Article:

Zhang, C. and Clough, P.D. (2020) Investigating clickbait in Chinese social media : a study of WeChat. Online Social Networks and Media, 19. 100095.

https://doi.org/10.1016/j.osnem.2020.100095

Investigating Clickbait in Chinese Social Media: A Study of WeChat

Canyu Zhang[a], Paul D. Clough[a,b,*]

[a] Information School, University of Sheffield, Sheffield, UK
[b] Peak Indicators, Chesterfield, UK

Abstract

Clickbait, the intentional use of exaggerated and misleading content to entice people to click on a link to a particular web page, is a phenomenon that has grown rapidly in recent years. Clickbait has become problematic in the post-truth era as it can de-value digital content and erode people's trust. The practice is especially common in social media where wider audiences can be reached more rapidly. Despite studies of clickbait being conducted on various social media sites, there has been little investigation of WeChat, the most popular social networking site in China. In this paper, we investigate clickbait behaviour in WeChat by analysing two samples (17,898 and 18,316 articles) for clickbait using supervised clickbait classifiers where an F_1-measure of 0.834 is obtained using Naïve Bayes. We train and test the classifier by manually annotating a sample of 3,000 examples for clickbait. Results show that approximately 70% of our WeChat examples are likely to be clickbait. We find that articles from publishers categorised as Funny, Anime, Entertainment and Culture exhibit the most clickbait, as well as posts from publishers in specific regions, such as Guangdong, Beijing and Jiangsu. We discuss the implications of results and provide recommendations. As far as we are aware, this is the first large-scale study of clickbait activity in WeChat.

Keywords: Clickbait, WeChat, Social Media, Machine Learning

*Paul Clough

Email addresses: zhangcanyu@mail.bnu.edu.cn (Canyu Zhang), p.d.clough@sheffield.ac.uk (Paul D. Clough)

1. Introduction

Clickbait is a term used to describe a phenomenon in which attractive, exaggerated and misleading titles are deliberately used in online content, such as social media, to entice people to click on a link to a particular web page in order to increase commercial revenue [1]. For example, sites such as Buzzfeed[1] and Upworthy[2] abound in clickbait content as their business model is based on generating page views, which in turn generates ad revenue. This is situated within the contemporary digital era of mass media production and consumption that must also deal with misinformation, disinformation and fake news [2, 3]. Additionally, generating disinformation, such as clickbait and fake news, is increasingly seen as source of profitable income[3].

Clickbait is *"a kind of internet content whose main purpose is to encourage users to follow a link to a web page especially where that web page is considered to be of low quality or value"* [4]. The links are typically versions of the headline or title of the linked web page. Examples of clickbait as used in headlines include *'The Hot New Phone Everybody Is Talking About!'*, *'10 things Apple didn't tell you about the new iPhone'*, and *'This guy went to hug an elephant. What happens next will blow your mind'*. Despite being commonplace in online content and part of the post-truth era, the practice of clickbait is considered problematic: undermining the credibility of media and contributing to the spread of rumours and misinformation online [5, 6, 7, 8]. Since mobile devices have become the most popular web browsing choices in recent years [9], clickbait is widespread in popular social media platforms, such as Facebook, Twitter and Instagram [10]. However, this has provided increased opportunities to spread misinformation, such as rumours, fake news and clickbait [11]. Indeed, Facebook twice updated

[1] `https://www.buzzfeed.com/`
[2] `https://www.upworthy.com/`
[3] `https://www.bbc.co.uk/news/technology-46136513`

its *News Feed* in 2017 to reduce the amount of clickbait headlines displayed on its platform [12].

However, to date there has been far less research into clickbait practices within Chinese social media, despite Chinese being the second most used language on the Internet [13] and where the Chinese Government has stressed the importance of reducing clickbait [14, 15, 16]. Furthermore, a 2018 survey conducted by China Youth Daily [17] found that 47.6% of respondents reported suffering from clickbait; therefore, this presents an important area of research. In this study we investigate clickbait activity in WeChat[4], the most popular social networking platform in China, which recently reached 1 billion monthly users[5] and exhibits a high growth rate year-on-year [18, 19]. However, as far as we are aware, there has been no prior large–scale study of clickbait in WeChat, despite its large volume of users and the high number of clickbait cases reported in the media [20]. The key outputs of this paper are:

1. We construct and analyse a dataset of 3,000 WeChat titles and articles and manually label them for clickbait.
2. Using the labelled dataset, we train supervised learning algorithms using various features to automatically classify clickbait in WeChat titles.
3. We apply the automated clickbait classifier to two WeChat data samples (14,898 and 18,316 articles), confirming approximately 70% clickbait.
4. We analyse the occurrence of clickbait in WeChat by metadata, including publisher region and article category.

The labelled dataset, the classifiers and the findings are all contributions of our work. For example, we identify the most common types of clickbait occurring in WeChat and important features for automatically classifying clickbait. These include the use of forward-references, specific words and phrases, use of punctuation and metadata. Further insights are gained from analysing the two

[4]`https://www.wechat.com/en/`

[5]`https://www.bbc.co.uk/news/business-43283690`

samples of data. For example, most clickbait occurs in articles categorised as Entertainment and Culture compared to Government and News; more clickbait is observed in specific locations, such as Guangdong and Beijing. Overall, our findings highlight and confirm the existence of clickbait activity in WeChat, which given the large numbers of WeChat users in China, is deserving of closer attention. In the study we address two main research questions: [RQ1]: What features can be used to automatically classify clickbait in WeChat? and [RQ2]: How often, and with what characteristics, does clickbait occur in WeChat?

The remainder of the paper is structured as follows. In Section 2, we review relevant literature on the notion of clickbait and its detection using computational methods. In Section 3, the functions and features of WeChat are introduced. In Section 4, we then describe the methodology used in our study, including the sample of WeChat articles gathered and analysed. Section 5 discusses the findings of the methods used to automatically identify clickbait and characteristics of clickbait articles found in WeChat. Section 6 then discusses our results in light of previous literature and our research questions, and provides avenues for future work. Finally, Section 7 concludes the paper.

2. Literature Review

The literature review first considers the context of clickbait and fabrication within the digital age (Section 2.1). This is followed by a summary of past work for automatically identifying clickbait (Section 2.2). We end by considering past studies of clickbait in Chinese media content (Section 2.3).

2.1. The Clickbait Phenomenon

Clickbait is commonplace on the Web and associated with the post-truth era, whereby alternative facts replace truth and feelings outweigh evidence [8]. In the new media era, the way people communicate and consume information has largely evolved and Online Social Networks (OSNs) have become important channels to share and receive messages. However, the OSNs have also turned into a mechanism for massive campaigns to spread false information, such as rumours, fake news, clickbait, and various other shenanigans [11].

Rubin [21] includes clickbait (along with satire and falsifications) as types of "fakes" compared to legitimate news. Zannettou et al. [11] also include clickbait as one of their eight categories of false information on the Web. More widely, the issue is one of disinformation (false information spread on purpose to deceive people) and misinformation (false or misleading information) [22]. Rumours can be diffused more easily across the web because of the absence of traditional fact-checking procedures used by news outlets [23]. The buzzword 'fake news' is narrowly defined by the New York Times as *'a made-up story with an intention to deceive, often geared towards getting clicks'*, and has gained increasing attention since the US Presidential Election 2016 [24, para. 11].

Clickbait occurs frequently online, especially across social media. According to a 2018 survey conducted by China Youth Daily, 47.6% of the participants reported suffering from clickbait frequently and 81.8% of them show antipathy towards it [17]. Past research into clickbait has studied social media sites, including Twitter [25, 26] and Facebook [27]. Clickbait strategies are also common in visual-centric social media, such as YouTube [11] and Instagram [10]. For example, Zannettou et al. [11] highlight the use of clickbait on YouTube where people upload videos to lure other users to click on their videos. This could include making eye-catching video thumbnails and intriguing video headlines. Clickbait is often compared to the notion of spam in email (perhaps more recently more akin to the notion of 'phishing') where users are led to malicious websites.

Publishers use strategies, such as building suspense, sensation or teasing, to invoke a 'curiosity gap' between the content and reader [25, 28]. In the case of clickbait, the curiosity is eased through clicking on a headline to navigate to the linked article[6]. People's curiosity can be explained by the 'information gap theory' established by Loewenstein [29, p. 87], who indicates that "deprivation

[6] `https://www.wired.com/2015/12/psychology-of-clickbait/`

labelled curiosity produced by an information gap arises when attention becomes focused on a gap in one's knowledge and the curious individual is motivated to obtain the missing information to reduce or eliminate the feeling of deprivation."

2.2. Automated Clickbait Detection

Various approaches to automate clickbait detection have been proposed, many of which are similar to approaches for detecting deception or false information, such as spam, fake news, fake websites, propaganda and rumours [3, 8]. Approaches typically comprise identifying (and engineering) features (e.g. the use of specific language, particular semantic and syntactic structures, etc.) that capture characteristics of clickbait and the use of supervised machine learning for classification (e.g. clickbait vs. non-clickbait) [30, 31, 25]. Alternatively, deep learning methods, such as Recurrent Neural Networks, have been shown to work successfully, whereby feature engineering is part of the machine learning process [32, 27, 33, 34]. Saquete et al. [8] provide a summary of resources and approaches used for identifying clickbait, noting small differences in performance between traditional machine learning and deep learning approaches. Often methods use features derived from the content of the headline or linked story, associated images, article metadata, and usage patterns. Some early research studies provide the basis for follow-up studies.

Chen et al. [3] propose a method of automated clickbait detection based on textual and non-textual cues. They divide the cue types into four categories: (i) lexical and semantic (e.g. specific word usage, punctuation, etc.), (ii) syntactic and pragmatic (e.g. forward-referencing and reverse narrative), (iii) images (e.g. emotional content) and (iv) user behaviour (e.g. reading time, shares, likes, etc.). The authors used a hybrid approach, whereby different machine learning methods are used to identify different cue bait types (e.g. lexical and semantic patterns identified with Support Vector Machines) and then combined. No evaluation data were provided.

Blom & Hansen [1] specifically focus on the lure of forward-reference in online news headlines. For example, *'This is the best news story you will ever read'* is defined as one form of forward-reference, due to the sentence using *'This'* to reference a forthcoming discourse (the full article) relative to the current location in the discourse (the headline). Forward-reference headlines are one of the main types of clickbait on news websites. Another early study investigated the phenomenon of 'listicles' [35]. There are many listicles on Buzzfeed, such as *'28 Underrated Desserts You Must Eat In NYC'*, such that the website has been criticised for the overuse of clickbait [36]. Inspired by these past two studies, the usage of pronouns and numerals will be introduced as important features to detect forward-reference and listicles in this study.

One of the first machine learning based approaches was suggested by Potthast et al. [25] to identify clickbait in Twitter. Tweets published by BBC News, Huffington Post, BuzzFeed and another top 20 most prolific accounts were sampled to build the Webis Clickbait Corpus of 38,517 tweets [37]. In the feature engineering stage, a total of 215 features were extracted with the majority being text-based cues, such as word occurrences, sentiment polarity, punctuation usage, etc. Features were divided into three main categories: (i) teaser message; (ii) linked page; and (iii) meta information. Logistic Regression (LR), Naïve Bayes (NB) and Random Forest (RF) algorithms were built and compared, with the RF classifier outperforming the rest with an F_1-score of 0.76. The initial work by Potthast et al. became the basis for the 2017 Clickbait Challenge[7] where participants had to classify tweets as not click-baiting, slightly click-baiting, considerably click-baiting and heavily click-baiting.

According to the primary evaluation measure Mean Squared Error (MSE), Omidvar et al. [38] ranked first amongst 16 teams using a Recurrent Neural Network (RNN) model to classify clickbait vs. non-clickbait using text-based cues only (MSE=0.0315; F_1=0.6703; accuracy=85.5%). Based on accuracy, the submission by Zhou [34] achieved the highest value (86%). He resolves this task as a multi-classification problem using a self-attentive neural network, previously introduced in [39, 40]. Although these top-ranked submissions utilised neural

[7] `https://clickbait-challenge.org`

network approaches rather than more traditional feature engineering methods, many of the other top-rated submissions used feature engineering and supervised machine learning [41, 42]. This is the approach we utilise in this work. Previous studies also helped to inform aspects of our study, such as feature selection/engineering, experimental setup and selection of machine learning algorithms.

2.3. Detecting Clickbait in Chinese Media

Despite many studies of mass communication and linguistics in the Chinese context, few have paid attention to clickbait. The earliest attempt to detect clickbait in Chinese was based on calculating textual similarity between headline and topic sentences, with low similarity indicating a high degree of clickbait [43]. The algorithm was previously used to detect academic plagiarism [44]. Subsequent studies used semantic similarity, including `Word2Vec` (an advanced approach for word embeddings) to detect clickbait on small samples of data [45, 46, 47]. However, in our study we focus on the headlines to detect clickbait; therefore, more advanced text similarity techniques are not further discussed.

Wei & Wan [48] divide clickbait headlines into two types (ambiguous and misleading) for which two different detection methods are designed. They use features from the headlines only to detect the ambiguous type, and add more features extracted from news article body to detect the misleading one. The ambiguous headlines can typically be identified without reading the body of the news article; however, for identifying misleading headlines the consistency between headlines and body text is important. For detecting the ambiguous headlines, which is more helpful in this study as the most common clickbait type, they specifically combine feature engineering with class sequential rule mining to build a classifier. The SVM machine learning algorithm was used and some of the important features extracted include number of words, numerals and clickbait words; together with use of internet slang, punctuation markers, and pronouns. We make use of these frequently-occurring cues in our study.

Zheng et al. [28] detect clickbait by first creating an initial model, then integrating this with a loss function according to user behaviour, thereby improving the performance of the initial prediction. Features used to build the initial model include the presence and the number of exclamation and question markers, the presence of pronouns and interrogatives, the number of words, ratio of stopwords and the TF-IDF weight of n-grams. The algorithms used to build the initial model included LR, NB, RF, SVM, and Gradient Boosted Decision Tree (GBDT), where a GBDT model performed best.

A more recent study using supervised learning methods to detect clickbait was conducted by Chen et al. [49]. However, there are several limitations of this research: the corpus is labelled according to the subjective judgement of the researcher (i.e. no multiple assessments); the scale of the clickbait corpus used to train the model is relatively small (in the order of 100s); finally, the model has not been applied beyond the training sample to explore clickbait on the web more generally. From all the past studies, only Zheng et al. [28] used WeChat articles as *one* of several resources for constructing a clickbait corpus. This highlights to date the limited study of clickbait activity in WeChat.

Existing models for detecting clickbait in Chinese content are summarised in Table 1. Based on *Precision*, *Recall*, and *F-score*, a GBDT classifier gives best performance. The first three classifiers were primarily built on the basis of calculating textual similarity between headline and body text; this is different from the other five classifiers, which are more similar to our model. Wei & Wan's SVM-based classifier [48] for detecting ambiguous clickbait is the most similar model to this study as the headline alone is used to detect clickbait. The GBDT model in Zheng et al.'s study [28] performs best, using feature engineering rather than deep learning. Therefore, we use feature engineering and include GBDT as one of the algorithms in this study.

Table 1: A comparison of existing clickbait classifiers for Chinese content.

Model	Precision	Recall	F-score
Topic sentence similarity-based model [43]	0.72	0.58	0.64
Latent semantic analysis-based model [47]	0.75	0.77	0.76
Semantic similarity-based model [46]	0.73	0.76	0.75
SVM-based ambiguous clickbait classifier [48]	0.71	0.80	0.75
SVM-based misleading clickbait classifier [48]	0.67	0.79	0.72
GBDT-based classifier [28]	0.75	0.81	0.78
DT-based classifier [49]	0.73	0.72	0.72
RF-based classifier [49]	0.71	0.71	0.71

3. Overview of WeChat

WeChat was launched in 2011 by Tencent and "has evolved into lifestyle platform for users in China. With approximately 850 million monthly active users, it now offers to its users what Facebook, WhatsApp, Messenger, Venmo, Grubhub, Amazon, Uber and Apple Pay together offer in the West[8]." In Q1 of 2020 the number of WeChat monthly active users rose to 1.17 billion[9].

WeChat is a closed social network and a contacts-based mobile-first application. It is used to communicate with friends and families, like Messenger, and has WeChat Moment (WM) in which people can share text, photos, videos and articles published by WeChat Official Accounts (WOA). These accounts are used to push out content, and on average, users follow between 10 and 50 accounts. Users can also interact with their friends, similar to Facebook. Browsing WeChat is a daily activity for millions of Chinese people and recently WeChat has added more ads in WM to increase revenue, similar to Facebook in their News Feed. The revenue model is such that the number of views defines advertising income[10].

[8] https://medium.com/harvard-business-school-digital-initiative/wechat-the-one-app-that-rules-them-all-38a876d04f3b

[9] https://www.businessofapps.com/data/wechat-statistics/

[10] https://www.abc.net.au/news/2018-07-21/nuclear-secrets-and-deadly-coffee-australian-fake-news-on-social/10002246

People can directly send a WM article to their friends from their contacts or share it to WM. When other people receive the article sent by their friends, or see it in their WM shared by friends, the information of the article they can see includes a title, an image and a piece of text (optional). Whether people click a WM article to read the story or not largely depends on the title content and is the reason why clickbait headlines prevail on WeChat. WeChat also has a way of protecting content - articles can be declared as 'original' meaning they should not be copied by another account and will likely reflect higher quality content.

In this study the data collected are the most popular articles over 30 days published by all WOAs. The WM articles published by WOAs consist of text information (the title, name of WOA, and content) and users' behaviour (reading number, thumbs-up number and report information). In this paper we refer to 'reading number' as 'Views' and 'thumbs-up number' as 'Likes' and use these to capture article popularity.

4. Methodology

4.1. WeChat Dataset Collection

The WeChat articles used in this study were collected from the *Qingbo WeChat Index*[11], which is the leading third-party new media data search engine in China [50]. The Index search engine provides lists of up-to-date WeChat articles. A custom-built crawler implemented in Python was used to collect two datasets of the top 25 WeChat articles published in the past 30 days: one from July 5th 2018 and the other from November 6th 2019, according to 24 different publisher categories with 34 different publisher locations. It is this popular content that tends to be viewed and shared by users. After filtering out articles written in Mongolian, Uighur, and Tibetan, this left two datasets of 17,898 articles (2018) and 18,316 articles (2019) written in simplified and

[11]Published on the following website: `http://www.gsdata.cn/`

Table 2: Summary of metadata collected for each WeChat article.

Feature	**Description**
URL	The web address of the WeChat article
Title	The headline of the WeChat article
Publisher	The publisher name of the WeChat article
Views	The number of views of the WeChat article
Likes	The number of likes of the WeChat article
Original	Whether the article is original or not
Video	Whether the article contains a video or not
Sound	Whether the article contains sounds or not
Publisher_category	The category of the article (one per article)
Publisher_location	The location of the article publisher

traditional Chinese. For each article we collected the metadata shown in Table 2. For preprocessing, the symbol of '10W+' in the 'views' and 'likes' columns was replaced with '100000' and missing values were replaced with '0'. The coverage of provider locations and categories is representative of WeChat as a whole. The datasets and other resources used in these experiments is available for download[12].

4.2. Clickbait Annotation in Sample Dataset

Through stratified sampling of publisher locations and categories in the 2018 WeChat dataset a random sample of 3,000 articles[13] was drawn to manually identify the existence and type of clickbait. Using a combination of convenience and snowball sampling methods, more than 100 adults were invited to participate in the annotation task through social media and face-to-face contact. No financial incentive was offered for the task. Participants were all native Chinese speakers, who could speak English fluently and were also regular WeChat users. The task was deployed on the *Figure Eight*, now called the *Appen*[14] data

[12]Data and resources available from: `https://tinyurl.com/y2l2ncwk` or `https://drive.google.com/drive/folders/13bR68hssHq7PmJkdbAtMXEcebZzMWQSV`

[13]Given the population size this is sufficient as using a 95% confidence level and 5% margin of error would give an ideal sample size of 377.

[14]`https://appen.com/`

Table 3: Categories of clickbait identified in WeChat articles (N=1,072).

Clickbait Type	**Description**	**Count (Percent)**
Ambiguous	Headline whose meaning is unclear relative to that of the content of the story	264 (24.6%)
Exaggeration	The title exaggerates the content on the landing page	241 (22.5%)
Formatting	Overuse of punctuation or some keywords, particularly exclamation marks	183 (17.1%)
Misleading	Headline whose meaning differs from that of the content of the story	98 (9.1%)
Teasing	Omission of details from title to build suspense	97 (9%)
Inflammatory	Either phrasing or use of inappropriate/vulgar language	93 (8.7%)
Graphic	Salacious, disturbing or unbelievable subject matter	56 (5.2%)
Wrong	Incorrect article or factually wrong	40 (3.8%)

annotation platform. Annotators were provided instructions before completing the task and examples of headlines of different clickbait types were also shown. Participants were then given WeChat headlines with a link to the full article and asked to judge them as clickbait or not.

If the annotators judged an article to be clickbait, they were asked to select from a list of clickbait types based on previous schemes [31, 48] - these are described in Table 3. Participants were also given the choice 'other' and asked to explain their selection. For each title, three people classified the headline for clickbait. A single clickbait type was assigned to each clickbait article - the most frequent of the three annotation results. If multiple types were assigned we used the Figure Eight confidence score to select one category based on how often the type occurred in the results and its position.

In order to measure the inter-rater reliability of the annotation results (for clickbait vs. non-clickbait classification), Fleiss' κ was used. We obtained an overall $\kappa = 0.35$. According to Landis & Koch's [51] interpretation, a value of κ between 0.21 and 0.40 represents a 'fair' inter-annotator agreement. Due to the level of agreement obtained, we used only the cases where all annotators agreed on the same category (for clickbait vs. non-clickbait), leaving a total of 1,595 articles (1,072 clickbait and 523 non-clickbait). In the case of assigning clickbait type we used a majority vote amongst annotators. Table 3 shows the proportion of clickbait articles categorised by type where Ambiguous (24.6%), Exaggeration (22.5%) and Formatting (17%) are the three most common types of clickbait in the sample.

4.3. Feature Engineering for Clickbait Detection

Given a WeChat headline the aim is to perform binary classification: clickbait or non-clickbait (i.e. estimate how clickbaity the headline appears). To achieve this we use supervised learning models trained on features derived from the annotated dataset (Section 4.2). The features arose from a review of previous literature and analysis of training examples in the annotated dataset (see Section 4.4). Although previous studies have created features based on the headline and the article (e.g. Biyani et al. [31]), in this study only features based on the headlines were used (similar to [48]). We did this to simplify processing and because not all participants in the annotation task reviewed the full article when judging clickbait. In the Chinese context, clickbait is more like tactics used for framing headlines. Also, previous works have explored the use of word embeddings as features and used deep learning [43]. However, these methods were not used in this study as simpler feature engineering approaches are more often used for a baseline [41, 42, 30, 25, 8]. To summarise key information from the short title sentences we used SimHash [52]. Overall, our approach is similar to [49, 28, 31].

To identify linguistic features (lexical and syntactic) from the headlines written in Chinese, we utilised the *jiebaR*[15] text mining package. Using this, we identified words and phrases from headlines and performed tasks, such as removal of stopwords, computing word counts, matching dictionary terms etc. In Chinese, word count refers to the number of phrases in a sentence (or headline). Chinese phrases consist of one or more Chinese characters and is the smallest unit of language that can be used independently. JiebaR cuts sentences into

[15]`https://cran.r-project.org/web/packages/jiebaR/index.html`

phrases (word segmentation) according to a word dictionary it uses and cutting rules. For example, "The University of Sheffield" is "谢菲尔德大学" in Chinese, and will be divided into two phrases: "谢菲尔德" (Sheffield) and "大学" (University), whose word count number (or phrase count) is 2. We also used the stopword list provided by jiebaR, which consists of 1,534 commonly-occurring symbols or characters (both Chinese phrases and English words).

Analysis of clickbait vs. non-clickbait headlines (using the training sample only) showed a statistically similar tendency in the use of punctuation symbols (using Wilcoxon signed-rank test Z= -0.928, p= 0.355). Differences in the use of punctuation markers is often a signal of clickbait [3, 28]. Although the general usage is similar, the Wilcoxon test cannot detect individual differences. Comparing punctuation usage between the clickbait and non-clickbait groups using log-likelihood (G^2) [53] we find that 5 punctuation markers ('.', '...', '?', '*', and ',') occur significantly more frequently (G^2 score $\geq$ 6.63 for significance $p < 0.01$ level[16]) in clickbait and 12 more frequently in non-clickbait. We group these into two features: clickbait and non-clickbait punctuation symbols (see Table 4).

From the training dataset we also identified 54 words and bigrams (mainly Chinese) commonly used in clickbait headlines and 347 commonly used in non-clickbait. These were used to perform dictionary lookup and identify the use of clickbaity and non-clickbaity words (a commonly used distinguishing feature of clickbait). The words where G^2 value are above 6.63 (a critical value for the difference to be significant at the $p < 0.01$) are highly used and the words where G^2 value are between 3.84 and 6.63 are less used ones. Words with the highest G^2 scores (translated into English) in clickbait headlines include: 'off', 'of', 'all', 'yes', 'I', 'this', 'you', 'no' and 'look'. This includes cues indicative of clickbait, such as personal pronouns, forward-references and curiosity stimulus. Common words of non-clickbait headlines tend to be more generic terms (e.g. '2018', 'month', 'recruitment', 'year' and 'release'). Common clickbait bigrams

[16] Computing log likelihood: `http://ucrel.lancs.ac.uk/llwizard.html`

include “的人” (kind of person) and “看完” (finish to read). Common bigrams in non-clickbait headlines include “权威发布” (Authoritative Release) and “的通知” (the Notice).

We also utilised six further dictionaries for identifying special word usage. Words regarding forward references were extracted from the clickbait examples and used to create a forward-reference dictionary. Dictionaries of personal pronouns, and interrogative pronouns were built using the *Xinhua Dictionary*[17]. The *Sogou Dictionary*[18] was used to capture the remaining three word types, including internet slang, celebrity names, and placenames. In addition to the word lookup approach mentioned above, the use of numerals and English words was also created for this feature type.

SimHash is an algorithm proposed by Charikar [52] to transform a document into an n-digit signature and has been widely used to compute textual similarity by comparing the signatures of documents. SimHash is performed using *jiebaR*, where the signature is generated based on a specified number of keywords in the text, with keywords selected based on *TF-IDF* weighting [54]. In this study, SimHash signatures are created for 1 to 10 keywords and used to represent each WeChat headline. In addition to SimHash, word count and stop word rate are also computed (referred to as *holistic information*).

A total of 62 features (see Tables 4 and 5) were extracted from the headlines for three categories: (i) punctuation symbol usage (32 features); (ii) special word usage (18 features); and (iii) holistic information (12 features). Table 4 shows the numeric features (e.g. Number of clickbait words) with the average values per title for the normal (non-clickbait) and clickbait cases. A Wilcoxon Rank Sum test (with continuity correction) is used to examine significance. For binary features (e.g. Presence of forward-reference words - True/False), we show the percentage of cases for normal and clickbait where the feature score is True (see Table 5). A Chi-square test (with Yates' continuity correction) is used to

[17]Xinhua Dictionary: `http://xh.5156edu.com/page/z2190m2907j18579.html`
[18]Sogou Dictionary: `https://pinyin.sogou.com/dict/`

Table 4: Numeric features extracted headlines with hypothesis testing results between clickbait and normal groups (N=1,275). Statistical significance: *$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$

Numeric feature	Normal (Avg)	Clickbait (Avg)	Wilcoxon	Sig
(i) Punctuation symbol usage				
Number of periods (in English)	0.074	0.361	176,490	***
Number of commas (in Chinese)	0.402	0.836	221,175	***
Number of ellipsis points	0.030	0.156	177,617	***
Number of question marks (in Chinese)	0.052	0.196	187,623	***
Number of periods (in Chinese)	0.003	0.039	168,162	*
Number of commas (in English)	0.003	0.031	169,611	**
Number of square brackets (in Chinese)	0.281	0.070	148,080	***
Number of dashes	0.099	0.007	157,414	***
Number of round brackets (in Chinese)	0.141	0.049	157,845	***
Number of vertical bars	0.099	0.029	154,012	***
Number of colons	0.105	0.071	160,421	0.067
Number of back-sloping commas (Chinese)	0.033	0.013	162,461	0.193
Number of backslashes	0.008	0.000	165,072	0.113
Number of angle brackets	0.013	0.033	163,881	0.097
Number of at marks	0.000	0.006	164,616	0.025
Number of clickbait punctuation symbols	0.565	1.618	250,484	***
Number of non-clickbait punctuation symbols	0.909	0.329	116,813	***
(ii) Special word usage				
Number of highly clickbaiting words	0.347	1.898	272,763	***
Number of lowly clickbaiting words	0.135	0.581	216,870	***
Number of clickbait words	0.482	2.479	279,577	***
Number of highly non-clickbaiting words	1.592	0.185	78,413	***
Number of lowly non-clickbaiting words	2.019	0.331	61,104	***
Number of non-clickbait words	3.612	0.156	43,060	***
(iii) Holistic information				
Number of words	14.0	15.2	18,8774	***
Ratio of stopwords	0.102	0.072	162,461	0.193
The SimHash signatures based on 1 to 10 keywords				

examine significance.

Except for the 10 SimHash features, 38 of the other 52 features are statistically significant between clickbait and non-clickbait headlines. These correspond to the lexis and syntactic categories of Chen et al's. [3] study. From Table 4, on average clickbait headlines are around 15.2 words (or symbols) in length and consist of 1.6 clickbait punctuation symbols and around 2.5 clickbait words. From Table 5, around 80.6% of clickbait headlines contain highly clickbaiting words, 35% forward-references and 22% first and second person pronouns.

Table 5: Binary features extracted headlines with hypothesis testing results between clickbait and normal groups (N=1,275). Statistical significance: $*p < 0.05$, $**p < 0.01$ and $***p < 0.001$

Binary feature	Normal (%)	Clickbait (%)	Chi-square	Sig
(i) Punctuation symbol usage				
Presence of periods (in English)	3.3	9.8	13.952	***
Presence of commas (in Chinese)	34.9	66.6	106.453	***
Presence of ellipsis points	1.9	9.2	19.691	***
Presence of question marks (in Chinese)	5.2	18.5	35.467	***
Presence of full stops (in Chinese)	0.3	1.9	3.636	0.056
Presence of commas (in English)	0.3	2.7	6.717	**
Presence of square brackets (in Chinese)	14.0	3.5	45.687	***
Presence of dashes	5.2	0.3	34.007	***
Presence of round brackets (in Chinese)	7.2	2.5	13.902	***
Presence of vertical bars	9.9	3.0	25.294	***
Presence of colons	10.2	7.1	2.912	0.088
Presence of back-sloping commas (Chinese)	6.9	5.0	1.345	0.246
Presence of backslashes	0.3	0.0	0.228	0.633
Presence of angle brackets	0.0	0.6	2.129	0.144
Presence of at marks	1.7	1.7	1.793	0.180
(ii) Special word usage				
Presence of highly clickbaiting words	26.7	80.6	329.969	***
Presence of lowly clickbaiting words	12.7	42.4	101.22	***
Presence of highly non-clickbaiting words	62.5	14.5	294.114	***
Presence of lowly non-clickbaiting words	78.5	23.2	331.110	***
Presence of forward-reference words	3.3	35.0	133.870	***
Presence of first and second person pronouns	9.1	22.0	28.195	***
Presence of interrogative pronouns	5.0	9.2	5.813	*
Presence of internet slang	3.3	2.1	1.161	0.281
Presence of English content	6.9	4.8	1.774	0.183
Presence of numerals	29.8	38.9	9.0554	**
Presence of celebrity names	2.2	8.6	15.644	***
Presence of place names	65.6	34.5	100.501	***

4.4. Experimental Setup

As previously mentioned a set of 3,000 articles have been labelled as ground truth (with 1,595 retained after passing inter-rater reliability checks). To investigate / engineer features (see Section 4.3) and train and test supervised learning methods the articles were randomly split (80:20) into a train / validation dataset (912 clickbait and 363 non-clickbait articles) to build models and the remainder a hold-out set (160 clickbait and 160 non-clickbait articles) to validate model performance on unseen data. We do this to improve generalisation of the classifier and establish a more accurate and realistic measure of performance.

Due to class imbalance in the training dataset, the minority non-clickbait class was oversampled using Synthetic Minority Over-sampling Technique or SMOTE [55]. This creates new instances based on existing instances and their nearest neighbours. In our case 5 nearest neighbours were considered and used to create new synthetic non-clickbait cases, giving a total of 912 non-clickbait examples. In the case of the numeric features shown in Table 4, we normalised the counts by number of words in the headline to reduce length effects. Values for other numeric variables were also normalised using Z-scores to fall with the same range. In addition to the features shown in Tables 4 and 5, we also computed ratios for the number of punctuation and special word usage. Features also include the metadata (Table 2), excluding title, publisher and URL. This resulted in a total of 95 features. Categorical features (e.g. publisher category) were label-encoded prior to training.

In order to obtain a best performing model we trained and compared seven popular machine learning algorithms: Logistic Regression, Naïve Bayes, Random Forest, SVM and GBDT are the five learning algorithms commonly utilised in the literature to detect clickbait in both Chinese and English. Since neural networks perform well in many studies, we also selected a MultiLayer Perceptron (MLP) and a feedforward Probabilistic Neural Network (PNN). In the model training phase the algorithms were run to obtain the best classification model (i.e. highest accuracy). First we performed feature selection to identify the least number of features that gave highest accuracy (e.g. for Logistic Regression this comprises 14 features). Features with low variance or highly correlated with other features were eliminated. We used 'backward feature elimination' to iteratively remove the feature of lowest importance at each step in the process. This was undertaken for each algorithm following which we trained classifiers using *stratified 10-fold cross-validation* [56]. Cross-validation is usually used to avoid the problem of overfitting [57], whilst taking advantage of the training set as much as possible. After performing cross-validation the entire training set was used to re-train the model.

For each algorithm we also performed hyperparameter optimisation using random search [58]. In this method parameter combinations are chosen at random and evaluated. The combination with the highest accuracy score is chosen based on using *nested cross-validation*. Through the process of feature selection, 10-fold cross-validation and hyperparameter optimisation, optimal classifiers were constructed for each machine learning algorithms. To determine how well the classification models are likely to perform in practice we tested the best performing models on the hold-out dataset. This helps to identify cases when the trained models are overfitting and we selected the 'best' model according to this score. To evaluate performance we used four measures from a two-dimensional *confusion matrix* (*Precision*, *Recall*, *Accuracy* and F_1*-measure*, *Cohen's* κ, and the area under the ROC curve *ROC-AUC* [59]. The best performing model based on the results of Accuracy, F_1-measure and Cohen's κ is selected and deployed on the larger datasets of 14,898 and 18,316 WeChat articles collected in 2018 and 2019. Experiments were performed using KNIME Analytics Platform 4.1 and data analysis using R version 3.4.4 and RStudio 1.1.442.

5. Results and Analysis

In this section we present the results of our study: Section 5.1 provides results of analysing the annotated sample dataset; Section 5.2 investigates the clickbait classifier trained on the labelled dataset; finally Section 5.3 discusses the results of classifying the larger samples to better understand the extent of clickbait within WeChat.

5.1. Analysis of Annotated Dataset

From the 3,000 annotated dataset we analyse the cases where annotators fully agreed on clickbait or not (1,595 headlines). In total, 67.2% are judged as clickbait. In the dataset, 18.8% of articles are 'original'. This means that the article comes from an Original Content account - WeChat public accounts for which the content is verified by Tencent as being unique and not infringing on the copyright of other accounts. We find that 20.2% of clickbait articles are original and 15.9% of non-clickbait are original. Few articles contain videos (0.69%) or sounds (0.13%).

With respect to popularity (views and likes), the average (median) number of likes per article is 99 (min=0, max=50,036, mad[19]=139) and views is 30,947 (min=1, max=100k, mad=40,769). Previous studies have shown that clickbait articles are typically liked and viewed more than non-clickbait. In our case, we find that likes for clickbait are on average (median) significantly higher than non-clickbait (141 vs. 40) and similar for views (39,296 vs. 15,320). Using a Mann-Whitney U (or Wilcoxon rank-sum) test between the groups we find that the differences are significant ($p < 0.01$). There is also a positive correlation[20] between the number of likes and views (rho =0.6554, $p < 0.001$).

Table 6 shows the average (median) number of likes and views, and mean headline length, by each clickbait category. The type with the most number of average likes (184) is formatting - the overuse of punctuation or some keywords, particularly exclamation points. This category also gets the greatest number of average views (45,894). Similarly, the least number of average likes (87) and views (24,793) is the teasing category - the omission of details from the title to

[19]Median Absolute Deviation (MAD) or Absolute Deviation Around the Median is a robust measure of central tendency that is less sensitive to outliers.

[20]Outliers are removed before computing correlation ($rho = 0.8710$ without removal)

Table 6: Categories of clickbait, average likes and views and percent original (N=1,595).

Clickbait Type	**Avg Likes (Median)**	**Avg Views (Median)**	**Avg Words (Mean)**	**%Orig.**	**Count**
Ambiguous	130	38,615	14.22	17.8%	264
Exaggeration	145	42,799	15.96	22.4%	241
Formatting	184	45,894	15.80	19.1%	183
Misleading	126	33,356	15.02	24.5%	98
Teasing	87	24,793	15.30	16.5%	97
Inflammatory	155	44,842	16.41	20.4%	93
Graphic	133	36,816	13.96	17.9%	56
Wrong	146	31,111	15.03	30.0%	40
None	40	15,320	13.95	15.9%	523

build suspense. Given the limited size of the annotated dataset we leave further analysis of clickbait patterns within WeChat to Section 5.3.

5.2. Clickbait Classifiers

Based on the labelled set of 1,575 WeChat training examples we train and test the clickbait classifiers (using oversampling on the non-clickbait cases). The results of the best models obtained on the hold-out data for 5 of the algorithms[21] are shown in Table 7. The highest score of each indicator is shown in bold. Naïve Bayes (NB) performs the best with the highest precision, recall, accuracy, F_1-measure and Cohen's κ. These results are comparable with approaches obtained in past studies of clickbait in Chinese media (see Table 1).

Table 8 shows the results of the NB algorithm on the training and hold-out datasets (for the best performing models). Compared to other algorithms, the difference between the training and hold-set sets is lowest suggesting the model is not overfitting as much as others on the training data. The performance on each class are also similar suggesting that the classifier may have benefitted from oversampling non-clickbait cases and training on balanced classes.

[21] We find that the performance of the SVM and PNN algorithms are poor on the non-clickbait examples and therefore do not report them here.

Table 7: Comparison of optimised classifiers on hold-out data (highest scores in bold).

Algorithm	**Precision**	**Recall**	**Accuracy**	F_1**-measure**	**Cohen's** κ	**ROC-AUC**
NB	**0.839**	**0.834**	**83.44%**	**0.834**	**0.669**	0.896
GBDT	0.8	0.781	78.13%	0.778	0.562	0.877
RF	0.804	0.766	76.56%	0.758	0.531	**0.901**
LR	0.806	0.8	80%	0.799	0.6	0.883
MLP	0.776	0.741	74.06%	0.732	0.481	0.83

Table 8: Results of the best performing NB model on training and hold-out datasets.

Class	**Precision**	**Recall**	F_1**-measure**
10-fold cross validation on training set			
Clickbait	0.883	0.895	0.889
Non-clickbait	0.894	0.882	0.888
Weighted average	0.889	0.888	0.888
Hold-out set			
Clickbait	0.799	0.894	0.844
Non-clickbait	0.880	0.775	0.824
Weighted average	0.839	0.834	0.834

For the best performing NB model we inspect feature importance by ranking features by their Information Gain (IG) score. The NB model selected 29 features during training and we show the top 17 (where IG > 0.05) most important features in Table 9. We observe that the presence (and amount) of specific words, especially from the custom-built dictionaries, is important for the classifier in distinguishing between clickbait and non-clickbait headlines. In addition, metadata, such as category and publisher location, are also useful information. Similar to existing studies, the presence of forward-references, stopwords and certain punctuation marks (e.g. '?') are also indicators of clickbaity headlines [26, 60].

Table 9: Ranking of feature importance for the best NB classifier.

Rank	**Feature**	**Inf. Gain**	**Feature type**
1	Ratio of non-clickbait words	0.65	Special word usage
2	Ratio of lowly non-clickbaiting words	0.48	Special word usage
3	Ratio of highly non-clickbaiting words	0.36	Special word usage
4	Presence of lowly non-clickbaiting words	0.31	Special word usage
5	Ratio of highly clickbaiting words	0.29	Special word usage
6	Ratio of stop words	0.28	Holistic information
7	Presence of highly clickbaiting words	0.25	Special word usage
8	Presence of highly non-clickbaiting words	0.24	Special word usage
9	Ratio of clickbait punctuation symbols	0.22	Punctuation and symbol usage
10	Ratio of lowly clickbaiting words	0.14	Special word usage
11	The category of the article publisher	0.13	Meta information
12	Presence of forward-reference words	0.12	Special word usage
13	Presence of lowly clickbaiting words	0.10	Special word usage
14	Presence of commas (in Chinese)	0.10	Punctuation and symbol usage
15	The location of the article publisher	0.09	Meta information
16	Presence of place names	0.08	Special word usage
17	The number of likes	0.06	Meta information

5.3. Clickbait in WeChat

Based on the results of evaluating the classifiers we used the best performing model (Naïve Bayes) to classify the larger WeChat datasets. These are used to estimate clickbait occurrences and provide insights into user activity in WeChat.

5.3.1. Overall Occurrences

Across all 33,214 cases the classifier judged 65.8% headlines to be clickbait, which is consistent across the years: 64.6% in 2018 and 66.7% in 2019. This is in line with the 67.2% cases identified in the annotated dataset. The NB model also produces a confidence value for each classification and using the first quartile as the lower bound score we consider cases when the classification confidence is 99.5% or greater. This reduces the dataset by 25% to 24,892 classified articles of which 17,199 (69.1%) were categorised as clickbait (66.9% in 2018 and 70.9% in 2019). We use this filtered dataset for the remainder of our analysis, where we mainly focus on features of the article metadata and clickbait.

5.3.2. Clickbait and Publisher Category

Metadata about each clickbait article includes the category of the publisher or Public Account[22]. Overall there are 24 categories with those represented the most in our samples being Livelihood (1,222 articles), Entertainment (1,211), Education (1,210), Culture (1,1199) and baby (1,184). With respect to clickbait, Figure 1 shows the proportion of clickbait across categories (ordered by the overall proportion of clickbait across all data points).

The top four categories with the highest proportion of clickbait are: Funny (95.8%), Anime (91%), Entertainment (89%) and Culture (86%). These categories are consistent across the years, although the ranking differs. The categories with the lowest levels of clickbait are: Government (41.4%), Education (44.1%), Game (50.1%) and Car (54.7%). This reflects that more stringent fact-checking procedures are most likely operating in the publisher accounts focusing on these topics.

We compute the correlation between the number of articles produced and number of publishers in a category ($rho = 0.2383$, $p = 0.26211$), number of articles and percent clickbait ($rho = 0.2614$, $p = 0.2173$) and number of publishers and percent clickbait ($rho = 0.1286$, $p = 0.3191$). Categories with more publishers do not necessarily produce more content. Also, categories with more publishers and articles does not result in higher proportions of clickbait. The occurrence of clickbait is therefore more likely due to specific publishers and the nature of the categories (e.g. news vs. entertainment).

We also compute the number of publisher locations that categories cover. All categories have content produced from >=29 regions, except Game (25 regions). This suggests that content in different categories is fairly evenly spread across publisher locations. Overall, results show that there are particular categories of publisher where users are more likely to encounter clickbait, especially those related to entertainment or amusement.

[22] Note that each publisher is assigned to only one publisher category.

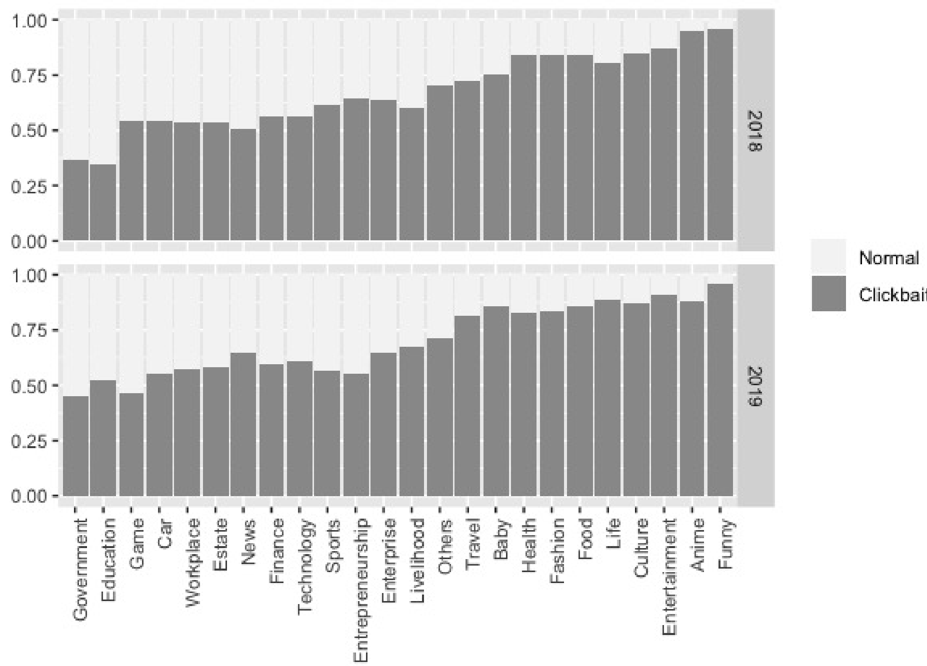

Figure 1: Proportion of normal and clickbait headlines by publisher category and year (ranked by overall proportion of clickbait).

5.3.3. Clickbait and Publisher Location

Similar to category, we can also analyse the proportion of clickbait by the publisher location. Overall, there are 34 locations (e.g. cities and provinces) of varying size and population. The proportion of clickbait for each location is shown in Figure 2. The top four locations with the most clickbait are: Guangdong (86.3%), Beijing (84%), Jiangsu (80.4%) and Shanghai (79.7%). These are major provinces and cities in mainland China. The locations with the lowest levels of clickbait are: Taiwan (31.7%), Macao (41.7%), Qinghai (42.8%) and Gansu (44.5%). Overall, the proportion of clickbait across years by location would appear consistent (Figure 2).

We compute the correlation between the number of articles produced and number of publishers in a location ($rho = 0.7512$, $p < 0.001$), number of articles and percent clickbait ($rho = 0.8662$, $p < 0.001$) and number of publishers and percent clickbait ($rho = 0.6583$, $p < 0.001$). This may suggest that locations with more publishers are likely to produce more articles, which may result in higher proportion of clickbait. This differs from publisher category. It is likely that population and editorial control may also affect the amount of clickbait emerging from specific locations. We also compute the number of categories that each publisher location covers. Results show that most locations include content from all categories, the main exceptions being Taiwan (6 categories), Macao (7 categories) and Tibet (18 categories).

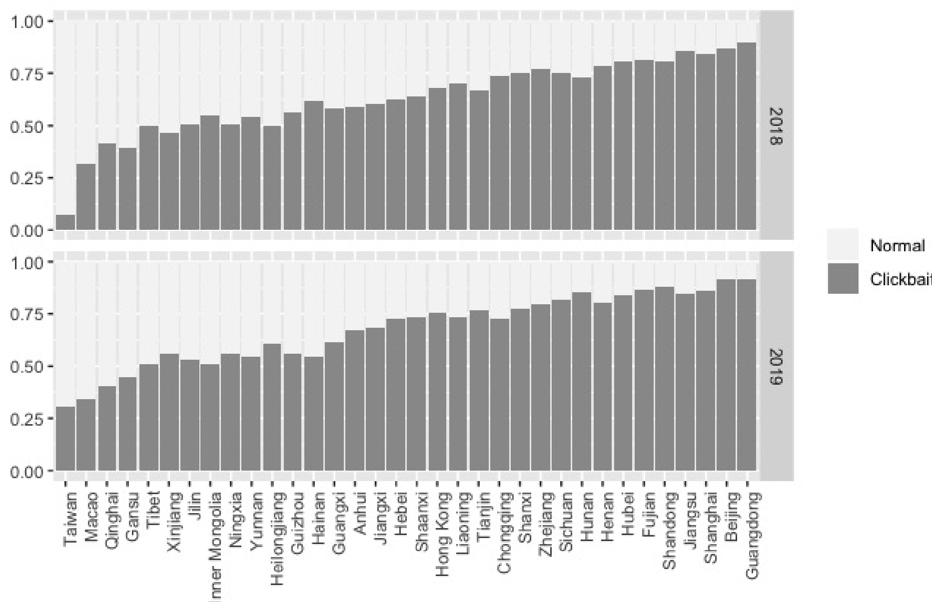

Figure 2: Proportion of normal and clickbait by publisher location and year (ranked by overall proportion of clickbait).

5.3.4. Clickbait and Publisher

The metadata also contains the names of the publisher generating the content. In total there are 4,424 publishers in the sample dataset producing an average (median) of 2 articles. The top 5 publishers producing the most content

Table 10: Top 5 publishers producing the most articles.

Publisher (English)	**Avg Likes (Median)**	**Avg Views (Median)**	**Avg Words (Mean)**	**%Orig.**	**%Clickb.**	**Count**
一禅小和尚(One Zen Little Monk)	3,360	100,000	8.43	95.7%	100%	46
新东北人(New Northeastern)	16	4,742	16.30	0%	100%	46
冷笑话精选(Selection of cold jokes)	1,936	100,000	17.89	20%	100%	45
十点读书(Ten o'clock reading)	16,464	100,000	10.24	22%	95.6%	45
少女兔(Maiden Rabbit)	8,558	100,000	17.22	8.9%	95.6%	45

are shown in Table 10. The output of 1,776 (40.2%) publishers were classified as being *all* clickbait (of these 855 publishers produced only 1 article). These publishers produced a total of 7,685 articles (30.9% of the total in this sample). This suggests that some publishers or accounts are focused on producing clickbait content.

5.3.5. Clickbait and Popularity

Metadata, such as likes and views, can be used to indicate the popularity of articles. The underlying purpose of clickbait is to encourage users to click the headline (or teaser message), thereby luring people to view the linked article (and generate revenue) Therefore, we would assume that clickbait articles will, on average, receive a higher number of views. In social media, the use of likes can also be used to surface and share content and similarly, due to the appeal of clickbait, we assume that clickbait will obtain more likes than normal (or non-clickbait) content.

Similar to the annotated dataset, the average (median) number of likes and views is higher for clickbait articles: 130 vs. 35 likes and 39,042 vs. 15,305 views. Using a Mann-Whitney U test between the groups we find that the differences are significant ($p < 0.001$). The correlation between likes and views is also significant (rho=0.6587, $p < 0.001$). A different view on popularity could also be the articles which receive >100,000 views. These are often referred to as '10W+' articles. Table 11 shows the 6 most and least popular publisher categories in the sample based on the median number of views from the average scores for each category. We also include the number of unique publishers

Table 11: Publisher category and average (mean) likes, views, percent original, count, number publishers, percent 10W+ ranked by popularity - median number of views - for top 6 highest and lowest scores (N=24,892).

Category	Avg Likes (Median)	Avg Views (Median)	Avg Words (Mean)	%Orig.	%Clickb.	#Pub.	Count	%10W+
Life	432	89,648	14.78	23.6%	85%	256	1,110	46.6
Government	374	84,027	14.73	7.7%	41.4%	371	1,140	43.4
Culture	536	76,152	13.07	29.4%	86%	177	1,199	45.1
News	142	56,003	16.57	20.6%	57.7%	320	1,011	30.5
Enterprise	108	53,668	14.26	10.2%	64.2%	268	997	28.8
Entertainment	395	51,783	14.94	28.7%	89%	146	1,211	38.4
...								
Technology	46	14,118	14.27	32.0%	58.7%	137	1,010	18.2
Baby	40	13,780	16.68	31.8%	81.8%	128	1,184	18.0
Sports	31	11,070	14.87	28.2 %	58.9%	132	1,061	15.9
Workplace	17	9,091	14.78	11.8%	55.7%	93	1,077	4.1
Game	18	7,621	14.42	18.9%	50.1%	77	623	19.6
Entreprenueurship	12	1,862	13.54	16.8%	59.2%	92	909	5.7

(#Pub.) producing content by category and percent of articles with 10W+ views (%10W+).

Overall, we find that in the top 6 most popular publisher categories, 3 out of 6 have over 80% of clickbait headlines. Life is the most popular category where each article has on average 432 ‘thumbs up’ or likes and 89,648 views. Government and News are also popular publisher categories. However, these articles may have less instances of clickbait due to stricter editorial policies. For example, the Cyberspace Administration of China (CAC) has taken measures to rectify clickbait practices since 2017 by issuing regulations on the use of news headlines and a total of five Chinese news sites were punished by the government in 2017[23]. Starting in March 2020, the upgraded regulation of the network information content ecology issued by the CAC has been in effect. The regulation emphasises that the producers of online information content shall not use exaggerated titles or create content seriously inconsistent with the title. Producers must also not hype rumors and scandals, and shall not incite

[23] http://www.cac.gov.cn/2017-01/13/c_1120302910.htm

discrimination amongst the population or regionally[24].

6. Discussion

Clickbait is a growing concern within online digital services, especially within Online social Networks, such as YouTube, Facebook, Instagram and Twitter. The results of our study show that clickbait is also prevalent in Chinese social media, which in our case is WeChat - the most popular platform in China. Our results estimate that around 70% of headlines exhibit clickbait tendencies, which may be far higher depending on the publisher, category and location. The consequence of this is that WeChat, similar to other social media platforms, may promote and encourage the spread of false and/or low quality information on the web [5, 6, 7, 8].

6.1. Research Question 1 - Clickbait Classification

Similar to existing work, we have shown that using a set of labelled examples and machine learning with selected features, we are able to develop an automated clickbait classifier. The F_1-measure score of 0.834 is comparable to the performance obtained in previous studies in clickbait detection, including in Chinese media (Table 1) and the 2017 Clickbait Challenge [34, 38]. In our study, the focus has been to identify headlines (or teaser messages) that exhibit clickbait characteristics, similar to [30, 48, 34], rather than additional analysis of the content of the linked article, which we leave for future work.

From the feature engineering stage (see Tables 5 and 4) and output of feature selection (see Table 9), clickbaity headlines can be distinguished from normal ones based on lexical and semantic features (e.g. specific term and punctuation usage), use of forward-references, publisher metadata (e.g. category of the article and account location) and social data (e.g. number of likes). This corresponds well the findings of previous work (see Section 2.2). In particular, the use of custom-built dictionaries that contain frequent words occurring

[24] http://www.cac.gov.cn/2020-03/06/c_1585041838522569.htm

in clickbait significantly more than non-clickbait articles (and vice-versa) using log likelihood appears an effective approach (see Tables 5 and 4). The dictionaries and the method used to construct them could be useful to others in WeChat clickbait detection.

From our analysis, example clickbait headlines include: "公开全院31位科主任私人号码，家属受不了了......" (The private numbers of the 31 directors of the whole hospital are disclosed, and the family members can't stand it ...) and "这构图绝了" (This composition is absolutely amazing!). An example of headlines categorised as normal are: "安徽这些名校高考喜报也来了" (Good news for college entrance exams in Anhui) and "新西兰数千银行账户被冻结，有华人$7000万资产被封" (Thousands of bank accounts in New Zealand were frozen, and $ 70 million of Chinese assets were blocked)

6.2. Research Question 2 - Clickbait in WeChat

Based on the results of this study we estimate approximately 70% of headlines exhibit characteristics of clickbait. Compared to figures of clickbait in other studies this could be an overestimate. For example, Zannettou et al. [11] estimate from 206k YouTube examples that 41% exhibit clickbait. For the benchmarks created by Potthast et al. [37] from Twitter, around 42% of clickbait is observed. Ha et al. [10] analysed fashion related posts in Instagram and found around 11% of posts contained clickbait. Clickbait is also common in online news sites and news shared on social media. For example, Rony et al. [27] collected 1.67 million Facebook posts created by 153 media organizations. The amount of clickbait varied, but across mainstream media this formed around 33.54% of headlines, rising to around 40% in unreliable media. An informal study estimates 63% articles on Buzzfeed are clickbait[25]. However, views drive revenue which can be increased through clickbait practices.

Users can access WeChat content from Public Accounts and from our results we see that the amount of clickbait may vary dramatically depending

[25] `https://keyhole.co/blog/buzzfeed-clickbait/`

on the category and location of the publisher (Sections 5.3.2 and 5.3.3). For example, 95.8% of articles produced by publishers categorised as Funny are estimated to be clickbait; whereas 41.4% of articles produced by Government exhibit clickbaity features. Similarly, 86.3% of headlines from publishers located in Guangdong are potential clickbait; whereas 31.7% of headlines from Taiwan are categorised as cickbait. The proportion of clickbait by location is likely due to the amount of content produced, rather than location itself; whereas the proportion of clickbait by publisher category is not as dependent on the amount produced and perhaps better reflects clickbait occurrence. We also find that all content from many publishers (40.2%) comprise entirely clickbait. We find that clickbait headlines are more likely to be viewed and liked than normal articles, which may increase the spread of clickbait (Section 5.3.5). All in all, clickbait poses a significant problem in digital content, including Chinese social media, such as WeChat.

6.3. Recommendations

Based on the popularity of WeChat in China, the use of clickbait by Public Accounts deserves attention. However, there is very little research in this area and many social media platforms in China have not yet begun to take measures to deal with this phenomenon. Many public accounts with social influence, even ones operated by the government, use clickbait - this seems to have become a necessary way of generating content to attract users attention and generate revenue. The articles that clickbait links refer to can be amusing and entertaining and thereby popular amongst groups of users (e.g. the most popular publishers in our study producing large amounts of clickbait are within entertainment categories, such as Funny). However, when the intent is to drive revenue through clicks that ultimately lead to low quality content (not the rich and detailed content that users may want) then clickbait can erode people's trust, polarise their views and limit satisfaction. This may be more problematic in categories such as news (Table 6 suggests that around 58% of articles produced by publishers categorised as News could be clickbait), as clickbait can often reflect article quality and be a source of false information.

Our results suggest that we can use a simple set of features (mainly lexical and syntactic) and machine learning methods to successfully model and identify clickbait headlines in WeChat. Such methods could be utilised by the users of WeChat or other Chinese websites in applications, such as browser add-ons [30, 27], to help signal potential cases of clickbait or help support community efforts to reduce clickbait, such as `stopclickbait.com`. Features associated with clickbait can also be used to educate and inform digital literacy initiatives around encountering false information. Such models could also be used by the providers of social media services to reduce the amount of clickbait proliferating the digital content they serve and reduce the spread of misinformation.

6.4. Limitations and Future Work

There are a number of limitations to our study. The estimates of clickbait are based on samples of articles from WeChat Public Accounts collected using the Qingbo WeChat Index search engine. Although taking two samples from a one month period in 2018 and 2019 to identify any year-on-year differences, the sample is insufficient to identify longer-term temporal trends around clickbait. In future work we plan to experiment with further samples across a wider time period.

The higher proportion of clickbait identified in this work could also, in part, be due to the data collection method where the Qingbo WeChat Index ranks articles based on their popularity. Our results suggest that popular content (based on the numbers of likes and views) may contain more clickbait. Therefore, we may have unintentionally collected more clickbait examples. However, given that users may encounter popular content more frequently then this figure may reflect more a realistic estimate of clickbait from a user's perspective. Another issue could be bot-generated content inflating the figures for potential clickbait, or bots influencing the popularity figures. We plan to investigate this in future work using methods for bot identification [61]. However, manual inspection of sample headlines and articles and writing style did not suggest large amounts of bot-generated content.

Currently we use feature engineering and machine learning methods to identify clickbait from labelled data. However, more recent studies have shown the benefits of using deep learning methods to identify more nuanced patterns of language that may help identify more clickbait examples, especially if more data are collected. As future work we also plan to conduct a more detailed failure-analysis to identify potential causes behind incorrect classification and implement enhancements to the classifier to deal with this, e.g. experiment with and engineer further features. We have also mainly analysed clickbait in WeChat articles based on available metadata, such as likes, views and publisher information. The publisher category would seem to provide useful insights into clickbait (especially more than location). However, at present we assume the article category reflects the publisher category. A more fine-grained and comprehensive analysis could be conducted through classifying the article based on its content (e.g. using topic modelling), rather than relying on the publisher metadata.

We would also like to undertake more studies with human annotators to generate more labelled examples and better understand the characteristics of clickbait. In our current study most annotators considered only the headline or teaser message rather than a deep analysis of the linked content. We would like to create a larger annotated dataset that takes into account the linked article, for example by displaying this more clearly in the crowdsourcing task interface. Rony et al. [27] highlight the complex nature of clickbait and whether it can be objectively determined. They highlight the need to consider the linked article as well as the headline to determine the quality and relevance of the linked article. In our own study, inter-annotator agreement was only 'fair' for whether a given headline was clickbait or not. This may suggest that clickbait is more of a "I know it when I see it" type of phenomenon and difficult to agree on a universal definition. We leave further analysis of the concept of clickbait in WeChat, and more detailed annotation of a larger number of samples, for future work.

7. Conclusions

In this paper, we have studied the occurrence of clickbait in WeChat, China's largest social media platform. We have collected a total of 36,214 articles from WeChat over two 30 day periods in July 2018 and November 2019. Using a crowdsourcing approach, a sample of 3,000 articles were manually annotated for the presence of clickbait. This dataset was analysed to identify characteristics of clickbait and create resources, such as custom dictionaries. These were used for feature engineering and with machine learning algorithms to train automated clickbait classifiers. Through empirical study we find that an optimised Naïve Bayes classifier gives highest performance on unseen test data with an F_1-measure of 0.834. Using this model to classify the data collected from WeChat we estimate that approximately 70% of headlines exhibit clickbaity characteristics. Through analysing the metadata information of the data collected we find that publisher categories, such as Funny, Anime, Entertainment and Culture exhibit the most clickbait, as well as posts from publishers in specific regions, such as Guangdong, Beijing and Jiangsu. Our findings highlight and confirm the presence and extent of clickbait within WeChat.

Acknowledgements

We thank the participants who contributed to the annotation task for their time and efforts. We also thank the journal reviewers who provided invaluable feedback on an initial version of the paper.

References

[1] J. N. Blom, K. R. Hansen, Click bait: Forward-reference as lure in online news headlines, Journal of Pragmatics 76 (2015) 87–100, `http://doi.org/10.1016/j.pragma.2014.11.010`. `doi:10.1016/j.pragma.2014.11.010`. URL `http://dx.doi.org/10.1016/j.pragma.2014.11.010`

[2] H. Örnebring, A. M. Jönsson, Tabloid journalism and the public sphere: a historical perspective on tabloid journalism, Journalism Studies 5 (3) (2004) 283–295, http://doi.org/10.1080/1461670042000246052. arXiv:https://doi.org/10.1080/1461670042000246052, doi: 10.1080/1461670042000246052.
URL https://doi.org/10.1080/1461670042000246052

[3] Y. Chen, N. J. Conroy, V. L. Rubin, Misleading Online Content: Recognizing Clickbait as "False News", In Proceedings of the 2015 ACM on Workshop on Multimodal Deception Detection. ACM. (NOVEMBER) (2015) 15–19, http://doi.org/10.1145/2823465.2823467. doi: 10.1145/2823465.2823467.
URL http://dl.acm.org/citation.cfm?doid=2823465.2823467

[4] Oxford English Dictionary, "clickbait, n.", retrieved 9 August, 2018, from http://www.oed.com/view/Entry/37263110 (2016).

[5] M. Luckie, Adele and the death of clickbait, retrieved 10 August, 2018, from http://www.niemanlab.org/2015/12/adele-and-the-death-of-clickbait/ (2015).

[6] J. M. Scacco, A. Muddiman, Investigating the influence of "clickbait" news headlines, Engaging News Project ReportRetrieved 10 August, 2018, from https://engagingnewsproject.org/wp-content/uploads/2016/08/ENP-Investigating-the-Influence-of-Clickbait-News-Headlines.pdf.

[7] C. Sutcliffe, Can publishers step away from the brink of peak content?, retrieved 10 August, 2018, from https://www.themediabriefing.com/analysis/can-publishers-step-away-from-the-brink-of-peak-content/ (2016).

[8] E. Saquete, D. Tomás, P. Moreda, P. Martínez-Barco, M. Palomar, Fighting post-truth using natural language processing: A review and

open challenges, Expert Systems with Applications 141 (2020) 112943. doi:https://doi.org/10.1016/j.eswa.2019.112943.
URL http://www.sciencedirect.com/science/article/pii/S095741741930661X

[9] S. Gibbs, Mobile web browsing overtakes desktop for the first time, retrieved 10 August, 2018, from https://www.theguardian.com/technology/2016/nov/02/mobile-web-browsing-desktop-smartphones-tablets (2016).

[10] Y. Ha, J. Kim, D. Won, M. Cha, J. Joo, Characterizing clickbaits on instagram, in: Twelfth International AAAI Conference on Web and Social Media, 2018, pp. 92–101.

[11] S. Zannettou, M. Sirivianos, J. Blackburn, N. Kourtellis, The Web of False Information: Rumors, Fake News, Hoaxes, Clickbait, and Various Other Shenanigans (2018) 1–26http://arxiv.org/abs/1804.03461. arXiv:1804.03461.
URL http://arxiv.org/abs/1804.03461

[12] A. Babu, A. Liu, J. Zhang, New Updates to Reduce Clickbait Headlines, retrieved 10 August, 2018, from https://newsroom.fb.com/news/2017/05/news-feed-fyi-new-updates-to-reduce-clickbait-headlines/ (2017).
URL https://newsroom.fb.com/news/2017/05/news-feed-fyi-new-updates-to-reduce-clickbait-headlines/

[13] Internet World Stats [IWS], Internet world users by language, retrieved 15 April, 2018, from https://www.internetworldstats.com/stats7.htm (2017).

[14] Cyberspace Administration of China [CAC], Provisions on the Administration of Internet News Information Services, retrieved 10 August, 2018, from http://www.cac.gov.cn/2017-05/02/c_1120902760.htm (2017).

[15] Global Times, News websites punished for making clickbait headlines, retrieved 10 August, 2018, from `http://www.globaltimes.cn/content/1028764.shtml` (2017).

[16] SBS News, Media can't just run clickbait: China govt, retrieved 10 August, 2018, from `https://www.sbs.com.au/news/media-can-t-just-run-clickbait-china-govt` (2016).

[17] China Youth Daily, 81.8% participants dislike "Clickbait", retrieved 12 August, 2018, from `http://zqb.cyol.com/html/2018-07/17/nw.D110000zgqnb_20180717_1-07.htm` (2018).

[18] A. Skuse, WeChat: The Chinese chat app stealing Weibo's thunder, retrieved 15 April, 2018, from `https://edition.cnn.com/2014/02/27/business/tencent-wechat-unseats-sina-weibo/` (2014).

[19] WeChat Blog: Chatterbox, The 2017 WeChat Data Report, retrieved 10 August, 2018, from `http://blog.wechat.com/2017/11/09/the-2017-wechat-data-report/` (2017).

[20] C. Custer, WeChat bans clickbait, rumors, and sex, retrieved 10 August, 2018, from `https://www.techinasia.com/wechat-bans-clickbait-sex-lots-kinds-content` (2016).

[21] V. Rubin, C. Brogly, N. Conroy, Y. Chen, S. Cornwell, T. Asubiaro, A news verification browser for the detection of clickbait, satire, and falsified news, Journal of Open Source Software 4 (35) (2019) 1208. `doi:10.21105/joss.01208`.
URL `https://doi.org/10.21105/joss.01208`

[22] D. M. J. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, M. Schudson, S. A. Sloman, C. R. Sunstein, E. A. Thorson, D. J. Watts, J. L. Zittrain, The science of fake news, Science 359 (6380) (2018) 1094–1096. `arXiv:https://science.sciencemag.org/content/`

359/6380/1094.full.pdf, doi:10.1126/science.aao2998.
URL https://science.sciencemag.org/content/359/6380/1094

[23] W. Ferreira, A. Vlachos, Emergent: a novel data-set for stance classification, in: Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: Human language technologies, 2016, pp. 1163–1168.

[24] S. Tavernise, As Fake News Spreads Lies, More Readers Shrug at the Truth, retrieved 12 August, 2018, from https://www.nytimes.com/2016/12/06/us/fake-news-partisan-republican-democrat.html (2016).

[25] M. Potthast, S. Köpsel, B. Stein, M. Hagen, Clickbait Detection, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 9626 (1) (2016) 810–817, http://doi.org/10.1007/978-3-319-30671-1_72. doi:10.1007/978-3-319-30671-1_72.

[26] A. Chakraborty, R. Sarkar, A. Mrigen, N. Ganguly, Tabloids in the Era of Social Media? Understanding the Production and Consumption of Clickbaits in Twitter 1 (1) (2017) 1–20, http://arxiv.org/abs/1709.02957. arXiv:1709.02957.
URL http://arxiv.org/abs/1709.02957

[27] M. M. U. Rony, N. Hassan, M. Yousuf, Diving Deep into Clickbaits: Who Use Them to What Extents in Which Topics with What Effects?http://doi.org/10.1145/3110025.3110054. arXiv:1703.09400, doi:10.1145/3110025.3110054.
URL http://arxiv.org/abs/1703.09400

[28] H.-T. Zheng, X. Yao, Y. Jiang, S.-T. Xia, X. Xiao, Boost clickbait detection based on user behavior analysis, in: L. Chen, C. S. Jensen, C. Shahabi, X. Yang, X. Lian (Eds.), Web and Big Data, Springer International Publishing, Cham, 2017, pp. 73–80, https://doi.org/10.1007/978-3-319-63564-4_6.

[29] G. Loewenstein, The psychology of curiosity: A review and reinterpretation., Psychological bulletin 116 (1) (1994) 75, http://dx.doi.org/10.1037/0033-2909.116.1.75.

[30] A. Chakraborty, B. Paranjape, S. Kakarla, N. Ganguly, Stop Clickbait: Detecting and preventing clickbaits in online news media, Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2016 (2016) 9–16http://doi.org/10.1109/ASONAM.2016.7752207. arXiv:1610.09786, doi:10.1109/ASONAM.2016.7752207.

[31] P. Biyani, K. Tsioutsiouliklis, J. Blackmer, " 8 Amazing Secrets for Getting More Clicks ": Detecting Clickbaits in News Streams Using Article Informality, Aaai (2016) 94–100https://pdfs.semanticscholar.org/dc0f/cfffa7fc9e2e84598fa318c78987b57de1ed.pdf.

[32] A. Anand, T. Chakraborty, N. Park, We used neural networks to detect clickbaits: You won't believe what happened next!, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 10193 LNCS (2017) 541–547, http://doi.org/10.1007/978-3-319-56608-5_46. arXiv:1612.01340, doi:10.1007/978-3-319-56608-5_46.

[33] S. Gairola, Y. K. Lal, V. Kumar, D. Khattar, A Neural Clickbait Detection EnginearXiv:1710.01507.
URL http://arxiv.org/abs/1710.01507

[34] Y. Zhou, Clickbait Detection in Tweets Using Self-attentive Networkhttp://arxiv.org/abs/1710.05364. arXiv:1710.05364.

[35] B. Vijgen, The Listicle: An Exploring Research On An Interesting Shareable New Media Phenomenon, Studia Universitatis Babes-Bolyai, Ephemerides 59 (1), http://studia.ubbcluj.ro/download/pdf/894.pdf.

[36] S. Ajani, A Full 63% of BuzzFeed's Posts are Clickbait, retrieved 13 August, 2018, from https://keyhole.co/blog/buzzfeed-clickbait/ (2014).

[37] M. Potthast, T. Gollub, K. Komlossy, S. Schuster, M. Wiegmann, E. P. Garces Fernandez, M. Hagen, B. Stein, Crowdsourcing a large corpus of clickbait on twitter, in: Proceedings of the 27th International Conference on Computational Linguistics, Association for Computational Linguistics, Santa Fe, New Mexico, USA, 2018, pp. 1498–1507.
URL https://www.aclweb.org/anthology/C18-1127

[38] A. Omidvar, H. Jiang, A. An, Using Neural Network for Identifying Clickbaits in Online News Mediahttp://arxiv.org/abs/1806.07713. arXiv:1806.07713.

[39] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, E. Hovy, Hierarchical Attention Networks for Document Classification, Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (2016) 1480–1489http://doi.org/10.18653/v1/N16-1174. arXiv:1606.02393, doi:10.18653/v1/N16-1174.
URL http://aclweb.org/anthology/N16-1174

[40] Z. Lin, M. Feng, C. N. dos Santos, M. Yu, B. Xiang, B. Zhou, Y. Bengio, A Structured Self-attentive Sentence Embedding (2017) 1–15http://arxiv.org/abs/1703.03130. arXiv:1703.03130, doi:10.1109/CVPR.2016.90.

[41] M. Glenski, E. Ayton, D. Arendt, S. Volkova, Fishing for Clickbaits in Social Images and Texts with Linguistically-Infused Neural Network Models (2017) 2–6http://arxiv.org/abs/1710.06390. arXiv:1710.06390.

[42] A. Grigorev, Identifying Clickbait Posts on Social Media with an Ensemble of Linear Modelshttp://arxiv.org/abs/1710.00399. arXiv:1710.00399.

[43] Z. Wang, N. Weng, Y. Wang, Research of Title Party News Identification Technology Based on Topic Sentence Similarity, Data Analysis and Knowledge Discovery (11) (2011) 48–53, http://doi.org/10.11925/infotech.1003-3513.2011.11.08. doi:10.11925/infotech.1003-3513.2011.11.08.
URL http://manu44.magtech.com.cn/Jwk_infotech_wk3/CN/abstract/article_3477.shtml

[44] S. Wang, Y. Wang, Algorithm of the Text Copy Detection Based on Text Structure Tree, Data Analysis and Knowledge Discovery (10) (2009) 50–55, http://doi.org/10.11925/infotech.1003-3513.2009.10.09. doi:10.11925/infotech.1003-3513.2009.10.09.
URL http://manu44.magtech.com.cn/Jwk_infotech_wk3/CN/abstract/article_842.shtml

[45] Google Code, word2vec, retrieved 17 August, 2018, from https://code.google.com/archive/p/word2vec/ (2013).

[46] J. Li, Research on the Calculation Method for Semantic Similarity of Sentence and Its Application, Master's thesis, Zhejiang University of Technology, Hangzhou, China, http://www.wanfangdata.com.cn/details/detail.do?_type=degree&id=D01164431 (2016).

[47] J. Luo, Research of Title Party News Identification Technology Based on Latent Semantic Analysis, Master's thesis, Hubei University of Technology, Wuhan, China, http://www.wanfangdata.com.cn/details/detail.do?_type=degree&id=XWC201510230000027310 (2015).

[48] W. Wei, X. Wan, Learning to identify ambiguous and misleading news headlines, IJCAI International Joint Conference on Artificial Intelligence (2017) 4172–4178https://arxiv.org/abs/1705.06031. arXiv:1705.06031.

[49] C. Chen, L. Yang, The research of "clickbaits" recognition technology

based on supervised learning, Sciencepaper Online (2018) 1–11http://www.paper.edu.cn/releasepaper/content/201804-262.

[50] H. Liu, W. Zhang, X. Xu, Y. Liu, Comparison of index analysis tools for big data platforms, DEStech Transactions on Computer Science and Engineering (iceiti), http://doi.org/10.12783/dtcse/iceiti2017/18843. doi:10.12783/dtcse/iceiti2017/18843.

[51] J. R. Landis, G. G. Koch, The measurement of observer agreement for categorical data, biometrics (1977) 159–174https://www.jstor.org/stable/pdf/2529310.pdf.

[52] M. Charikar, Similarity estimation techniques from rounding algorithms, 2002, pp. 380–388, http://doi.org/10.1145/509907.509965. doi:10.1145/509907.509965.

[53] P. Rayson, D. Berridge, B. Francis, Extending the cochran rule for the comparison of word frequencies between corpora, 2004, purnelle G., Fairon C., Dister A. (eds.) Le poids des mots. Published by Presses universitaires de Louvain. ISBN 2-930344-50-4.; 7th International Conference on Statistical analysis of textual data (JADT 2004) ; Conference date: 10-03-2004 Through 12-03-2004.

[54] X. Jin, S. Zhang, J. Liu, H. Guan, Research on similarity detection of massive text based on semantic fingerprinthttps://pos.sissa.it/300/009/pdf.

[55] N. V. Chawla, K. W. Bowyer, L. O. Hall, W. P. Kegelmeyer, Smote: Synthetic minority over-sampling technique, J. Artif. Int. Res. 16 (1) (2002) 321–357.

[56] R. Kohavi, et al., A study of cross-validation and bootstrap for accuracy estimation and model selection, in: Ijcai, Vol. 14, Montreal, Canada, 1995, pp. 1137–1145, http://robotics.stanford.edu/~ronnyk/accEst.pdf.

[57] D. M. Hawkins, The problem of overfitting, Journal of Chemical Information and Computer Sciences 44 (1) (2004) 1–12, http://doi.org/10.1021/ci0342472. arXiv:https://doi.org/10.1021/ci0342472, doi:10.1021/ci0342472.
URL https://doi.org/10.1021/ci0342472

[58] J. Bergstra, Y. Bengio, Random search for hyper-parameter optimization, Journal of machine learning research 13 (Feb) (2012) 281–305.

[59] I. H. Witten, E. Frank, M. A. Hall, C. J. Pal, Data Mining: Practical machine learning tools and techniques, Morgan Kaufmann, 2016.

[60] X. Chen, Comparison of Weibo and WeChat Public Platform as News Delivery Channels in Web 2.0 era, Master's thesis, Hong Kong Baptist University, Kowloon Tong, Hong Kong, retrieved 10 August, 2018, from http://comd.hkbu.edu.hk/macomm/maproject/image/projects/2016_project/15412164_chen.pdf (2016).

[61] K.-C. Yang, O. Varol, C. A. Davis, E. Ferrara, A. Flammini, F. Menczer, Arming the public with artificial intelligence to counter social bots, Human Behavior and Emerging Technologies 1 (1) (2019) 48–61. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/hbe2.115, doi:10.1002/hbe2.115.
URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hbe2.115