IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 6, 2019, accepted March 26, 2019, date of publication April 3, 2019, date of current version May 2, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2909054

Stories That Big Danmaku Data Can Tell as a New Media

QINGCHUN BAI[1], QINMIN VIVIAN HU[2], LINHUI GE[3], AND LIANG HE[1]
[1]School of Computer Science and Software Engineering, East China Normal University, Shanghai, China
[2]Department of Computer Science, Ryerson University, Toronto, ON, Canada
[3]Amazon.com, Inc., Seattle, WA, USA

Corresponding author: Liang He (lhe@cs.ecnu.edu.cn)

This work was supported in part by the Natural Science Foundation of Shanghai under Grant 17ZR1444900, and in part by the Xiaoi Research.

ABSTRACT In online video communities, an emerging source of comment, such as danmaku, allows viewers to interact when watching. Prior work discussed the feasibility of application using danmaku, while the comprehensive analysis of large-scale data is vacant to be filled in. We here release our danmaku data collection and report interesting observed phenomena in the danmaku comments. This analysis reveals the temporal distributions and user's access patterns of online time-sync comments. In particular, we distribute two novel natural language processing (NLP) tasks based on our danmaku dataset and provide baseline models. In the first task, we show how the naive models predict positive or negative sentiment given a danmaku comment, which effectively extends the real applications such as opinion poll prediction and marketing investigation. In the second task, we propose to use the NLP summarization model to make video tagging and summarization. The experimental results suggest that danmaku can not only support deeper and richer interactions between viewers and videos but also with high research value.

INDEX TERMS Danmaku, HCI, big danmaku data, text tagging, sentiment analysis, summarization.

I. INTRODUCTION

User-generated danmaku comment, which originated in Niconico,[1] is exceedingly popular among China video sharing websites such as YouKu,[2] iQIYI[3] and Bilibili[4] [1]–[3]. Figure 1 shows a typical example of a video with time-sync danmaku comments. We then bring two definitions to introduce danmaku as a new medium in Definition 1.1 and a big data resource in Definition 1.2.

Definition 1.1: Danmaku is a new online emerging medium that allows viewers to generate scrolling marquee comments on videos, and to instantly share and exchange small elements of a video content while watching, such as feelings and opinions at a moment in a video, personal features in close up, which encourages viewers to participate into videos and join a specific community anonymously.

- The sender of a danmaku is anonymous.
- A danmaku is usually very short.
- A danmaku is along with a video.
- A danmaku differs from traditional comments in that it is posted by anonymous viewers, and meaningless if not associated with a point-in-time of a video.
- A danmaku differs from a microblog in that its content is sent out anonymously, typically shorter, full of noises and attached with the point-in-time of a video, while microblogs can present the whole picture independently.

Definition 1.2: Big danmaku data represents the danmaku assets which not only inherit the characters of danmaku comments in Definition 1.1, are also characterized by big data.

- Big danmaku data is larger, more complex data sets, with a high volume, velocity, variety and veracity to require specific technology and analytical methods for its transformation into value.
- Big danmaku data is as short as short is average 4 Chinese characters.
- Big danmaku data is sent out anonymously which makes the contents be more important than the senders.

The associate editor coordinating the review of this manuscript and approving it for publication was Arif Ur Rahman.

[1]www.nicovideo.jp
[2]http://www.youku.com/
[3]http://www.iqiyi.com/
[4]https://www.bilibili.com/

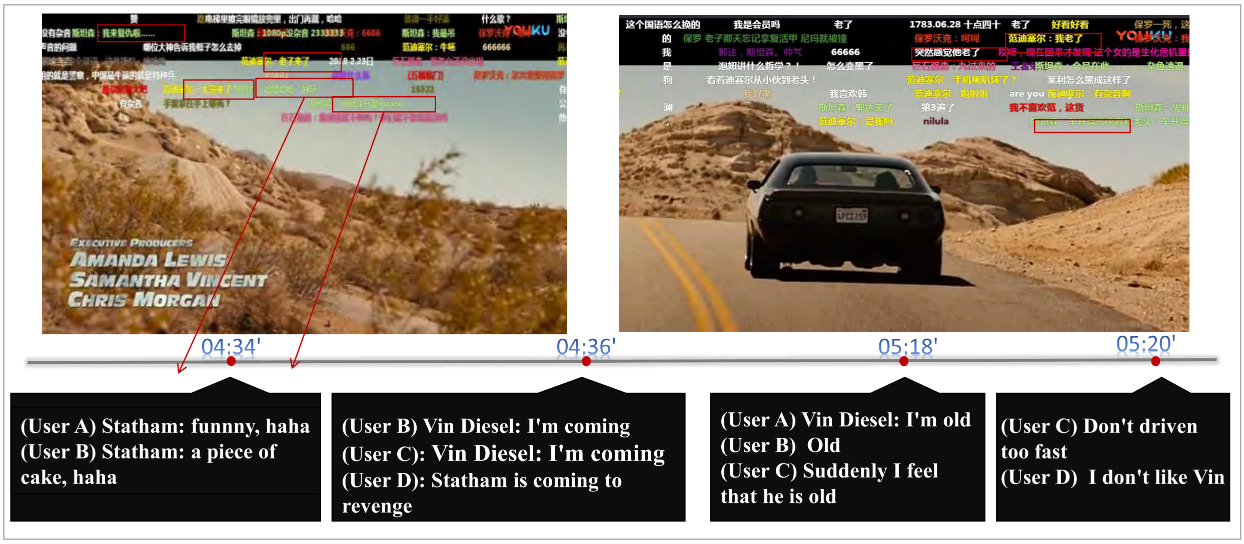

FIGURE 1. An example video named "Fast Furious 7", ongoing discussion an funny shot of "Statham (an actor in the video) driving a car quickly" with time-sync Danmaku comments. Reviewers can show their opinions along the playback time while watching which enhances the watching experience of participation.

- Big danmaku data is based on video time, where danmaku can be understood associated with the point-in-video-time, instead of our clock time.

Previous work on Danmaku mostly focuses on the potential value of danmaku data and viewer's motivation of interactions from the perspective of social networking [4]–[8] . For example, Ma and Cao [1] analyzed danmaku usage through the viewer's viewpoints and motivations with semi-structured interviews, Yao *et al.* [4] discussed Danmaku's potentiality and proposed to apply Danmaku in online video learning. He *et al.* [9] explored the leading danmaku and its herding effect phenomena to illustrate the unique characteristics of danmaku. Olsen and Moon [10] in 2011 discusses video summarization using user interactions, Danmaku interactions would be ideal to improve the summarization. However, these work can only make use of a small part of danmaku data and treat danmaku data as a complimentary to other data sources as interviews and comment comments. Therefore, we are urged to investigate what danmaku data itself can tell us, based on its strong characters and its abounding resources as a new medium.

First of all, we introduce our big danmaku data with its statistics, in which we present the characteristics of danmaku in variety, volume, velocity, veracity, length, anonymous space and time synchronization.

Second, we perform sentiment analysis on big danmaku data with a sentiment classification task. This intuitive idea comes from that danmaku often reflects viewer's main views and emotions for the current video clips, which makes danmaku a high research value of online communities behavior problem, where the annotation process, the problem definition and the benchmark experiments are presented.

Third, we propose video tagging and summarization task to show how to model the textual features with time-sync danmaku. Here we recruit five state-of-the-art summarization methods as our benchmarks on our big danmaku data for evaluation purpose. Experimental results on the above two natural language processing (NLP) tasks suggest that danmaku can support deeper and richer interactions between viewers and videos, with high research value.

In summary, our contributions can be summarized as follows.

- We define big danmaku data a new medium with its unique characters as variety, volume, velocity, veracity, length, anonymous space and time synchronization.
- We explore the uniqueness social sharing features of danmaku by analyzing the distribution of danmaku-based users and the distribution of features in data sets.
- We provide a large scale danmaku data sets with emotional polarity tagging, thus to help us understand the emotional outbreaks and negative emotions of the video clip. Based on the dataset, further applications and analysis of the danmaku can be more convenient.
- We put forward two NLP cases and corresponding evaluation criteria of danmaku by using the manual labeling of videos, meanwhile, state-of-the-art benchmark baselines are provided on our dataset, which can be used as the baselines for future algorithm development.

II. RELATED WORK

Recently, as a potential data source, danmaku data has attracted the attentions from academic research. A summary is as below.

A. INFERENCE ON VIDEO CONTENT

Researchers have been mining information from danmaku comments by applying text processing technologies to the research field of video analysis.

1) PREDICTION OF VIDEO POPULARITY AND RECOMMENDATION.

Since danmaku provides rich time-synchronized text information for the video, many researchers [11], [12] applied it in predicting the prevalence of videos by adopting statistical information of danmaku. Based on the statistics of the emotions in danmaku, Wu *et al.* [13] proposed to establish the correlation between danmaku emotion number and video playback times. He *et al.* [11] focused on predicting the popularity of videos by leveraging video information and danmaku comment number as important features. The prediction of video popularity usually needs to combine the information of the video, such as the title, uploads, collections, etc., to predict the amount of video playing or the trend of collections.

2) EXTRACTION OF VIDEO HIGHLIGHT SEGMENTS

To utilize statistical information of danmaku, text information and timing characteristics are also included for the extraction of video highlights [10], [14]–[16]. Firstly, the concentration of the topics contained in the danmaku can be used to dynamically segment the video; secondly, the number of danmaku that changes over time also reflects the trend of the highlights on different parts of the video. Therefore, researchers extracted videos by combining these two aspects of highlights, and managed to get corresponding segmented text topics.

3) AUTOMATIC TAGGING OF VIDEOS

According to previous reports [13], [17]–[20], danmaku can be used as time-sync text information flow over videos as a research method of video tagging. Automatic video tagging generates text tags or descriptions on videos based on the content, to reduce time-consuming manual labeling. With the synchronization of danmaku and video content, researchers focused on using machine learning algorithms to model danmaku for video segmentation and content analysis.

B. EXPLORING USER BEHAVIOURS FROM THE PERSPECTIVE OF SOCIAL INTERACTION

Another area of work focused on danmaku's potential value and user's motivation of interactions from the perspective of social networking since danmaku is an emerging comment type on videos in recent years [4]–[8], [21]. Ma and Cao [1] analyzed danmaku usage through the users' viewpoints and motivations with semi-structured interviews. Yao *et al.* [4] discussed Danmaku's potentiality and proposed to apply Danmaku in online video learning. He *et al.* [9] explored the leading danmaku and its herding effect phenomena to illustrate the unique characteristics of danmaku. Chen *et al.* [21]. analyzed reasons for watching or not watching Danmaku videos and identified underlying factor structures of motivations for the participants.

C. DATA ANALYSIS ABOUT DANMAKU

The value and the effectiveness of danmaku in promoting research on videos has been proved in [13], [17], [18]. Few of prior works have comprehensive analysis of the data, and in-depth to study how danmaku data is shared, and what danmaku tell us. Research on Danmaku is currently limited to comment analysis to assist data processing of video content. Furthermore, the full set of Danmaku features such as redundant information or short comments are not efficiently used with non-linearly increasing level of difficulties in data analysis.

Therefore, we think it is highly valuable to extend Danmaku analysis to additional areas with more characteristics of Danmaku included.

III. CHARACTERISTICS OF DANMUKU COMMENT

A. VARIETY OF DANMAKU COMMENT DATASET

As a new type of instant interactive video comment, danmaku has unique features different from traditional comments: 1) Danmaku allows users to share their opinions during watching, while video descriptions of the danmaku will be abundant in complex scenes, and 2) user-generated danmaku comments represent viewer's sentiments about the segment of video as it is synchronized to the video timeline. 3) When users are watching a video, the danmaku comments can have impacts on the viewing experience of other users. Time-synchronized danmaku comments can convey interesting and useful content information about videos, enhancing watching experience through diversified perspectives. YouKu[5] website is one of the fastest-growing websites [22], [23], which allows users to send danmaku while watching. As a new media which differs to texts, multimedia and microblogs, variety of the Meta-data we formatted as shown in Table 1.

TABLE 1. The meta-data of a danmaku comment.

danmaku comment	notes
id	danmaku id
play time	timestamp related to the video
play mode	location overlaid on the video screen
content color	the color of the danmaku
danmaku pool	Whether or not the senior danmaku
sent time	exact sent time
user id	the user who sent the danmaku
content	danmaku content
TV type	type of the video
TV name	name of the video
TV id	id of the video
Season id	id of the season video
Tags	tags of the video

We have built a danmaku crawler and collected the YouKu videos' danmaku information through a combination of YouKu API and scrapes of YouKu video webpage. Figure 2 shows an overview of how crawler gets danmaku data from Youku.com. We use three approaches to construct our dataset: 1) We crawl both publicly available traditional video comments and time-sync danmaku comments. 2) We crawl users and video information for further analysis such as timestamps and video tags. 3) We further annotate the sentiment label of danmaku comments to better understand the emotion of the users.

[5]http://www.youku.com/

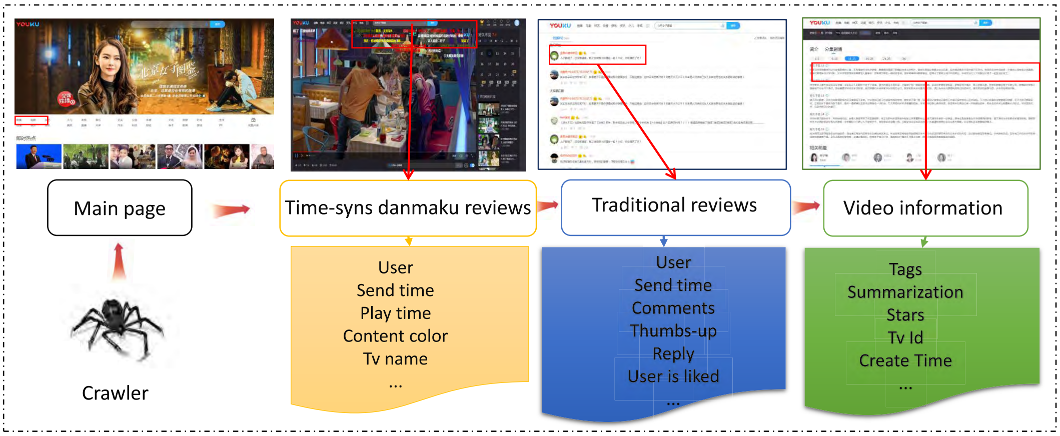

FIGURE 2. An overview of how crawler get the Danmaku data from Youku.com.

TABLE 2. The statistics of danmaku datasets. (1) #Seasons,#Videos, #Danmaku, #Users, #Tags: the number of seasons,videos, danmaku comments, uses and tags respectly; (2)#comments: the number of traditional comments about the video; (3) #UCS : the number of users of traditional comments; (4) Summarization: video summarization; (5)Sentiment Tag: the labeled sentiment class.

Data	Source	# Seasons	# Videos	# Danmaku	# Users	#Tags	Likes	#Comments	#Users of Comments	Summarization	Sentiment Lable
Paper [9]	acfun.tv	×	6,506	1,704,930	320,000	×	×	×	×	×	×
Paper [24]	acfun.tv	×	6	234,003	×	×	×	×	×	×	×
Paper [12]	iqiyi	2	7,166	11,842,166	1,133,750	×	×	×	×	×	×
Paper [16]	bilibili	×	6	52,174	×	×	×	×	×	×	×
Paper [17]	acfun.tv	×	120	227,780	×	3	√	×	×	×	×
Paper [25]	bilibili	64	716	7,413,517	1,482,120	42	√	×	×	×	×
Paper [13]	acfun.tv	×	16,414	1,103,884	382,752	×	×	×	×	×	×
Paper [15]	bilibili	×	×	133,250	×	×	×	×	×	×	×
Paper [11]	bilibili	×	3,623	60,956	278,520	×	×	×	×	×	×
DR_Four	youku	×	4,517	32,950,000	5,412,000	×	×	×	×	×	×
DR_E	youku	612	**8,156**	**57,176,457**	**6,259,558**	**17**	√	**2,170,033**	**1,264,811**	√	√

B. VOLUME OF DATASET

Table 2 illustrates the scale of our datasets and provides comparison results against previous datasets. We also included video summarization and sentiment tag annotation data from our DR_E dataset. This is so far the most comprehensive datasets with a rich set of data types that make it possible to use NPL research technologies to better understand videos.

Compared with the data collected from popular websites, our datasets come with limitations as follows: 1) The size of video summarization information is not large enough, especially for the purpose of exploring automatic summarization, where larger training data set will improve the accuracy of model building; 2) Most of existing data points are lack of sentiment tags, which hinders better understanding of user behaviors.

As shown in Table 2, paper [25], Paper [13], Paper [15] and Paper [11] have involved a large number of danmaku comments to understand video content. Although current data sets have made significant progress in understanding video content and user's watching behaviors, it is still not close to being widely applied in real scenarios.

Targeted to address the limitations above, we collected a new dataset from the video website Youku,[6] one of the largest video websites in China. We extracted danmaku comments from October, 2017 to April, 2018 and contributed to two datasets: one with four categories of danmaku comments (DR_Four) and another one with additional extension (DR_E). Some statistics of our two datasets are shown in Table 3.

DR_Four is a large-scale dataset containing four video categories: Cartoon, Entertainment, Movies and Opera, data ranging from 5rd Oct. to 30rd Nov., 2017.

While DR_Four does not cover a number of in-depth information, major refactoring efforts have been applied to the extension dataset DR_E from 5rd. Jan. to 9rd Apr.,2018. The data is constantly updated. DR_E data set also includes detailed information on: (1) play time and play mode of the video; (2) creation time and upload time by users; (3) tags and summarization of TV shows; (4) counts of replies and thumbs-up's of each traditional comment.

Comparing with existing datasets, we summarized the new features in *DR* dataset from the following aspects:

1) The largest dataset of traditional video comments, danmaku comments and users. To the best of our knowledge, *DR* is so far the most comprehensive danmaku dataset. As shown in table 3, *DR_E* contains 57,176,457 danmaku comments and 6,259,558 users, which are significantly larger than any other danmaku datasets.

[6]http://www.youku.com/

TABLE 3. The details of danmaku comments datasets: (1) #Video, #Danmaku, #Users: the number of videos, danmaku comments and uses respectly; (2) Mean_v: the mean number of comments per video; (3) Mean_c: the mean length per comments and (4) Max_user: The maximum number of an user.

Data set	Video Category	#Video	#Danmaku	Mean_v	Mean_c	#Users	Max_user	#Year
DR_Four	Cartoon	1628	3559258	2186.276	10.116	401880	6846	2017
	Entertainment	1010	6688902	6622.675	10.718	1016476	112	2017
	Movies	867	11697826	13492.302	8.584	2043899	11760	2017
	Opera	1012	10984756	10854.502	9.922	1950501	66905	2017
DR_E	Opera	8156	57176457	7010.355	9.158	6259558	47977	2018

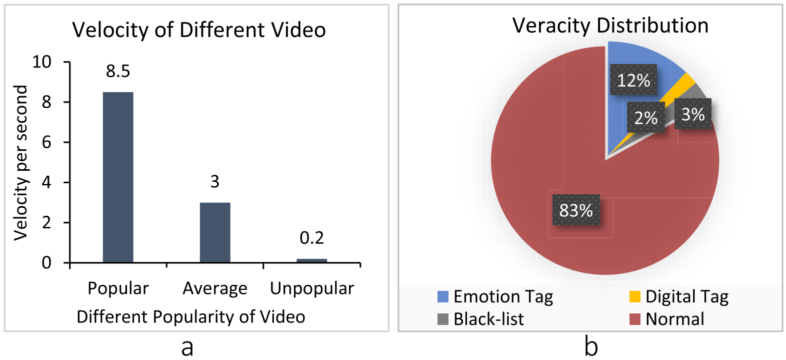

FIGURE 3. Velocity and veracity distribution of danmaku comments.

2) With multiple sentiment labels. *DR_E* contains the largest number of sentiment labels, which has not been considered in previous datasets.

3) Synopsis and summarization of videos. Due to the complexity of video data, *DR_E* considered summarization information effective for researchers to analyze videos on a large scale.

C. VELOCITY AND VERACITY OF DANMAKU

Big danamku data velocity deals with the playback speed of the danmku screen, which depending on the user's personal experience. As shown in Figure 3-a, we calculated the frequency of danmaku screens per second in different videos. In some hit episodes, the frequency of danmaku screens will reach at 8.5, which means the danmaku will cover the screen if the speed of previous is too slow. At the same time, the danmaku presents a lot of features, short, and full of noise, which hindered the further analysis. We further analyzed the veracity of the danmaku in Figure 3-b, the danmaku context presents a lot of features, such as emoticons, digital symbols, special words, which are usually the expression of the user's personal emotions. We define Veracity refers to the biases, noise and abnormality in danmaku data. Notice that normality data occupy advantage point, with 83% of the total data, which indicate that danmaku is an meaningful data.

D. EPHEMERAL AND SHORT DANMUKU COMMENTS

Here we would like to demonstrate the context characteristics of danmaku comments, specifically on the length distribution of danmaku comment as shown in Figure 4-a. Since each video has Chinese tags, we calculated the distribution of Chinese tags as shown in Figure 4-b.

From Figure 4, we have two observations: 1) danmaku comments are ephemeral and short compared with traditional media comments. Most users send danmaku comments in the length of around 4 Chinese characters. This is mainly due to the limit of 40 words imposed by YouKu website on regular danmaku comments which aims to make interaction experience convenient and interesting. 2) Users prefer sending danmaku comments when they are watching dramas, implying that users incline to send entertainment-like messages in favor of novelty, improvisation and self-expression.

E. ANONYMOUS SPACE– USER ACCESS PATTERN

The interaction of danmaku is an anonymous interaction. That is to say, the interaction of users always self-identity, which encourage users more honest and novelty self-expression.

TABLE 4. Distribution of danmaku numbers per user.

#Danmaku	#User	Pct.
>10000	4	0.00%
1000-10000	1,194	0.02%
200-1000	22,623	0.36%
101-200	48,971	0.78%
4-100	1,969,621	31.47%
2-4	223,821	3.58%
1	3,993,324	63.80%

Table 4 lists the numbers and percentages of all danmaku comments. From our datasets, we have seen that the distribution of danmaku comments is highly skewed: 63.8% of users send only 1 danmaku comments, and 98.84% of users sent less than 100 danmaku. This phenomena indicate that even

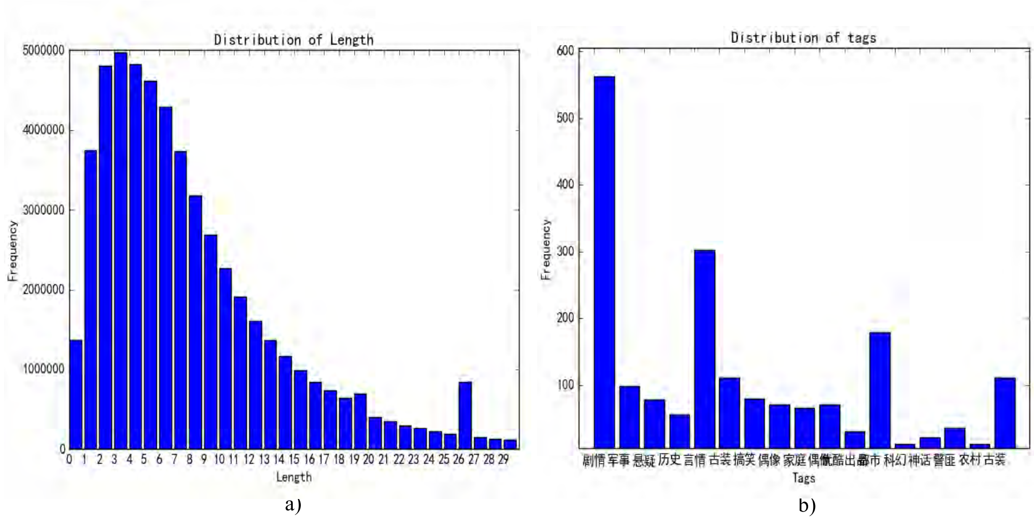

FIGURE 4. Length distribution of danmaku comment. a) is the relationship between the number and length of danmaku comments, and b) is the relationship between the number of tags and videos. (a) Distribution of danmaku length. (b) Distribution of Chinese tags.

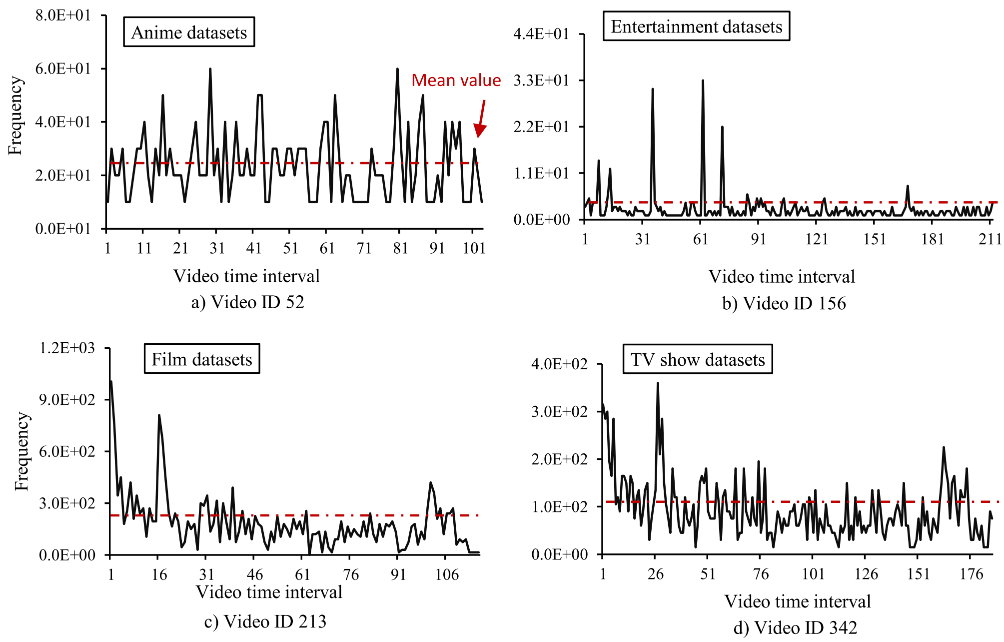

FIGURE 5. Examples of temporal distribution. For TV show datasets, the peaks correspond to the start. For the Entertainment datasets, the peaks correspond to multiple time steps for the video. The red dotted line represents the mean frequency of the comments. (a) Video ID 52. (b) Video ID 156. (c) Video ID 213. (d) Video ID 342.

in the anonymous space, users will not sent a large quantity of danmaku comments. As a part of video, most of users just watch the danmaku, and some of them will reply when they interested in the topic at current screen.

F. TIME-SYNCHRONIZATION BASED ON A VIDEO

Users can overlay anonymous comments on the video content which are aligned with the video timeline, regardless of the actual submission time. Figure 5 shows four examples of temporal distribution of Danmaku data, where x-axis denotes time intervals related to the video, and y-axis represents the frequency of danmaku comments. The red dotted line stands for the mean frequency of the comments.

From the distribution, we have observed that the growth of danmaku comments is mostly related to the video slices and the peak appears anytime, which may be caused by emotional resonance. At the same time, for different types of videos, frequency distribution shows different trends, for example, for

Class	Category	Selection Criteria	Example
Neutral	Hint the next screen	This type of danmaku is the most representative of the "high-energy ahead", more used in horror films and comedies, horror films often appear full screen "high-energy ahead" barrage, mainly the screen users with a barrage mask terror screen, reduce the degree of horror.	蛋白质含量最高的弹幕来了。 The most protein-content danmaku is coming.
Neutral	Observation	Present some facts on video and life	警告！我方失守！我方失守！ Warning! We are lost! We are lost!
Neutral	Need	Ask some questions	所以……后来凌云和谁一起了？ Then, who is together with LingYun?
Neutral	Explaining the plot	Explain the plot, or some detail in the video to help others understand the video	他们是现实中的夫妻They're couples in real life.
Neutral	Soliloquize	Meaningless review	一个人都没有；Nobody here;默默地看，不想说话； Watch silently, do not want to speak;
Positive	Spoof	Push danmaku with positive words for fun	可怜的水桶啊--这桶满分真结实 Poor bucket. This is a really solid barrel.
Positive	Thought	Personal positive opinion	bgm蜜汁好听。Very nice background music .
Negative	Spoof	Push danmaku with negative words for fun	啥玩意儿？镜头晃得我要吐了 What the hell? I'm going to puke on the camera.
Negative	Thought	Personal negative opinion	讨厌郑直，一点都不正直I hate Zheng Zhi, it's not honest. 凌云这样真讨厌Lingyun, that's disgusting.
Negative	Quarrel	Users quarrel with each other	你们是不是傻.一群疯子。。。。一群键盘出来出来了 Are you stupid? A bunch of lunatics. A bunch of keyboard-warriors came out.

FIGURE 6. Coding scheme for categorizing different types of danmaku comments.

the category of "Film" and "TV show" the peaks correspond to the start, while for the category of "Anime" the peaks correspond to multiple timestamps of the video. The reason being is viewers share empathy on video content on different levels. A large portion of "Entertainment" videos are talk shows, which by no means will attract full attentions all the time. "Film" and "TV show" have a similar distribution since videos keep users more engaged.

IV. SENTIMENT ANALYSIS WITH DANMAKU

Users share their watching experience with others by sending their comments directly on the screen when they are watching the same video. Synchronized comments floating on the screen express the feeling about particular scenes, with hints about the emotional development behind the video [1]. In this section, we will focus on the NLP Applications, starting from introducing in sentiment classification based on the *DR* dataset.

A. SENTIMENT CLASSIFICATION TASK

We frame the sentiment understanding task as a sentiment classification task. However, it is not always obvious whether a danmaku comment contains a sentiment. For these cases, we use the following litmus test: if the comment is related to only one scene of the video, it belongs to the neutral class. For example, the following comment "I'm the first one to watch the movie" can be classified onto the neutral class.

The length of a danmaku comment is usually as short as less than 20 words, and each video can easily contains thousands of danmaku comments. To prevent the interference of meaningless noisy data, we removed the comments with the length less than 5, driving the number of danmaku comments down to 34, 441, 296.

B. ANNOTATION PROCESS

Our annotators are undergraduate students in a Chinese research university, with a total of 5 students involved in the annotation task. Since it is difficult to use manual annotation method for large-scale danmaku data, open-source tools of emotion dictionaries were used to complete the construction of labeling work (e.g. Hownet emotion dictionary,[7] SnowNlp,[8] National Taiwan University NTUSD simplified Chinese effective dictionary, Chinese commendatory and derogatory Dictionary by Jun Li at Tsinghua University.)[9] At the same time, word frequency was also used to calculate the most frequent words for manual annotation.

Figure 6 shows the coding book for categorizing different types of Danmaku, where we randomly sampled 10% Danmuku comments from our datasets with a uniform distribution across time.

We measured the inter-annotator agreement using Krippendorff's alpha coefficient metrics [26], which has been

[7] http://www.keenage.com/html/c_index.html
[8] https://pypi.org/project/snownlp/
[9] http://yynl.jsnu.edu.cn/_t307/0c/b4/c541a3252/page.html

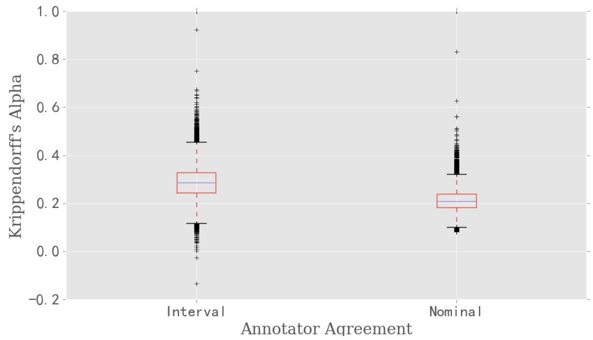

FIGURE 7. The annotator agreement result.

TABLE 5. The rule of sentiment classification.

sentiment-level	threshold
negative	0.0-0.3
neutral	0.3-0.7
positive	0.7-1.0

established as a standard disagreement measurement. Figure 7 shows the inter-annotator agreement results.

C. PROBLEM FORMULATION OF SENTIMENT CLASSIFICATION

The purpose of sentiment analysis on web comments is to categorize emotional tendency of comments based on textual contents. We classified sentiment in 3 classes, namely positive, negative or neutral feeling, defined as $(-1, 1, 0)$, respectively. Given a train data set $T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$, where $x_i \in \Re$ denotes Danmaku comments, $y_i \in \Re$ represent sentiment labels. For a video V, there are m Danmaku comments with the sharp of $m * n$, where n represents that a comment r_i comprises n words $\{w_1, w_2, \ldots, w_n\}$. Sentiment classification task aims to learn a classifier f: $X \rightarrow Y$. The Loss function is defined as $L(w) = \frac{1}{2}\sum_{i=1}^{N}(f(x_i, w) - y_i)^2$ with w_i as the learning parameter.

D. CLASSIFICATION EXPERIMENTS

1) EXPERIMENT SETTING

The goal of our research can be transformed into a three-dimensional classification task, and each Danmaku comment is labeled with a specific category. For example, we set Danmaku sentiment with values of positive, negative, or neutral. By combining the positive effective probability of SnowNlp[10] with Chinese text prediction, we divided the effect polar direction as $(-1, 0, 1)$.

In combination with Danmaku comments, it is hypothesized that a large number of comments are with neutral sentiment. The degree of danmaku sentiment are ranging from 0 to 1. The rules of the classification in the experiment datasets is defined in table 5.

[10]https://pypi.org/project/snownlp/

2) TEXT PRE-PROCESSING

Text pre-processing step ensures the danmaku comments are interesting and the initial classification data in the tasks are divided into train data and test data according to the proportion of 8:2 [27]. First, we extract all the data to build our training set from train data with the tag of -1 as Neg.csv and tag values 1 as Pos.csv, respectively. Second, all data in test data are stored as test.csv to be test dataset. The Chinese words segmentation utilities used in this paper is JieBa word segmentation tool, and then emotional stop words are used according to the vocabulary. The original corpus of Chinese affective classification is not specific to the film field, so it is obviously inaccurate for the emotional classification of the network terms in the danmaku data.

3) TRAINING THE CLASSIFIERS

We conduct experiments on our *DR_E* danmaku dataset and compare four baselines in sentiment classification.

We evaluate the state of art sentiment classification method on *DR_E* data as following. We first take 75% of the training set as training sets, another 25% as a validation set to start training the classifier, and features derived from the words appeared in the danmaku, characterized by the bi-gram of the word. We employed both traditional baselines and Deep Neural Network methods which have been shown powerful for text analysis. The following methods are include.

- LR [28]. Logistic regression LR, where LR models have subsequently been shown to be effective for text classification and have achieved excellent results, here we use the tool of sklearn.
- SVM [28]. The n-gram features are used to train a SVM classifier, here we use the tool of libSVM.
- RNN [29]. Trains a classifier using RNN to Learns sentence representation.
- CNN [30]. Convolutional neural networks (CNN) utilize layers with convolving filters for sentiment analysis.
- LSTM [27], [31]. Embeds words sentiment in the word level and implements LSTM classifier.

4) EVALUATION OF THE CLASSFIERS

To evaluate the performance of the model, we use the metrics of Precision, Recall and F-measure, which are the primary metrics used in sentiment classification.

To make a comparison of different features, we report the sentiment recognition results in table6. Note that for the danmaku dataset, the combination of different features make a good performance. The value of precision, recall rate and the comprehensive F-measure based on the emotion features of the comprehensive dictionary have obtained the best effect, which also shows that the lexical library emotion features have a significant effect on the recognition of the sentiment sentence.

E. CASE STUDY

To have an intuitive understanding of the sentiment in the video, we draw a sentiment curves for a case based on the

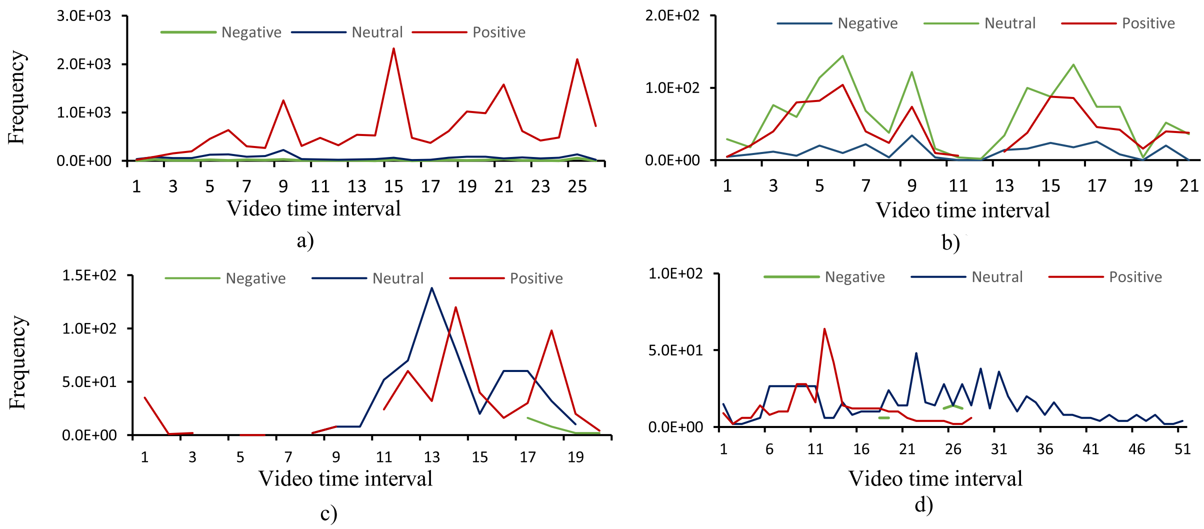

FIGURE 8. Sentiment trends in each video. The trends shows the mood of the users towards to the video, which can help predict user's likelihood of trolling. (a) Video ID 313095784. (b) Video ID 313099899. (c) Video ID 313102803. (d) Video ID 313106153.

TABLE 6. The results of the recognition of sentiment sentence with different features.

Method/Features	Precision	Recall	F-Measure
TF-IDF	0.5862	0.6264	0.6166
HowNet	0.6134	0.6793	0.6698
SnowNlp	0.6247	0.6345	0.6201
Combine	0.6334	0.7154	0.7106
LR	0.6099	0.7742	0.6125
SVM	0.5753	0.6264	0.6166
RNN	0.6120	0.7712	0.6154
CNN	0.6238	0.7691	0.6232
LSTM	0.6552	0.7670	0.7105

video time-line in Figure8. Obviously, danmaku comments with lots of personal views and user's emotional stance on different videos is also different. In Figure8 a), most of the danmaku comments are positive, which indicates that users have a better react to this kind of video. However, in Figure 8 d), user's emotions became very controversial, eventually negative emotions affected the spread, and began to dominate the screen.

V. VIDEO TAGGING AND SUMMARIZATION

Since danmaku is interesting and challenging because each text content is short hence is different from e.g., a news article. It is similar to chat room, but it has many more users watching the video at the same time. In addition, the comments are more or less related to the video on display.

In this section, we focus on using will introduce the task of video summarization. We first define the problem of the summarization task, second, we show how to model the textual features with time-sync danmaku comments, as well as the evaluation results of the dataset. Finally, we present some implementation details of the pre-process of danmaku comments.

A. PROBLEM FORMULATION

The problem of video summarization using danmaku comment can be formulated as follows. First, we define the input training set is a set of time-sync comments $C_{train} = \{c_1, c_2, \ldots, c_T\}$ with time stamps $T = \{t_1, t_2, \ldots, t_T\}$ of a video v, the label summary note as $L_{train} = s_v$. Second, for the test set C_{test}, our task is to generate the summaries $Sum(v_{test}) = \{s_1, s_2, \ldots, s_T\}$ as closely as the ground-truth of video v_{test}. Then, the problem of video summarization can be turn to an NLP task.

B. ALGORITHMS FOR COMPARISON

To comprehensively evaluate the state of the art summarization methods, we evaluate it on DR_E data by comparing with following baselines:

- Random, we developed the random baseline method that selects sentences randomly for each video as the summary.
- TextRank [32]. The TextRank algorithm is a graph model based on PageRank, which is used to generate keywords and summaries for text. The algorithm uses the local lexical relations (co-occurrence windows) to sort the subsequent keywords and extract them from the text directly.
- Seq2Seq model [33]. A general-purpose encoder-decoder framework, which uses the RNNs as encoder to predict the probability of a target sequence. Here an attention function is used to calculate an unnormalized alignment score.
- Seq2Seq + TextRank. In order to make sure the data quantity of gold label, Textrank model is used to extract the key sentence of the danmaku first, then Seq2Seq model with attention function are used for generate the summarization.

- Seq2Seq + Comments. A Seq2Seq model, which uses the video comments information to enhance the summarization results.

C. EVALUATIONS

Following the previous works of automatic text summarization, we adopt the evaluation metrics of ROUGE score [34], such as ROUGE-1, ROUGE-2 and ROUGE-L. ROUGE-N is an widely used metric to measure the quality of summarization, which computed as follows:

$$ROUGE - N = \frac{\sum_{S\epsilon ref} \sum_{gram_n \epsilon S} Count_{match(gram_n)}}{\sum_{S\epsilon ref} \sum_{gram_n \epsilon S} Count(gram_n)} \quad (1)$$

where n represents the length of n-gram, ref is the reference summaries, and $Count_{match(gram_n)}$ stands for the max number of n-grams co-occurring in a reference summary.

D. RESULTS AND ANALYSIS

Table 7 shows the video summarization result based on our danmaku data. From Table 7, we also observe an interesting phenomenon, the model combine with the comments data achieves better results. This might be due to two reasons: 1) the length of traditional comments are much longer than danmaku, thus provide more important information, and 2) the combined traditional comments dataset has more training samples, which helps to train a better network.

TABLE 7. The result of video summarization.

Method	rouge-1	rouge-2
Random	0.0045	0.0045
TextRank	0.1560	0.0180
Seq2Seq+Att	0.1772	0.0250
TextRank+Seq2Seq+Att	0.1460	0.0168
Seq2seq+Att+comments	0.1825	0.0347

VI. DISCUSSIONS

We now discuss the feasibility of using danmaku comments for guidance and influence user behavior based on findings of the archival analysis. Then we reflect on potential service based on danmaku for future application.

In the past, HCI has researched the online communities behavior problem, and there often exist some antisocial behaviors in the network disrupts constructive discussion [35]–[37]. Since the negative mood increases a user's likelihood of trolling, danmaku video platforms in China require users to verify their identity before they register the website. Trolling discussion context will be filtered by the website directly, which may have a dampening effect.

Our result echoed those found prior in that viewers send danmaku with more positive mood than negative after filtered by video platform. Since behavior can spread from user to user, negative mood and negative behavior increase trolling behavior, guidance and inflect user's behavior is crucial for online communities. To apply some NLP technologies guiding the discussion, large-scale data is needed since the quality of the danmaku sentiment detection depends largely on the results of data quality. When a user is watching a video, guidance leader provides suitable topics at appropriate times. User will follow the guidance's leading and send comments of the topics provided by guidance.

VII. CONCLUSIONS

We empirically analyzed online time-sync video comments and contributed a large-scale danmaku dataset : *DR* dataset, the first publicly available video comments for research purposes, containing 51,977K danmaku comments and 2,434K traditional comments. We analyzed the dataset, showing unique features and interesting trends such as sentiment analysis and summarization on videos. We defined two novel tasks based on *DR* dataset and proposed some baselines for further research: (i) sentiment classification based on danamku comments features and (ii) summarization using the danamku comments and traditional comments contents. Our experiments show that danmaku is an potential data for further analysis. Our work also outlined the importance of applications thinking about the user behavior and guidance.

REFERENCES

[1] X. Ma and N. Cao, "Video-based evanescent, anonymous, asynchronous social interaction: Motivation and adaption to medium," in *Proc. ACM Conf. Comput. Supported Cooperat. Work Social Comput.*, 2017, pp. 770–782.
[2] Y. Tang *et al.*, "Is Danmaku an effective way for promoting event based social network?" in *Proc. Companion ACM Conf. Comput. Supported Cooperat. Work Social Comput.*, 2017, pp. 319–322.
[3] Y. Zhao, X. Peng, J. Tang, and S. Song, "Understanding young people's we-intention to contribute in Danmaku Websites: Motivational, social, and subculture influence," in *Proc. iConf.*, 2017, pp. 359–369.
[4] Y. Yao, J. Bort, and Y. Huang, "Understanding Danmaku's potential in online video learning," in *Proc. CHI Conf. Extended Abstracts Hum. Factors Comput. Syst.*, 2017, pp. 3034–3040.
[5] Q. Wu, Y. Sang, S. Zhang, and Y. Huang, "Danmaku vs. forum comments: Understanding user participation and knowledge sharing in online videos," in *Proc. ACM Conf. Supporting Groupwork*, 2018, pp. 209–218.
[6] L. Liu, A. Suh, and C. Wagner, "Who is with you? Integrating a play experience into online video watching via Danmaku technology," in *Proc. Int. Conf. Hum.-Comput. Interact.*, 2017, pp. 63–73.
[7] X. Qian, X. Liu, X. Ma, D. Lu, and C. Xu, "Who is with you? Integrating a play experience into online video watching via Danmaku technology," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 9, pp. 1746–1757, 2016.
[8] X. Yang, W. Binglu, H. Junjie, and L. Shuwen, "Natural language processing in 'Bullet Screen' application," in *Proc. IEEE Int. Conf. Service Syst. Service Manage. (ICSSSM)*, Jun. 2017, pp. 1–6.
[9] M. He, Y. Ge, E. Chen, Q. Liu, and X. Wang, "Exploring the emerging type of comment for online videos: *Danmu*," *ACM Trans. Web*, vol. 12, no. 1, 2017, Art. no. 1.
[10] D. R. Olsen and B. Moon, "Video summarization based on user interaction," in *Proc. ACM 9th Eur. Conf. Interact. TV Video*, 2011, pp. 115–122.
[11] M. He, Y. Ge, L. Wu, E. Chen, and C. Tan, "Predicting the popularity of *DanMu*-enabled videos: A multi-factor view," in *Database Systems for Advanced Applications*. Cham, Switzerland: Springer, 2016.
[12] X. Chen, Y. Zhang, Q. Ai, H. Xu, J. Yan, and Z. Qin, "Personalized key frame recommendation," in *Proc. 40th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2017, pp. 315–324.
[13] B. Wu, E. Zhong, B. Tan, A. Horner, and Q. Yang, "Crowdsourced time-sync video tagging using temporal and personalized topic modeling," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2014, pp. 721–730.
[14] Y. Xian, J. Li, C. Zhang, and Z. Liao, "Video highlight shot extraction with time-sync comment," in *Proc. Int. Workshop Hot Topics Planet-Scale Mobile Comput. Online Social Netw.*, 2015, pp. 31–36.

[15] G. Lv, T. Xu, E. Chen, Q. Liu, and Y. Zheng, "Reading the videos: Temporal labeling for crowdsourced time-sync videos based on semantic embedding," in *Proc. 13th AAAI Conf. Artif. Intell.*, 2016, pp. 3000–3006.
[16] S. Sun, F. Wang, and L. He, "Movie summarization using bullet screen comments," *Multimedia Tools Appl.*, vol. 77, no. 7, pp. 9093–9110, 2017.
[17] W. Yang, N. Ruan, W. Gao, K. Wang, W. Ran, and W. Jia, "Crowdsourced time-sync video tagging using semantic association graph," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2017, pp. 547–552.
[18] Q. Bai, Q. Hu, F. Fang, and L. He, "Topic detection with Danmaku: A time-sync joint NMF approach," in *Proc. 29th Int. Conf. Database Expert Syst. Appl.* Cham, Switzerland: Springer, 2018, pp. 428–435.
[19] S. Ma, L. Cui, D. Dai, F. Wei, and X. Sun. (2018). "LiveBot: Generating live video comments based on visual and textual contexts." [Online]. Available: https://arxiv.org/abs/1809.04938
[20] X. Wang, Y.-G. Jiang, Z. Chai, Z. Gu, X. Du, and D. Wang, "Real-time summarization of user-generated videos based on semantic recognition," in *Proc. 22nd ACM Int. Conf. Multimedia*, 2014, pp. 849–852.
[21] Y. Chen, Q. Gao, and P.-L. P. Rau, "Watching a movie alone yet together: Understanding reasons for watching Danmaku videos," *Int. J. Hum.-Comput. Interact.*, vol. 33, no. 9, pp. 731–743, 2017.
[22] Z. Li, R. Gu, and G. Xie, "Measuring and enhancing the social connectivity of UGC video systems: A case study of YouKu," in *Proc. IEEE 19th Int. Workshop Qual. Service*, Jun. 2011, Art. no. 36.
[23] C. Li and J. Liu, "Large-scale characterization of comprehensive online video service in mobile network," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–7.
[24] L. Xu and C. Zhang, "Bridging video content and comments: Synchronized video description with temporal summarization of crowdsourced time-sync comments," in *Proc. AAAI*, 2017, pp. 1611–1617.
[25] Z. Liao, Y. Xian, X. Yang, Q. Zhao, C. Zhang, and J. Li, "TSCSet: A crowdsourced time-sync comment dataset for exploration of user experience improvement," in *Proc. ACM 23rd Int. Conf. Intell. User Interfaces*, 2018, pp. 641–652.
[26] M. W. Bauer, "Content analysis. An introduction to its methodology—By Klaus Krippendorff from words to numbers. Narrative, data and social science—By Roberto Franzosi," *Brit. J. Sociol.*, vol. 58, no. 2, pp. 329–331, 2007.
[27] Q. Hu, J. Zhou, Q. Chen, and L. He, "SNNN: Promoting word sentiment and negation in neural sentiment classification," in *Proc. 32nd AAAI Conf. Artif. Intell.*, New Orleans, LA, USA, Feb. 2018.
[28] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "LIBLINEAR: A library for large linear classification," *J. Mach. Learn. Res.*, vol. 9, pp. 1871–1874, Jun. 2008.
[29] R. Socher *et al.*, "Recursive deep models for semantic compositionality over a sentiment treebank," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2013, pp. 1631–1642.
[30] Y. Kim. (2014). "Convolutional neural networks for sentence classification." [Online]. Available: https://arxiv.org/abs/1408.5882
[31] Q. Hu, Y. Pei, Q. Chen, and L. He, "SG++: Word representation with sentiment and negation for Twitter sentiment classification," in *Proc. 39th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr. (SIGIR)*, New York, NY, USA, 2016, pp. 997–1000.
[32] R. Mihalcea and P. Tarau, "TextRank: Bringing order into text," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2004, pp. 1–8.
[33] D. Bahdanau, K. Cho, and Y. Bengio. (2014). "Neural machine translation by jointly learning to align and translate." [Online]. Available: https://arxiv.org/abs/1409.0473
[34] C.-Y. Lin, "ROUGE: A package for automatic evaluation of summaries," in *Proc. Text Summarization Branches Out*, 2004, pp. 1–8.
[35] J. Cheng, C. Danescu-Niculescu-Mizil, J. Leskovec, and M. Bernstein, "Anyone can become a troll," *Amer. Sci.*, vol. 105, no. 3, p. 152, 2017.
[36] C. Fiesler and B. Hallinan, "'We are the product': Public reactions to online data sharing and privacy controversies in the media," in *Proc. CHI Conf. Hum. Factors Comput. Syst. (CHI)*, Montreal, QC, Canada, Apr. 2018, Art. no. 53. doi: 10.1145/3173574.3173627.
[37] D. Lottridge and F. R. Bentley, "Let's hate together: How people share news in messaging, social, and public networks," in *Proc. CHI Conf. Hum. Factors Comput. Syst. (CHI)*, New York, NY, USA, 2018, pp. 60-1–60-13.

QINGCHUN BAI received the M.S. degree in computer science from the Zhengzhou University of Light Industry, Zhengzhou, in 2016. She is currently pursuing the Ph.D. degree with the Department of Computer Science and Technology, East China Normal University, Shanghai, China. Her research interests include social computing, text mining, context-aware computing, and neural networks.

QINMIN VIVIAN HU received the Ph.D. degree in computer science from York University, Toronto, Canada. She was an Associate Professor with East China Normal University, Shanghai, China, and a Postdoctoral Fellow at the MRI research facility, the Wayne State University, USA. She is currently an Assistant Professor with the Department of Computer Science, Ryerson University, Toronto. She has published more than 30 referred papers in top tier journals, such as the IEEE Transactions on Knowledge and Data Engineering and *ACM Transactions on Intelligent Systems and Technology*, and conferences, such as AAAI and ACM SIGIR. She received the National NSERC Postdoctoral Fellowship as one of the best Ph.D. fellows in Canada, in 2013.

LINHUI GE received the master's degree in electrical engineering from Wuhan University, China. He has been working in the software industry for more than 15 years, focusing on distributed system and data analysis. He is currently a Senior Engineering Manager with Amazon.com, Inc., Seattle, WA, USA, where he led a number of product initiatives and launches across search experience, advertisement, and voice assistant (Alexa) in the past eight years. One shared theme of his work in these different domains is observed as how to leverage large data sets to identify user patterns to provide guidance on product definition and customer experience improvements.

LIANG HE received the bachelor's and Ph.D. degrees from the Department of Computer Science and Technology, East China Normal University, China, where he is currently a Professor and the Director. In recent years, he has hosted a number of National Science and Technology Support and has participated in the National 13th "Five-Year" Technology Support Programs, the Shanghai Science and Technology Long-Term Development Plan, and the Shanghai 13th "Five-Year" Science and Technology plan. He has obtained more than 10 patents and has published two monographs and more than 70 refereed papers in national and international journals and conference proceedings. His current research interests include knowledge processing, user behavior analysis, and context-aware computing. He is also a Council Member of the Shanghai Computer Society, a member of the Academic Committee, the Director of the Technical Committee of the Shanghai Engineering Research Center of Intelligent Service Robot, and a technology foresight expert of the Shanghai Science and Technology in focus areas. He has been awarded the Star of the Talent in Shanghai. He received the Shanghai Science and Technology Progress Award for five times and the First Prize, in 2013, and the second prize, in 2015.