Against Neural Fake News. *CoRR* abs/1905.12616 (2019). arXiv:1905.12616 http://arxiv.org/abs/1905.12616
[70] Álvaro Figueira and Luciana Oliveira. 2017. The current state of fake news: challenges and opportunities. *Procedia Computer Science* 121 (2017), 817–825. https://doi.org/10.1016/j.procs.2017.11.106 CENTERIS 2017 - International Conference on ENTERprise Information Systems / ProjMAN 2017 - International Conference on Project MANagement / HCist 2017 - International Conference on Health and Social Care Information Systems and Technologies, CENTERIS/ProjMAN/HCist 2017.

A APPENDIX A