Fake News Detection in the Framework of Decision-Making System through Graph Neural Network

Ihor Pilkevych
Research Department
Korolyov Zhytomyr
Military Institute
Zhytomyr, Ukraine
igor.pilkevych@meta.ua

Dmytro Fedorchuk
Research Department
Korolyov Zhytomyr
Military Institute
Zhytomyr, Ukraine
fedor4uk.d@gmail.com

Olena Naumchak
Research Department
Korolyov Zhytomyr
Military Institute
Zhytomyr, Ukraine
olenanau@gmail.com

Mykola Romanchuk
Research Department
Korolyov Zhytomyr
Military Institute
Zhytomyr, Ukraine
romannik@ukr.net

2021 IEEE 4th International Conference on Advanced Information and Communication Technologies (AICT) | 978-1-6654-0618-5/21/$31.00 ©2021 IEEE | DOI: 10.1109/AICT52120.2021.9628907

***Abstract* – The rapid growth of the number of fake news has become a serious threat to the credibility of the governments of many countries. The extensive use of social platforms contributes to the creation and dissemination of fake news. This is actualize the task of fake news detection. Much research has already focused on this. Modern methods of detecting fake news rely heavily on textual information, examining the extracted news content or writing style based on internal knowledge. However, intentional rumors can mask the style of writing, bypassing languages and text patterns. In order to combat the spread of fake news, methods of automatic detection based on artificial intelligence and machine learning were studied. In a world where millions of articles are deleted and published every minute, this cannot be done effectively manually. The solution may be to develop a system to provide reliable automated assessment systems or to assess the reliability of different publishers and news contexts using deep learning techniques, namely graph neural networks with an inductive structure.**

***Keywords* — fake news detection, graph neural networks, graph convolution network, graph attention network, graph sample and aggregate.**

I. Introduction

The availability of the online media, the rapidly growing number of sources of information (such as news sites, social medias, blogs, websites, etc.) and the ease of using them to quickly disseminate information entails the viral spread of fake news. The popularization of social media has exacerbated the longstanding problem of fake news. Now fake news has grown into a big problem for societies and individuals, as well as for organizations and authorities that counter disinformation and propaganda. This phenomenon harms democratic elections, the reputation of individuals or organizations, and negatively affects citizens. The use of digital communications and social networks as weapons for conducting disinformation campaigns, their scale, speed and engagement of target audiences make the task of detecting and countering fake news very relevant.

II. Problem Statement in General Terms

It is essential to consider that fake news is clearly different from false information with a several of distinctions.

First, the authors deliberately and intendently create fake news, to mislead readers. The peculiarity is that news about the same event, published by different authors, may have very similar content, but the actual news contains objective statements, and fake news additionally contains malicious content intentionally added to it. While the proportion of this malicious content may be negligible, it is often enough to make news an information weapon.

Second, for example, spam is a classic example of information with malicious content, but people instinctively handle it with caution. But users, as a rule, on the contrary, actively search for, receive and share news among themselves, not caring about their reliability.

Third, spam is easier to detect due to the number of regular messages; and fake news is harder to recognize because it is time sensitive. The evidence gathered about past news cannot help identify fake news. All of these factors make it difficult to identify fake news.

The fact of Internet's and social media's popularizing is leading reason for fake news spreading in an enormous scale. Because the phenomenon of fake news is a global problem today, flexible and reliable solutions are needed to address the problem of counterfeiting fake news.

Fake news can be identified by content analyzing. Between the approaches are digital processing of images, text, network data and web content, ad less used the reputation of the author / source.

According to Google Scholar, the number of posts with the concept of "fake news" is growing rapidly (Figure 1).

Comparison of search queries "fake news" and "propaganda" for the period since 2008 using the Google Trends platform is shown in fig. 2. It can be seen that over the past few years the problem of fake news has become more discussed, while there is a constant interest in propaganda.

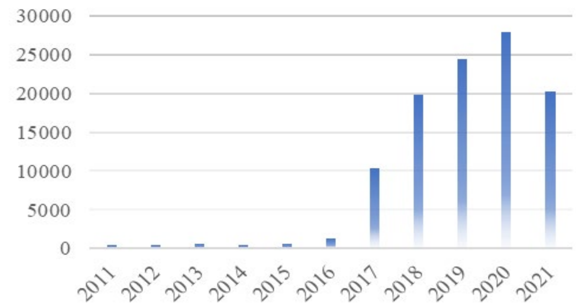

Fig. 1. The increasing of a number of posts about fake news

Fig. 2. Comparison of search queries "fake news" and "propaganda"

The data from e-print repository of scientific articles arXiv.org also shows a growing interest in the problem of fake news. As can be seen in Fig. 3, the number of scientific articles for the first half of 2021 exceeds the indicator for 2019 and is almost equal to the indicator for 2020. This indicates a high scientific interest in problem of detecting fake news.

Numerous techniques and algorithms in the machine learning and natural language processing fields are already realized in many projects for identifying and classifying fake news. The number of projects on GitHub repository in this field shown on Fig.3.

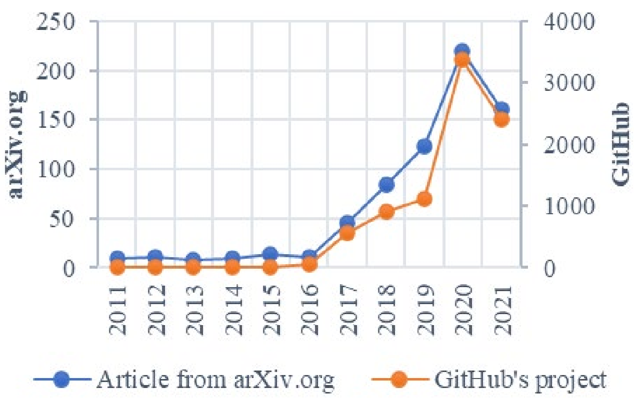

Fig. 3. Evolution of scientific research and practical solutions for fake news

As shown on fig. 2-3, a sharp increase in interest in the topic of fake news occurred in 2016. Most likely, the reason for this was the elections in the United States, which were accompanied by active information campaigns by various interested organizations and entire states. Although disinformation and propaganda have been around since ancient times, their importance and impact in the age of social media is still not clear enough. This only confirms the importance of research in the field of information operations, their influence on the course of in various fields, from political to military. Fake news as part of mass digital disinformation is a serious technological and geopolitical risks.

The most common and apparent approach to automated fake news detection is text analysis (or natural language processing). There are several main trends in this area. The simplest approach is the analysis of the presentation of a text without a linguistic context. Often used the form of a bag of words or N-grams. Among other approaches - the psycholinguistic factors analysis, syntactic and semantic analysis, and some non-linguistic methods. But these approaches are limited and do not allow to take into account the peculiarities of natural language, so the results may be radically different from the truth. Cause the social context of the text message transmitted in electronic media is a highly relevant factor.

Since 2016, there has been a sharp increase in approaches that use, instead of purely analyzing the content of information messages, their relationship with each other. The best solutions for analyzing fake news are approaches that use graph models of neural networks that solve the transduction problem.

III. Review of the Latest Research and Published Materials

There were considered some approaches to fake news detection. The knowledge-based approach of Fake News Detection that aims to ascertain the news authenticity by matching the knowledge extracted from the news contents with real knowledge described in [1]. Another typical way that based on writing style analysis, such as discourse level by employing rhetorical structure theory was presented in [2]. Authors of [3] based their research on studying the sentiment & readability of news. Another way, based on relationships among news articles, users (spreaders) and user posts that used matrix factorization declared in [4]. Among other bases for fake news detection were reviewed tensor factorization [5] and hierarchical word encoder [6].

In [7] authors proposed use attributes and peculiarities of the falsification content for analyzing the quantitative content of false tweets. In [8] group of authors found out that the model for fake news detection seems to predictive analysis model. They use the hybrid classification model that combine the three parts: processing, feature extraction and classification for fake news detection. [9] detected fake news using common methods: Naïve Bayes, Neural Network and Support Vector Machine (SVM). [10] examined the carrying out of 8 supervised machine learning models for classifying messages about 2012 Dutch elections from Twitter datasets. For fake news detection in [11] and [12] authors used Naive Bayes classifier. This method was implemented as a framework that proceed the data from the Facebook and Twitter. The [13] presented the methodology to create a model that will detect if an article is authentic or fake based on its words, phrases, sources and titles, by applying supervised machine learning algorithms.

Next models used in task Fake Detection were analyzed: CSI [14] proposed model based on LSTM that encode information about the news content with aim to detect fake news. In [15] SAFE uses the TextCNN from [16] for encoding textual data from electronic news resources. Proposed GCNFN framework in [17] based on graph convolution network (GCN) as distinguished from other cause the firstly using the news propagation graph [18]. It takes as input features embeddings of user information (the data from user profiles). Graph neural network (GNN) in GNN-CL [19] used DiffPool [20], a GNN designed for graph classification.

But the best results are shown by solutions built on the basis of 3 popular models for fake news detection: GCN, Graph ATtention Network (GAT) [21] and Graph SAmple and aGgregatE (GraphSAGE) [22].

IV. Core Material Summary

The basis of these models is a graph neural networks. The main idea of which is to built a computational graph for each node, the features of which are determined by the features of its neighbors through a nonlinear aggregator. An aggregator is a nonlinear, preferably differentiated function that combines node's neighbor information with node information.

This exchange is called neural message transmission and it is carried out simultaneously by all nodes of the graph for each layer that transmits messages.

That is, a node having features $\left(x_{11}, x_{12}, x_{13}, \ldots, x_{1m}\right)$ receives information from neighboring nodes about its properties, the same data sets $\left(x_{n1}, x_{n2}, x_{n3}, \ldots, x_{nm}\right)$. There is a generalization of information about the state of neighbors in generalized vector $h_i^{(k)}$, which denotes the state of node i at step k. This process takes place until all nodes will know the information about each node in the graph.

Therefore, the formal definition of a graph neural network can be presented as follows [23]:

$$h_u^{(k+1)} = UPDATE^{(k)}\left(h_u^{(k)}, AGGAREGATE^{(k)}\left(\left\{h_v^{(k)}, \forall v \in \mathcal{N}(u)\right\}\right)\right) \quad (1)$$

where $\mathcal{N}(u)$ – set of neighbors for node u.

The literature in the graph theory considers two approaches to the use of convolutions on graphs - spectral and non-spectral (spatial).

Spectral approaches work with spectral representation of graphs. The main disadvantage of spectral approaches is the dependence on the structure of the graph. This means that a model taught on a certain structure cannot be also applied to a graph with another structure. And the main problem of non-spectral approaches is the execution of the convolution operation on neighbors of different sizes.

Among the spatial approaches discussed in this article are: GraphSAGE and GAT. A typical spectral method is GCN. Other spectral and non-spectral approaches are discussed in [24].

GCN – Graph Convolution Network – is semi-supervised learning model on graph-structured data. In [25] presented end-to-end framework that based on GCN and called The Multi-Depth Graph Convolution Networks (M-GCN).

For GCN models the main purpose is to learn a function of features on a graph $G = (V, E)$. As input data it can take the follows:

1) a feature description x_i for every node i summarized in a feature matrix $X[N \times D]$ (N: number of nodes, D: number of input features);

2) an adjacency matrix A as a representative description of structure of the graph.

The model produces output feature matrix $Z[N \times F]$ on a node-level (F : the number of node's output features).

Then, every layer of the neural network can be described as a non-linear function:

$$H^{(l+1)} = f\left(H^{(l)}, A\right), \quad (2)$$

where $H^{(0)} = X$ and $H^{(L)} = Z$ (L – is a number of layers).

Layer-wise propagation can be written in the simplest form of a rule as:

$$f\left(H^{(l)}, A\right) = \sigma\left(AH^{(l)}W^{(l)}\right), \quad (3)$$

where $W^{(l)}$ is a weight matrix for the l-th layer of neural network and $\sigma(\cdot)$ is a non-linear activation function (like the ReLU). Notwithstanding the simplicity of this model, it is already quite powerful.

Graph Convolutional layer-wise propagation rule in vector form can be defined as:

$$h_{v_i}^{(l+1)} = \sigma\left(\sum_j \frac{1}{c_{ij}} h_{v_j}^{(l)} W^{(l)}\right). \quad (4)$$

GCN at general uses the sum of normalized neighbors embeddings:

$$h_v^{(k)} = \sigma\left(w^{(k)} \sum_{v \in \mathcal{N}(u) \cup \{u\}} \frac{h_v}{\sqrt{|\mathcal{N}(u)||\mathcal{N}(v)|}}\right). \quad (5)$$

The GCNs has several weaknesses and shortcomings especially in the part of usage train methods and optimization techniques. The first, GCN demands complete Laplacian matrix for the whole graph, which is consuming for giant graphs. Therefore, it lacks the capability for inductive learning cause the GCN had been trained independently for a fixed graph.

GAT – Graph Attention Network – neural network architectures that works on data with a graph structure described in [21]. The basic idea is to assign an attention weight or importance to each neighbor, which is used to weigh this neighbor's influence during the aggregation step.

A single graph attentional layer is the unique layer utilized throughout all of the GAT architectures. There is used the particular attentional setup but the framework is agnostic to the particular choice of attention mechanism.

A linear transformation with parameterizing by a weight matrix $W \in \mathbb{R}^{F'} \times \mathbb{R}^{F}$ is applied to each node as an initial step. After that self-attention on the nodes has been performed – a shared attention mechanism:

$$\alpha : W \in \mathbb{R}^{F'} \times \mathbb{R}^{F'} \to \mathbb{R}, \tag{6}$$

and computes attention coefficients that demonstrate the features significance of node v to node u . (These coefficients obtained at once and after being normalized used to compute final output features):

$$e_{uv} = a\left(W\overrightarrow{h_u}, W\overrightarrow{h_v}\right). \tag{7}$$

The benefit of the model is the knowledge about each node on every other node, without structural information. The graph structure injected into this mechanism by carrying out the masked attention. The model computes just coefficients e_{uv} for nodes $v \in \mathcal{N}_u$ where $\mathcal{N}_u$ is neighborhood around node i.

GAT is a single-layer feedforward neural network with an attention mechanism. Such network is parametrized by a weight vector, and normalized with SoftMax function across all variants of v and applying the LeakyReLU nonlinearity.

Weighted sum of the neighbors defined as:

$$m_{\mathcal{N}(u)} = \sum_{v \in \mathcal{N}(u)} a_{u,v} h_v, \tag{8}$$

where $a_{u,v}$ – the attention weights that denotes the attention on neighbor $v \in \mathcal{N}(u)$ when we are aggregating information at node u and defined as follows:

$$a_{u,v} = \frac{\exp\left(a^T\left[wh_u \oplus wh_v\right]\right)}{\sum_{v' \in \mathcal{N}(u)} \exp\left(a^T\left[wh_u \oplus wh_v\right]\right)}. \tag{9}$$

Then after applying a nonlinearity σ the resulting output features of every node can be received by:

$$h_u^{'} = \sigma\left(\sum_{v \in \mathcal{N}_u} a_{uv} W h_v\right). \tag{10}$$

GraphSAGE – is a general inductive framework based on a graph neural network. It generates embeddings obtained by sampling and aggregating features from a node's neighborhood [22].

GraphSAGE is a exhaustive enhancement of initial GCN. To solve the problem with a big Laplacians for a giant real graphs GraphSAGE replaced it with learnable aggregation functions. This function is a key to carry out message passing and generalize to hidden nodes. GraphSAGE could create embeddings for hidden nodes using learned aggregation and propagation functions. Also, GraphSAGE can expanse the alleviate receptive field by using a sampling on neighbors. The process of embeddings generation take as input data graph $G = (V, E)$ and all node's features $\{x_1, ..., x_v\}$. There are vector representations of all nodes as output z_v.

On 1 step – GraphSAGE aggregating the neighbors feature vectors for each node $h_{\mathcal{N}(v)}^k$;

2 step – concatenating currents node's representations h_v^k with the aggregated vector $h_{\mathcal{N}(v)}^k$ for node's neighborhood;

3 step – concatenated vector is fed through a fully connected layer with nonlinear activation function σ.

GraphSAGE does not use a full set of neighbors, but uses a set of neighbors of a fixed size by uniform sampling.

$$h_{\mathcal{N}_v}^t = AGGREGATE_t\left(\left\{h_u^{t-1}, \forall u \in \mathcal{N}_v\right\}\right), \tag{11}$$

$$h_v^t = \sigma\left(W^t \cdot \left[h_v^{t-1} \,||\, h_{\mathcal{N}_v}^t\right]\right). \tag{12}$$

The aggregation of the neighbor representations can be performed by a some of aggregator structures: MEAN, LSTM, Pooling.

V. Experimental results

The User Preference-aware Fake News Detection dataset (UPFD) was used to study the application of the proposed models and qualitative modeling of the fake news detection process [26].

It has been used the dataset which includes networks on Twitter for propagation fake&real news. It is built according to verification information from Politifact and Gossipcop [27]. The graphs of retweeting news were obtained by FakeNewsNet. There were pulled around 20 million tweets from users who engaged in fake news propagation in FakeNewsNet and then it was generated node features.

The data set consists of four types of node characteristics:

768-dimensional BERT features obtained through the use of pre-trained NLP-model BERT;

300-dimensional SpaCy features obtained through the use of pre-trained NLP-model SpaCy word2vec;

10-dimensional feature profiles derived from a Twitter account profile;

310-dimensional content features consisting of a 300-dimensional representation of a user's word2vec comment along with a 10-dimensional profile feature.

The simulation was performed using the PyTorch Geometric framework, which was used to develop the graph neural models proposed above. В якості функції активації використовувалась RELU. The Adam algorithm was used to train the models, which is better suited for sparse data and with dynamic learning speed (lr = 0.001, weight_decay = 0.01). As a function of losses – negative log likelihood loss. The result is passed to the logarithmic softmax activation function. GNN models have been train for 100 epochs. The results of training of the considered models for comparison of fake news detection are given in table 1.

The obtained results show the most up-to-date indicators on the UPFD data set, obtained by the models according to the discussion in Section 4. GAT and GraphSAGE models due to improvements and demonstrating an increase in accuracy by 3.3-3.45% due to context-based similarity, information aggregation.

TABLE I. COMPARING ACCURACY OF FAKE NEWS DETECTION

	GCN		GAT		GraphSAGE	
Politifact	Loss	Test_acc	Loss	Test_acc	Loss	Test_acc
profile	0.2587	0.7873	0.1544	0.7557	0.0476	0.8009
spaCy	0.0417	0.7907	0.0415	0.7919	0.0266	0.8100
BERT	0.0079	0.8371	0.0071	0.8326	0.0013	0.8462
content	0.0560	0.8869	0.0363	0.8959	0.0180	0.8978
Gossipcop	Loss	Test_acc	Loss	Test_acc	Loss	Test_acc
profile	0.2441	0.9038	0.1890	0.9140	0.1633	0.9258
spaCy	0.1010	0.9634	0.1129	0.9597	0.0584	0.9681
BERT	0.0347	0.9660	0.0170	0.9698	0.0135	0.9757
content	0.1082	0.9663	0.0822	0.9773	0.0698	0.9801

The results obtained in the experiment demonstrate that the suggested approach has the potential for application in inductive conditions and that more predictiveness can be used by observing the entire neighborhood.

CONCLUSIONS

We have introduced advanced networks focusing on graphs that work on structured graph data using masked layers of self-awareness. The level of attention to graphs used in these networks is computate productivity. It doesn't necessitate extensive operations with large matrix and it can be paralleled for full graph. Proposed network allows (not clearly) to attribute various values to different nodes in the neighborhood, interacting with neighborhoods of different sizes. It does not require to know the entire structure of the graphs in advance – thus solving many hypothetical issues using previous spectrum-based approaches. Our attention-grabbing models have successfully achieved or matched modern performance on the UPFD dataset.

In addition, usage of extending method which take performing graph classification instead of node classification can also be appropriate for application. And lastly, model improvement by including the boundary functions (that possible to point out a nodes links) would provide us to solve a wider range of problems.

REFERENCES

[1] G.L. Ciampaglia, P. Shiralkar, L.M. Rocha, J. Bollen, F. Menczer, A. Flammini. "Computational fact checking from knowledge networks". arXiv:1501.03471, 2015.

[2] V.L. Rubin, T. Lukoianova. "Truth and deception at the rhetorical structure level". Journal of the Association for Information Science and Technology, vol. 66, no.5, 2015, pp. 905-917.

[3] V. P´erez-Rosas, B. Kleinberg, A. Lefevre, R. Mihalcea. "Automatic detection of fake news". arXiv:1708.07104, 2017.

[4] K. Shu, S. Wang, H. Liu. "Beyond news contents: The role of social context for fake news detection". In: Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, 2019.

[5] S. Gupta, R. Thirukovalluru, M. Sinha, S. Mannarswamy. "Cimtdete A community infused matrix-tensor coupled factor- ization based method for fake news detection". arXiv:1809.05252, 2018.

[6] L. Cui, K. Shu, S. Wang, D. Lee, H. Liu. "Defend: A system for explainable fake news detection". In: Proceedings of the 28th ACM International Conference on Information and Knowledg Management, 2019.

[7] K. Sharma, F. Qian, H. Jiang, N. Ruchansky, M. Zhang, Y. Liu. "Combating fake news: A survey on identification and mitigation techniques". ACM TIST, 10(3), 2019. pp. 1-42.

[8] M. S. Looijenga. "The Detection of Fake Messages using Machine Learning". 29 Twente Student Conference on IT, Netherlands: essay.utwente.nl, 2018.

[9] Prabhjot Kaur et al. "Hybrid Text Classification Method for Fake News Detection." International Journal of Engineering and Advanced Technology (IJEAT) , 2019, pp. 2388-2392,

[10] Z. Khanam, S. Alkhaldi. "An Intelligent Recommendation Engine for Selecting the University for Graduate Courses in KSA: SARS Student Admission Recommender System." In: Smys S., Bestak R., Rocha Á. (eds) Inventive Computation Technologies. ICICIT 2019. Lecture Notes in Networks and Systems, vol 98. Springer, 2019.

[11] S. Gilda. "Evaluating machine learning algorithms for fake news detection." 15th Student Conference on Research and Development (SCOReD), IEEE, 2017, pp. 110-115.

[12] Yumeng Qin et al. "Predicting Future Rumours". Chinese Journal of Electronics, vol. 27, 2018, pp. 514 - 520.

[13] Z. Khanam, B. Alwasel, H. Sirafi, M. Rashid. "Fake News Detection Using Machine Learning Approaches". IOP Conf. Series: Materials Science and Engineering, doi:10.1088/1757-899X/1099/1/012040, 2021.

[14] N. Ruchansky, S. Seo, and Y. Liu. "A hybrid deep model for fake news detection". arXiv:1703.06959, 2017.

[15] Xinyi Zhou, Jindi Wu, and Reza Zafarani. "SAFE: Similarity-Aware Multimodal Fake News Detection". In Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer, 2020, pp. 354–367.

[16] Ye Zhang and Byron Wallace. "A sensitivity analysis of (and practitioners'guide to) convolutional neural networks for sentence classification". arXiv:1510.03820, 2015.

[17] F. Monti, F. Frasca, D. Eynard, D. Mannion, M. Bronstein. "Fake news detection on social media using geometric deep learning". ICLR Workshop, 2019.

[18] Th. N. Kipf and M. Welling. "Semi-supervised classification with graph convolutional networks". arXiv:1609.02907, 2017.

[19] Y. Han, Sh. Karunasekera, Ch. Leckie. "Graph Neural Networks with Continual Learning for Fake News Detection from Social Media". arXiv:2007.03316, 2020.

[20] Zh. Ying, J. You, Ch. Morris, X. Ren, W. Hamilton, J. Leskovec. "Hierarchical graph representation learning with differentiable pooling. In Advances in neural information processing systems". arXiv:1806.08804, 2018.

[21] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, Yo. Bengio. "Graph Attention Networks". arXiv:1710.10903, 2018.

[22] W. L. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," In NIPS 2017, 2017, pp. 1024–1034.

[23] Guoyong Hu, Ye Ding , Shuhan Qi, Xuan Wang, Qing Liao. "Multi-Depth Graph Convolutional Networks for Fake News Detection", In Computer Science, NLPCC, 2019.

[24] Zh. Ying, J. You, Ch. Morris, X. Ren, W. Hamilton, J. Leskovec. "Hierarchical graph representation learning with differentiable pooling. In Advances in neural information processing systems". arXiv:1806.08804, 2018

[25] W. L. Hamilton. "Graph Representation Learning". Synthesis Lectures on Arti_cial Intelligence and Machine Learning, vol. 14, no. 3, P. 159, 2020.

[26] Yingtong Dou, Kai Shu, Congying Xia, Philip S. Yu, Lichao Sun: "User Preference-aware Fake News Detection". arXiv:2104.12259, 2021.

[27] PolitiFact Dataset, Available at: https://paperswithcode.com/dataset/politifact.