IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

Summarization of Scientific Paper through Reinforcement Ranking on Semantic Link Network

Xiaoping Sun and **Hai Zhuge***, **Senior Member, IEEE**

Laboratory of Cyber-Physical-Social Intelligence, Guangzhou University, China
Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, University of Chinese Academy of Sciences, Chinese Academy of Sciences, China
System Analytics Research Institute, Aston University, UK

*Corresponding author: Hai Zhuge (e-mail: h.zhuge@aston.ac.uk).

This work was supported by the National Science Foundation of China (No. 61075074 and No. 61640212), National Science & Technology Development Plan (No. 2016YFB1000505), and Research Funding of Guangzhou University.

ABSTRACT The Semantic Link Network is a semantics modeling method for effective information services. This paper proposes a new text summarization approach that extracts Semantic Link Network from scientific paper consisting of language units of different granularities as nodes and semantic links between the nodes, and then ranks the nodes to select Top-k sentences to compose summary. A set of assumptions for reinforcing representative nodes is set to reflect the core of paper. Then, Semantic Link Networks with different types of node and links are constructed with different combinations of the assumptions. Finally, an iterative ranking algorithm is designed for calculating the weight vectors of the nodes in a converged iteration process. The iteration approximately approaches a stable weight vector of sentence nodes, which is ranked to select Top-k high-rank nodes for composing summary. We designed six types of ranking models on Semantic Link Networks for evaluation. Both objective assessment and intuitive assessment show that ranking Semantic Link Network of language units can significantly help identify the representative sentences. This work not only provides a new approach to summarizing text based on extraction of semantic links from text but also verifies the effectiveness of adopting the Semantic Link Network in rendering the core of text. The proposed approach can be applied to implementing other summarization applications such as generating an extended abstract, the mind map and the bulletin points for making the slides of a given paper. It can be easily extended by incorporating more semantic links to improve text summarization and other information services.

INDEX TERMS Semantics Modeling, Natural Language Processing, Text Summarization, Reinforcement.

I. INTRODUCTION

Summarization is one of the challenges of natural language processing and understanding [1].

A. MOTIVATION

Automatic text summarization has been studied for over a half century, but how to generate a summary based on semantics is still a challenge [2]. Finding a proper semantic representation of text is the key to making a summary that can represent the core of text.

The motivation of this paper concerns two parts: (1) Propose an effective summarization approach based the Semantic Link Network [3]. We will investigate the construction of Semantic Link Network consisting of various language units (word, sentence, paragraph and section) extracted from a scientific paper as a semantic representation of paper structure for ranking top-k sentences as a summary. By investigating summarization based on various types of ranking models on Semantic Link Network, we can identify which kinds of Semantic Link Network is more appropriate for representing and understanding text. (2) Verify the effectiveness of summarization based on the semantic links extracted from text. It provides an evidence for adopting Semantic Link Network as an effective semantics modelling method for information services.

B. METHOD AND RESULT

This paper focuses on constructing *is-part-of* link, *similar-to* link and *co-occurrence* link between nodes of various language units (words, sentences, paragraphs and sections) extracted from a scientific paper to build a Semantic Link Network for modelling the basic semantic structure of the paper and then ranking the sentences. Sub-paragraphs and section titles of a paper are also extracted as larger granularities of nodes of the Semantic Link Network.

Summary is composed by selecting Top-k highly ranked sentences from the source paper. To rank the sentences, the weights of sentences are calculated with an iterative graph-ranking algorithm on a Semantic Link Network consisting of various language units of the paper as nodes. A weight vector of one type of node is defined by the weights of all nodes of that type.

To build an instance of Semantic Link Network, we first set a set of reinforcement assumptions on the relations among different types of nodes of paper, and then a Semantic Link Network is constructed according to a combination of those assumptions. To rank nodes on a Semantic Link Network, we formalize a set of iterative functions of the weight vectors of nodes, and finally an iteration algorithm is applied to those iterative functions to approximate stationary distributions of node weight vectors that are finally used to rank sentences.

Combing different assumptions, we design six different instance models consisting of different instances of Semantic Link Network with different iterative functions of nodes' weights for testing the roles of different semantic links. Three models using the TFIDF and sentence similarity information are compared.

A set of 175 papers was collected from the proceedings of ACL 2014 for experiment. The combination of abstract and conclusion is the natural summarization of paper so it is sound to use them as the gold standard for assessing the generated summary. ROUGE scores are used for the objective comparison metrics [4]. Intuitive observation on ten papers is conducted. We also test those models on a classical benchmark documents in DUC 2002 (http://www-nlpir.nist.gov/projects/duc/data.html) to show how the models perform on short news document.

Finally, we apply the models on a paper that is much longer than any papers in the ACL proceedings and evaluate the extracted sentences by comparing the sentences with a word tree composed manually according to a MindMap approach [5].

The experimental results demonstrate the effectiveness of adopting the Semantic Link Network of various language units in representation and understanding. When using the *is-part-of* link to incorporate section, paragraph, and section titles into ranking sentences, the models achieved better scores than those models used less types of structure units and links on the papers in the ACL 2014 proceedings with stop words being removed. Specifically, the Hybrid model and the model using section title (Sectitle in short) achieve Top-2 best performance on the ACL 2014 papers without stop words. The Hybrid model uses sections, paragraphs, sentences, and words as nodes along with the sentence similarity network for ranking sentences. The Sectitle model includes sections, paragraphs, sentences, words, and section titles as nodes. The model on the paragraph and sentence nodes (Para model in short) achieves the best score on the news report texts of DUC 2002 without stop words. On the long paper, both the keyword tree evaluation and the ROUGE score evaluation show that the model with sub-paragraphs and section nodes achieves much better performance than others, demonstrating that using richer semantic links to represent document structure can improve the quality of summarization, especially for long papers.

II. MODELLING THE SEMANTICS OF PAPER

A. SEMANTIC LINK NETWORK OF PAPER

A scientific paper can be represented as a Semantic Link Network of semantic nodes (e.g., sections, sub-sections, paragraphs, sentences and words) and semantic links (e.g., *is-part-of*, *similar-to* and *co-occurrence*) between nodes. An example is shown in Figure 1. All those components can be easily parsed out from the paper in HTML.

A semantic link *a*—*r*—*b* represents that a semantic node *a* links to a semantic node *b* with *relation r*. As shown in

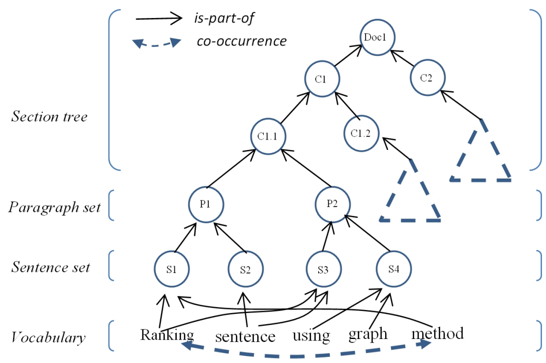


FIGURE. 1. **A Semantic Link Network of a paper.**

Figure 1, p_1—*is-part-of*—c_1 represents that p_1 (e.g., a paragraph) is a part of c_1 (e.g., a section), and w_1—*co-occurrence*—w_2 represents that word w_1 and word w_2 co-occur within one sentence. A semantic node can be anything including a concept, a set, and a physical object. Various semantic links, category hierarchies, and the reasoning rules on semantic links constitute the basic model of the Semantic Link Network [6], which is more capable of modeling semantics than traditional graph structure.

1) SEMANTIC NODES

Semantic nodes are identified by strings and interpreted by a semantic space [2]. It is natural to assume that each such string unit has a unique ID. Such an ID can be assigned by the author (for example, a section index number within a paper) or by an indexing algorithm (such as a sequential indexing schema).

The following vectors are used to represent the weight of different types of semantic nodes within a paper:

w: the word vector that contains the vocabulary of the whole paper.

s: the sentence vector that contains all sentences appeared in sequence. That is, two sentences will have different IDs even when they have the same character sequence. So the position of sentences within the paper determines the sentence IDs.

p: the paragraph vector that contains the paragraphs that appear in the sequence.

c: the section vector that contains the sections appeared in sequence. A section can contain sub-sections.

t: the sub-paragraph vector that represents the subparagraphs that contain similar sentences in the sequence of appearance.

h: the section title vector containing the string of section titles appeared in the paper.

We use these vector symbols to represent the ranking weight vector of the corresponding structural units within a paper when the context is clear. So text summarization can be implemented by sorting the structural units according to their weight vectors and composing top-k language units.

2) SEMANTIC LINKS

Here focuses on the following three types of semantic links:

is-part-of: If a string is a sub-string of another string, an *is-part-of* semantic link exists between the two semantic nodes. Although a sentence can be a sub-string of a sentence in different sequence positions within a paper, links will not be set between them if one is not a sub-string of another according to their sequence positions within a paper.

similar-to: the similarity between two semantic nodes o_1 and o_2 can be measured by a distance metric function $dist(o_1, o_2)$, e.g., the Jaccard distance between two sentence strings.

co-occurrence: it links two words appear in the same sentence within a paper. A *co-occurrence* link between two words w_1 and w_2 can be derived from the *is-part-of* links with a simple semantic link reasoning rule: w_1—*is-part-of*—s_1 and w_2—*is-part-of*—s_1 implies that w_1—*co-occurrence*—w_2. A set of rules for reasoning on semantic links was introduced in [6].

B. ITERATIVE RANKING MODEL

1) REEINFORCEMENT ASSUMPTION

It is hard to determine an exact weight that reflects the absolute importance of a language unit within a paper. The self-organization nature of Semantic Link Network enables semantic nodes to reinforce each other in rendering meaning, which can be used to define iterative functions [2]. The relative weights of nodes can be differentiated through a reinforcement process. The classical algorithms for ranking network concern PageRank [7] and HITS [8]. The PageRank algorithm assumes that a node has more important in-links gains higher ranks while the HITS algorithm assumes that a node has many important out links is also important.

A sentence is naturally composed of words, and authors often use those highly representative words to compose an important sentence. For example, the weight of the title sentence "*Dimensionality of summarization*" is determined by the weights of words "*dimensionality*" and "*summarization*", which are extensively discussed in [1]. Therefore, we have the following assumption.

Assumption 1. *The weight of a sentence is determined by the weight of its words.*

An important paragraph often has many important sentences within a paper. For example, the abstract of a paper is often deemed as an important and representative paragraph of the paper because we believe that authors will deliver important ideas using representative sentences in abstract. Therefore we have the following assumption.

Assumption 2. *The weight of a paragraph is determined by the weight of its sentences.*

Similarly, if a section has many important paragraphs, we deem that section as an important one.

Assumption 3. *The weight of a section is determined by the weight of its paragraphs.*

On the other hand, it is natural that a word located in an important sentence is deemed important. An example is that

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2856530, IEEE Access

people usually pay more attention to the words appeared in the title of paper, the title of section and the title of subsections when reading a paper. Therefore, we have the following assumption to make a connected reinforcement relationships among words, sentences and paragraphs.

Assumption 4. *The weight of a word is determined by the weight of the sentence that contains the word, by the weight of the paragraphs that contain the word, and also by the sections that contain the word*.

Authors often use related sentences to render an idea (e.g., with emphasis, different aspects or further explanation), which produces similar sentences. Therefore we have the following assumption.

Assumption 5. *A sentence that has more similar sentences has a higher rank or is more important*.

A long paragraph may contain several topics. Analogy to assumption 2, we have the following assumption.

Assumption 6. *The weight of a paragraph is determined by its sub-paragraph, and the weight of a sub-paragraph is determined by its sentences*.

Similar to the importance of the title of paper, we have the following assumption.

Assumption 7. *A word in the title of a section is important*.

Each assumption corresponds to one iteration relationship of the weight vectors of two types of nodes, which forms a set of links of the instance of Semantic Link Network. Different instances with corresponding iterative algorithms can be built on different combinations of the above reinforcement assumptions for calculating the weight vectors of nodes. A basic rule is to form a closed loop of reinforcement relationships, corresponding to a strongly connected Semantic Link Network instance. Different combinations of the assumptions can be formed to build different networks, which in turn generate different ranking results for different applications.

2) BASIC ITERATIVE MODEL

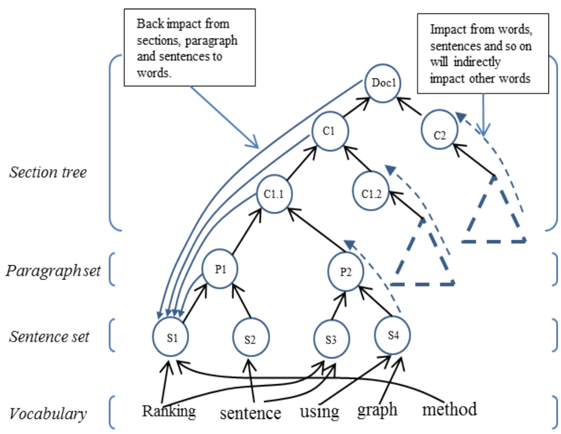


FIGURE 2. Reinforced loops among words and their sentences, paragraphs and sections.

We combine the first four assumptions to build an iterative function as a basic iterative model, based on which variants of the model can be made.

Adjacent matrices $\boldsymbol{W_S}$, $\boldsymbol{S_P}$, and $\boldsymbol{P_C}$ are used to describe the word-sentence, sentence-paragraph and paragraph-section links respectively. Matrix multiplication by the adjacent matrix is used to calculate the weight summation relation between weight vectors of different nodes.

According to Assumption 1, formula (1) calculates the sentence weight vector $\boldsymbol{s}$ by the summation of the weights of the words it contains. Similarly, formula (2) and formula (3) are based on Assumption 2 and Assumption 3 respectively. According to Assumption 4, the word weight is calculated by the summation of the weight of sentences, paragraphs and sections it is located in, which is determined by the matrix multiplication of corresponding adjacent matrices $\boldsymbol{W_S}$, $\boldsymbol{S_P}$, and $\boldsymbol{P_C}$ in formula (4). The weight vectors on the left hand side of formula are normalized after summations.

$$s^{(n)} = W_S \times w^{(n)} \tag{1}$$

$$p^{(n)} = S_P^T \times s^{(n)} \tag{2}$$

$$c^{(n)} = P_C^T \times p^{(n)} \tag{3}$$

$$\begin{aligned} w^{(n+1)} = W_S \times s^{(n)} + W_S \times S_P \times p^{(n)} \\ + W_S \times S_P \times P_C \times c^{(n)} \end{aligned} \tag{4}$$

We finally obtain an iterative function defined on the word weight vector when merging all previous four equations:

$$\begin{aligned} w^{(n+1)} = (W_S \times W_S^T + W_S \times S_P \times S_P^T \times W_S^T \\ + W_S \times S_P \times P_C \times P_C^T \times S_P^T \times W_S^T) \times w^{(n)} \end{aligned} \tag{5}$$

Let $A = W_S \times W_S^T + W_S \times S_P \times S_P^T \times W_S^T + W_S \times S_P \times P_C \times P_C^T \times S_P^T \times W_S^T$.

We have:

$$w^{(n+1)} = A w^{(n)}. \tag{6}$$

Assumption 4 makes the whole iterative function closed because one keyword in general appears in different sentences within a paper and a sentence often contains multiple keywords. The first three assumptions are based on the facts that the structure on sentences within a graph of a paper is a tree, which means that the weight of a section is directly determined by the summation of the weights of its paragraph and the weights of a paragraph is just the summation of its sentences. However, in Assumption 4, the weight of a word is determined not only by its sentences but also by its paragraph and section, which finally makes those nodes to influence others through an indirect path as shown in Figure 2.

ALGORITHM A: *Iteration*(W_S, S_P, P_C)

RETURNS: The weight vectors w, s, p, c..

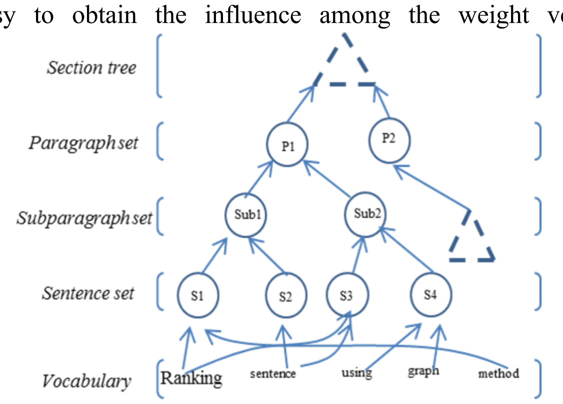
n_w=|w|
w=rand(1,n_w);//Randomly initialize w using real number in **[0, 1]**
n=1
While Converged($w^{(n)}$, $w^{(n+1)}$) != 0
  $w^{(n)}$=w.
  $s^{(n)} = W_S \times w^{(n)}$;
  $p^{(n)} = S_P^T \times s^{(n)}$;
  $c^{(n)} = P_C^T \times p^{(n)}$
  $w^{(n+1)} = W_S \times s^{(n)} + W_S \times S_P \times p^{(n)} + W_S \times S_P \times P_C \times c^{(n)}$
  w= *Normalize*($w^{(n+1)}$).
  n=n+1
return w,s,p,c

FIGURE 3. **The iteration algorithm consisting of words, sentences, paragraphs and sections.**

For example, word w_1 appears in sentence s_j, which contains word w_2, then w_1 is also indirectly determined by w_2 because the weight of sentence s_j is determined by word w_2 (according to Assumption 1) and w_2 is partially determined by sentence s_j (according to Assumption 4). So, w_1 is partially influenced by w_2 through s_j. Similarly, if sentence s_k also contains w_2, which means that the weight of w_2 is determined by s_k and this weight will be transferred to s_j, then the weight of sentence s_j is also indirectly determined by sentence s_k. If s_j is in the paragraph p_1 and s_k is in paragraph p_2, then p_1 is also indirectly determined by p_2 (according to Assumption 1, 2, and 4), the dashed arrows in Figure 2.

Property 1. *Let* G *be the adjacent matrix consisting of the nodes of all types and their is-part-of links. The weight of a node v in* G *is directly and indirectly determined by those nodes that have a directed path to v in the iterative process through formula (1) to (5).*

Proof. Just simply replacing the right hand side variables with its definition formula and applying the chain rule, it is easy to obtain the influence among the weight vector variables.



NOTES: Subparagraph is a partition of a paragraph.

FIGURE 4. **The structure of a paper with sub-paragraphs.**

Figure 3 shows the iterative algorithm designed based on formula (1)-(4). The algorithm in Figure 3 can be reduced to the formula (6). When matrix W_S, S_P, and P_C are normalized, the iterative function in (6) is a Markov process approaching its stationary distribution w of A. Other weight vectors s, p and c can be consequently computed using formulas (1), (2), and (3).

Property 2. *When the adjacent matrix* G *formed by* W_S, S_P, *and* P_C *is strongly connected, the iterative process will be converged to its stationary distribution* w, *i.e., the word weight vector is achieved.*

Proof. This can be simply derived from the *Perron–Frobenius* theorem's application in the Markov process [9].

After obtaining the sentence weight vector, text summarization can be implemented by ranking the sentence weight vector in descent order and extracting *Top-k* sentences to compose the summary.

When treating the matrix A in formula (6) as the adjacent matrix of a graph, the iterative process is the same as the PageRank algorithm. The edges in A can be interpreted as: If there is a path from a word to another word through the *is-part-of* relation in G, then there is an edge in A. Thus, similar to the PageRank algorithm, if A is strongly connected, the algorithm will converge. Moreover, dangling nodes (the nodes have no output links) can be technically connected to other nodes to make a connected graph. The convergence is determined by the vector distance that can be easily computed using the L1 norm [9].

C. VARIANTS OF SEMANTIC MODEL

When combining different assumptions, we can construct different Semantic Link Networks and corresponding iterative processes. For example, we can omit the impact of section on the important word with the assumption that the weight of a word is determined by its sentences and paragraphs. We can also incorporate sentence similarity into the network. The following are different models we designed.

1). PARAGRAPH MODEL

We change formula (4) by removing the section node part:

$$w^{(n+1)} = W_S \times s^{(n)} + W_S \times S_P \times p^{(n)} \tag{7}$$

Then, the section weight vector computing can be omitted and the final iterative function of word vector will be:

$$w^{(n+1)} = (W_S \times W_S^T + W_S \times S_P \times S_P^T \times W_S^T) \times w^{(n)} \tag{8}$$

2). SUB-PARAGRAPH MODEL

We can get a different Semantic Link Network by adding more elements according to Assumption 6. We use an incremental similarity computing algorithm to group sentences within a paragraph into another layer of structure, named sub-paragraph. Then, the parent nodes of sentences are sub-paragraphs and sub-paragraphs are child nodes of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2856530, IEEE Access

ALGORITHM B: *BuildSubpara (P)*

INPUT: Sentences set S = [s_1, s_2, ..., s_n] in one paragraph P.

RETURNS: Subparagraph set B={b_1, b_2, ..., b_k}

```
initialize b_i = {} for i = 1.
B=[ b_i]
for k = 1 to n:
    // compute distance from current sentence to the current context.
    d = ComputeDist(b_i, s_k);
    if d > threshold,
        B.append(b_i)
        b_i = {s_k}
        i=i+1
    else:
        b_i.append(s_k)
    endif;
return B
```

FIGURE 5. Building sub-paragraphs from a paragraph.

paragraphs (as shown in Figure 4). Figure 5 shows such a procedure to produce sub-paragraphs within a paragraph.

Sentences within a paragraph are added to a sub-paragraph one by one. That is, when the next sentence is quite similar to the sentence set within the current sub-paragraph, the sentence is added to the current context. Otherwise, a new sub-paragraph is created by adding the current sentence to it. A Jaccard distance function is used to compute the average distance from the current sentence to the following sentences within a sub-paragraph. A network is built similar to the previous model by using *is-part-of* link except adding a layer between the sentence layer and the paragraph layer. Consequently, a sentence-subparagraph matrix $\boldsymbol{S_T}$ and subparagraph-paragraph matrix $\boldsymbol{T_P}$ are added to the iterative model.

Then, formula (2) will be replaced by:

$$t^{(n)} = S_T^T \times s^{(n)} \text{ and } p^{(n)} = T_P^T \times t^{(n)} \tag{9}$$

The weight vector of word in formula (4) becomes the following formula:

$$\begin{aligned} w^{(n+1)} &= W_S \times s^{(n)} + W_S \times S_T \times t^{(n)} \\ &+ W_S \times S_T \times T_P \times p^{(n)} + W_S \times S_T \times T_P \times P_C \times c^{(n)} \end{aligned} \tag{10}$$

ALGORITHM C: *ComputeDist (b, s)*

INPUT: Subparagraph *b* containing a set of sentence strings; *s* is a sentence string that is to be computed with *b*.

RETURNS: distance from *s* to context *b*.

```
d=[]
for each sentence t in b:
    d[i] = Jaccard(t, s);
    i = i + 1
return avg(d)
```

FIGURE 6. Computing distance from a sentence to a sub-paragraph.

3). SENTENCE SIMILARITY MODEL

Sentence similarity is used to build the graph for ranking in TextRank model [10]. Specifically, a Jaccard function $d(s_1, s_2)$ is used to measure the similarity between two sentences s_1 and s_2 within the full-text of a paper. Matrix $\boldsymbol{S_s}$ is built to represent the similarity distance between any two sentences within a paper. Then, sentence weight is determined equally by iteration function in formula (11).

$$s^{(n)} = 0.5 \times s^{(n-1)} \times S_s + 0.5 \times [\frac{1}{n}] \tag{11}$$

$$s^{(n)} = s^{(n-1)} S_s \tag{12}$$

Formula (11) ranks the sentence using the similarity matrix $\boldsymbol{S_s}$. It adopts the PageRank model that uses a dangling weight factor to ensure the connectivity of the similarity matrix, which is in accordance with Assumption 5. It can be simply treated as matrix iteration in formula (12) if matrix $\boldsymbol{S_s}$ is connected. In this case, the sentence weight is mainly determined by the similarity between this sentence and the other sentences.

4). WORD GRAPH MODEL

If we can compute the word weight, we can use Assumption 1 to directly obtain the sentence weight. We build a word graph by using the *co-occurrence* semantic link among words. Then, a word-word graph $\boldsymbol{W_w}$ is constructed, the PageRank algorithm is applied to the graph to calculate the weight vector of words and the sentence weight is calculated as following:

$$w^{(n+1)} = W_w \times w^{(n)} \tag{13}$$

$$s = W_s \times w^{(n)} \tag{14}$$

Formula (14) calculates the sentence weight by summing the weights of its words. In another word graph model, we update the sentence weight by the following formula:

$$s = \text{Mean}(\{w_i \mid w_i \text{ is in s}\} \times \text{Max}(\{w_i \mid w_i \text{ is in s}\}) \tag{15}$$

Formula (14) implies that a longer sentence could have a higher weight than a shorter sentence. Formula (15) avoids this bias by choosing the average weight of words (*Mean*() function in the formula) as the weight of a sentence. The *Max*() function is to select the maximum weight from the words it contains. So the final function is to synthesize the average weight and the maximum the weight of words together as the sentence weight.

5). HYBRID MODEL COMBINING IS-PART-OF LINK AND SIMILAR-TO LINK

We combine the structural model with the sentence similarity model to build a hybrid graph for calculating sentence ranks. That is, we consider both the tree structure of paper and the sentence similarity according to Assumption 5. We modify the formula (11) by adding word weights to the sentence weight as shown in the following formula.

$$s^{(n)} = 0.5 \times r^{(n-1)} + 0.5 \times W_s \times w^{(n-1)}, \quad where\ r^{(n)} = r^{(n-1)} S_s. \tag{16}$$

Since word weight is calculated by formula (7), the sentence weight combines both weights from the *similar-to* link and the *is-part-of* link. In the experiment, we combine the similarity network with the subparagraph, paragraph and section nodes.

6). SECTION TITLE MODEL

Based on Assumption 6, we build the Section Title model by extracting section titles from each section as a set of special nodes ***h*** in a Semantic Link Network consisting of sections, paragraphs, sentences and words. In general, section titles are very short description of section content, either related to the technique details or just a category title. We assumed that a section title is related to all sentences in the main body text of its section and the weight of words in the title string also determines the weight of the title. Thus, we compose two new iterative functions for calculating section title weight vector ***h*** and word weight vector ***w***.

$$h^{(n+1)} = H_s \times s^{(n)} \tag{17}$$

$$w^{(n+1)} = \frac{1}{2} W_h \times h^{(n)} + \frac{1}{2} W_s \times s^{(n)} \tag{18}$$

H_s is the matrix of the adjacent graph of connecting a section title to all the sentences in the corresponding section. W_h is the adjacent matrix of connections between the keywords and the section title. Then, formula (17) ensures that the weight of a section title is the summation of all sentences in that section while formula (18) is to integrate the weight of the section title with the weight of sentence for computing the word weight, which finally makes the whole iteration complete.

III. EXPERIMENTS

A. DATA SET

A set of 175 scientific papers is extracted from the proceedings of ACL2014 (https://aclweb.org/anthology/P/P14/) of the ACL Anthology as the benchmark document set. Most papers are regular papers with moderate lengths. Words, sentences, paragraphs and sections are parsed from the texts for building the Semantic Link Network. Abstract texts and conclusion texts are extracted as the gold standards.

Two types of tests are conducted on the dataset of ACL2014: Exclusion Test (*Exc-Paper-Test*) excludes the abstract and conclusion from text; and Inclusion Test (*Inc-Paper-Test*) keeps abstract and conclusion texts, therefore sentences from abstract or conclusion sections could be directly extracted by summarization models.

We also use the DUC 2002 news document collection as the benchmark for evaluation (*DUC-Test*). News report documents in the *DUC-Test* collection are shorter than the ACL papers and have no sections, and each has about ten sentences of a news report. We only test the models without section nodes.

Finally, stop-words are always major consideration in text processing. Our models link words to sentences to give the weights of sentences. To evaluate how stop-words affect the extraction of Semantic Link Network, we conduct experiments on networks removing stop-words and networks including stop-words for comparison.

ROUGE-scores are used as the objective metric to compare these models [4]. Specifically, we use ROUGE-1 to measure how these models perform on the ACL benchmark and on the DUC 2002 benchmark because in our experiments, score ranks are stable among ROUGE-1, ROUGE-2, ROUGE-3, ROUGE-4 and ROUGE-L. A comparison between ROUGE-1 and ROUGE-2 scores is given in Appendix.

When computing the ROUGE score, each paper will have one or several standard summarization text given by humans and the ROUGE score computes how many *n*-gram words of text of the summary are in the standards (known as precision) and how many *n*-grams of standard texts are included in the summary, i.e., recall rate. That is, The ROUGE score is to compute *n*-gram recall rate, precision and *f*-score of summary. Note that each paper can have more than one standard summary. In DUC 2002, two standards are provided. In our ACL 2014 test, we use *abstract* and *conclusion* as two standards for one paper. Finally, since all the summaries are limited to a predefined length, for example, 100 words in DUC 2002. Thus, the recall rate is the major metric to be evaluated for comparing different models. In experiments, we use the Top k sentences and the max-100-words sentences of summarization to compare the performance of models.

We conducted human observation tests on ten ACL 2014 papers. We manually score the extracted sentences by different models and check if the extracted sentence is related to the abstract of the selected paper by manually scoring every extracted sentence. Another experimental data is a long paper for comparing the performance of the models.

B. MODELS TO BE EVALUATED

To facilitate understanding in the following discussion, we list all shorthand and descriptions in Table 1. For comparison, we implemented the *TextRank* model [10], which uses the sentence similarity graph to rank sentences.

TABLE 1.
SHORTHAND AND DESCRIPTION

Notation	Description
Para	Combing word, sentence, and paragraph.
Tfidf	The model using sentence tf-idf score to rank sentence
Simgraph	The model using sentence-sentence similarity matrix and apply PageRank on it (TextRank)
Wordgraph	The model using neighboring words in sentences to build a graph and use formula (13) to compute sentence rank.
Subpara	The model combing word, sentence, sub-paragraph, and paragraph

Parasec	The model combing word, sentence, paragraph and sections.
Subparasec	The model combing word, sentence, sub-paragraph, paragraph and sections.
Hybrid	The model combing **Subparasec** with **Simgraph**
Sectitle	The model combing section title with Parasec
Mywordgraph	Using formula (14) to compute the rank of sentence.
Inc-Paper Test	Tests on ACL papers with abstract and conclusion being kept.
Exc-Paper-Test	Tests on ACL papers with abstract and conclusion being excluded.
DUC-Test	Tests on Single document summarization benchmark from DUC2002.
Top-k	Tests that select top-k sentences as the summary
Max-100	Tests that select top sentences with 100 words limitation as final summary.

TABLE 2.
MAX-100 INC-PAPER-TEST (STOP WORDS ARE REMOVED)

	Average_P	Average_R	Average_F
Para	0.43593	0.34201	0.36680
Tfidf	0.40209	0.30920	0.33119
Simgraph	0.43833	0.32315	0.35147
Wordgraph	0.41936	0.30305	0.32580
Subpara	0.47271	0.37475	0.39979
Parasec	0.44976	0.35285	0.37746
Subparasec	0.45623	0.36181	0.38493
Hybrid	**0.49107**	**0.38547**	**0.41333**
Sectitle	0.47392	0.38354	0.40293

TABLE 3.
MAX-100 EXC-PAPER-TEST (STOP WORDS ARE REMOVED)

	Average_P	Average_R	Average_F
Para	0.40417	0.31130	0.33424
Tfidf	0.35763	0.26576	0.28768
Simgraph	0.38742	0.27524	0.30309
Wordgraph	0.37707	0.26873	0.28917
Subpara	0.39509	0.30124	0.32482
Parasec	0.39080	0.30848	0.33028
Subparasec	0.39665	0.30863	0.33056
Hybrid	0.41583	0.31721	0.34317
Sectitle	**0.41626**	**0.32530**	**0.34665**

C. EXPERIMENTAL RESULTS

Table 2-4 show the ROUGE−1 scores (at the 95% confidence interval) of the extracted sentences with the maximal 100 words on *Inc-Paper-Test*, *Exc-Paper-Test* and *DUC-Test* excluding stop words. In the tables, Average_P is the precision score, Average_R for recall-rate, and Average_F is for *f*-score. In this test, we select top-k sentences with the whole length of selected sentences less than 100 words. We mark the best scores in bold.

TABLE 4.
MAX-100 DUC-TEST (STOP WORDS ARE REMOVED).

Model Name	Average_P	Average_R	Average_F
Para	**0.46027**	**0.45984**	**0.45484**
Tfidf	0.42936	0.41321	0.41193
Simgraph	0.45589	0.44438	0.44366
Wordgraph	0.44479	0.42447	0.42610
Subpara	0.44591	0.43952	0.43728
Hybrid	0.45736	0.44802	0.44740

The results show that the model *Hybrid* obtains the best overall performance. It ranked Top 1 in *Inc-Paper-Test* and Top 2 in *Exc-Paper-Test* and *DUC-Test*. The model *Sectitle* ranks Top 1 in *Exc-Paper-Test* and the *Para* model ranks Top 1 in *DUC-Test*. All models work better in *Inc-Paper-Test* than in *Exc-Paper-Test* because sentences from abstract and conclusion text can be selected in the *Top-k* sentences by models in *Inc-Paper-Test*.

TABLE 5.
TOP-K INC-PAPER-TEST (STOP WORDS ARE REMOVED)

Model Name	Average_P	Average_R	Average_F
Para	0.39540	0.37235	0.35948
Tfidf	0.33576	0.50990	0.38830
Simgraph	0.41250	0.37619	0.37388
Wordgraph	0.32032	**0.52118**	0.38008
Subpara	0.43595	0.41228	0.40175
Parasec	0.41781	0.38192	0.37517
Subparasec	0.42153	0.39560	0.38529
Hybrid	**0.44229**	0.43791	**0.41744**
Sectitle	0.43777	0.43029	0.40955

In the *Subpara* and *Subparasec* model, we use 0.06 as the threshold to extract sub-paragraphs from paragraph of ACL paper and DUC document.

In Table 5, 6, and 7, Top 5 sentences are used as the summarization result without the word count limitation and the summarization result are longer than the *max-100* tests that have length limitation. Similarly, the model *Hybrid* and the model *Sectitle* have Top 2 scores in ROUGE-1 evaluation on the *Inc-Paper-Test* and *Exc-Paper-Test* while *Para* and *Simgraph* are Top 2 models in *DUC-Test*. Note that model *Tfidf* and model *Wordgraph* can have a very high recall-rate score in the *Top-k* experiments because in the *Top-k* test they tend to select longer sentences than other models, which results in a high recall and low precision. Their *f-scores* are lower than others.

In the experiments shown in Table 8-10, the stop words will contribute to the weight of sentences that contain them. We perform a *max-100* test on three benchmark data sets and all the performances degrade. The model *Simgraph* performs the best and the *Hybrid* model performs the second best in *Inc-Paper-Test* and *Exc-Paper-Test*. The model *Para* performs the second best in *Duc-Test*. It is mainly because the Jaccard distance of sentences is less sensitive to the stop-words in *Simgraph* model than the word-sentence link relationships in the Semantic Link Network.

TABLE 6.
TOP-K EXC-PAPER-TEST (STOP WORDS ARE REMOVED)

Model Name	Average_P	Average_R	Average_F
Para	0.36660	0.33622	0.32719
Tfidf	0.29609	0.44908	0.34113
Simgraph	0.36147	0.32742	0.32673
Wordgraph	0.28384	**0.46902**	0.33875
Subpara	0.36421	0.33452	0.32705
Parasec	0.36835	0.32919	0.32443
Subparasec	0.36665	0.33663	0.33033
Hybrid	0.37312	0.36352	0.34580
Sectitle	**0.38142**	0.36301	**0.34942**

TABLE 7.
TOP-K DUC-TEST (STOP WORDS ARE REMOVED)

Model Name	Average_P	Average_R	Average_F
Para	**0.40295**	0.55608	**0.45894**
Tfidf	0.35420	0.61962	0.44616
Simgraph	0.39107	0.56628	0.45528
Wordgraph	0.35162	**0.63336**	0.44754
Subpara	0.39798	0.52135	0.44252
Hybrid	0.40402	0.53970	0.45350

TABLE 8.
MAX-100 INC-PAPER-TEST (STOP WORDS ARE KEPT)

Model Name	Average_P	Average_R	Average_F
Para	0.38938	0.28693	0.30978
Tfidf	0.40062	0.29081	0.31354
Simgraph	**0.45869**	0.33191	**0.36621**
Wordgraph	0.35817	0.25002	0.26954
Subpara	0.42401	0.32670	0.34878
Parasec	0.41432	0.30512	0.32940
Subparasec	0.41615	0.31358	0.33732
Hybrid	0.43881	**0.33321**	0.35818
Sectitle	0.43262	0.31714	0.34199

TABLE 9.
MAX-100 EXC-PAPER-TEST (STOP WORDS ARE KEPT)

Model Name	Average_P	Average_R	Average_F
Para	0.36925	0.27244	0.29398
Tfidf	0.37108	0.26385	0.28553
Simgraph	**0.40565**	**0.30109**	**0.32818**
Wordgraph	0.34590	0.24512	0.26343
Subpara	0.38947	0.28496	0.30880
Parasec	0.38317	0.27962	0.30200
Subparasec	0.38364	0.28612	0.30757
Hybrid	0.39842	0.29587	0.31903
Sectitle	0.39479	0.28834	0.31166

TABLE 10.
MAX-100 DUC-TEST (STOP WORDS ARE KEPT)

Model Name	Average_P	Average_R	Average_F
Para	0.44136	0.43273	0.42884
Tfidf	0.41586	0.38928	0.39280
Simgraph	**0.44819**	**0.44279**	**0.43869**
Wordgraph	0.42124	0.40195	0.40195
Subpara	0.42648	0.41860	0.41464
Hybrid	0.43432	0.43487	0.42779

D. ROUGE RESULTS ANALYSIS

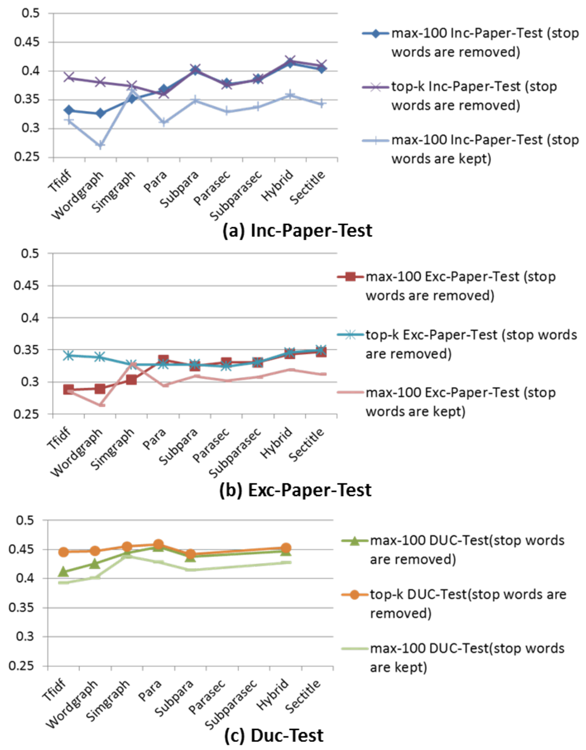
The following are observations:

(1) *The is-part-of link is helpful for summarization* Most of our models have better scores than the *Tfidf* model and the *Wordgraph* model as shown Figure 7.

(2) *The is-part-of link can be more helpful in making short summaries than in making long summaries*. Results of models in the *max-100* tests are shorter than *Top-k* results. Models using *is-part-of* links outperform others more significantly in the *max-100* tests than in the *Top-k* tests.

(3) *The is-part-of links performs better in longer papers with more structural information*. Models perform differently in different documents sets. In general, there is a trend that when the structural information is more obvious, the models that involve more structural links work better. The DUC text contains no section and in many articles one paragraph contains only one sentence. So the structural information is not so obvious. The ACL 2014 papers are conference paper of moderate length (in general 6 pages), and most contains well organized sections and paragraphs, but not so regular. Some papers are short and some papers have few sections, for example, only three or four sections. The long paper we selected for a specific test is mostly well organized. In general, it shows that with more structural information, our *Subparasec* model and *Parasec* model work better as shown in Figure 7. This can be observed from the experiment: The *f-score* is lower than the *Subpara* model that only used *paragraph* nodes and *sub-paragraph* nodes when using section nodes in the *Subparasec* model in *Inc-Papaer-Test*. Moreover, Figure 7 shows the *f-scores* of those models with different semantic links on three benchmarks. We list the models in X-axis in an order of increasing types of semantic links involved in the model. It can be observed as more structural information are involved, the model can have a better performance in the ACL 2014 test document set. So, we believe that it is mainly the organizational structure that enables the models to rank abstract sentence higher.

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal



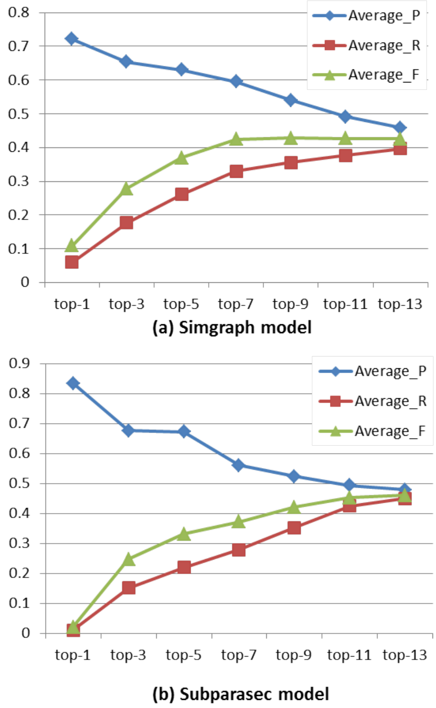


FIGURE 8. Performances of the Simggraph and the Subparasec model with selecting top-1, top-3, ..., top-13 sentences as a summary.

(4) The *Wordgraph* model uses the average and the maximal weight of words of a sentence as the weight of that sentence, which is try to directly bias the sentence length factor. However, in the Top-k experiments where the *Wordgraph* model has a very high recall rate, indicating that the model still tends to rank longer sentences a higher score. The *Simgraph* model uses the Jaccard distance to measure the similarity between sentences, which can also avoid the affect from long sentences. On the contrary, in our other proposed models, sentence length factor is not explicitly modeled. They still can work better than the *Wordgraph* model and the *Simgraph* model.

(5) Stop-words are often removed in the text pre-processing stage in many other models. We are interested in studying how the graph model can cope with this problem when stop-words are not removed. The performances degrade but the *Simgraph* model become better than others in coping with the stop words. This opened a problem that how to use a structural graph to bias the stop words affects.

E. TOP-K PERFORMANCE ON LONG PAPER

The ROUGE-1 scores of Top-k (k=1, 3, 5, …, 13) sentences of the model *Subparasec* and the model *Simgraph* (denoted as Top-1 to Top-13 in x-axis in Figure 8) on a longer paper [1] are shown in Figure 8. The precision and recall-rate scores change when k changes. The precision score of Top-1 summary of the model *Subparasec* is significantly higher than others because it ranks the title of the paper as its first sentence (See Figure 14 for the Top-1 sentences of two models). This is significant as the title of the paper is the best sentence that reflects the core of the paper. The Top-1 sentence of the *Simgraph* model is not as good as *Subparasec*. The ability to select the title sentence into the summary demonstrates the performance of the *Subparasec* model and the *Parasec* model. When adding more sentences, precision scores are decreasing and recall rate are increasing, which is naturally because more sentences means more coverage of the papers' words and consequently increase the recall rate of the ROUGE-1 score. Decreasing precisions means more irrelevant words are also incorporated when more sentences are selected. When *f-score* is increasing, it shows that the model is still achieving a better performance. Starting from Top-9, the *f*-score of *Simgraph* no longer increases while the *f*-score of *Subparasec* still increases because more abstraction and conclusion sentences are selected by the *Subparasec* model after Top-9, Top-10 and Top-13 (See the sentences in red color in Figure 14).

F. COMPARISON WITH MIND MAP

To observe the result on long paper, we further conducted an experiment on [1], a much longer paper than those in the ACL2014 proceedings. The paper contains an extensive survey on relevant research works and the abstract is more general than technical papers in the ACL proceedings. We manually build another summarization benchmark for this paper: Constructing a word tree ***T*** using a classical MindMap method [5], which consists of words or sentences as nodes representing a topic at a certain abstraction level. Figure 9 shows the Top-3 levels of the tree where the sub-trees of leaf

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2018.2856530, IEEE Access

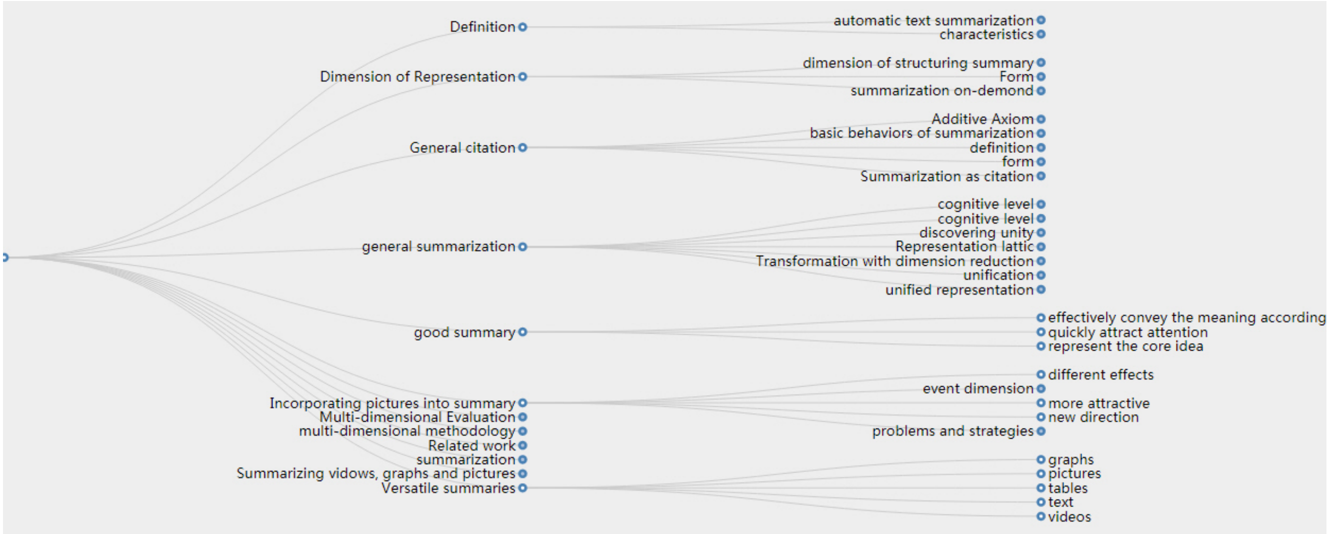
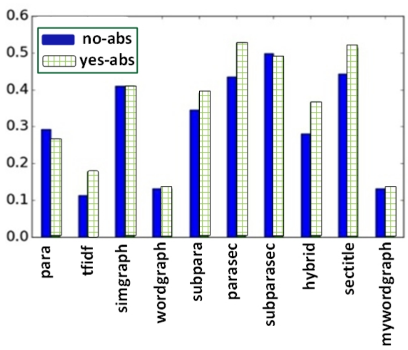


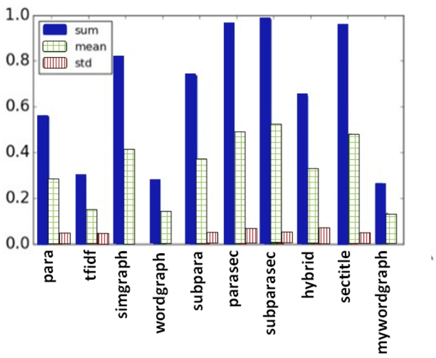
FIGURE 9. Part of the word tree of "Dimensionality on summarization"[1]

nodes are all folded. The first layer of sub-tree nodes are given by section titles and then the following sub-trees are composed by choosing the representative words or sentences while reading the section and paragraph text of the paper.

We calculate the similarity between the summary and the word tree. Since tree nodes are of different weights in nature, we first use the PageRank algorithm to rank the nodes on the tree. Then, for each sentence *s*, we compute its average similarity to ***T*** as follows:



(a) Similarity of summary to the keyword tree



(b) Average Similarity of summary to the keyword tree

$Rank(s) = \frac{1}{n}\sum_{i=1}^{n} w_i \times DistCos(s,T_i)$, where $DistCos(s, T_i)$ is the *Cosine* similarity (the higher, the more similar) between the string of sentence ***s*** and the string of the ***ith*** node in the tree of *n* nodes. Symbol $\boldsymbol{w_i}$ denotes the weight of the ***ith*** node. Finally, the similarity between a summary and the word tree *T* is calculated by averaging ranks of all sentences in summary.

TABLE 11.
ROUGE-1 SCORE ON PAPER "DIMENSIONALITY OF SUMMARIZATION"

	Average_P	Average_R	Average_F
para	0.4005	0.39918	0.39978
tfidf	0.37671	0.3959	0.38605
simgraph	0.4345	0.43355	0.43402
wordgraph	0.36364	0.38344	0.37328
subpara	0.36744	0.38778	0.37734
parasec	**0.48136**	**0.48303**	**0.4821**
subparasec	0.45695	0.48684	0.47123
hybrid	0.38595	0.38069	0.38329
sectitle	0.46411	0.46778	0.46594
mywordgraph	0.36364	0.38344	0.37328

Figure 10 shows the similarity between the model results and the word tree. Two versions of papers are tested, one keeps abstraction and conclusion sentences(*yes_abs* in Figure 10(a)) and another removed those sentences being removed (*no_abs* in Figure 10(a)). The average (*mean*), summation (*sum*) and standard deviation (*std*) of the scores of two versions are shown in Figure 10(b). *Subparasec* and *Parasec* are two best scores in this experiment.

Table 11 shows the ROUGE-1 score of the ten models on the two versions of the paper. *Parasec* and *Subparasec* achieve the Top-2 ranks in *f-score*, which demonstrates that models using *is-part-of* links with section and sub-paragraph nodes outperform others. Figure 14 lists top-11 sentences of three models on the paper of *yes_abs* version. The paper title

"*Dimensionality on Summarization*" is ranked Top 1 by both the *Subparasec* model and the *Parasec* model. We mark the sentences extracted from the abstraction and conclusion section in red color. There are totally five sentences selected from these two sections by the *Subparasec* model and the *Parasec* model. Blue sentences in Figure 14 are sentences shared by these two models and they have only one different sentence marked in green color. So these two models achieved a close rank score on this paper. The model *Simgraph* cannot rank abstraction and conclusion sentences as Top 11 and only two sentences are shared by the model *Subparasec* and the model *Parasec*. Thus, the score of model *Simgraph* is significantly lower than that of model *Subparasec* and model *Parasec* in both the word tree test and the ROUGE-1 test.

TABLE 12.
SCORE OF 10 SELECTED PAPERS IN INC-PAPER-TEST JUDGED BY RESEARCHERS.

DOC ID	SECTITLE	SIMGRAPH	HYBRID
DOC 1	3.5/4	3.5/5	3.0/3
DOC 2	4.5/8	4.0/7	5.0/8
DOC 3	4.0/6	2.0/7	7.0/7
DOC 4	7.5/11	6.0/12	10.0/11
DOC 5	6.0/11	2.0/9	6.5/11
DOC 6	4.0/6	2.5/5	5.0/5
DOC 7	8.5/9	7.5/9	8.5/9
DOC 8	2.5/6	2.5/3	2.5/5
DOC 9	3.0/3	3.0/3	3.0/3
DOC 10	2.0/3	3.0/3	2.0/2
SUM	7.02	6.44	8.56
AVERAGE	0.70	0.64	0.86

G. Observation

To further observe the proposed models, we select the summaries of ten papers from the results of the *Inc-Paper-Test* on the ACL 2014 proceedings and manually score each sentence. If a sentence in the summarization is closely related to one sentence in the abstract, it is given 1 score. If a sentence is indirectly related to one of the sentences in the abstract text, it is given 0.5 score. If a sentence has no direct or indirect relation to any sentence in the abstract, it is given 0 score. The result is shown in Table 12. Each score has a denominator representing the number of sentences in its summary and the numerator is the total score of the sentences in the summary. To show the significance of differences between the three models, we compute the *p-value* of two models. *T-test* shows that the p-value between the *Sectitle* model and the *Simgraph* model is 0.0035, which shows the difference is greatly significant, *p-value* between *Hybrid* and *Simgraph* is 0.027, which is still lower than 0.05, while the *p-value* between *Sectitle* and *Simgraph* is 0.25, which does not support a meaningful difference. Thus, the *Hybrid* model significantly outperforms the *Simgraph* model, which in turn confirms the ROUGE-score tests in previous sections.

IV. RELATED WORK

Texts can be viewed from different dimensions and scales [2]. A basic strategy is to represent text in a structure that reflects the importance of sentences at a certain dimension so that the weights of sentences can be approximately calculated from the text structure [11]. Graph structures have been widely used as a specific text representation model for sentence ranking and extractive summarization. For example, a sentence similarity network was built to represent a document and a PageRank algorithm is used to rank sentences [10]. The weights of sentence features were incorporated in a sentence similarity network for ranking sentences [12]. Semantic graphs were investigated for ranking sentences with an accurate semantics representation [13]. Sentences were ranked by their nested discourse tree structure within a single document for single document summarization [14]. Recognizing Textual Entailment relationship was built as a network to rank sentences for summarization [15]. A hyper graph was leveraged to model groups of sentences for ranking and summarization [16]. An entity-sentence structure was used to rank sentences according to the importance, coherence and non-redundancy, where entities are nouns extracted from text [17]. In these methods, a graph with sentences as nodes was constructed as the representation of a document, then an iterative ranking algorithm was applied to the graph to obtain the weights of sentences, and finally a selecting schema is used to select sentences according to the weights of sentences as well as other features. Machine learning methods were used to learn to produce summarization [18]. Rhetorical structures of sentences from scientific documents were investigated in a supervised model for extractive summarization [19]. However, they lack extensibility facing various domains and topics.

Graph-based methods are flexible and extensible in that they in general do not need pre-training samples. PageRank [9] or HITS [10] can be applied to rank graph for obtaining the weight vector of nodes. The centrality property of a graph of lexicon can be used for ranking sentences [20]. In previous methods, graphs were mostly built based on sentence-sentence relations and sentence-word relations. They seldom concern structural information such as paragraph and section. The importance of paragraph was computed and extracted for summarization [21], but their concept of paragraph is just the same as sentence. Recently, researchers focused on the extension of documents using information from outside networks. For example, citation networks were leveraged to build the summarization of document [22]. Related works of papers were generated by producing summarization [23]. Wiki and social context knowledge were investigated for single document summarization [24].

The Semantic Link Network is a self-organized network that consists of semantic nodes, semantic links and rules on semantic links. A semantic space defines the semantics of semantic nodes, semantic links and rules for reasoning on semantic links [2]. Semantic link network is a desirable semantics modeling technique for implementing advanced services based on text such as question answering, recommendation and information extraction [25] and intelligent knowledge services in the emerging Cyber-Physical Society [26].

Researches on Semantic Link Network can be traced to the definition of inheritance rules [6] and the implementation of Active Document Framework [27]. It was then used for supporting intelligent Web applications by extending hyperlink [28]. The Semantic Link Network emerges semantic communities in a different way as the general social network does [29]. The interactive semantics and the semantic base were proposed for establishing the semantic basis of understanding [30]. Semantic Link was used to model the basic structure of Cyber-Physical Society [31]. The Semantic Link Network model is able to cope with dynamicity of Cyber-Physical Society [32]. Decentralized semantic overlay networks were studied to support high-level semantics-based applications [33]. Distributed Semantic Link Network query process is also supported [34]. Research also concerns automatic construction of Semantic Link Networks on various resources like texts [35]. The integration of Semantic Link Network and Multi-Dimensional Resource Space forms a new semantic model with stronger expression ability [36]. Recently, a general summarization approach based on the definition of general citation was proposed [1]. Semantic Link Network has become a stream of studying semantics modeling.

A scientific paper normally consists of sections, which consists of paragraphs. This structure helps readers easily read and understand the content. Paragraphs and sections indicate the topics to be delivered. Previous models mostly work on vector space features [37] or a graph containing sentences and words [38] without considering larger representation units such as paragraph or section. We used several structure graphs of a paper for ranking sentences [39].

The semantics of text can be modeled by a process of forming its structure. Different formation processes determine different semantic structures. The emerging semantic structure of text was studied as a special case of near decomposable complex system [2].

Different from previous work, this work proposes a systematic method for summarizing scientific paper by ranking sentences, proves the convergence of the iterative ranking algorithm, and conducts extensive experiments to verify the roles of semantic links. Current techniques are still far from implementing an expert-level summarization. A proper semantics representation method is important for making a satisfactory summarization. The extractive summarization of single document can be easily extended to summarize multiple documents [40].

V. ADVANTAGES AND IMPLICATIONS

The proposed method has the following advantages:

(1) *It is extensible*. Different instances of Semantic Link Network can be built by incorporating different types of nodes and semantic links with different reinforcement assumptions.
(2) *It does not need any preprocessing or entity-concept extraction from document*. The instances of Semantic Link Network can be directly constructed on word, sentence and paragraph, which is easy to parse out from texts. Once a Semantic Link Network is built, the iterative function can be easily implemented within the proposed framework after selecting the appropriate assumptions to form iteration relations on the instance Semantic Link Network.
(3) *Its convergence is quick*. The iterative function can quickly converge because we reinforce strong link connection when making iterative relations among nodes.

This work also draws some implications.

(1) The *is-part-of* link plays an important role in rendering the core of paper, especially when the paper is long. This is an effect of emerging semantics introduced in [2].
(2) Combining the *is-part-of* link and the *similar-to* link of sentence generates the better solution for papers of moderate length. This can be interpreted as the effect of increasing the connectivity of the Semantic Link Network brought by incorporating more semantic links.
(3) Sub-paragraph information is also important in summarizing long documents.
(4) Too macro or too micro structural information may not help improve the quality of summarization when the paper length is moderate.
(5) The models can be extended and leveraged to automatically generate the MindMap of paper, just as what we have demonstrated that the generated summary is highly similar to manually built Mind Map.

VI. CONCLUSION

The reinforcement ranking on the Semantic Link Network of various representation units within scientific paper can significantly improve extractive summarization of paper. It not only provides an approach for summarization based on semantics modeling but also verifies the significance of Semantic Link Network in representing and understanding the content of paper.

The proposed approach has stable quality in single document summarization on both scientific papers and short news text in DUC 2002 test documents and perform better when documents has more structural information modelled by Semantic Link Network.

The proposed approach can be applied to any structural text and the provision of various summarization services such as automatically generating the Mind Map of scientific paper, slides for a given paper, and extended abstract for a long scientific paper or book to give readers a quick impression of the core content.

Acknowledgment

Research was supported by the International Research Network on Cyber-Physical-Social Intelligence consisting of Guangzhou University, Aston University, Key Laboratory of Intelligent Information Processing at the Institute of Computing Technology in Chinese Academy of Sciences, and University of Chinese Academy of Sciences (all units are equal). Guangzhou University supports the open-access publication of this paper. Professor Hai Zhuge is the corresponding author.

APPENDIX

ROUGE-1 scores and ROUGE-2 scores are consistent in evaluating the proposed models. We show three comparisons in Appendix. Figure 11 contains two metrics obtained in MAX-100 INC-Paper-Test; Figure 12 lists results of MAX-100 EXC-Paper-Test; Figure 13 shows the scores of MAX-100 DUC-Test.

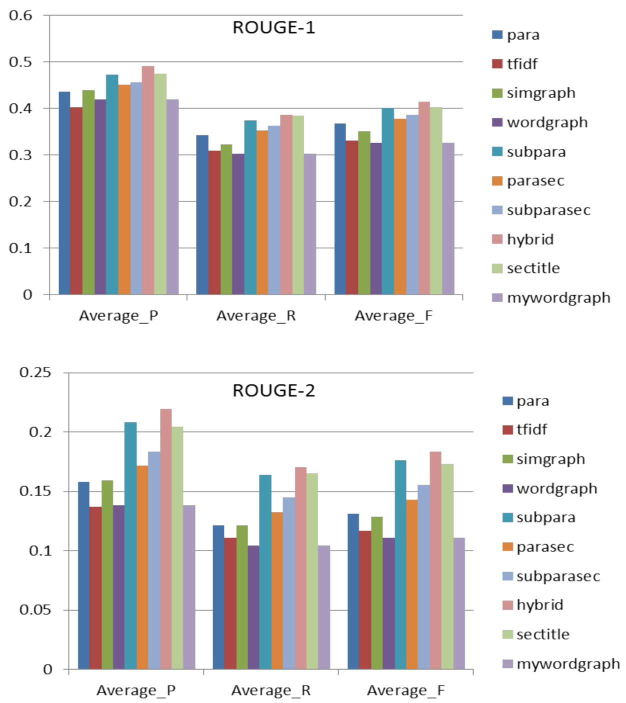


FIGURE 11. ROUGE-1 and ROUGE-2 scores on I MAX-100 INC-PAPER-TEST (Stop words are removed)

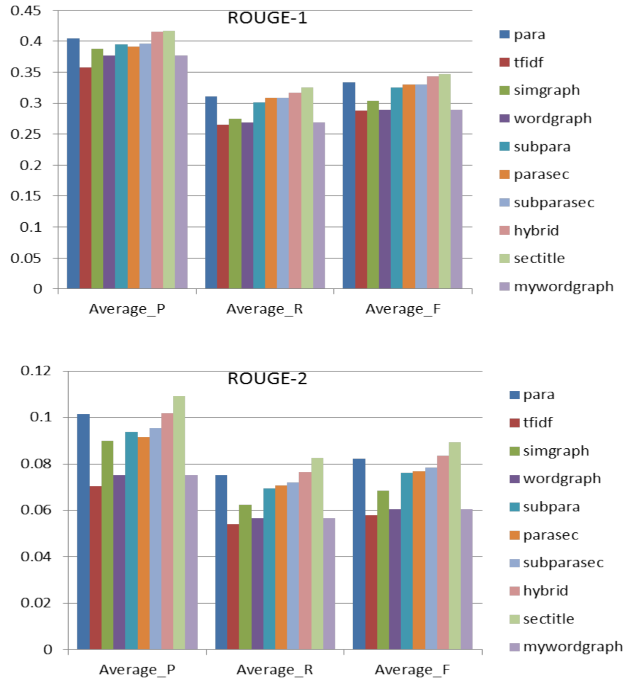


FIGURE 12. ROUGE-1 and ROUGE-2 scores on I MAX-100 EXC-PAPER-TEST (Stop words are removed)

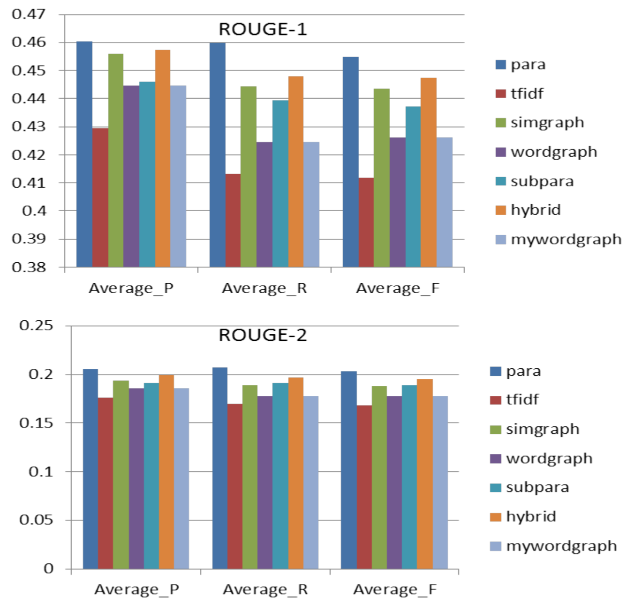


FIGURE 13. ROUGE-1 and ROUGE-2 scores on I MAX-100 DUC-TEST (Stop words are removed)

REFERENCES

[1] H. Zhuge, "Dimensionality on summarization" arXiv preprint arXiv:1507.00209, 2015.

[2] H. Zhuge, "Multi-Dimensional Summarization in Cyber-Physical Society". Morgan Kaufmann, 2016.

[3] H. Zhuge and J. Zhang, "Automatically constructing Semantic Link Network on documents" Concurrency and Computation: Practice and Experience, vol. 23, no. 9, pp. 956-971, 2011.

[4] C.-Y. Lin and E. Hovy, "Automatic evaluation of summaries using n-gram co-occurrence statistics," in Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1, 2003, pp. 71-78: Association for Computational Linguistics.

[5] M. Davies, "Concept mapping, mind mapping and argument mapping: what are the differences and do they matter? " Higher education, vol. 62, no. 3, pp. 279-301, 2011.

[6] H. Zhuge, "Inheritance rules for flexible model retrieval," Decision Support Systems, vol. 22, no. 4, pp. 379-390, 1998.

[7] L. Page, S. Brin, R. Motwani, and T. Winograd, "The PageRank citation ranking: Bringing order to the web" Stanford InfoLab1999.

[8] J. M. Kleinberg, "Hubs, authorities, and communities" ACM computing surveys (CSUR), vol. 31, no. 4es, p. 5, 1999.

[9] R. S. Varga, "Matrix iterative analysis". Springer Science & Business Media, 2009.

[10] R. Mihalcea and P. Tarau, "Textrank: Bringing order into text" in Proceedings of the 2004 conference on empirical methods in natural language processing, Barcelona, Spain, 2004, pp. 404–411.

[11] A. Nenkova and K. McKeown, "Automatic summarization" Foundations and Trends in Information Retrieval, vol. 5, no. 2–3, pp. 103-233, 2011.

[12] J.-Y. Yeh, H.-R. Ke, and W.-P. Yang, "iSpreadRank: Ranking sentences for extraction-based summarization using feature weight propagation in the sentence similarity network" Expert Systems with Applications, vol. 35, no. 3, pp. 1451-1462, 2008.

[13] M. Joshi, H. Wang, and S. McClean, "Dense semantic graph and its application in single document summarisation," in Emerging Ideas on Information Filtering and Retrieval: Springer, 2018, pp. 55-67.

[14] Y. Kikuchi, T. Hirao, H. Takamura, M. Okumura, and M. Nagata, "Single Document Summarization based on Nested Tree Structure," in ACL (2), 2014, pp. 315-320.

[15] M.-T. Nguyen and M.-L. Nguyen, "Sortesum: A social context framework for single-document summarization" in European Conference on Information Retrieval, 2016, pp. 3-14: Springer.

[16] W. Wang, F. Wei, W. Li, and S. Li, "Hypersum: hypergraph based semi-supervised sentence ranking for query-oriented summarization" in Proceedings of the 18th ACM conference on information and knowledge management, 2009, pp. 1855-1858: ACM.

[17] D. Parveen and M. Strube, "Integrating Importance, Non-Redundancy and Coherence in Graph-Based Extractive Summarization" in Proceedings of the 24th International Confer-ence on Artificial Intelligence, 2015, pp. 1298-1304.

[18] G. Durrett, T. Berg-Kirkpatrick, and D. Klein, "Learning-based single-document summarization with compression and anaphoricity constraints," arXiv preprint arXiv:1603.08887, 2016.

[19] S. Teufel and M. Moens, "Summarizing scientific articles: experiments with relevance and rhetorical status" Computational linguistics, vol. 28, no. 4, pp. 409-445, 2002.

[20] G. Erkan and D. R. Radev, "Lexrank: Graph-based lexical centrality as salience in text summarization" Journal of Artificial Intelligence Research, vol. 22, pp. 457-479, 2004.

[21] M. Mitrat, A. Singhal, and C. Buckleytt, "Automatic text summarization by paragraph extraction," presented at the Fifth Workshop on Very Large Corpora, Beijing, China, 1997.

[22] Q. Mei and C. Zhai, "Generating Impact-Based Summaries for Scientific Literature" in ACL, 2008, vol. 8, pp. 816-824; A. Abu-Jbara and D. Radev, "Coherent citation-based summarization of scientific papers," in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1, 2011, pp. 500-509: Association for Computational Linguistics.

[23] J. Chen, H. Zhuge. "Automatic generation of related work through summarizing citations". Concurrency & Computation Practice & Experience, 2017(1):e4261, online.

[24] X. Wan and J. Xiao, "Exploiting neighborhood knowledge for single document summarization and keyphrase extraction," ACM Transactions on Information Systems (TOIS), vol. 28, no. 2, p. 8, 2010.

[25] H. Zhuge, "Semantic linking through spaces for cyber-physical-socio intelligence: A methodology" Artificial Intelligence, vol. 175, no. 5-6, pp. 988-1019, 2011.

[26] H. Zhuge, The Knowledge Grid: Toward Cyber-Physical Society. World Scientific, 2nd edition, 2012.

[27] H. Zhuge, "Active e-document framework ADF: model and tool," Information & management, vol. 41, no. 1, pp. 87-97, 2003.

[28] H. Zhuge, R. Jia, and J. Liu, "Semantic Link Network builder and intelligent semantic browser," Concurrency and Computation: Practice and Experience, vol. 16, no. 14, pp. 1453-1476, 2004.

[29] H. Zhuge, "Communities and emerging semantics in Se-mantic Link Network: Discovery and learning," IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 6, pp. 785-799, 2009.

[30] H. Zhuge, "Interactive semantics," Artificial Intelligence, vol. 174, no. 2, pp. 190-204, 2010.

[31] H. Zhuge, "Socio-natural thought Semantic Link Network: A method of semantic networking in the cyber physical society," in Advanced Information Networking and Applications (AINA), 2010 24th IEEE International Conference on, 2010, pp. 19-26: IEEE.

[32] H. Zhuge and B. Xu, "Basic operations, completeness and dynamicity of cyber physical socio Semantic Link Network CPSocio‐SLN," Concurrency and Computation: Practice and Experience, vol. 23, no. 9, pp. 924-939, 2011.

[33] H. Zhuge and X. Li, "Peer-to-Peer in Metric Space and Semantic Space" IEEE TKDE, vol. 19, no. 6, p. 759-771.

[34] H. Zhuge, J. Liu, L.Feng, X.Sun, and C. He, "Query Routing in a Peer-to-Peer Semantic Link Network" Computational Intelligence, vol. 21, no. 2, p. 197-216, 2005.

[35] H. Zhuge and J. Zhang, "Automatically constructing Se-mantic Link Network on documents," Concurrency and Computation: Practice and Experience, vol. 23, no. 9, pp. 956-971, 2011.

[36] H. Zhuge and Y. Xing, "Probabilistic Resource Space Model for Managing Resources in Cyber-Physical Society," IEEE Transactions on Services Computing, vol. 5, no.3, p. 404-21. 2012.

[37] J. Conroy and S. Davis, "Vector space and language models for scientific document summarization," in Proceedings of NAACL-HLT, 2015, pp. 186-191.

[38] Y. Yoshida, J. Suzuki, T. Hirao, and M. Nagata, "Dependency based Discourse Parser for Single-Document Summarization," in EMNLP, 2014, pp. 1834-1839.

[39] J. Tian, M. Cao, J. Liu, X. Sun, and H. Zhuge, "Sentence Ranking with the Semantic Link Network in Scientific Paper" in Semantics, Knowledge and Grids (SKG), 2015 11th International Conference on, 2015, pp. 73-80: IEEE.

[40] L. Marujo, R. Ribeiro, D. M. de Matos, J. P. Neto, A. Gershman, and J. Carbonell, "Extending a single-document summarizer to multi-document: a hierarchical approach" arXiv preprint arXiv:1507.02907, 2015.

Xiaoping SUN is an associative research professor of Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, University of Chinese Academy of Sciences, Chinese Academy of Sciences. His research interests include semantic computing, graph computing, P2P computing and decentralized information sharing. His work appeared in several international journals including IEEE TKDE and IEEE TPDS.

Hai Zhuge is a Distinguished Scientist of the ACM (Association of Computer Machinery), a Fellow of British Computer Society, and a Senior Member of IEEE. He has made a systematic contribution to semantics and knowledge modeling through lasting fundamental research on the Semantic Link Network and the Resource Space Model based on multi-dimensional methodology. He is leading research towards Cyber-Physical Society through methodological, theoretical and technical innovation. He gave over 20 keynotes at international conferences and invited lectures in universities of many countries as a Distinguished Speaker of the ACM. He leads the International Research Network on Cyber-Physical-Social Intelligence consisting of Aston University, Guangzhou University, Key Laboratory of Intelligent Information Processing at Institute of Computing Technology in Chinese Academy of Sciences, and University of Chinese Academy of Sciences. He was a Distinguished Visiting Fellow of Royal Academy of Engineering. He is the author of three monographs: Multi-Dimensional Summarization in Cyber-Physical Society (Morgan Kaufmann, 20016), The Knowledge Grid: Toward Cyber-Physical Society (World Scientific, 2012 2nd Ed), and The Web Resource Space Model (Springer, 2008). He is serving as an associate editor of IEEE Intelligent Systems.
Homepage: http://www.knowledgegrid.net/~h.zhuge.

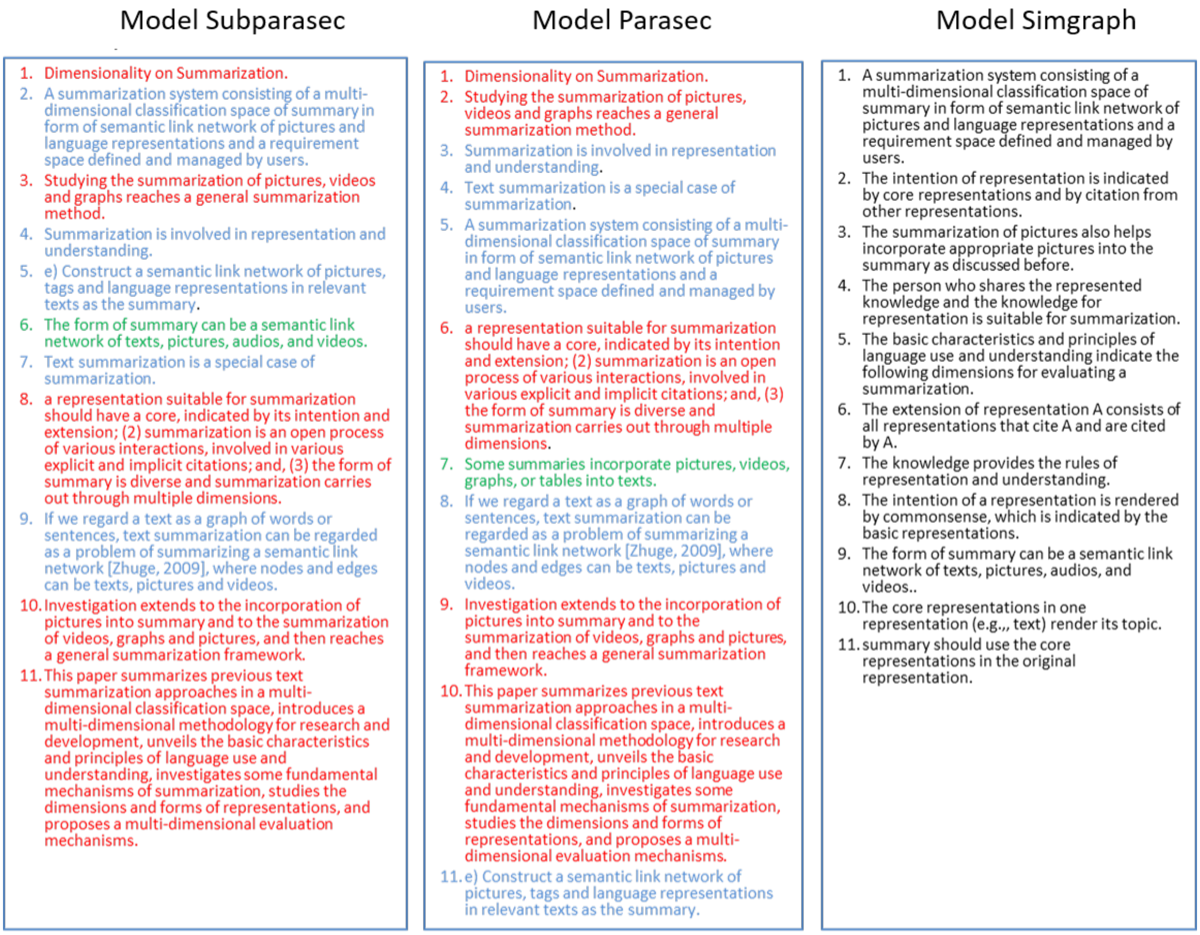


FIGURE 14. Top-11 sentences on paper "Dimensionality of Summarization".