World Wide Web (2022) 25:1243–1258
https://doi.org/10.1007/s11280-022-01032-3

Leverage knowledge graph and GCN for fine-grained-level clickbait detection

Mengxi Zhou[1] **· Wei Xu**[1] **· Wenping Zhang**[1] **· Qiqi Jiang**[2]

Accepted: 24 February 2022 / Published online: 16 March 2022

Abstract

Clickbait is the use of an enticing title as bait to deceive users to click. However, the corresponding content is often disappointing, infuriating or even deceitful. This practice has brought serious damage to our social trust, especially to online media, which is one of the most important channels for information acquisition in our daily life. Currently, clickbait is spreading on the internet and causing serious damage to society. However, research on clickbait detection has not yet been well performed. Almost all existing research treats clickbait detection as a binary classification task and only uses the title as the input. This shallow usage of information and detection technology not only suffers from low performance in real detection (e.g., it is easy to bypass) but is also difficult to use in further research (e.g., potential empirical studies). In this work, we proposed a novel clickbait detection model that incorporated a knowledge graph, a graph convolutional network and a graph attention network to conduct fine-grained-level clickbait detection. According to experiments using a real dataset, our novel proposed model outperformed classical and state-of-the-art baselines. In addition, certain explainability can also be achieved in our model through the graph attention network. Our fine-grained-level results can provide a measurement foundation for future empirical study. To the best of our knowledge, this is the first attempt to incorporate a knowledge graph and deep learning technique to detect clickbait and achieve explainability.

Keywords Knowledge graph · Graph convolutional network · Graph attention network · Clickbait detection

This article is part of the Topical Collection: *Special Issue on Web Intelligence = Artificial Intelligence in the Connected World*
Guest Editors: Yuefeng Li, Amit Sheth, Athena Vakali, and Xiaohui Tao

✉ Wenping Zhang
wpzhang@ruc.edu.cn

Extended author information available on the last page of the article

1 Introduction

Benefiting from the development of the modern internet, especially the mobile internet, online media has become one of the most important information channels for people. This trend has prompted the emergence of many online media giants, such as Twitter, Facebook and TikTok. For these online media giants, network flow (e.g., visits and readership) is one of the most precious fortunes that they are willing to do their best to pursue. Thus, the ability to attract network flow has become a widely used performance indicator for online media workers. In recent years, the market for online media has become saturated, which makes the battle for network flow increasingly fierce. From the perspective of communication and journalism, high-quality media content (e.g., news articles) leads to successful information diffusion and ultimately high network flow. High-quality news materials and professional editing are two irreplaceable prerequisites for high-quality content. However, given that high-quality news materials are limited and professional editing is very energy- and time-consuming, the quantity of high-quality content is also doomed to be limited and far from sufficient. Given that most platforms take “clicks” as the measure of network flow, the process of click induction, known as “clickbait”, has become widespread. Instead of providing high-quality content, clickbait employs enticing headlines to persuade users to click. Usually, relatively low-quality content is observed after clicking, which makes users feel disappointed or fooled. From the perspective of the information-gap theory of curiosity [23], clickbait stimulates curiosity when a user perceives a gap between his or her knowledge and attention. Such a gap generates a “feeling of deprivation labeled curiosity” and motivates the user to acquire the missing information to alleviate the “feeling of deprivation” through clicking the headline (e.g., hyperlink) and reading the content. Given that the primary goal of clickbait is to bait users to click, it is understandable that no quality content is provided when users have already bitten the bait. Thus, in most cases, the quality of the corresponding content is rather low and cannot fulfill the curiosity and alleviate the feeling of deprivation.

Currently, clickbait is spreading on online media worldwide and causing serious damage to the economy and society. On the one hand, clickbait could facilitate the spreading of fake information. One obvious aim of clickbait is to attract users’ attention, which is critical for fast information diffusion in a social system. To reach the highest spreading speed, it is intuitive for conspirators to use clickbait when diffusing fake information (e.g., fake news). Multiple examples have been observed during the COVID-19 pandemic in the past few months. On the other hand, clickbait may decrease social trust, especially the public credibility of media. The media play significant roles in our society, such as information communication and social supervision. Public credibility acts as the foundation of this role, which has been corroded by clickbait in recent years. When induced by clickbait and disappointed by the subsequent low-quality content, the trust between people and media is weakened. Accumulatively, public credibility will be seriously damaged. Due to the cry wolf effect [3], it will be extremely dangerous when people’s attention is truly needed (e.g., public crisis) but people have been exhausted by prior clickbait.

Attempts have been made to detect clickbait in online media. However, most existing research suffers from at least three disadvantages. First, many attempts mixed the concept of clickbait and fake news [1]. Fake news can be defined as false or misleading information presented as news, usually motivated by economic or political benefits. In this process, false or misleading information carried by the news is the key. Information can exert its power only through audiences. To achieve this goal, the content of fake news needs to be meticulously designed and prepared so that audiences can catch the information they hope to transfer easily. In contrast, the aim of clickbait is merely to induce people to click; thus, the aim will have been fulfilled after the clicking. In other words, there is no need for audiences to read any of the subsequent content, and the information carried by the content does not matter at all after the clicking. One possible reason why many researchers mixed these two concepts is the fact that the most fake news employs the trick of clickbait to accelerate the speed of spread, as mentioned above. The mixture of these two concepts will lead to unreliable detection results. Second, previous studies on clickbait have only examined titles [25]. According to the definition, the essence of clickbait is to construct a curiosity gap in the title that cannot be fulfilled by the corresponding content. This indicates that the key characteristic of clickbait is the difference between the "promise" the title makes and the extent of the following redemption. Apparently, only using the title as the input is insufficient and incorrect for clickbait detection. Third, most existing research considers clickbait detection as a binary classification task [18]. Previous research has revealed that the attributes (e.g., degree and type) of clickbait vary [30]. For instance, a low degree of clickbait may enliven the atmosphere under certain conditions, while a high degree of clickbait may anger the reader. Thus, it is necessary and important to perform fine-grained-level clickbait detection other than binary classification.

According to cognition theory [4], prior knowledge plays an important role in human cognition, such as reading. For instance, metaphor is a common rhetorical strategy in literary creation and is also widely adopted by clickbait. Prior knowledge of related concepts is indispensable to comprehending metaphors. Thus, it is necessary and important to adopt prior knowledge in clickbait detection. One practical problem is how to effectively incorporate prior knowledge in limited computation. One reason for many previous studies only using article titles in clickbait detection lies in the difficulty of making use of long and complex article content. Inspired by human memory and thinking, a knowledge graph (KG), which conceptualizes knowledge as entities and the relationships among them, has been proposed [11]. Previous research has verified the effective knowledge representation of KG in various fields [39]. In our research, we constructed a global KG to conceptualize knowledge in online media. Then, a graph convolutional network (GCN) was adopted to predict the clickbait level of a given article. The aim of clickbait is to attract people's attention. Thus, it is intuitive to take attention into consideration in our design. In this research, we developed a graph attention model to enhance our GCN on the learning and prediction process based on KG. Moreover, the combination of GCN and graph attention also brings certain interpretability to our results.

The contributions of our research are fourfold. First, we utilized a knowledge graph to make an enriched representation of online articles and introduced external background knowledge to facilitate content comprehension, which would alleviate the challenge brought by the informality and dynamics of the online media content. Second, we developed a graph convolutional network (GCN) to make deep use of the information provided by the knowledge graph. Compared with a convolutional neural network (CNN), a GCN directly performs convolution operations on irregular networks. It could make use of the information beyond direct spatial connections, simultaneously reducing the computational cost since knowledge presented by KG is much denser than plain images. Third, we proposed a graph attention model to enhance the performance of our GCN, which would also bring certain explainability to our results. Fourth, instead of treating clickbait detection as a binary classification task, we performed fine-grained-level detection, which would be helpful for further related studies, such as investigating the influence of various levels of clickbait on article readership.

The rest of the paper is organized as follows. In Section 2, we review existing research on clickbait detection and summarize related techniques. Our proposed model is described in Section 3. In Section 4, we verify the effectiveness of our proposed model on a real dataset and compare its performance with classical and state-of-the-art algorithms. Finally, we summarize our research and propose future work in Section 5.

2 Related work

2.1 Clickbait detection

In the era of online media, network flow is one of the most valuable resources that platforms pursue. As the battle for network flow is becoming increasingly fierce, clickbait has been widely used to grab users' attention and obtain network flow. However, the side effects of clickbait along with helping platforms obtain network flow have been proven to be very harmful to our society. Thus, many attempts have been made to detect and prevent clickbait. For instance, Chakraborty et al. [5] developed a browser plug-in named "Stop Clickbait" based on a machine learning classifier. They extracted 14 features from headlines and then employed a support vector machine (SVM) to conduct the detection task. Rony et al. [32] designed an embedded model of distributed expressions to recognize clickbait from a large corpus. Dong et al. [10] tested a method that used similarity and quality features for clickbait detection. Pujahari and Sisodia [30] developed a hybrid classification framework that combined different features, sentence structure, and clustering to distinguish clickbait and nonclickbait. Zheng et al. [44] formed a deep model based on the combination of lure and similarity and made reasonable predictions using an adaptive prediction mechanism.

Recently, inspired by the good performance of deep learning on natural language processing (NLP) tasks, researchers have tried to introduce deep learning techniques in the clickbait detection task. For instance, Chawda et al. [6] adopted an RCNN model and enhanced it with LSTM and GRU to capture long-term dependency to achieve high detection performance. Kaur et al. [18] used the embedding model (GloVe) to make rich presentations from clickbait headlines and then utilized CNN-LSTM to detect clickbait.

There is an obvious shortcoming of existing research on clickbait detection—most of them only utilized the headlines and ignored the article content. As mentioned above, it is unreliable to classify an article as a clickbait without considering the article content. Some headlines may look like very "unrealistic" or "shocking" since they create a huge curiosity gap that seems difficult or even impossible to fulfill. However, if their article contents are "rich" and "hot" enough to fill the curiosity gap, they cannot be recognized as clickbait. Worse still, such "hot" articles are usually very important and serious (e.g., breaking news on sudden crisis). If these articles were classified as clickbait and blocked by a clickbait detection system only because the headlines were too "unrealistic" or "shocking", the diffusion of this valuable and important information would be blocked, which would ultimately lead to serious consequences. Thus, in our research, we will return to the definition of clickbait and make more reliable detections.

One possible reason why most existing research only used headlines lies in the difficult utilization of the article content. Compared with the summarized, short and brief headline, article content is scattered, long, complex and full of noise. Without proper presentation, introducing article content into the detection model may seriously decrease the detection accuracy [29]. Thus, one of the most serious challenges we need to face is how to present the article content well before we adopt it into the detection model. In our research, we incorporate a knowledge graph (KG) to introduce the information embedded in the article content. Moreover, a large KG can also yield rich background information to facilitate the performance of our detection.

2.2 Knowledge graph

A knowledge graph describes the concepts, entities, and relationships in a structured form to understand and manage massive information in a brief way. It was first proposed by Google in 2012 to optimize the results returned by search engines [35]. The great value possessed by the KG was soon recognized by academia and industry. The KG was developed and expanded quickly from that point on. Currently, it can be seen in various areas, such as e-commerce, finance, and health care [36], to help achieve conversational agents, intelligent financial fraud detection, health risk prediction, and recommendation systems.

From 1991 to 2020, interest in knowledge graph research has increased significantly and is constantly expanding [8]. Recent research on KG concentrates on knowledge graph refinement, knowledge graph construction, knowledge reasoning, and knowledge representation learning [16]. Knowledge representation learning lies at the center of knowledge graph research and is of great significance to many knowledge acquisition tasks and subsequent applications [16]. Knowledge representation learning can significantly improve computational efficiency, effectively alleviate data sparsity, and achieve heterogeneous information fusion. In recent years, with the emergence of embedding technology such as word-to-vector (Word2Vec) in natural language processing, using continuous vectors to represent knowledge has gradually replaced integration with the symbolic logic-based knowledge representation method [16]. Additionally, knowledge graph embedding is an effective method to introduce prior knowledge into the input for many deep neural network models to constrain and supervise the training process of neural networks [34].

Illuminated by the Word2Vec model, some researchers used distribution representation to represent entities and relationships in the knowledge graph. Given a triple (head, relation, tail) in the knowledge graph, previous work such as TransE [2], TransH [12], and TransR [22] regarded the relationship as the translation from head to tail. Ji et al. [15] proposed a more fine-grained model TransD to improve existing models. TransD uses two vectors to represent each entity and relationship. The first vector represents the meaning of the entity or relationship, and the second vector is used to construct the mapping matrix. Inspired by the overwhelming performance to present complex KG in recent research [46], we also incorporate TransD in our research.

2.3 Graph convolutional network

Convolutional neural networks (CNNs) have shown good performance on spatial dependence tasks, such as object detection [31], semantic image segmentation [24], and image recognition [14]. CNN requires input data in a grid-like matrix. There is an obvious disadvantage that data represented in the form of grid-like matrices are usually very sparse, which leads to serious computational waste. Although methods (e.g., pooling) have been proposed to alleviate this disadvantage, the computational cost is still very expensive, which seriously limits its applications in many cases, such as text mining. Moreover, data in many real settings are not in grid-like form.

Graphs have been proven to be an effective form to represent rich information and have great potential to enhance the performance of learning and prediction tasks. For example, Li et al. [21] constructed a two-step graph-based model for purchase prediction tasks and achieved notable improvement. Certain spatial dependence exists in the graph. However, the irregular shape makes it almost impossible to adopt CNN directly to handle the graph form data. To solve this challenge, a graph convolutional network (GCN) was proposed [19]. GCN is a kind of graph neural network that can conduct convolution operations on a network [19]. It has shown remarkable advantages in various domains, such as computer vision, natural language processing, and recommender systems [41, 42].

Recently, GCN has obtained great achievements in text classification. For instance, Lai et al. [20] designed a syntax-based GCN model to classify Chinese microblogs with sentiment. The model utilized a single-layer GCN to analyze the word features and dependency among microblog contents. Lu et al. [26] combined BERT with a vocabulary graph convolutional network (VGCN) to classify multiple text datasets and achieved higher performance. Some research has also tried to introduce external information by knowledge graphs and GCNs to enhance the final performance. For instance, Wang et al. [40] took both the global semantic relation and local sequence into consideration. Leveraging the self-attention GCN and 1D CNN, they achieved good performance in fake news detection. Their work could be a good foundation for our application of KG and GCN.

2.4 Graph attention network

GCN can convolute on the graph and allow us to apply deep learning to the graph structure. Although GCN has shown good performance in many tasks, it still has some defects. First, it depends on the Laplacian matrix and cannot be used in directed graphs. Second, the model training relies on the whole graph structure and cannot analyze the dynamic graph. Moreover, there is no way to assign different weights to neighboring nodes in convolution [38]. To solve these challenges, a graph attention network (GAT) was proposed [38].

Attention mechanisms have been proven to be of great power to enhance the performance of deep learning, including GCNs. Based on previous work, Veličković et al. [38] proposed GAT to assign different weights to different nodes in the neighborhood when dealing with varying sizes of the neighborhood to achieve node classification and performance improvements. Since then, GAT and its variants have been widely adopted in various tasks. For instance, Zhong et al. [45] designed novel hybrid GCNs with multihead attention for recommending points of interest. A deep GCN model with temporal attention presented by Zi et al. [47] helped to rebalance bike sharing by accurately predicting the bike number of each station. GAT also contributes to the field of biomedicine. Researchers utilized an attention-based GCN to extract drug-drug interaction relationships [28].

In particular, Zw et al. [48] proposed a propagation graph neural network with an attention mechanism to detect rumors on social media. The attention mechanism helped to adjust the weight of each node in the graph dynamically and achieved better results than advanced algorithms. Attempts have also been made to combine GAT and various KGs. Some researchers developed a hierarchical attention graph convolutional network to mine users' potential preferences from heterogeneous knowledge graphs. Experiments have demonstrated outstanding performance and certain explainability [43]. These attempts could act as good foundations for our design.

3 Model design

In our work, we used a knowledge graph to conceptualize text content and employed a GCN with a graph attention model to perform fine-grained-level clickbait detection. Our process includes two primary steps. First, we constructed a global KG to represent the rich information embedded in a large dataset and a local KG for a given article. Then, we represent a given article, including headlines and corresponding content, with graph embedding in the global KG. Second, we developed a GCN to predict the clickbait level of the article based on the representation of the article. A graph attention model is added to enhance the performance of GCN and achieve certain explainability. The process is illustrated in Figure 1.

3.1 Knowledge graph construction

There are three primary tasks in this step. First, we constructed a global KG based on a large dataset. Our large dataset includes news articles from several widely used online media platforms for half a year. This global KG could provide rich background information that assists comprehension of certain terms, such as metaphors. Second, we represented a given article as a small local KG. Intuitively, this local KG is a brief representation of the key content of the article. We treated the title and body of the article as two independent texts. Therefore, each small local KG has two parts: title KG and body KG. Third, we locate the graph embedding of the local KG (e.g., entity and relationship) in the global KG. The embedding of the title KG is expanded and spliced with the embedding of the body KG to meet the input requirement.

Generally, there are two operations in KG construction—entity (node) recognition and relationship (edge) identification. Since most entities are nouns, the simplest way to extract entities is by employing part-of-speech recognition of meaningful noun terms in the text. However, an entity may refer to a sequence of terms that are not limited to nouns. Thus, a better way to extract entities is to utilize named entity recognition (NER) techniques. NER can recognize embedded named entities, such as human names and place names. In our work, we adopt a widely used Chinese NER package provided by HIT [7]. We also used noun terms recognized by part-of-speech as a supplement to enrich the information represented by the graph.

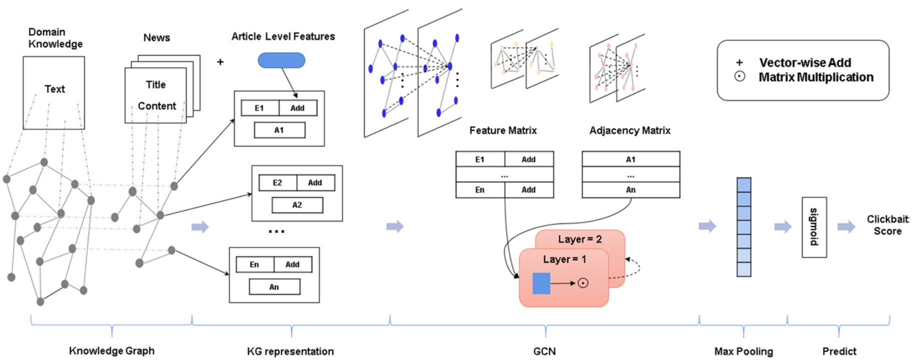

Figure 1 Architecture of the KG-GCN + ATT model

For relationship identification, we utilized pointwise mutual information (PMI), a popular algorithm to calculate the degree of association between two words [9]. The PMI of two words in a given text can be calculated as follows:

$$\mathrm{PMI}\left(word_1, word_2\right) = log_2\left(\frac{P\left(word_1 \cap word_2\right)}{P\left(word_1\right)P\left(word_2\right)}\right) \tag{1}$$

where $P\left(word_1\right)$ and $P\left(word_2\right)$ denote the probabilities of $word_1$ and $word_2$, respectively. $P\left(word_1 \cap word_2\right)$ is the joint probability that is determined by the co-occurrence of $word_1$ and $word_2$ in a text. Usually, the given text can be a sentence or a predefined window.

In our clickbait detection task, entities and associative relationships may be insufficient. For instance, exaggeration is one of the most widely used tricks adopted by clickbait. However, most exaggeration terms are adjectives or adverbs associated with a subject, which is usually a noun. Thus, even though we used noun terms recognized by part-of-speech as supplements to named entities, some crucial information was ignored. To solve this challenge, we added an extra operation in addition to entity (node) recognition and relation (edge) identification—attribute extraction. In attribute extraction, we extracted associative terms, such as adjectives, adverbs and auxiliary words, as attributes to corresponding entities. These attributes will also be represented in the embedding and participate in the calculation.

To enhance the performance of these three operations, we also adopted the phrase mining technique [33] to obtain meaningful short phrases. Phrase mining allows us to extract quality phrases that may be associated with extra information, such as taxonomy and topic, from a text corpus. For instance, it could help us to identify the specific relationship (e.g., father-in-law) between two entities other than a simple association value. These phrases are expected to be information-enriched and to be more contributive in the learning and prediction tasks.

In addition to this textual information, some article-level features also play important roles in clickbait detection. For instance, interrogative sentences are a widely used rhetoric in clickbait. It is also popular to include attractive pictures to attract users' attention. These features are difficult to capture by textual representation. To solve this problem, we constructed a set of categorical features to describe these informative characteristics. These features will be augmented to the representation of the article before the convolution operation.

3.2 Graph convolutional network design

After the construction of KG and the representation of articles, a GCN was adopted for learning and prediction. To avoid GCN's trend that prefers nodes with more neighbors and ignores the information carried by the node itself, the representation matrix A for an article was normalized. A one-layer GCN encodes only the information about the nearest neighbor, while an L-layer GCN can aggregate the L-order neighborhood [27]. Through L-layer GCNs, graph KG_i (the local KG for article i) can gain L representations. $H_i^{(l)}$ indicates the representation of graph KG_i after l-layer GCNs:

$$H_i^{(l)} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H_i^{(l-1)}W^{(l-1)} + b^{(l-1)}\right) \tag{2}$$

where $\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}$ is the renormalization process of the representation matrix A, $H_i^{(l-1)}$ denotes the output of graph KG_i for a given article i after $(l-1)$-layer GCNs, $H_i^{(0)}$ is X, that is, the feature matrix is the input of the first layer of GCN, $W^{(l-1)}$ and $b^{(l-1)}$ are the weight matrix and bias in the $(l-1)$-th GCN layer and σ is an activation function (ReLU).

We added a pooling layer after each graph convolutional layer. To extract the most important features from the original ones, we adopted maximum pooling.

3.3 Graph attention model development

To utilize the attention mechanism, we added two steps of computation based on the above direct convolution. First, the attention coefficient was needed. For the nodes in the graph, we calculated the similarity coefficient between its neighbors and itself one by one. Then, the activation function was used to transform the correlation coefficient into the attention coefficient.

$$\alpha_{ij} = \frac{exp\left(LeakyRelu\left(a^T\left[Wh_i \parallel Wh_j\right]\right)\right)}{\sum_{k\epsilon N_i} exp\left(LeakyRelu\left(a^T\left[Wh_i \parallel Wh_k\right]\right)\right)} \tag{3}$$

where W is a shared parameter that is used to increase the dimension of the node's features, $[\cdot \parallel \cdot]$ means concatenating the transformed features of node i and its neighboring node j, and *LeakyRelu* is the activation function.

Second, we aggregated the features according to the calculated attention coefficient.

$$h_i' = \sigma\left(\sum_{j\epsilon N_i} \alpha_{ij} W h_j\right) \tag{4}$$

where output h_i' is the new feature of each node fused with neighborhood information and σ is the activation function.

To further improve the utility of the attention mechanism, we used multihead attention to gain the final feature of the node.

$$h_i'(K) = concat\left(\sigma\left(\sum_{j\epsilon N_i} \alpha_{ij}^k W^k h_j\right)\right) \tag{5}$$

4 Experiments

4.1 Dataset construction

To the best of our knowledge, there are few public datasets available for clickbait detection, and even fewer in Chinese. The widely used dataset for clickbait detection is the "clickbait challenge 2017" dataset. Given the characteristics of tweets, e.g., their limited length, this dataset is unsuitable for the development of general detection models. Thus, we constructed our own dataset to verify the effectiveness of our model. To guarantee the diversity and representativeness of the data, we collected news from several popular online news platforms in China, such as Sina News and Surging News, from March 12, 2021, to April 12, 2021. After preprocessing, e.g., removing articles shorter than 200 bytes, we constructed a dataset that contains 8212 articles. Then, seven experts were invited to grade the clickbait level from 0 (nonclickbait) to 5 (strong clickbait) after specific training. The final scores were determined by majority voting. The statistics of the dataset are as follows:

4.2 Comparison experiments

We adopted widely used classical models, such as decision trees, random forests, support vector machines, and state-of-art models, such as TextCNN and LSTM, as our baselines. Given that the performance of the classical models relies heavily on the quality and quantity of input features, we first constructed an enriched feature set based on previous research [5, 25, 29]. The tf-idf algorithm [13] was used to obtain the vector representation of the text. Finally, we obtained a 300-dimensional representation of a given article. Given that some research argued that the difference between the title and article content meant a lot in the clickbait detection task [17], we also calculated the cosine similarity between the title vector and content vector and augmented the score as an additional dimension in the input.

Given that most existing research treated clickbait detection as a binary classification task, we first compared our proposed model with existing models on the binary detection task. Given that most classical models, e.g., SVM, are very sensitive to data balance, we intended to construct a relatively balanced dataset. According to Table 1, our dataset contains 4139 clickbait items and 4073 nonclickbait items. The widely used evaluation matrix, accuracy, precision, recall and F1-score, was used for comparison. These measures can be calculated as follows:

$$\text{accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{6}$$

$$\text{precision} = \frac{TP}{TP + FP} \tag{7}$$

$$\text{recall} = \frac{TP}{TP + FN} \tag{8}$$

$$\text{F1} - \text{score} = 2 \times \frac{precision \times recall}{precision + recall} \tag{9}$$

where TP, FP, TN and FN represent the number of true-positive instances, number of false-positive instances, number of true-negative instances and number of false-negative instances.

The results of each model on binary detection are summarized in Table 2, and the model comparison is illustrated in Figure 2.

From Table 2; Figure 2, we can first observe that the performance of each model is relatively low. This proved the difficulty of detecting clickbait in real settings. Nevertheless, deep learning-based models (TextCNN, LSTM, KG-GCN, and KG-GCN+ATT) achieved higher

Table 1 Statistics of our clickbait dataset

	Non-clickbait	Clickbait				
Clickbait Level	0	1	2	3	4	5
Number	4073	1540	1086	617	316	580

Table 2 Performance of each model for binary clickbait detection

	accuracy	precision	recall	f1-score
Features-DT	0.65	0.65	0.65	0.65
Features-RF	0.65	0.66	0.65	0.66
Features-SVM	0.60	0.60	0.60	0.60
TextCNN	0.67	0.60	0.60	0.60
LSTM	0.67	0.67	0.67	0.67
KG-GCN	0.67	0.70	0.67	0.68
KG-GCN + ATT	0.69	0.69	0.68	0.69

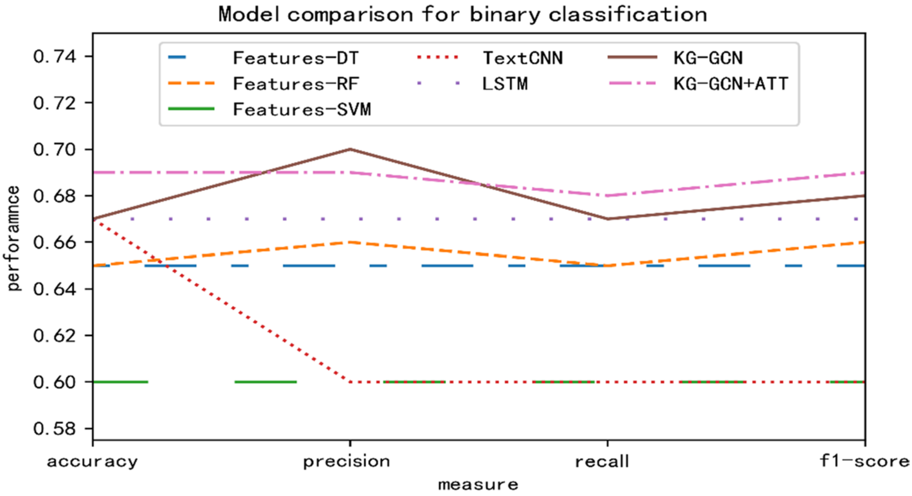

Figure 2 Illustration of the model comparison over each measure for binary classification

performance than traditional models (Features-DT, Features-RF and Features-SVM) since deep learning models could make use of the information embedded in the text more effectively as representation learning methods. Moreover, our novel proposed model (KG-GCN + ATT) made further improvements on other deep learning models.

In the second experiment, we compared the performance of each model for fine-grained-level clickbait detection. Widely used evaluation metrics for prediction, including the mean square error (MSE), the root mean square error (RMSE) and the mean absolute error (MAE), were used for comparison. MSE uses the square of the difference between predicted and true values. RMSE demonstrates the square root of the second sample moment of differences between anticipated values and observed values or the quadratic mean of these differences. MAE indicates the average error of the predicted values by calculating the absolute values of the differences between the predicted value and the corresponding true value. Predictions with smaller MSE, RMSE and MAE can be considered better.

These metrics are defined as follows:

$$\mathrm{MSE} = \frac{1}{m}\sum_{i=1}^{m}\left(y_i - \hat{y}_i\right)^2 \tag{10}$$

Table 3 Performance of various methods for fine-grained-level clickbait detection

Model	MSE	RMSE	MAE
Features-DT	0.0108	0.1041	0.0383
Features-RF	0.0125	0.1119	0.0449
Features-SVM	0.0114	0.1065	0.0199
TextCNN	0.0103	0.1014	0.0323
LSTM	0.0099	0.0996	0.0954
KG-GCN	0.0097	0.0983	0.0265
KG-GCN + ATT	0.0079	0.0889	0.0244

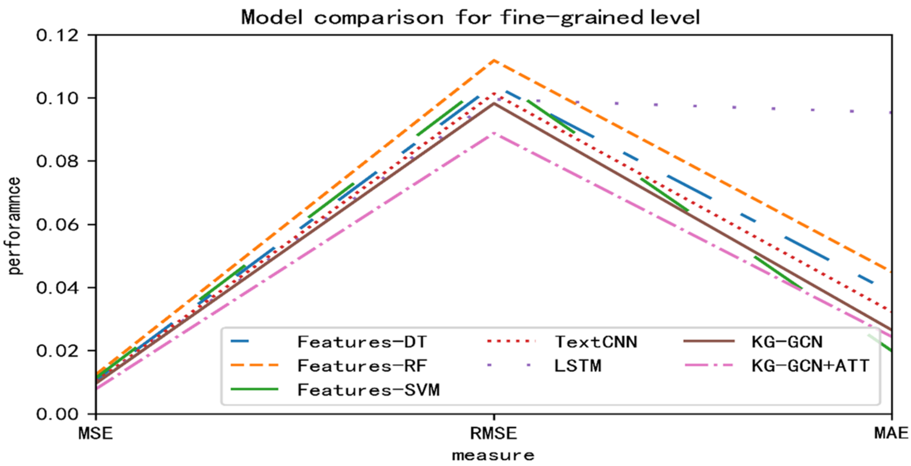

Figure 3 Illustration of the model comparison over each measure for fine-grained-level detection

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(y_i - \hat{y}_i\right)^2} \tag{11}$$

$$\mathrm{MAE} = \frac{1}{m}\sum_{i=1}^{m}\left|\left(y_i - \hat{y}_i\right)\right| \tag{12}$$

where $\hat{y}_i$ is the predicted value, y_i is the actual value, and m is the number of documents in the dataset.

In Table 3; Figure 3, we can observe that these advanced techniques showed obvious advantages over classical techniques, although the feature set for these classical algorithms had been carefully designed and enriched. Moreover, we can also see obvious enhancements by our KG-GCN and KG-GCN + ATT. These results verified the effectiveness of our design for fine-grained-level clickbait detection.

4.3 Explainability achievement

Clickbait is strongly related to human cognition. Thus, it is necessary to make the detection results understandable for humans. In our study, we tried to achieve certain explainability by tracking the contributions of each term in the final results from the graph attention network [37]. An example is shown in Figure 4. For the simplicity of visualization, only a few elements in the graph are selected. The darker the color is, the more important the elements (entity, attribute or features) are in the final determination. From Figure 4, we can see that some meaningful terms, including entities and corresponding attributes, play a significant role in the detection. Article-level features also contributed to the detection but were not as determinative as we supposed them to be.

5 Conclusion and future work

Network flow has become a critical resource developing a competitive advantage in online media. As a result, the battle for network flow is becoming increasingly fierce. Clickbait is considered an inexpensive but effective way to attract user attention and obtain network flow. Currently, clickbait is overflowing on the internet and causing serious damage to society. In this paper, we constructed a GCN-based model to conduct fine-grained-level clickbait detection. First, we used a knowledge graph (KG) to make brief but rich representations of articles. The global KG could also bring rich background information that facilitates the comprehension of the content (e.g., metaphors). Then, we developed a GCN to conduct fine-grained-level prediction. A graph attention model was also designed to enhance the performance of GCN and bring certain explainability to the detection results. Our experimental results showed that our novel proposed model outperformed classical and state-of-the-art models.

Our work could make both academic and application contributions. First, our attempt to use KG to present the articles and bring background knowledge could provide a new direction for future clickbait detection research. Second, our fine-grained-level clickbait detection could provide a measurement foundation for possible empirical study. For instance, different levels of clickbait (from weak to strong) may cast different influences on readers. The findings of such studies may act as the foundation of certain business strategies (e.g., control the level of clickbait to enliven the atmosphere while avoiding making users feel fooled). Third, clickbait detection and filter systems can be developed based on the model we designed to control the spread of clickbait.

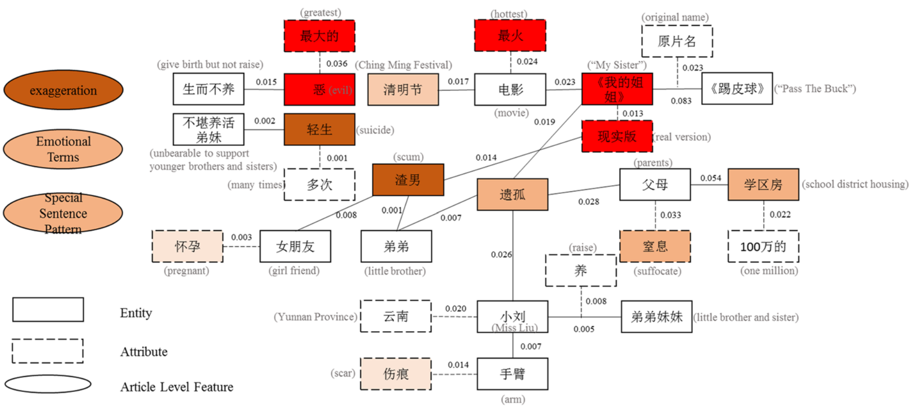

Figure 4 Illustration of the explainability achieved with our model

In the future, our work can be further improved from four aspects. First, as mentioned above, the traditional KG cannot sufficiently present the article needed for clickbait detection. In our design, we added attribute- and augmented article-level features. In the future, more effective presentation methods based on KG will be explored. Second, we will continue to improve the structure of the GCN model to strengthen its detection power. Third, we will also improve our graph attention model to achieve both performance and explainability enhancement. Fourth, we will adopt data augmentation methods to improve the robustness of our model.

Acknowledgements This work is supported by the Fundamental Research Funds for the Central Universities, and the Research Funds of Renmin University of China (No. 21XNA035).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Bondielli, A., Marcelloni, F.: A survey on fake news and rumour detection techniques. Inf. Sci. **497**, 38–55 (2019)
2. Bordes, A., Usunier, N., García-Durán, A., Weston, J., Yakhnenko, O.: Translating embeddings for modeling multi-relational data, pp. 2787–2795. NIPS (2013)
3. Breznitz, S.: Cry Wolf: The Psychology of False Alarms. Psychology Press, Hove (2013)
4. Cartwright, K.B.: Cognitive developmental theory and spiritual development. J. Adult Dev. **8**(4), 213–220 (2001)
5. Chakraborty, A., Paranjape, B., Kakarla, S., et al.: Stop Clickbait: Detecting and preventing clickbaits in online news media. In: 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), ACM (2016)
6. Chawda, S., Patil, A., Singh, A., Save, A.: A novel approach for clickbait detection. In: 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), pp. 1318–1321. IEEE (2019)
7. Che, W., Feng, Y., Qin, L., et al.: N-LTP: A open-source neural chinese language technology platform for Chinese. arXiv preprint arXiv:2009. **11616** (2020)
8. Chen, X., Xie, H., Li, Z., et al.: Topic analysis and development in knowledge graph research: A bibliometric review on three decades. Neurocomputing (2021). https://doi.org/10.1016/j.neucom.2021.02.098
9. Church, K., Hanks, P.: Word association norms, mutual information and lexicography. In: Proceedings of the 27nd Annual Meeting of the Association for Computational Linguistics (1989)
10. Dong, M., Yao, L., Wang, X., Benatallah, B., Huang, C.: Similarity-aware deep attentive model for clickbait detection. In: PAKDD (2), 56–69 (2019)
11. Dong, X., Gabrilovich, E., Heitz, G., et al.: Knowledge vault: a web-scale approach to probabilistic knowledge fusion. ACM, New York (2014)
12. Feng, J.: Knowledge graph embedding by translating on hyperplanes. AAAI (2014)
13. Hakim, A.A., Erwin, A., Eng, K.I., et al.: Automated document classification for news article in Bahasa Indonesia based on term frequency inverse document frequency (TF-IDF) approach. International Conference on Information Technology & Electrical Engineering. IEEE (2015)
14. He, K., Zhang, X., Ren, S., et al.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Recognition, P. (CVPR), IEEE (2016)
15. Ji, G., He, S., Xu, L., et al.: Knowledge graph embedding via dynamic mapping matrix. Meeting of the Association for Computational Linguistics & the International Joint Conference on Natural Language Processing (2015)
16. Ji, S., Pan, S., Cambria, E., et al.: A survey on knowledge graphs: Representation, acquisition and applications. arXiv preprint arXiv:2002.00388 (2020)
17. Karadzhov, G., Gencheva, P., Nakov, P., et al.: We built a fake news & click-bait filter: what happened next will blow your mind! RANLP 2017 - Recent Advances in Natural Language Processing Meet Deep Learning (2017)

18. Kaur, S., Kumar, P., Kumaraguru, P.: Detecting clickbaits using two-phase hybrid CNN-LSTM biterm model. Expert Syst. Appl. **151**(CSCW), 113350 (2020)
19. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. In: Proc. of ICLR (2017)
20. Lai, Y., Zhang, L., et al.: Fine-grained emotion classification of Chinese microblogs based on graph convolution networks. World Wide Web. **23**(4) (2020)
21. Li, Z., Xie, H., Xu, G., et al.: Towards purchase prediction: A transaction-based setting and a graph-based method leveraging price information. Pattern Recogn. **113**, 107824 (2021)
22. Lin, Y., Liu, Z., Sun, M., Liu, Y., Zhu, X.: Learning entity and relation embeddings for knowledge graph completion, pp. 2181–2187. AAAI (2015)
23. Loewenstein, G.: The psychology of curiosity: a review and reinterpretation. Psychol. Bull. **116**(1), 75–98 (1994)
24. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **39**(4), 640–651 (2015)
25. López-Sánchez, D., Herrero, J.R., Arrieta, A.G., et al.: Hybridizing metric learning and case-based reasoning for adaptable clickbait detection. Appl. Intell. **48**(9), 2967–2982 (2018)
26. Lu, Z., Du, P., Nie, J.Y.: VGCN-BERT: augmenting BERT with graph embedding for text classification. Adv. Inform. Retr. **12035**, 369 (2020)
27. Marcheggiani, D., Titov, I.: Encoding sentences with graph convolutional networks for semantic role labeling. In: Proceedings of the: 2017 Conference on Empirical Methods in Natural Language Processing (2017)
28. Park, C., Park, J., Park, S.: AGCN: Attention-based graph convolutional networks for drug-drug interaction extraction. Expert Syst. Appl. **159**, 113538 (2020)
29. Potthast, M., Köpsel, S., Stein, B., et al.: Clickbait detection. European Conference on Information Retrieval, Springer, Cham, 810–817 (2016)
30. Pujahari, A., Sisodia, D.S.: Clickbait detection using multiple categorisation techniques. J. Inform. Sci. **47**(1), 118–128 (2021)
31. Ren, S., He, K., Girshick, R., et al.: Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell. **39**(6), 1137–1149 (2017)
32. Rony, M., Hassan, N., Yousuf, M.: Diving deep into clickbaits: who use them to what extents in which topics with what effects? ACM (2017)
33. Shang, J., Liu, J., Jiang, M., Ren, X., Voss, C.R., Han, J.: Automated phrase mining from massive text corpora. IEEE Trans. Knowl. Data Eng. **30**(10), 1825–1837 (2018)
34. Shang, C., Tang, Y., Huang, J., et al.: End-to-end structure-aware convolutional networks for knowledge base completion. Proceedings of the AAAI Conference on Artificial Intelligence, 33, 3060–3067 (2019)
35. Singhal, A.: Introducing the knowledge graph: Things, not strings. http://googleblog.blogspot.com/2012/05/introducing-knowledge-graph-things-not.html (2012). Accessed May 2012
36. Tao, X., Pham, T., Zhang, J., et al.: Mining health knowledge graph for health risk prediction. World Wide Web **23**(5) (2020)
37. Vashishth, S., Upadhyay, S., Tomar, G.S., Faruqui, M.: Attention interpretability across nlp tasks. arXiv:1909.11218 (2019)
38. Veličković, P., Cucurull, G., Casanova, A., et al.: Graph attention networks. arXiv preprint arXiv:1710.10903 (2017)
39. Wang, X., Gao, T., Zhu, Z., et al.: KEPLER: a unified model for knowledge embedding and pretrained language representation. Trans. Assoc. Comput. Linguist. **9**(11), 176–194 (2021)
40. Wang, Y., Wang, L., Yang, Y., et al.: SemSeq4FD: Integrating global semantic relationship and local sequential order to enhance text representation for fake news detection. Expert Syst. Appl. **166** (2021)
41. Wei, X., Yu, R., Sun, J.: View-GCN: View-based graph convolutional network for 3D shape analysis. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 1850–1859 (2020)
42. Wu, T., Qi, G., Cheng, L., et al.: A survey of techniques for constructing Chinese knowledge graphs and their applications. Sustainability. **10**(9), 3245 (2018)
43. Yang, Z., Dong, S.: HAGERec: Hierarchical attention graph convolutional network incorporating knowledge graph for explainable recommendation. Knowl. Based Syst. **204**, 106194 (2020)
44. Zheng, J., Yu, K., Wu, X.: A deep model based on lure and similarity for adaptive clickbait detection. Knowl. Based Syst. **214**(5–6), 106714 (2021)
45. Zhong, T., Zhang, S., Zhou, F., et al.: Hybrid graph convolutional networks with multi-head attention for location recommendation. World Wide Web. **23**(3) (2020)

46. Zhu, Y., Lin, Q., Lu, H., et al.: Recommending scientific paper via heterogeneous knowledge embedding based attentive recurrent neural networks. Knowl. Based Syst. **215**, 106744 (2021)
47. Zi, W., Xiong, W., Chen, H., et al.: TAGCN: station-level demand prediction for bike-sharing system via a temporal attention graph convolution network. Inf. Sci. **561**, 274–285 (2021)
48. Zw, A., Dp, A., Jc, A., et al.: Rumor detection based on propagation graph neural network with attention mechanism. Expert Syst. Appl. **158** (2020)

Authors and Affiliations

Mengxi Zhou[1] · Wei Xu[1] · Wenping Zhang[1] · Qiqi Jiang[2]

Mengxi Zhou
zhoumx0808@ruc.edu.cn

Wei Xu
weixu@ruc.edu.cn

Qiqi Jiang
qj.digi@cbs.dk

1 School of Information, Renmin University of China, Beijing, China

2 Department of Digitalization, Copenhagen Business School, Frederiksberg, Denmark