Pattern Recognition Letters 151 (2021) 135–139

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Effective fake news detection using graph and summarization techniques☆

Gihwan Kim, Youngjoong Ko*

Department of Computer Science and Engineering, Sungkyunkwan University, 2066, Seobu-ro, Jangan-gu, Suwon-si, Gyeonggi-do, Republic of Korea

ARTICLE INFO

Article history:
Received 19 October 2020
Revised 17 June 2021
Accepted 31 July 2021
Available online 10 August 2021

Edited by : Jiwen Lu

MSC:
41A05
41A10
65D05
65D17

Keywords:
Fake news detection
Graph neural networks
Summarization
Deep neural networks

ABSTRACT

Nowadays, fake news is widely spreading in various media, and this fake information is causing serious damage in many areas. Therefore, there is an increasing need to accurately detect fake news to prevent such damage. In this paper, we propose a novel method that uses graph and summarization techniques for fake news detection. Our proposed method represents the relationship of all sentences in a graph structure to accurately understand the context information of the document. Accordingly, the relationship between sentences in the graph is calculated as a score through the attention mechanism. Then, the summarization technique is used to reflect the sentence subject information in the graph update process. Our proposed method shows better performance than Karimi's and BERT based models by approximately 10.34%p and 3.72%p, respectively.

1. Introduction

Nowadays, people receive a large amount of information through various media. However, some of this information are fake with a different purpose other than telling the truth, and people fail to validate such information and identify it as true. This fake information confuses people and causes social, economic, and national damages. Therefore, the need to detect fake information is of paramount importance to prevent such damages. In fact, fake news detection has been researched industrially and academically [1,21,22].

Previous studies on fake news detection have used internal and external information in documents. The detection of fake news using internal information is mainly performed by analyzing linguistic features in news, such as contextual information [9], writing styles and consistency [14], and relational structure between sentences [7]. By contrast, the detection of fake news using external information is performed by analyzing metadata, such as aspects of news [12] and user profiles of people spreading the news [10]. However, the process of getting all external information on the news involves considerable time and cost. Meanwhile, understanding the internal information of documents is more basic and important. Therefore, we propose an effective fake news detection method using graph and summarization techniques. We represent the relationship of all sentences in a document in a graph structure and update them to accurately understand contextual information. When updating the graph, we use the summarization technique to weigh the important sentence information containing the subject.

All sentences in a document are closely related to one another. Therefore, to better understand the document, we construct a context graph with all sentences that are connected to one another and generate a contextualized sentence embedding that reflects other sentences information. At this time, because the relationship ratios of all sentences are different, the ratio of sentence information reflection is calculated through the attention mechanism. In the context graph, nodes consist of initial sentence embeddings for all the sentences in a document, and the weight of an edge is estimated by an attention score between two end nodes of the edge. We assume that all the nodes are connected because all the sentences in a document are strongly related to one another. Thus, each node is updated to contextualize the sentence embedding, which reflects the information of the connected nodes sentence embeddings. The contextualized sentence embedding of a node is

☆ Handling Editor Emmanouil Benetos

* Corresponding author.
E-mail address: yjko@skku.edu (Y. Ko).

https://doi.org/10.1016/j.patrec.2021.07.020
0167-8655/

computed by the sum of the products of the attention score between the node and its connected node and the sentence embedding of the connected node. By iterating the process of updating nodes in the context graph, an optimized contextual information embedding is generated.

The subject of a document is a very important information in identifying the content of the document, and fake information is generally relevant to the subject. We use the summarization technique [6] to effectively identify effective subject information and detect fake news. By using the summarization technique, we create ranking and ranking scores based on sentences that contain most information about the subject. The ranking scores consequently influence the edges attention score for the construction of the context graph.

For a performance comparison with the proposed method, we implemented a baseline model that detects fake news with the sum of all sentence embeddings in the document. Our proposed model shows a performance accuracy of 92.53%, which is 12.68%p better than the performance of the baseline model. It also shows 10.34%p and 3.72%p better performances compared to other models, Karimi and Tang [7] and BERT based model, that use the same dataset and a dependency tree structure among sentences in a document.

The remainder of the paper is organized as follows: Section 2 briefly introduces related work. Section 3 describes the no-tation used. Section 4 describes our proposed method. Section 5 reports our experimental results. Section 6 reports the analysis of our experiment. Finally, Section 7 draws the conclusions.

2. Related work

Internal information-based approaches rely on the text con-tent to detect the truth of news articles, which are usually comprised of long texts. To learn internal information, various methods have been studied. Castillo et al. [1] detected fake news by extracting topic information using Term Frequency - Inverse Document Frequency (TF-IDF). Potthast et al. [14] detected fake news by analyzing the writing style and consistency of content between real and fake news. Karimi and Tang [7] examined the content of a document by constructing the relationship of the sentences in the document into a dependency tree to detect fake news. Silva et al. [17] proposed a novel framework that jointly preserved domain-specific and cross-domain knowledge to detect fake news from different domains, and used an unsupervised technique to select a set of unlabelled informative news records for manual labelling.

External information-based approaches model the process of spreading contents, which typically consists of long text, the user characteristics that propagate the contents, or the response and comments of users who read the contents. Qian et al. [15], Ruchansky et al. [16] used news contents, user response, comments, and SNS data to detect fake news while [13,18] did an embedding graph to detect the fake news more accurately. Monti et al. [12] attempted to detect fake news by analyzing metadata, such as aspects of news propagation and [10] analyzed the information of users who spread the news to detect fake news. Cui et al. [2] utilized correlation between news articles and text of user comments to detect fake news and generate reasonable explanation.

Hybrid information-based approaches model to detect fake news suggested to use both text and image feature information in the news article. Wang et al. [20] constructed an end-to-end multi-modal model for fake news detection for extracting the textual and visual features from news article. Besides, Cui et al. [3] presented a fake news detection method based on news contents, images, and sentiments of user comments.

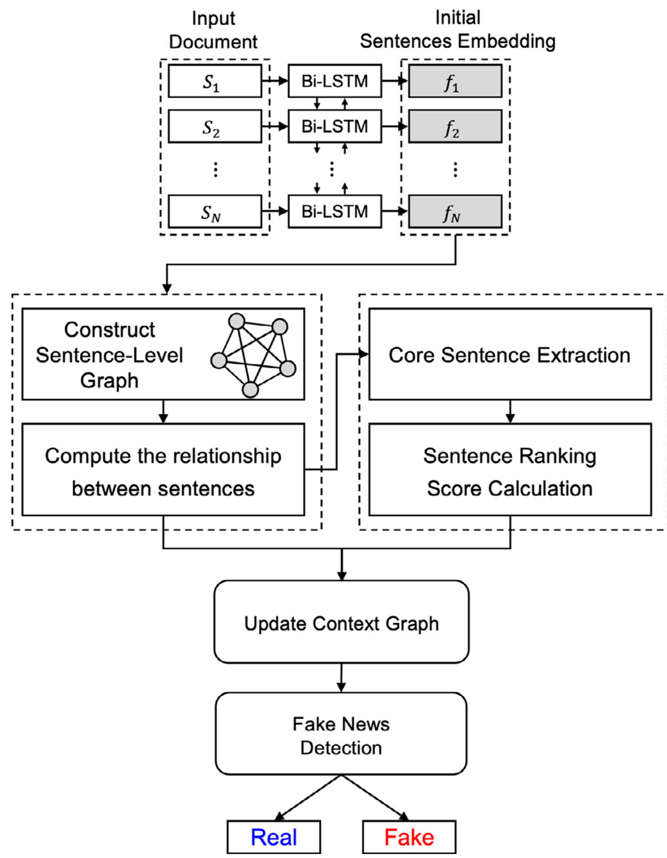

Fig. 1. Architecture of our proposed model.

In the present study, we compare the performance of our system with that of the system proposed by Karimi and Tang [7], which detects fake news by constructing documents among previous studies using internal information.

3. Notations

We have a corpus D of fake and real news documents. Let a document $d \in D$ contain N sentences $s_1, s_2, \ldots s_N$ and each sentence $s_i \in d$ includes words, i.e., $word_1, word_2, \ldots word_l$, where l denotes the number of words in sentence s_i. By using bi-directional long/short term memory (Bi-LSTM), we can obtain the initial sentence embeddings of all sentences in a document, i.e., $\mathcal{H} = \{h_1, h_2, \ldots h_N\} \in \mathbb{R}^{N \times dim}$.

4. Proposed method

We propose a novel fake news detection method using graph and summarization techniques (see Fig. 1). Our method consists of three components. The first one is *graph construction using the attention mechanism* for representing the relationship between all sentences using a graph. The second one is *core sentence extraction* for ranking sentences with subject information using the summarization technique. The third one is *update context graph* for calculating more accurate contextual information. The fourth one is *fake news detection* for discriminating fake news from several documents.

4.1. Graph construction using the attention mechanism

To model the relationship between sentences and their contextual information, we construct a context graph $\mathcal{G} = (F, E)$. The graph $\mathcal{G}$ is composed of a node (*i.e.*, $F = \{f_1, \ f_2, \ \dots f_N\}$) and the edge (*i.e.*, $E = \{e_{1,2}, \ e_{ij}, \ \dots e_{N-1,N}\}$). Each node is represented by sentence embedding H, and the edge between the ith node and the jth node is represented by $e_{i,j}$, and its weight $w_{i,j}$ represents the relational strength of the ith sentence and jth sentence. Because all sentences in a document are strongly related to one another, we consider that $\mathcal{G}$ is a fully connected graph. To reflect the different relational strengths between sentences, each edge $e_{i,j}$ is associated with a weight $w_{i,j}$, and the weight is derived by the attention mechanism between sentence embeddings h_i and h_j of two nodes f_i and f_j as follows (Eqs. (1)–(3)):

$$x_i = ReLU(W_1 h_i + bias_1) \tag{1}$$

$$x_j = ReLU(W_2 h_j + bias_2) \tag{2}$$

$$w_{i,j} = x_i U x_j \tag{3}$$

where $W_1, W_2 \in \mathbb{R}^{m \times dim}$, and $x \in \mathbb{R}^m$. The dimension size of sentence embeddings can be reduced to avoid overfitting without weakening the capacity of the LSTM [5] using Eqs. (1) and (2). U is the weight matrix to compute the relational strength between x_i and x_j. $w_{i,j}$ represent an edge weight between pair of nodes f_i and f_j in the graph $\mathcal{G}$.

4.2. Core sentence extraction

Based on the constructed context graph, the subject information of a sentence is estimated by summing the attention scores (edge scores) between the node of the sentence and its adjacent nodes, and a sentence that contains the most subject information is extracted as a core sentence. *Core_Sent* represents the extracted sentence that contains the most subject information of the subject.

$$Core\_Sent = argmax_{1 \le i \le N} \sum_{j=1, j \ne i}^{N} w_{i,j} \tag{4}$$

Based on the cosine similarity values between the core sentence and all other sentences, all sentences in the document are divided into a subject relevant sentence set and an irrelevant sentence set. The subject relevance score of each word is based on the frequency of appearance in the subject relevant and irrelevant sentence sets, and then the word relevance score RS_d is calculated by Eq. (5).

$$\begin{aligned} RS_d(word) &= \log \frac{p_d \times (1-q_d)}{(1-p_d) \times q_d} \\ &= \log \frac{(r_d+0.5)(S_d-s_d+0.5)}{(R_d-r_d+0.5)(s_d+0.5)} \end{aligned} \tag{5}$$

where p_d and q_d are the probabilities that a word appears in the subject relevant and irrelevant sentences set in a document d. Respectively, R_d and S_d are the number of subject relevant sentences and irrelevant sentences in document d, respectively. r_d and s_d are the number of subject relevant and irrelevant sentences that include the word in a document d, and 0.5 is a naòve smoothing factor used to avoid a zero-denominator or log-zero [6]. In this study, the top 30% of sentences with subject similarity were selected as the relevant sentence set and the others as the irrelevant sentence set [6]. Afterward, the sentence score is calculated with the sum of the relevance score of words included in the sentence [6] (see Fig. 2).

$$Score(s_i) = \sum_{word \in s_i} RS_d(word) \tag{6}$$

4.3. Update context graph

We use the sentence ranking to calculate the sentence score and then use it to update the weight of the edge in the context graph. The weight of the edge represents the reflection rate of the subject information and context information. *Sentence_Rank* is the order of sentences by the sum of the attention scores between each sentence and other sentences; the first ranked sentence has the highest sum of attention scores.

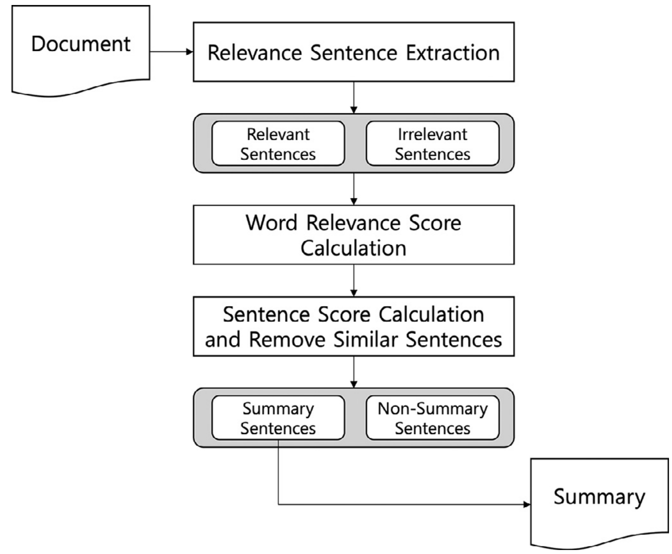

Fig. 2. Summarization techniques used in the core sentence extraction.

$$Rank\_Score(s_i) = 1 - \frac{Sentence\_Rank(s_i) - 1}{N} \tag{7}$$

$$\begin{aligned} &w'_{i,j} = w_{i,j} * Rank\_Score(s_i) \\ &(for\ i, j \ = \ 1, \ 2, \ \dots, \ N\ and\ i \ne j) \end{aligned} \tag{8}$$

The sentence ranking score is calculated by Eq. (7), and the weight of all edges is updated by Eq. (8). Subsequently, we can construct subject contextualized sentence embeddings using the updated context graph. The subject contextualized sentence embedding (i.e., h'_i) is created by the weighted sum of the initial sentence embedding of the current sentence (i.e., h_i) and those of the other adjacent sentences (i.e., h_j).

$$h'_i = h_i + \sum_{j=1, j \ne i}^{N} w'_{i,j} * h_j \tag{9}$$

To set up the number of iterations, we conducted experiments for various models according to the number of iterations of the updated context graph, and the best performance was obtained when the number of iterations was 2.

4.4. Fake news detection

Finally, we created a document embedding doc with the average of all nodes (the contextualized sentence embedding) in the context graph. The document embedding doc is was then fed into a multi-layer feedforward neural network to predict its label for the binary classification, i.e., real or fake, as a vector $\hat{y}$.

$$doc = \frac{1}{N} \sum_{i=1}^{N} h'_i \tag{10}$$

$$\hat{y} = Softmax(sigmoid(W_3 doc + bias_3)) \tag{11}$$

The cross-entropy loss function was used to optimize our neural network.

$$loss = -(y \log(p_{fake}) + (1-y)\log(p_{real})) \tag{12}$$

Table 1
Size of the HDSF dataset.

Label	HDSF dataset		
	Train	Validation	Test
Fake	3226	67	67
Real	3226	67	67

where y is the golden standard data to in the document and p_{fake} and p_{real} are the probabilities of a document being fake or real of $\hat{y}$, respectively.

5. Experiments

5.1. Dataset

Because the proposed model uses a graph model and internal information, its target application domain to detect fake news is selected as news articles. Thus we chose the HDFS dataset created by Karimi and Tang [7] for our experiments. Table 1 shows the size of the dataset. The HDSF dataset is larger than other datasets because it was created by a combination of several datasets, such as Kaggle, Buzzfeed, PolitiFact. The dataset consists of 3360 real documents and 3360 fake documents. We followed the HDFS data split: 6452 documents for the training data, 134 for the validation set, and 134 ones for the test set. Each data split contains an even number of fake and real documents. In addition, we performed the experiments on a five-fold cross-validation because the original HDFS dataset contains too small test documents.

5.2. Experimental settings

We used word2vec embeddings [11], which are pre-trained by Google as initial word embeddings, and set the Bi-LSTM hidden unit size to 200. We set the maximum number of epoch to 300, mini-batch size as 40, and dropout as 30% in each experiment. We used the Adam optimizer [8] and set the initial learning rate as 0.001 and reduced it by 10 times for every 50 epochs. In addition, we used early stopping to maximum number of epoch, 300, to prevent overfitting. Accuracy is employed as the metric of performance.

5.3. Models

The proposed model is compared to the baseline model and Karimis model [7] to prove a superiority of our proposed model in this subsection.

- **Karimi's Model** [7] A fake detection method to predicts whether a document is fake or real by constructing relationships of each sentence in a document using the hierarchical discourse-level dependency tree.
- **Baseline Model** The baseline model obtains the document embedding only using Bi-LSTM and predicts whether it is fake or real through the feedforward neural network.
- **Graph Model** This graph model connects all sentences in a document by constructing a graph structure. Fake news is detected by creating sentence embedding that only reflects contextual information on a graph.
- **Graph + Summarization Model (Proposed)** The graph + summarization model is our final proposed model. It classifies fake or real documents after representing a document embedding that reflect the contextual and subject information on the graph structure through the summarization technique.

Table 2
Comparison results of the proposed model and other models, such as Karimi's one and BERT. The $*$ and $\star$ denote the differences between the proposed model and Karimi's or BERT.

Model	Data	
	HDSF	Cross validation
RST	67.68	–
LIWC	70.26	–
N-grams	72.37	–
BiGRNN-CNN	77.06	–
BERT	<u>88.81</u>	–
Karimi's	<u>82.19</u>	<u>81.71</u>
Baseline	79.85	79.27
Graph Model	88.81	85.44
Graph + Summarization (Proposed)	**92.53** (**+10.34%p**$*$) (**+3.72%p**$\star$)	**89.4** (**+7.69%p**$*$)

Table 3
Model performance comparison by iterations.

Data	Model and iteration		
	Graph + summarization (iteration 1)	Graph + summarization (iteration 2)	Graph + summarization (iteration 3)
HDSF	91.04	**92.53**	89.55
Cross-validation	88.33	**89.4**	88.05

5.4. Experiment results

Table 2 shows a comparison of the experiment results for the proposed model and other models. The experiment results of RST, LIWC, N-grams, and BiGRNN-CNN were reported as comparison models in Karimi and Tang [7] and the BERT model was implemented by ourselves according to Devlin et al. [4]. The proposed model achieved a better performance than baseline, the Karimi's model, and the BERT model by approximately 12.68%p, 10.34%p, and 3.72%p, respectively. In addition, we obtained a 3.72%p improvement when summarization technique was used in the graph model. Moreover, the final proposed model showed the best performance in the five-fold cross-validation experiments as well.

In addition, we did several experiments for the proposed model based on the number of iterations of context graph updates. Table 3 shows the performance changes of the models based on the number of iterations. The best performance was obtained when the number of iterations was 2 as follows:

As the graph was updated by increasing the number of iterations, the performance of our model becomes worse. It is because too many graph updates make a spread of noisy features in some sentences to other sentences.

6. Analysis

Herein, we introduce two analyses by an additional experiment for subject consistency detection in our model and dataset. The proposed method attempted to effectively detect subject information in fake news and the consistency of subject information is also important to detect fake news. In the first analysis, we verify our process of updating the graph in the proposed model, which uses a fully connected graph, by comparing other model using a 50% connected graph, which uses edges with only the top 50% with high attention scores. As a result, the proposed model with a fully connect graph showed 1.29%p higher performance than the model with a 50% connected graph. We think it means that a fully connect graph is more useful to detect fake news. Secondly, we observed the variances of the attention scores in the fake and real

Table 4
Variance/standard deviation of attention scores.

	Variance	Standard deviation
Fake	8.6	2.93
Real	5.41	2.33

documents (see Table 4). The variance in the fake documents is higher than that of the real documents and it means that the fake documents has an inconsistent subject distribution.

7. Conclusion

In this study, we propose a novel graph-based fake news detection method using graph and summarization techniques. Our model shows that contextual and subject information in documents are helpful in detecting fake news. Our final proposed model achieves better performance than the baseline and comparison models by approximately 12.68%p and 10.34%p, respectively.

As our future work, we plan to conduct further experiments with much larger datasets. In addition, we plan to apply the transformer-based language models [4,19] and a variety of summarization techniques to detect fake news.

Declaration of Competing Interest

Authors declare that they have no conflict of interest.

Acknowledgments

This work was supported in part by Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea Government (MSIT) (No. 2020-0-00368, A Neural-Symbolic Model for Knowledge Acquisition and Inference Techniques), in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2020R1A2C2100362), and in part by the MSIT (Ministry of Science and ICT), Korea, under the ICT Creative Consilience program (IITP-2020-0-01821) supervised by the IITP.

References

[1] C. Castillo, M. Mendoza, B. Poblete, Information credibility on twitter, in: Proceedings of the 20th International Conference on World Wide Web, 2011, pp. 675–684.
[2] 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 Knowledge Management, 2019, pp. 2961–2964.
[3] L. Cui, S. Wang, D. Lee, Same: sentiment-aware multi-modal embedding for detecting fake news, in: Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 2019, pp. 41–48.
[4] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: pre-training of deep bidirectional transformers for language understanding, in: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 2019, pp. 4171–4186.
[5] T. Dozat, C.D. Manning, Simpler but more accurate semantic dependency parsing, in: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), Association for Computational Linguistics, Melbourne, Australia, 2018, pp. 484–490, doi:10.18653/v1/P18-2077.
[6] H. Jeong, Y. Ko, J. Seo, How to improve text summarization and classification by mutual cooperation on an integrated framework, Expert Syst. Appl. 60 (2016) 222–233.
[7] H. Karimi, J. Tang, Learning hierarchical discourse-level structure for fake news detection, in: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), Association for Computational Linguistics, Minneapolis, Minnesota, 2019, pp. 3432–3442, doi:10.18653/v1/N19-1347.
[8] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, arXiv preprint arXiv:1412.6980(2014).
[9] O. Levi, P. Hosseini, M. Diab, D. Broniatowski, Identifying nuances in fake news vs. satire: using semantic and linguistic cues, in: Proceedings of the Second Workshop on Natural Language Processing for Internet Freedom: Censorship, Disinformation, and Propaganda, Association for Computational Linguistics, Hong Kong, China, 2019, pp. 31–35, doi:10.18653/v1/D19-5004.
[10] Y.-J. Lu, C.-T. Li, GCAN: graph-aware co-attention networks for explainable fake news detection on social media, in: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Online, 2020, pp. 505–514.
[11] T. Mikolov, K. Chen, G. Corrado, J. Dean, Efficient estimation of word representations in vector space, arXiv preprint arXiv:1301.3781(2013).
[12] F. Monti, F. Frasca, D. Eynard, D. Mannion, M.M. Bronstein, Fake news detection on social media using geometric deep learning, arXiv preprint arXiv:1902.06673(2019).
[13] J.Z. Pan, S. Pavlova, C. Li, N. Li, Y. Li, J. Liu, Content based fake news detection using knowledge graphs, in: International Semantic Web Conference, Springer, 2018, pp. 669–683.
[14] M. Potthast, J. Kiesel, K. Reinartz, J. Bevendorff, B. Stein, A stylometric inquiry into hyperpartisan and fake news, in: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Melbourne, Australia, 2018, pp. 231–240, doi:10.18653/v1/P18-1022.
[15] F. Qian, C. Gong, K. Sharma, Y. Liu, Neural user response generator: fake news detection with collective user intelligence.
[16] N. Ruchansky, S. Seo, Y. Liu, CSI: a hybrid deep model for fake news detection, in: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, 2017, pp. 797–806.
[17] A. Silva, L. Luo, S. Karunasekera, C. Leckie, Embracing domain differences in fake news: cross-domain fake news detection using multi-modal data(2021).
[18] A. Silvaa, Y. Hana, L. Luoa, S. Karunasekeraa, C. Leckiea, Embedding partial propagation network for fake news early detection.
[19] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need, in: Advances in Neural Information Processing Systems, 2017, pp. 5998–6008.
[20] Y. Wang, F. Ma, Z. Jin, Y. Yuan, G. Xun, K. Jha, L. Su, J. Gao, EANN: event adversarial neural networks for multi-modal fake news detection, in: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2018, pp. 849–857.
[21] R. Yan, I.E. Yen, C.-T. Li, S. Zhao, X. Hu, Tackling the achilles heel of social networks: influence propagation based language model smoothing, in: Proceedings of the 24th International Conference on World Wide Web, 2015, pp. 1318–1328.
[22] F. Yang, Y. Liu, X. Yu, M. Yang, Automatic detection of rumor on Sina Weibo, in: Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics, 2012, pp. 1–7.