IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 7, 2021, accepted March 12, 2021, date of publication March 17, 2021, date of current version March 25, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3066484

A Combined Extractive With Abstractive Model for Summarization

WENFENG LIU, YALING GAO, JINMING LI, AND YUZHEN YANG
School of Computer, Heze University, Heze 274015, China

Corresponding author: Yaling Gao (gaoyaling@hezeu.edu.cn)

This work was supported in part by the Shandong Province Social Science Popularization and Application Research Project under Grant 2020-SKZZ-51, in part by The Social Science Planning Office of Heze City under Grant 2020_zz_55, in part by the Shandong Province Social Science Planning Project, and in part by Heze University Doctoral Research and Development Fund under Grant XY20BS19, and in part by Shandong Province Educational Science Planning under Grant BYZN201910.

ABSTRACT Aiming at the difficulties in document-level summarization, this paper presents a two-stage, extractive and then abstractive summarization model. In the first stage, we extract the important sentences by combining sentences similarity matrix (only used for the first time) or pseudo-title, which takes full account of the features (such as sentence position, paragraph position, and more.). To extract coarse-grained sentences from a document, and considers the sentence differentiation for the most important sentences in the document. The second stage is abstractive, and we use beam search algorithm to restructure and rewrite these syntactic blocks of these extracted sentences. Newly generated summary sentence serves as the pseudo-summary of the next round. Globally optimal pseudo-title acts as the final summarization. Extensive experiments have been performed on the corresponding data set, and the results show our model can obtain better results.

INDEX TERMS Extractive summarization, abstractive summarization, beam search, word embeddings.

I. INTRODUCTION

With the explosive growth of text data on the web, how quickly obtain the nut graph or thematic meaning of long text is a vitally important research in natural language processing. This task is further referred as the text summarization or text semantic extraction and generation [1], [2]. Depending on the form of obtaining and outputting, text summarization is generally divided into the following two paradigms, extractive and abstractive [3], [4]. Extractive selects important sentences as the summary of the document, while abstractive mainly obtain the summarization by generating, rewriting, similar to the knowledge extraction of the human brain [5], [6].

The features commonly used in extractive summarization are mainly sentence position, part of speech, word frequency, sentence length, etc [7]. Nonetheless, vector space representation is to map text to vector space, use matrix decomposition, dimensionality reduction or by calculating the similarity of sentences in documents, and then choose the optimal sentence as the summary [8], [9]. The graph-based method is to treat the content of the sentence in each document as the representative structure of the graph, use the word or sentence as the node in the graph, and utilize the relationship in the graph for the summary [10], [11]. The method of combinatorial optimization can acquire the optimal solution by using methods such as integer linear programming, submodular function or the appropriate combinatorial optimization through the coverage and diversity of the summarization.

Words and sentences in the summary obtained by the abstractive methods maybe not in the document. The mainstreams are based on encoder-decoder or sequence-to-sequence models. Encoder-decoder can be assigned to sequence-to-sequence models too. For the past few years, to solve the problem of Out Of Vocabulary (OOV), there are two typical summarization methods in the abstractive summary, one is the copy network (CopyNet) [12], and the other is the pointer network structure (Ptr-Net) [13]. CopyNet's network structure uses an end-to-end training model. Its framework is based on the encoder-decoder model of RNN. There are two sub-modules for addressing OOV. One is the generation module and the other is the copy module. This model deals with the process of generating words differently, and it owns a state update mechanism.

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wang.

However, content generated by abstractive methods are often suffering issues such as poor readability, data redundancy, and large semantic deviations from source. So they cannot convey the semantics of the document [14], [15]. While the extractive ones often face one-sidedness, coverage narrow, so they cannot convey the overall semantics of the document [16]. In order to obtain a better summary of the document, we propose a two-stage, hybrid abstractive and extractive, summarization that combined the advantages of the two methods. This method does not prescribe the document-summarization pairs, so it is an unsupervised learning model. Firstly, the extractive model selects the most significant sentences in the document. Taking into account the differences or features of sentences (such as word embeddings, sentence position, paragraph position, etc.), we extract several notable ones. Embeddings of extracting sentences are constructed based on the syntactic dependent blocks. The second step is the abstractive model that rewrites these sentences as the final summarization. We use the beam search algorithm on the syntactic blocks of extracted sentences. The sentence that has the highest score will serve as the pseudo-summary. The next cycle will be executed. The best pseudo-summary is the last summarization. Extensive experiments have been conducted on duc2004 and Chinese dataset. Ultimately, we have achieved better results.

II. RELATED WORKS

Summarization is a hot topic in current natural language processing. This task has two main paradigms: extractive and abstractive. According to the number of source documents, it can be divided into Single Document Summary (SDS) and Multi-Document Summary (MDS) [17]. According to the adoptive technology, it has a tripartition: graph-based (such as TextRank) [18], neural-network-based, clustering-based [19], [20]. Early research mainly focused on extractive [21], [22]. In recent years, with the increase of deep learning and large corpora, many scholars have conducted extensive research on document-level summary.

Deep learning based is a very popular method in recent years. It uses sequence models for processing such as RNN, LSTM, BERT, etc. Then uses attention mechanism [23], [24], fine-grained [25], SummaRuNNer [26], etc. Abstractive utilizes encoder-decoder or sequence model (Seq2Seq) as input, including convolutional neural networks, recurrent neural networks, long-term and short-term memory networks, and gated networks. Qian *et al.* [27] use multiple neural network models to choose important words. In particular, it has been used to optimize non-differential measures of language generation and reduce exposure bias [28], [29]. Henß *et al.* [30] exploit Q-learning-based reinforcement learning models that are suitable for single-document and multi-document summaries. Paulus *et al.* [31] utilize weighted machine learning and reinforcement learning (ML+RL), mix loss functions to achieve the stability and linguistic fluency. These methods mainly use reinforcement learning, bridge the sentence extractor for end-to-end training.

Ling and Rush proposed a coarse-to-fine method that firstly extracted a sentence, and then used reinforcement learning to bridge the factorized representation, and finally generated the answer. Zhou *et al.* [32] proposed selective gates to increase abstract generalizations. Tan *et al.* [33] used the extraction and synthesis on Question Answering system (QA). To solve the problems in the extended auto-summarization method, the literature [34] proposed a two-phase auto-summarization named TP-AS. It combines pointer and attention mechanism. For the high cost of large-scale corpus tagging, many experts have adopted unsupervised methods, such as TF-IDF which utilizes the statistical feature-based, cluster-based, and graph-based [35], [36]. However, the current summarization rarely involves the syntactic structure, and often uses a large number of parameters, so those models are more complex.

Our proposed method, which utilizes extractive and abstractive, is an emerging way. Integrating extractive and abstractive can produce better quality. In a general way, extractive will extract a certain number of significant sentences, and then execute abstractive method on the extracted sentences.

III. A HYBRID EXTRACTIVE-ABSTRACTIVE TEXT SUMMARY MODEL

A. BASIC FRAMEWORK

This paper proposes a hybrid extractive-abstractive two-stage summary generation model. As shown in Figure 1, in the first stage, the neural network attention mechanism and sentence display features are used to extract the most important sentences (as candidate summary sentences), which is to extract top-k sentences by segmenting (paragraphs, sentences and words), and use attention mechanism (including sentence position, paragraph position, keywords, sentence relations, etc.). In the second phase, the extracted sentences are to the dependent syntactic analysis, and then they will be divided into different syntactic component blocks. For each word in the syntactic blocks, a distributed vector representation and attention mechanism are used to construct a simplified syntactic block vector. Beam search algorithm is performed for syntactic recombination. The best one will act as the document summarization.

This paper implements the two stages and combines the two abstract parts. For document-summary pairs $\{x_i, y_i\}_{i=1}^{N}$, N is the number of pairs. The target is to construct an approximate:

$X \xrightarrow{h} Z \xrightarrow{f} Y$, $X = \{x_i\}_{i=1}^{N}$, $Z = \{z_i\}_{i=1}^{N}$, $Y = \{y_i\}_{i=1}^{N}$, $1 \leq i \leq N$. X is the set of original document sentences, Z corresponds to several sentences extracted from the source document, and Y is the summarization, $h(x_i) = z_i, f(z_i) = y_i$, $1 \leq i \leq N$. h and f correspond to the extractive model and the abstractive model, respectively. Our method has combined both extractive and generative methods. In Figure 1, the most important top-k sentences of the document (the dark blue frame) are extracted based on the similarity between

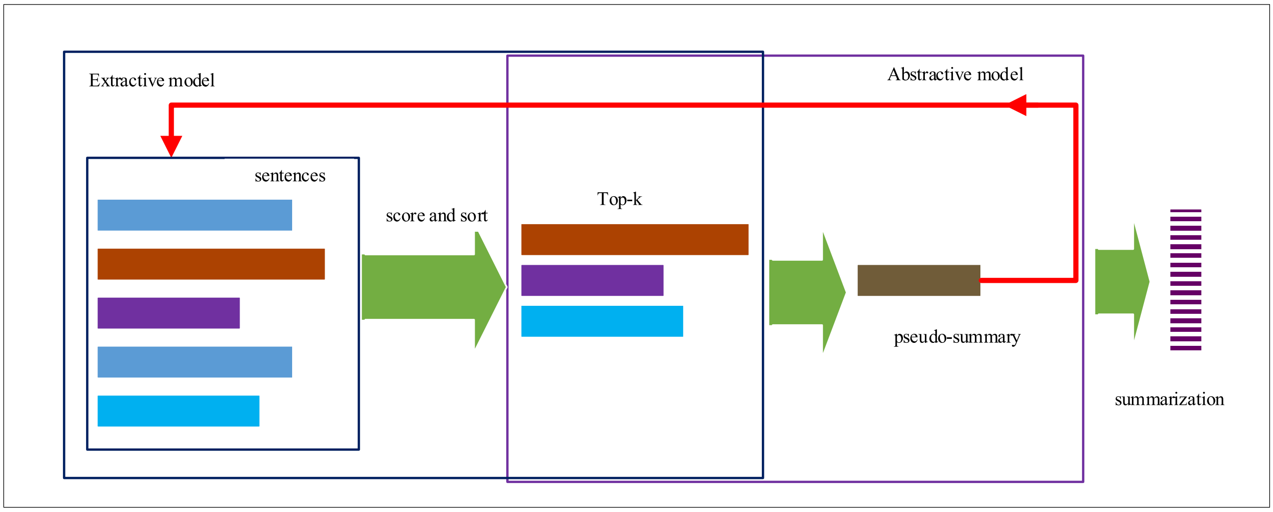


FIGURE 1. The basic framework of our model.

sentences in document. The bright blue frame is the abstractive module.

B. SENTENCE REPRESENTATION

In this section, sentence vector representation is constructed by the dependent syntax and word embeddings in the sentence. After word segmentation, word embeddings of each word is obtained through pre-training. The sentences are sent to the syntactic analysis module for obtaining the dependency tree. According to the dependency relationship, we separate them for different syntactic blocks. The stop words inside the syntactic block are deleted, and the syntactic block dependency relationship label is kept unchanged. For the interior of syntactic blocks, the syntactic representation is constructed by combining the dependent distances from the core word.

C. SYNTACTIC BLOCK ALIGNMENT

In order to avoid the differences caused by the syntactic structure in sentence representation, we use syntactic block alignment to ensure that the same syntactic components appear at the same position in different sentences. We adjust the syntax according to the dependency tags or the order of the blocks (the passive sentence should be changed into the active sentence uniformly, and different syntactic components are normalized). Finally, we concatenate the block vectors of the sentence for the sentence representation.

D. EXTRACT IMPORTANT SENTENCES

The explicit features of sentences mainly include sentence position, paragraph position, key words, key sentence, and others. We can precisely observe those features from the document. This information plays a very important role in the extractive phase. In addition, there are certain specific relationships between sentences, that is, the logical relationship of theirs. We can seek the semantic evolution from the sentence context. The semantical coherence of logical content is evident in a structured hierarchy. Therefore, this section models the relationship between sentences on basis of the above. Figure 2 depicts a tree-like sentence relationship for modeling sentence relationships. We score the sentence according to the connectives contained sentence position, paragraph position or other features.

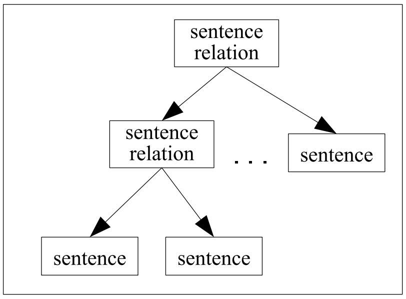


FIGURE 2. The diagram of sentence relation.

1) CONSTRUCT LOGICAL CONNECTIVE DICTIONARY

Combining the commonly used sentence connectives, we construct and build a logical connective dictionary. Some connectives are given in Table 1. Conjunctions are utilized to concatenate words, phrases, or sentences, and to represent certain logical relationships. This section focuses on conjunctions that appear between sentences after segmentation. Conjunctions indicate relationships such as juxtaposition, transition, cause and effect, choice, assumption, comparison, concession, etc. We weight summative, progressive, and turning sentences to increase the importance of the homologous ones. The main reason is that the model we proposed combines a variety of features including deep learning features (word embeddings, Beam search), and text display features (paragraph location features, sentence location features, etc.), which has more semantic representation capabilities.

TABLE 1. Weight of Chinese connectives.

Logical Type	Examples	Sentence Weight
Summative	总之, 总而言之, 综上所述, 简言之等(in short, in summary, to sum up, in a nutshell etc.).	0.5
Adversative	却,但是,然而,而,偏偏,只是,不过,至于,致,不料,岂知等(but, yet, however, on the contrary, turned out, although, nevertheless, while etc.).	0.3
Causal	原来,因为,由于,以便,因此,所以,是故,以致等(for, because, since, as, due to, because of, the consequence of, thus, so, etc.).	0.3
Parallel	首先,其次,并且,第一, 第二,和,与等(and, both and, as well as, firstly, secondly, etc.).	0.2
Selective	或,抑,非…即,不是…就是等(or, either or, or else, otherwise, etc.).	0.2
Assumptive	若,如果,若是,假如,假使,倘若,要是,譬如等(if, provided, supposing, only if, as long as, unless, on condition, but for, etc.).	0.2
Comparative	像,好比,如同,似乎,等于,不如,不及,与其…不如,若…则,虽然…可是等(likewise, similarly, in the same way, like, not as good as, if then, etc.).	0.2
Concessive	虽然,固然,尽管,纵然,即使等(though, even if, even though, as, while, though, despite, whereas, etc.).	0.2
Others		0.1

2) WEIGHTING SENTENCE BY LOGICAL CONNECTIVES

The weight of logical connectives is defined by W_l. *Condition_1* denotes that the sentence contains summary words. *Condition_2* represent that the sentence contains words of progressive or transition type, and *Condition_3* denotes sentences contain other general connectives. *Condition_4* represents general sentences. Some examples of logical words are shown in the Table 1.

$$W_l = \begin{cases} 0.5, & Condition\_1 \\ 0.3, & Condition\_2 \\ 0.2, & Condition\_3 \\ 0.1, & Condition\_4 \end{cases} \tag{1}$$

3) THE PARAGRAPH WEIGHT

The paragraph weight can be calculated as:

$$W_{(d_i)} = \begin{cases} 0.5^{n-1}, & n \in \left[1, \dfrac{m}{2}\right] \\ 0.5^{|m-n|}, & n \in \left[\dfrac{m}{2}, m\right] \end{cases} \tag{2}$$

where W_{d_i} represents the weight of the i-th paragraph, and m is the total number of paragraphs in the document. As can be seen, the weight of the first and last paragraphs is relatively high, and reduced by the second paragraph and the penultimate paragraph, etc. i takes an integer from 1 to m, and the weight of m paragraphs is normalized as:

$$W'_{d_i} = \frac{W_{d_i}}{\sum_{j=1}^{m} W_{d_j}} \tag{3}$$

4) SENTENCE WEIGHT IN A PARAGRAPH

The weight of the position of sentence in paragraph is similar to the weight of paragraph in document, and it can be represented as:

$$W_{s_i} = \begin{cases} 0.5^{n-1}, & n \in \left[1, \dfrac{|ds|}{2}\right] \\ 0.5^{||ds|-n|}, & n \in \left[\dfrac{|ds|}{2}, |ds|\right] \end{cases} \tag{4}$$

where W_{s_i} represents the weight of the i-th sentence in paragraph, and $|ds|$ denotes the total number of sentences in a paragraph, and we normalize the weight is:

$$W'_{s_i} = \frac{W_{s_i}}{\sum_{j=1}^{m} W_{s_j}} \tag{5}$$

5) WORDS WEIGHT IN A SENTENCE

For all words in a sentence, we weight words in the light of importance in the document after removing the stop words. For reducing the impact of word frequency differences, we use an improved TF-IDF version to calculate the weight of words.

$$W_{c_i} = \frac{lg\left(F_{c_i}+1\right) \times lg\frac{N}{N_{c_{ik}}}}{\sqrt{\sum_{j=1}^{h}\left(\lg\left(F_{c_j}+1\right) \times lg\frac{N}{N_{c_{ik}}}\right)^2}} \tag{6}$$

N is the number of words after segmentation excluding the stop words. $N_{c_{ik}}$ denotes the number of word c_i that appeared in document. And W_{c_i} represents the result about word c_i. We score the sentence according to words appeared in the sentence.

$$W'_{c_i} = \frac{\sum_{j=1}^{|s|} W_{c_j}}{|s|} \tag{7}$$

where $|s|$ represents the number of words in the sentence.

6) SIMILARITY MATRIX OF SENTENCES

We dynamically constructed the similarity matrix based on the euclidean distance between sentences in a document. Assuming that the document has N sentences, the similarity of the sentences is a matrix of $N{*}N$. a_{ij} is the similarity value between sentence i and sentence j. In the matrix, the sum of row i (or column j) indicates the importance of the i-sentence. The sum of every row (or column) will be sorted. Top-k (or k-highest similarity with pseudo-title in the next round) sentences are used as candidate summaries which are used as input for the abstractive model.

$$SIM = \begin{bmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{bmatrix} \tag{8}$$

$$Wsim_{s_i} = \frac{\sum_{j=1}^{|D|} a_{ij}}{\sum_{k=1}^{|D|} \sum_{j=1}^{|D|} a_{kj}} \tag{9}$$

where $|D|$ represents the number of sentences in document, and $Wsim_{s_i}$ denotes the normalized weight of i-th sentence. If the similarity of two sentences is too high, and one sentence is already in the candidate set. We add a threshold that can ensure redundancy of the candidate sentences. The comprehensive score of the i-th sentence can be calculated as:

$$W_S = \alpha W_{l_i} + \beta W'_{d_i} + \gamma W'_{s_i} + \delta W'_{c_i} + \mu Wsim_{s_i} \tag{10}$$

where $\alpha + \beta + \gamma + \delta + \mu = 1$, $\alpha, \beta, \gamma, \delta$ and μ are regulatory factors.

Extractive model sorts the top-k sentences scored by Eq.(10). The sentence with the highest weight is used as the initial pseudo-title.

E. REDUNDANCY OF CANDIDATE SENTENCES

The extractive model scores the sentences based on the previous content, and obtains the top-k important sentences in the document. If there is already a sentence in the candidate summary sentences. In order to obtain better top-k sentences and remove very similar sentences, we set a threshold to prevent completely consistent or very similar sentences from adding to the candidate sentences. The limit threshold is as follows:

$$W_s^h = \begin{cases} W_S, & Wsim_{s_{ih}} < 0.8 \\ 0, & others \end{cases} \tag{11}$$

where $Wsim_{s_{ih}}$ is the maximum similarity between the current sentence with the candidate ones. If their similarity value is greater than 0.8, it cannot be put into the candidate set which serves as input to the next stage.

F. SUMMARIZATION GENERATION

We syntactically analyze the extracted sentences, divide them into different dependent syntactic blocks, and then use beam search algorithm on the syntactic blocks. The best results act as the pseudo-summary in the next round. The pseudo-summary generated by the algorithm in this round is used as a pseudo-summary for the next round of loops until it converges to the optimal value which acts as the final summarization. The pseudo codes are shown in Algorithm 1.

G. BEAM SEARCH

Traditional broad search algorithm can find the optimal result by traversing all the nodes, but the consumption of resources increases exponentially, if the searching-space is comparatively large, it will cause excessive or insufficient memory consumption. The Beam Search algorithm can optimize and greatly reduce the searching-space and time cost, and it only maintains the specified number of nodes. Initially, only the starting nodes are stored and subsequently we add other related ones to the ordered sequence. We only preserve the specified numbers of nodes during processing until all subsequent nodes visited. This section utilizes Beam Search algorithm to combine the syntactic components and constitute a sentence. In order to prevent the deviation of syntactic components, it is necessary to align the syntactic components and perform anaphora resolution.

As shown in Figure 3, we perform column search on the syntactic blocks of candidate summary sentences, the green syntactic blocks are the syntactic component of sentence_1, the blue ones are the syntactic component of sentence_2, and the yellow ones are of sentence_3. Since the number of syntactic component blocks in various sentences may be different, some syntactic blocks may be nonexistent. To ensure the effectiveness of the final summarization, the same column is the same syntactic components. In addition, disambiguation of referential pronouns in adjacent sentences will be solved.

Algorithm 1 Abstractive Summarization Model

```
Input: $m$ sentences $S = \{s_1, s_2, s_3, \ldots, s_m\}$;
Output: a summary sentence $S_{summary}$;
1: Syntactically analyze each sentence to obtain the syntactic dependency of its words in $S$.
2: According to the dependency relationship of the words in the sentences, we divide the corresponding sentence into different dependency syntactic blocks, and record the dependency relationship of the syntactic block:
   $s_{i\_block} = \{block_1, block_2, block_3, \ldots, block_n\}$;
3: Normalize the syntactic blocks, and fulfill the missing syntax chunks according to the context;
4: Use beam search algorithm on the generated syntactic blocks;
5: Scoring the generated summary sentences, the highest $S_{summary}$ serves as summarization.
6: The result will be put into the extractive model for the next round.
```

IV. EXPERIMENT

A. EXPERIMENTAL DATA

In order to verify the effect of our model, we do a lot of experiments on Chinese data and DUC-2004. Chinese data that we used are collected about 1200 news texts from large portals such as Sina and NetEase. We utilize the genism software package for Word2Vec tool (https://radimrehurek.com/gensim/). Pre-trained word embeddings are obtained by training of the People's Daily Corpus and Wikipedia. The dimension of word embeddings is 300, other parameters are used by default parameters of gensim. The title of the corresponding article plays as the standard summary for evaluations. We manually construct more reference summaries that needed with corresponding text.

B. BASELINES

- TextRank [10], similar to PageRank, is just an unsupervised algorithm. It is primarily used for keyword extraction or summarization, and mainly treats the

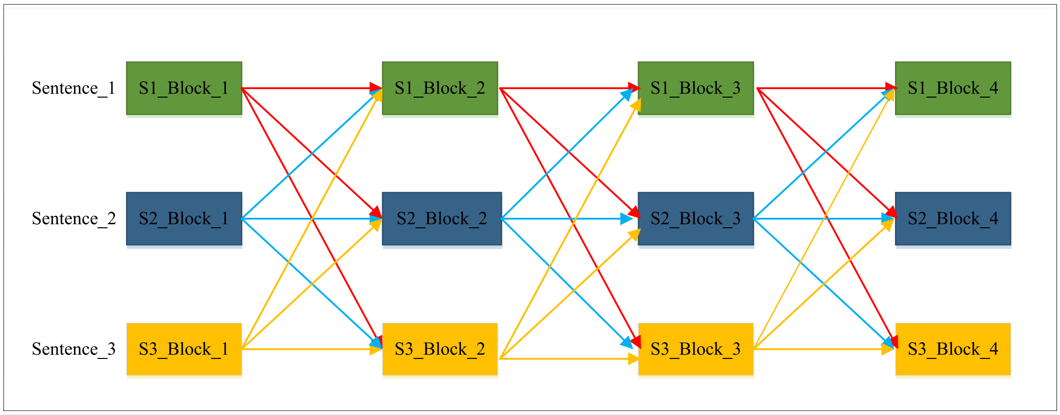


FIGURE 3. **Beam search model.**

Algorithm 2 BeamSearch

```
Input: syntactic blocks of $i$ sentences $\{S_{1\_block}, S_{2\_block}, S_{3\_block}, \ldots, S_{i\_block}\}$;
Output: the optimal sequence order of the syntactic blocks and the final summary $S_{summary}$;
$M$ = number of standardized syntactic blocks in a sentence;
$N$ = number of Beams;
1: Align the syntactic blocks of the input sentences according to the dependency patterns of theirs;
2: Pronouns in the context are replaced with their real content;
3: Add the first syntactic chunks of these sentences to the hash table and the first column of Beam;
4: Append subsequent syntax blocks of these sentences to the ordered sequence, processing all node in turns;
5: For $i$ in $M$:
    Combine the $N$-optimal sequence before $i$-$1$ blocks with the $i$-th column;
    And still retain the optimal $N$ sequences;
6: The optimal sequence of syntactic blocks is used as the final result, and we concatenate the optimal syntactic blocks sequentially as the final result $S_{summary}$.
```

relationship of sentences as a voting system which is used to construct a TextRank network.

- GraphSum [37] has integrated the information such as headings, word frequency, and other features into the summarization, and it can construct a network graph of text.
- CSAE [38] incorporates syntactic information, semantics, statistics, and ranks for the summarization.
- ADOAT [39] has fully considered the similarity between sentences, the structural information, and the core sentence.
- PGNet & PGNet+ [6] are models with a hybrid pointer-generator network. The PGNet+ model additionally utilizes coverage mechanism to address the word repetition in the generated sentence.
- KEDBS [40] is an abstractive model that exploits the content-introducing approach to neural text generation.

C. EVALUATION

We have made use of the average accuracy P, recall R, and average $F1$ as well as ROUGE for evaluating our model.

$$\mathrm{P} = \frac{1}{n}\sum_{i=1}^{n}\frac{|a_i \cap b_i|}{|b_i|} \tag{12}$$

$$\mathrm{R} = \frac{1}{n}\sum_{i=1}^{n}\frac{|a_i \cap b_i|}{|a_i|} \tag{13}$$

$$\mathrm{F1} = \frac{2PR}{P+\mathrm{R}} \tag{14}$$

In the above formulas, a_i represents the summary generated by the corresponding model. b_i denotes the summarization of the real title or artificial markup. Moreover, we also use ROUGE (ROUGE-1, ROUGE-2, ROUGE-L) to evaluate our model. The formula is as follows.

$$ROUGE_N = \frac{\sum_{s\in ref}\sum_{N-gram\in s} count_{match}\,(\text{N-gram})}{\sum_{s\in ref}\sum_{N-gram\in s} count\,(\text{N-gram})} \tag{15}$$

ref represents the reference summaries. $count_{match}$ (N_gram) denotes the number of matches generated by the algorithms and the reference summaries. $count_{match}(N\_gram)$ denotes the number of matches generated by the algorithms with the reference summaries. $count$ (N_gram) represents the number of N_grams in the reference summaries.

D. EXPERIMENTAL RESULTS AND ANALYSIS

As is shown in Table 2, S2, S3, S4, and S5 respectively indicate the number of generated sentences. For example, S2 indicates two sentence summary results, etc. GraphSum, CSAE, and ADOAT are better than the results of the TextRank. It indicates that the more features contained, the better result.

TABLE 2. Multi-sentence examples of summarization.

Models	Evaluation Indexes	S2	S3	S4	S5
TextRank	P	0.369	0.344	0.322	0.325
	R	0.328	0.346	0.352	0.381
	F1	0.347	0.345	0.342	0.351
GraphSum	P	0.388	0.334	0.330	0.338
	R	0.389	0.389	0.409	0.424
	F1	0.388	0.359	0.365	0.376
CSAE	P	0.375	0.326	0.332	0.340
	R	0.393	0.410	0.432	0.455
	F1	0.384	0.363	0.375	0.389
ADOAT	P	0.453	0.437	0.422	0.413
	R	0.435	0.471	0.493	0.530
	F1	0.444	0.453	0.455	0.464
Our model	P	0.512	0.509	0.513	0.510
	R	0.507	0.511	0.512	0.512
	F1	0.508	0.510	0.512	0.511

TABLE 3. ROUGE on Chinese data and DUC-2004.

	DUC-2004			Chinese dataset		
Model	ROUGE-1	ROUGE-2	ROUGE-L	ROUGE-1	ROUGE-2	ROUGE-L
GraphSum	15.26	6.27	12.61	8.82	4.81	6.27
CSAE	21.43	8.72	18.34	17.54	6.07	8.71
ADOAT	26.17	9.86	21.52	18.02	7.11	11.85
PGNet	27.54	11.23	23.85	21.67	7.06	12.03
PGNet+	29.59	12.03	24.05	22.06	8.28	12.76
KEDBS	32.28	12.42	25.86	24.00	8.75	13.72
Our model	36.06	13.61	29.21	26.56	10.08	15.25

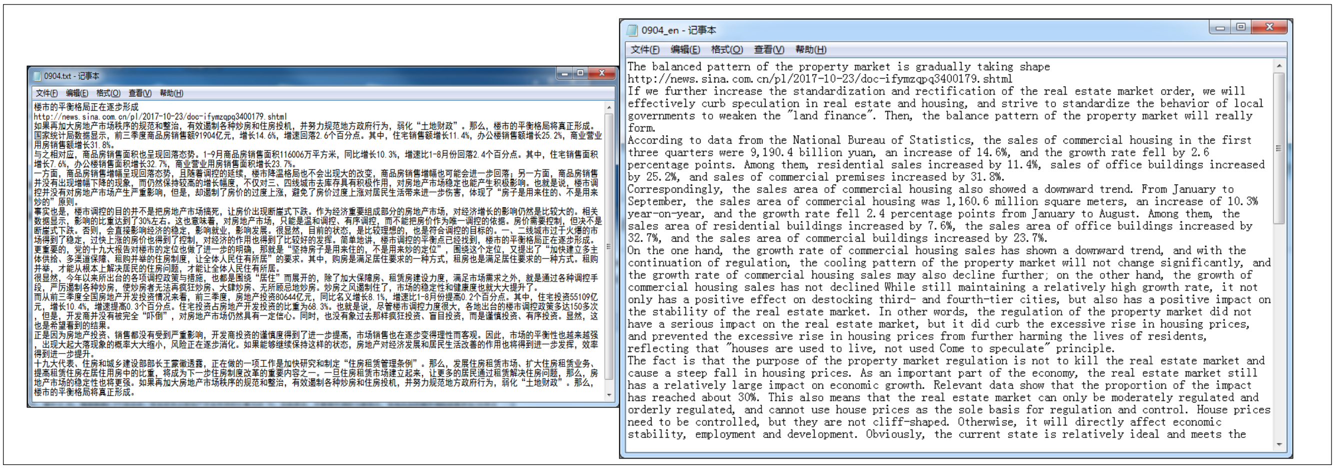
After incorporating statistical information such as syntactic information and semantics, the CSAE model obtains better results than GraphSum which only takes into account title and sentence position. Since the ADOAT algorithm further reduces and optimizes the generated summaries, the result that obtained is better than that of CSAE. Our model has taken into account the dependency syntactic relations and various information, furthermore it constructs dependent syntactic blocks which have taken full advantage of the attention mechanism. Experimental results of our model have greatly improved in terms of accuracy, recall, and $F1$ value.

The effects of the abstractive models are better than these extractive ones from Table 3. The main reason is that the abstractive can fully express the semantics of the document. PGNet+ employs coverage mechanism to avoid word redundant. Therefore it gets better results than PGNet. KEDBS model exploits the content-introduced for text generation, and it has stronger expressive ability than pointer-generator network (PGNet & PGNet+). Our model combines the advantages of extractive and abstractive summaries. First, we obtain important sentences through the extractive model, and then the syntactic blocks of these sentences are reorganized by beam search. The best one acts as the final summary. Our model achieves an improvement of more than 3% on DUC-2004(in ROUGE-1 and ROUGEL), and on the Chinese dataset, it has greatly improved (+ 2.56 ROUGE-1, + 1.33 ROUGE-2, + 1.53 ROUGE-L). The main reason is that the model we proposed combines a variety of features including deep learning features (word embeddings and beam search), and text display features (paragraph location features, sentence location features, etc.), which has more semantic representation capabilities.

E. AN EXTRACTIVE EXAMPLE

Figure 4 is a news document titled "楼市的平衡格局正在逐步形成" (The balanced pattern of the property market is gradually taking shape) which is collected from the Sina.com.cn (http://news.sina.com.cn/pl/2017-10-23/doc-ifymzqpq3400179.shtml).

Figure 4 is an extractive example, the extractive sentences are shown in Table 4. Our extractive model has achieved satisfactory results compared with the title.



楼市的平衡格局正在逐步形成

http://news.sina.com.cn/pl/2017-10-23/doc-ifymzqpq3400179.shtml

如果再加大房地产市场秩序的规范和整治，有效遏制各种炒房和住房投机，并努力规范地方政府行为，弱化“土地财政”。那么，楼市的平衡格局将真正形成。

国家统计局数据显示，前三季度商品房销售额91904亿元，增长14.6%，增速回落2.6个百分点。其中，住宅销售额增长11.4%，办公楼销售额增长25.2%，商业营业用房销售额增长31.8%。

与之相对应，商品房销售面积也呈现回落态势。1-9月商品房销售面积116006万平方米，同比增长10.3%，增速比1-8月份回落2.4个百分点。其中，住宅销售面积增长7.6%，办公楼销售面积增长32.7%，商业营业用房销售面积增长23.7%。

一方面，商品房销售增幅呈现回落态势，且随着调控的延续，楼市降温格局也不会出现大的改变，商品房销售增幅也可能会进一步回落；另一方面，商品房销售并没有出现增幅下降的现象，而仍然保持较高的增长幅度，不仅对三、四线城市去库存具有积极作用，对房地产市场稳定也能产生积极影响。也就是说，楼市调控并没有对房地产市场产生严重影响，但是，却遏制了房价的过度上涨，避免了房价过度上涨对居民生活带来进一步伤害，体现了“房子是用来住的、不是用来炒的”原则。

事实也是，楼市调控的目的并不是把房地产市场搞死，让房价出现断崖式下跌。作为经济重要组成部分的房地产市场，对经济增长的影响仍然是比较大的。相关数据显示，影响的比重达到了30%左右。这也意味着，对房地产市场，只能是温和调控、有序调控，而不能把房价作为唯一调控的依据。房价需要控制，但决不是断崖式下跌。否则，会直接影响经济的稳定，影响就业，影响发展。很显然，目前的状态，是比较理想的，也是符合调控的目标的。一、二线城市过于火爆的市场得到了稳定，过快上涨的房价也得到了控制，对经济的作用也得到了比较好的发挥。简单地讲，楼市调控的平衡点已经找到，楼市的平衡格局正在逐步形成。

更重要的，党的十九大报告对楼市的定位也做了进一步的明确，那就是“坚持房子是用来住的，不是用来炒的定位”，围绕这个定位，又提出了“加快建立多主体供给、多渠道保障、租购并举的住房制度，让全体人民住有所居”的要求。其中，购房是满足居住要求的一种方式，租房也是满足居住要求的一种方式。租购并举，才能从根本上解决居民的住房问题，才能让全体人民住有所居。

很显然，今年以来所出台的各项调控政策与措施，也都是围绕“居住”而展开的，除了加大保障房、租赁房建设力度，满足市场需求之外，就是通过各种调控手段，严厉遏制各种炒房，使炒房者无法再疯狂炒房、大肆炒房、无所顾忌地炒房。炒房之风遏制住了，市场的稳定性和健康度也就大大提升了。

而从前三季度全国房地产开发投资情况来看，前三季度，房地产投资80644亿元，同比名义增长8.1%，增速比1-8月份提高0.2个百分点。其中，住宅投资55109亿元，增长10.4%，增速提高0.3个百分点。住宅投资占房地产开发投资的比重为68.3%。也就是说，尽管楼市调控力度很大，各地出台的楼市调控政策多达150多次，但是，开发商并没有被完全“吓倒”，对房地产市场仍然具有一定信心。同时，也没有象过去那样疯狂投资、盲目投资，而是谨慎投资、有序投资。显然，这也是希望看到的结果。

正是因为房地产投资、销售都没有受到严重影响，开发商投资的谨慎度得到了进一步提高，市场销售也在逐步变得理性而客观。因此，市场的平衡性也越来越强，出现大起大落现象的概率大大缩小，风险正在逐步消化。如果能够继续保持这样的状态，房地产对经济发展和居民生活改善的作用也将得到进一步发挥，效率得到进一步提升。

十九大代表、住房和城乡建设部部长王蒙徽透露，正在做的一项工作是加快研究和制定“住房租赁管理条例”。那么，发展住房租赁市场、扩大住房租赁业务、提高租赁住房在居住用房中的比重，将成为下一步住房制度改革的重要内容之一。一旦住房租赁市场建立起来，让更多的居民通过租赁解决住房问题，那么，房地产市场的稳定性也将更强。如果再加大房地产市场秩序的规范和整治，有效遏制各种炒房和住房投机，并努力规范地方政府行为，弱化“土地财政”。那么，楼市的平衡格局将真正形成。

The balanced pattern of the property market is gradually taking shape

http://news.sina.com.cn/pl/2017-10-23/doc-ifymzqpq3400179.shtml

If we further increase the standardization and rectification of the real estate market order, we will effectively curb speculation in real estate and housing, and strive to standardize the behavior of local governments to weaken the ″land finance″. Then, the balance pattern of the property market will really form.

According to data from the National Bureau of Statistics, the sales of commercial housing in the first three quarters were 9,190.4 billion yuan, an increase of 14.6%, and the growth rate fell by 2.6 percentage points. Among them, residential sales increased by 11.4%, sales of office buildings increased by 25.2%, and sales of commercial premises increased by 31.8%.

Correspondingly, the sales area of commercial housing also showed a downward trend. From January to September, the sales area of commercial housing was 1,160.6 million square meters, an increase of 10.3% year-on-year, and the growth rate fell 2.4 percentage points from January to August. Among them, the sales area of residential buildings increased by 7.6%, the sales area of office buildings increased by 32.7%, and the sales area of commercial buildings increased by 23.7%.

On the one hand, the growth rate of commercial housing sales has shown a downward trend, and with the continuation of regulation, the cooling pattern of the property market will not change significantly, and the growth rate of commercial housing sales may also decline further; on the other hand, the growth of commercial housing sales has not declined While still maintaining a relatively high growth rate, it not only has a positive effect on destocking third- and fourth-tier cities, but also has a positive impact on the stability of the real estate market. In other words, the regulation of the property market did not have a serious impact on the real estate market, but it did curb the excessive rise in housing prices, and prevented the excessive rise in housing prices from further harming the lives of residents, reflecting that ″houses are used to live, not used Come to speculate″ principle.

The fact is that the purpose of the property market regulation is not to kill the real estate market and cause a steep fall in housing prices. As an important part of the economy, the real estate market still has a relatively large impact on economic growth. Relevant data show that the proportion of the impact has reached about 30%. This also means that the real estate market can only be moderately regulated and orderly regulated, and cannot use house prices as the sole basis for regulation and control. House prices need to be controlled, but they are not cliff-shaped. Otherwise, it will directly affect economic stability, employment and development. Obviously, the current state is relatively ideal and meets the

FIGURE 4. An extractive example.

TABLE 4. Chinese experimental results.

Sentences Number	Summary Content
Five results	1.楼市的平衡格局将真正形成(The balance pattern of the real estate market will really be formed). 2.楼市调控的平衡点已经找到, 楼市的平衡格局正在逐步形成(The balance point of the regulation and control has been found, and the balance pattern of the real estate market is gradually being formed). 3.坚持房子是用来住的, 不是用来炒的定位(Insist that the houses are used for living, not for speculation). 4.市场的稳定性和健康度也就大大提升了(Stability of the market and health will be greatly improved). 5.弱化“土地财政”, 那么, 楼市的平衡格局将真正形成(Weakening the "land finance", then the balance of the property market will really form).
Four results	1. 楼市调控的平衡点已经找到, 楼市的平衡格局正在逐步形成(The balance point of the regulation and control has been found, and the balance pattern of the real estate market is gradually being formed). 2. 坚持房子是用来住的, 不是用来炒的定位(Insist that the houses are used for living, not for speculation). 3. 市场的稳定性和健康度也就大大提升了(Stability of the market and health will be greatly improved). 4.谨慎投资,有序投资, 弱化“土地财政”, 楼市的平衡格局将真正形成(Prudent investment and orderly investment will weaken the "land finance", and the balance pattern of the real estate market will form).
Three results	1. 坚持房子是用来住的, 不是用来炒的定位(Insist that the houses are used for living, not for speculation). 2. 楼市调控的平衡点已经找到, 楼市的平衡格局正在逐步形成(The balance point of the regulation and control has been found, and the balance pattern of the real estate market is gradually being formed). 3.谨慎投资,有序投资, 弱化“土地财政”, 楼市的平衡格局将真正形成(Prudent investment and orderly investment will weaken the "land finance", and the balance pattern of the real estate market will really form).
Two results	1. 楼市调控的平衡点已经找到, 楼市的平衡格局正在逐步形成(The balance point of the regulation and control has been found, and the balance pattern of the real estate market is gradually being formed). 2.谨慎投资,有序投资, 弱化“土地财政”, 楼市的平衡格局将真正形成(Prudent investment and orderly investment will weaken the "land finance", and the balance pattern of the real estate market will really form).
One results	楼市调控的平衡点已经找到, 楼市的平衡格局正在逐步形成(The balance point of the regulation and control has been found, and the balance pattern of the real estate market is gradually being formed).
title	楼市的平衡格局正在逐步形成(The balanced pattern of the property market is gradually taking shape).

F. AN ABSTRACTIVE EXAMPLE

First, all sentences in the document are preprocessed, and then sorted by extractive model (only senven sentences are listed):

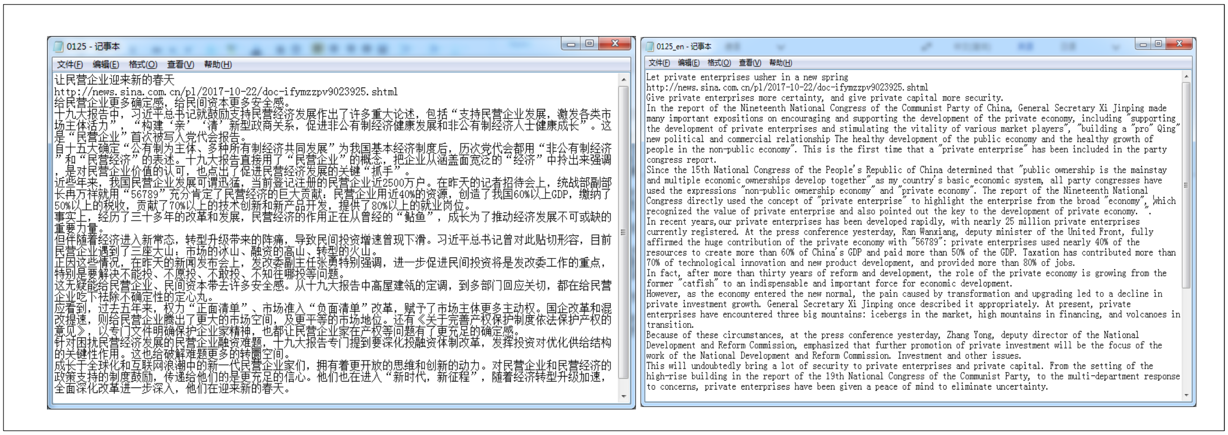
Sentence_1: 近些年来我国民营企业发展可谓迅猛 (In recent years, our private enterprises have been developed rapidly).

Sentence_2: 民营企业和民营经济在迎来新的春天 (Private enterprises and private economy are ushering in a new spring).

Sentence_3: 给民营企业更多确定感(Give private enterprises more certainty).

Sentence_4: 把企业从涵盖面宽泛的、“经济”中拎出来强调，是对民营企业价值的认可，也点出了促进民营经济发展的关键抓手” (Carrying out the enterprise from the broad “economy” and emphasizing it is the recognition for the private enterprises, and also points out the key “hands” to promote the development of private economy).

Sentence_5: 从十九大报告中高屋建瓴的定调 到多部门回应关切、都在给民营企业吃下祛除不确定性的定心丸 (From the fixed tone of the high-rise building in the report of the 19th National Congress to the multiple departments responding to concerns, they are giving private enterprises a peace of mind to eliminate uncertainty);


0125 - 记事本
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)
0125_en - 记事本
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

让民营企业迎来新的春天
http://news.sina.com.cn/pl/2017-10-22/doc-ifymzzpv9023925.shtml
给民营企业更多确定感，给民间资本更多安全感。
十九大报告中，习近平总书记就鼓励支持民营经济发展作出了许多重大论述，包括“支持民营企业发展，激发各类市场主体活力”，“构建‘亲’‘清’新型政商关系，促进非公有制经济健康发展和非公有制经济人士健康成长”。这是“民营企业”首次被写入党代会报告。
自十五大确定“公有制为主体、多种所有制经济共同发展”为我国基本经济制度后，历次党代会都用“非公有制经济”和“民营经济”的表述。十九大报告直接用了“民营企业”的概念，把企业从涵盖面宽泛的“经济”中拎出来强调，是对民营企业价值的认可，也点出了促进民营经济发展的关键“抓手”。
近些年来，我国民营企业发展可谓迅猛，当前登记注册的民营企业近2500万户。在昨天的记者招待会上，统战部副部长冉万祥就用“56789”充分肯定了民营经济的巨大贡献：民营企业用近40%的资源，创造了我国60%以上GDP，缴纳了50%以上的税收，贡献了70%以上的技术创新和新产品开发，提供了80%以上的就业岗位。
事实上，经历了三十多年的改革和发展，民营经济的作用正在从曾经的“鲇鱼”，成长为了推动经济发展不可或缺的重要力量。
但伴随着经济进入新常态，转型升级带来的阵痛，导致民间投资增速曾现下滑。习近平总书记曾对此贴切形容，目前民营企业遇到了三座大山：市场的冰山、融资的高山、转型的火山。
正因这些情况，在昨天的新闻发布会上，发改委副主任张勇特别强调，进一步促进民间投资将是发改委工作的重点，特别是要解决不能投、不愿投、不敢投、不知往哪投等问题。
这无疑能给民营企业、民间资本带去许多安全感。从十九大报告中高屋建瓴的定调，到多部门回应关切，都在给民营企业吃下祛除不确定性的定心丸。
应看到，过去五年来，权力“正面清单”、市场准入“负面清单”改革，赋予了市场主体更多主动权。国企改革和混改提速，则给民营企业腾出了更大的市场空间，及更平等的市场地位。还有《关于完善产权保护制度依法保护产权的意见》，以专门文件明确保护企业家精神，也都让民营企业家在产权等问题有了更充足的确定感。
针对困扰民营经济发展的民营企业融资难题，十九大报告专门提到要深化投融资体制改革，发挥投资对优化供给结构的关键性作用。这也给破解难题更多的转圜空间。
成长于全球化和互联网浪潮中的新一代民营企业家们，拥有着更开放的思维和创新的动力。对民营企业和民营经济的政策支持的制度鼓励，传递给他们的是更充足的信心。他们也在进入“新时代，新征程”，随着经济转型升级加速，全面深化改革进一步深入，他们在迎来新的春天。

Let private enterprises usher in a new spring
http://news.sina.com.cn/pl/2017-10-22/doc-ifymzzpv9023925.shtml
Give private enterprises more certainty, and give private capital more security.
In the report of the Nineteenth National Congress of the Communist Party of China, General Secretary Xi Jinping made many important expositions on encouraging and supporting the development of the private economy, including "supporting the development of private enterprises and stimulating the vitality of various market players", "building a "pro" Qing" new political and commercial relationship The healthy development of the public economy and the healthy growth of people in the non-public economy". This is the first time that a "private enterprise" has been included in the party congress report.
Since the 15th National Congress of the People's Republic of China determined that "public ownership is the mainstay and multiple economic ownerships develop together" as my country's basic economic system, all party congresses have used the expressions "non-public ownership economy" and "private economy". The report of the Nineteenth National Congress directly used the concept of "private enterprise" to highlight the enterprise from the broad "economy", which recognized the value of private enterprise and also pointed out the key to the development of private economy. ".
In recent years, our private enterprises has been developed rapidly, with nearly 25 million private enterprises currently registered. At the press conference yesterday, Ran Wanxiang, deputy minister of the United Front, fully affirmed the huge contribution of the private economy with "56789": private enterprises used nearly 40% of the resources to create more than 60% of China's GDP and paid more than 50% of the GDP. Taxation has contributed more than 70% of technological innovation and new product development, and provided more than 80% of jobs.
In fact, after more than thirty years of reform and development, the role of the private economy is growing from the former "catfish" to an indispensable and important force for economic development.
However, as the economy entered the new normal, the pain caused by transformation and upgrading led to a decline in private investment growth. General Secretary Xi Jinping once described it appropriately. At present, private enterprises have encountered three big mountains: icebergs in the market, high mountains in financing, and volcanoes in transition.
Because of these circumstances, at the press conference yesterday, Zhang Yong, deputy director of the National Development and Reform Commission, emphasized that further promotion of private investment will be the focus of the work of the National Development and Reform Commission. Investment and other issues.
This will undoubtedly bring a lot of security to private enterprises and private capital. From the setting of the high-rise building in the report of the 19th National Congress of the Communist Party, to the multi-department response to concerns, private enterprises have been given a peace of mind to eliminate uncertainty.

FIGURE 5. An abstractive example.

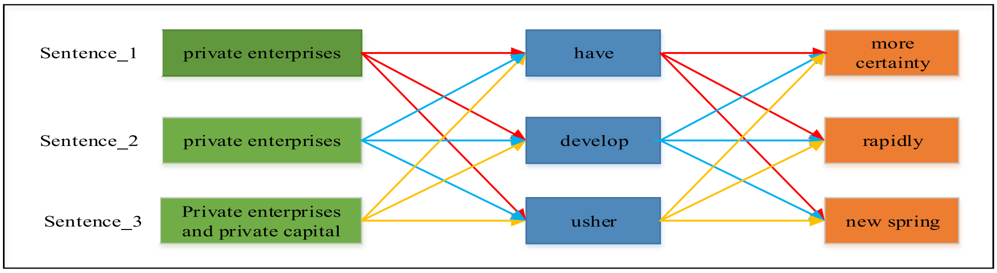


FIGURE 6. Beam search of syntactic blocks.

Sentence_6: 权力则给民营企业腾出了更大的市场空间，及更平等的市场地位 (Power gives private enterprises greater market space and a more equal market position);

Sentence_7: 发挥投资对优化供给结构的关键性作用 (Give play to the key role of investment in optimizing the supply structure). 这也给破解难题更多的转圜空间.

Suppose the number of sentences extracted by the extractive model is three. Then we perform a generative summary on the first three sentences. First, according to the content of the document, we have appropriately preprocessed the sentences, filled in the syntactic components and performed anaphora resolution. After the above processing, the three sentences are as following:

Sentence_1: 民营企业有更多确定感 (private enterprises have more certainty). Sentence_2: 民营企业发展迅猛 (Private enterprises have developed rapidly). Sentence_3: 民营企业和民营经济迎来新的春天 (Private enterprises and private economy are ushering in a new spring).

Beam Search algorithm is Performed as figure 6, sentences have multiple combinations. The top four sentence combinations are [*民营企业(private enterprises) 迎来(usher) 新的春天(new spring)*], [*民营企业(private enterprises) 发展(develop) 迅猛(rapidly)*], [*民营企业 发展 新的春天*] and [*民营企业和民营资本(Private enterprises* and *private capital) 发展(develop) 迅猛(rapidly)*]. *民营企业(private enterprises) 迎来(usher) 新的春天(new spring)]* is slightly higher than [*民营企业(private enterprises)发展(develop) 迅猛(rapidly)*]. After in-depth analysis, the main reason is the appearance of the sentence [*民营企业(private enterprises) 迎来(usher) 新的春天(new spring)*] is in the last sentence of the last paragraph. While the sentence [*民营企业(private enterprises) 发展(develop) 迅猛(rapidly)]* appears in the middle of the document.

The abstractive summarization [*民营企业(private enterprises) 迎来(usher) 新的春天(new spring)*] has the same semantics as the real title of this document [*让民营企业迎来新的春天(Let private enterprises usher in a new spring)*]. This example further proves that fusion of more information (such as syntactic components, semantic location information, etc.) can achieve better results. In addition, Beam Search based on syntactic structure has ensured language fluency, semantic consistency, and readability of summarization.

V. CONCLUSION

At present, there are two main methods of summarization. One is extractive and the other is abstractive. In order to solve some problems in document-level summarization, this paper combines the advantages of the two methods, and proposes a two-phase, hybrid extractive and abstractive, abstract generation method. This method does not require document-summary pairs that are indispensable in neural network models. Therefore our model is an unsupervised approach. By virtue of standardization and reorganization for syntactic components, we have solved the poor readability of the generated summarization to a certain extent. Resulting from the complexity, diversity, and cross-document semantic differentiation of multi-documents, this manuscript does not involve multi-document summaries. Our plan of next steps is about multi-document and cross-document text summarization.

REFERENCES

[1] J. N. Madhuri and R. G. Kumar, "Extractive text summarization using sentence ranking," in *Proc. Int. Conf. Data Sci. Commun.*, Mar. 2019, pp. 1–3.

[2] E. Reategui, M. Klemann, and M. D. Finco, "Using a text mining tool to support text summarization," in *Proc. IEEE 12th Int. Conf. Adv. Learn. Technol.*, Jul. 2012, pp. 607–609.

[3] S. R. Rahimi, A. T. Mozhdehi, and M. Abdolahi, "An overview on extractive text summarization," in *Proc. IEEE 4th Int. Conf. Knowl.-Based Eng. Innov. (KBEI)*, Dec. 2017, pp. 54–62.

[4] A. M. Rush, S. Chopra, and J. Weston, "A neural attention model for abstractive sentence summarization," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2015, pp. 379–389.

[5] H. Kim and S. Lee, "A context based coverage model for abstractive document summarization," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2019, pp. 1129–1132.

[6] A. See, P. J. Liu, and C. D. Manning, "Get to the point: Summarization with pointer-generator networks," in *Proc. 55th Annu. Meeting Assoc. Comput. Linguistics*, 2017, pp. 1073–1083.

[7] R. Nallapati, B. Zhou, C. Nogueira dos santos, C. Gulcehre, and B. Xiang, "Abstractive text summarization using sequence-to-sequence RNNs and beyond," 2016, *arXiv:1602.06023*. [Online]. Available: http://arxiv.org/abs/1602.06023

[8] J. Conroy, S. Davis, J. Kubina, Y. Liu, D. P. O'leary, and J. Schlesinger, "Multilingual summarization: Dimensionality reduction and a step towards optimal term coverage," in *Proc. Conf. Multi-Document Summar.*, 2013, pp. 55–63.

[9] T. He, L. Hu, K. C. C. Chan, and P. Hu, "Learning latent factors for community identification and summarization," *IEEE Access*, vol. 6, pp. 30137–30148, 2018.

[10] R. Mihalcea and P. Tarau, "Textrank: Bringing order into text," in *Proc. Empirical Methods Natural Lang. Process.*, 2004, pp. 404–411.

[11] A. Alzuhair and M. Al-Dhelaan, "An approach for combining multiple weighting schemes and ranking methods in graph-based multi-document summarization," *IEEE Access*, vol. 7, pp. 120375–120386, 2019.

[12] J. Gu, Z. Lu, H. Li, and V. Li, "Incorporating copying mechanism in sequence-to-sequence learning," in *Proc. 54th Associat. Comput. Ling.*, 2015, pp. 1–10.

[13] O. Vinyals, M. Frotunato, and N. Jaitly, "Pointer networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 2692–2700.

[14] Z. Cao, W. Li, S. Li, and F. Wei, "Retrieve, rerank and rewrite: Soft template based neural summarization," in *Proc. 56th Annu. Meeting Assoc. Comput. Linguistics*, vol. 1, 2018, pp. 152–161.

[15] M. Yang, Q. Qu, W. Tu, Y. Shen, Z. Zhao, and X. Chen, "Exploring humanlike reading strategy for abstractive text summarization," in *Proc. 33rd AAAI Conf. Artif. Intell.*, vol. 33, Aug. 2019, pp. 7362–7369.

[16] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 5998–6008.

[17] P. Li, W. Lam, L. Bing, and Z. Wang, "Deep recurrent generative decoder for abstractive text summarization," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2017, pp. 2091–2100.

[18] C. Hark, T. Uckan, E. Seyyarer, and A. Karci, "Graph-based suggestion for text summarization," in *Proc. Int. Conf. Artif. Intell. Data Process. (IDAP)*, Sep. 2018, pp. 1–6.

[19] J. Tan, X. Wan, and J. Xiao, "Abstractive document summarization with a graph-based attentional neural model," in *Proc. 55th Annu. Meeting Assoc. Comput. Linguistics*, vol. 1, Jul. 2017, pp. 1171–1181.

[20] P. Yang, W. Li, and G. Zhao, "Language model-driven topic clustering and summarization for news articles," *IEEE Access*, vol. 7, pp. 185506–185519, 2019.

[21] J. Clarke and M. Lapata, "Discourse constraints for document compression," *Comput. Linguistics*, vol. 36, no. 3, pp. 411–441, Sep. 2010.

[22] K. Filippova, E. Alfonseca, C. A. Colmenares, L. Kaiser, and O. Vinyals, "Sentence compression by deletion with LSTMs," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2015, pp. 360–368.

[23] S. Gehrmann, Y. Deng, and A. Rush, "Bottom-up abstractive summarization," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2018, pp. 4098–4109.

[24] Z. Li, Z. Peng, S. Tang, C. Zhang, and H. Ma, "Text summarization method based on double attention pointer network," *IEEE Access*, vol. 8, pp. 11279–11288, 2020.

[25] Y. Liu, "Fine-tune BERT for extractive summarization," 2019, *arXiv:1903.10318*. [Online]. Available: http://arxiv.org/abs/1903.10318

[26] R. Nallapati, F. Zhai, and B. Zhou, "SummaRuNNer: A recurrent neural network based sequence model for extractive summarization of documents," in *Proc. 31st AAAI*, 2017, pp. 3075–3081.

[27] Q. Chen, X. D. Zhu, Z. H. Ling, S. Wei, and H. Jiang, "Distraction-based neural networks for modeling document," in *Proc. Int. Conf. Artif. Intell.*, 2016, pp. 2754–2760.

[28] M. Ranzato, S. Chopra, M. Auli, and W. Zaremba, "Sequence level training with recurrent neural networks," in *Proc. 4th Int. Conf. Learn. Represent.*, 2016, pp. 1–16.

[29] D. Bahdanau, P. Brakel, K. Xu, A. Goyal, R. Lowe, J. Pineau, A. Courville, and Y. Bengio, "An actor-critic algorithm for sequence prediction," 2016, *arXiv:1607.07086*. [Online]. Available: http://arxiv.org/abs/1607.07086

[30] S. Henß, M. Mieskes, and I. Gurevych, "A reinforcement learning approach for adaptive single-and multi-document summarization," in *Proc. Conf. Comput. Ling. Lang. Tech.*, 2015, pp. 3–12.

[31] R. Paulus, C. Xiong, and R. Socher, "A deep reinforced model for abstractive summarization," in *Proc. 6th Int. Conf. Learn. Represent.*, 2018, pp. 1–12.

[32] Q. Zhou, N. Yang, F. Wei, and M. Zhou, "Selective encoding for abstractive sentence summarization," 2017, *arXiv:1704.07073*. [Online]. Available: http://arxiv.org/abs/1704.07073

[33] C. Tan, F. Wei, N. Yang, B. Du, W. Lv, and M. Zhou, "S-net: From answer extraction to answer generation for machine reading comprehension," 2017, *arXiv:1706.04815*. [Online]. Available: http://arxiv.org/abs/1706.04815

[34] S. Wang, X. Zhao, B. Li, B. Ge, and D. Q. Tang, "TP-AS: A two-phase approach to long text automatic summarization," *J. Chin. Inf. Process.*, vol. 32, no. 6, pp. 71–79, 2018.

[35] R. Blanco and C. Lioma, "Graph-based term weighting for information retrieval," *Inf. Retr.*, vol. 15, no. 1, pp. 54–92, 2012.

[36] C. Fang, D. Mu, Z. Deng, and Z. Wu, "Word-sentence co-ranking for automatic extractive text summarization," *Expert Syst. Appl.*, vol. 72, pp. 189–195, Apr. 2017.

[37] S. S. Yu, J. D. Su, and P. F. Li, "Improved textrank-based method for automatic summarization," *Comput. Sci.*, vol. 43, no. 6, pp. 240–247, 2016.

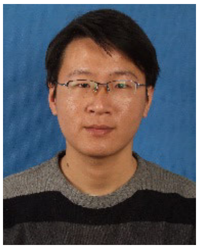
[38] B. Araly and R. Verma, "Combining syntax and semantics for automatic extractive single-document summarization," in *Proc. Int. Conf. Intell. Text Process. Comput. Linguistics*, 2012, pp. 366–377.

[39] N. N. Li, P. Y. Liu, W. F. Liu, and W. T. Liu, "Automatic digest optimization algorithm based on textrank," *Appl. Res. Comput.*, vol. 36, no. 4, pp. 1045–1050, 2019.

[40] X. Chen, J. Li, and H. Wang, "Keyphrase enhanced diverse beam search: A content-introducing approach to neural text generation," *IEEE Access*, vol. 7, pp. 72716–72725, 2019.

WENFENG LIU received the B.S. degree in computer science and technology from Ludong University, Yantai, China, in 2005, the M.S. degree in computer technique from the University of Science and Technology Liaoning, Anshan, China, in 2011, and the Ph.D. degree from the College of Computer Science and Engineering, Shandong Normal University, China, in 2020. He is currently an Associate Professor with the School of Computer Science, Heze University, China. His research interests include natural language processing, machine learning, and deep learning.

YALING GAO is currently a Network Engineer with Heze University, China. Her research interests include network information security, information retrieval, natural language processing, and artificial intelligence. She obtained the most Senior Network Engineer Certificate of the National Computer Rank Examination.



JINMING LI received the Ph.D. degree from Chongqing University, China, in 2015. He is currently an Associate Professor with Heze University, China. His current research interests include information acquiring, image processing, machine learning, and deep learning.

YUZHEN YANG received the Ph.D. degree from Shandong Normal University, in 2014. She is currently an Associate Professor with the School of Computer Science, Heze University, China. Her research interests include natural language processing, machine learning, and deep learning.