2013 IEEE/WIC/ACM International Conferences on Web Intelligence (WI) and Intelligent Agent Technology (IAT)

A Four Dimension Graph Model for Automatic Text Summarization

Rafael Ferreira*†, Frederico Freitas*, Luciano de Souza Cabral*, Rafael Dueire Lins*, Rinaldo Lima*, Gabriel França*, Steven J. Simske‡, and Luciano Favaro§

*Informatics Center, Federal University of Pernambuco, Recife, Pernambuco, Brazil

†Department of statistics and informatics, Federal Rural University of Pernambuco, Recife, Pernambuco, Brazil

‡Hewlett-Packard Labs., Fort Collins, CO 80528, USA

§Hewlett-Packard Brazil, Barueri, São Paulo, Brazil

***Abstract*—Text summarization is the process of automatically creating a shorter version of one or more text documents. In this context, word-based, sentence-based and graph-based methods approaches are largely used. Among these, graph-based methods for automatic text summarization produce summaries based on the relationships between sentences. These relationships may also support the creation of several text processing applications such as extractive and abstractive summaries, question-answering and information retrieval systems, among others. A new graph model for text processing applications is proposed in this paper. It relies on four dimensions (similarity, semantic similarity, coreference, discourse information) to create the graph. The rationale behind the proposal presented here is resorting to more dimensions than previous works, and taking into account coreference resolution, taking into account to the role of pronouns in connecting the sentences. Coreference was not used in any previous graph-based summarization technique. An experiment was performed using the TextRank algorithm with the presented approach, on the CNN corpus. The results show that the model proposed here outperforms the current approaches both quantitatively and qualitatively.**

***Keywords*-Graph-Model; Summarization; TextRank**

I. Introduction

The rapid growth of the Internet has yielded an unprecedented amount of information available, especially regarding text documents (e.g. news articles, electronic books, scientific papers, blogs, etc.). Such paramount volume of information on the Internet made unfeasible to efficiently sieve useful information. Automatic methods to index, classify and present all information in a clear and concise way, allowing users to save time and resources, are needed today, more than ever.

In such a scenario, text summarization arises as a possible solution. Text summarization (TS) aims at automatically creating a compressed version of one or more documents, extracting the essential information in them. TS techniques are classified as *Extractive* and *Abstractive* [1]. Extractive systems select a set of the most significant sentences from a document, exactly as they appear, to form the summary. Abstractive systems attempt to improve the coherence among the sentences in the summary by eliminating redundancies and clarifying their context. In the latter technique, sentences that are not in the original text may be introduced. Extractive summaries are of more widespread use today, as less computational resources are required.

Graph-based approaches for automatic summarization have provided encouraging results [2], [3], [4]. In such methods, the importance of each sentence is regarded in relation to the other sentences in the document. When a sentence refers to another one a link (edge) with an associated weight between them is created. The weights are used to generate the scores of a sentence. The TextRank [2] algorithm order sentences using graphs, and is widely referenced in the literature. Such an approach is analogous to the PageRank algorithm [5] to rank sentences in a document or in a collection of documents.

Despite the very promising results presented by some graph-based approaches in literature, they do not take into account the inner document representation [6], [4], [7], [4]. In general, those approaches use sentences as nodes in graphs and the similarity values between sentences as edges. The similarity values are often measured by the content overlap between sentences. The process is often accomplished in two steps. First, to create an edge between two sentences, stop words are removed, and stemming and lemmatization techniques [8] are executed to preprocess sentences. Then, the words remaining from the sentences are matched against the words forming other sentences in text. Such representation does not take into account important aspects such as:

- **Semantic similarity**: Traditional approaches to find the similarity between sentences rely only in the lexical and syntactical aspects of words in sentences. The semantic similarity, may also include synonyms and relations such as hyponym hypernym[9];
- **Coreference resolution**: This task tries to map expressions in a document that refer to the same noun. It is useful to link sentences that are about the same subject [10];
- **Discourse relations**: The Discourse relations presents in what way parts of a text relate to each other to form a coherent whole. For this reason, discourse structure evidences the important relationship in the text, that shall not be neglected [11].

978-1-4799-2902-3/13 $31.00 © 2013 IEEE
DOI 10.1109/WI-IAT.2013.55

A four dimension graph model for text processing applications is presented here. The rationale of the proposal is combining the sentence similarity strategy used by previous work with semantic similarity, coreference resolution and discourse relations. Section 3 presents the details on the importance of encompassing all those aspects. The approach presented here differs from related work in two main aspects: (i) It combines more dimensions than the other approaches. In general, the related works apply only one or two techniques to create the graph-model; and (ii) It additionally relies on coreference resolution, a feature not found in any previous work that makes use of graph-models. It is important to notice that no new algorithm to select sentences to an extractive summary, such as TextRank, is created. A graph model that could support TextRank and other algorithms to create several text processing applications is created instead (for example query-answering and information retrieval systems), not only summarization. This paper focuses exclusively on text summarization application, however.

The proposal presented was evaluated using the TextRank algorithm and the CNN corpus. The results obtained showed that the proposed model outperforms its predecessors. In particular, the TextRank algorithm with the four dimension graph model achieved better scores both in terms of quantitative (F-measure) and qualitative evaluations, improving the results obtained by some traditional methods up to 34%.

The rest of this paper is organized as follows: Section II presents the related works and compares them with proposal presented. Section III details the new graph-based representation proposed here. Section IV shows the results obtained using the textrank algorithm in CNN news dataset. Finally, some conclusions and discussion of possible future work are presented in Section V.

II. Related Work

There are several papers in the technical literature that uses sentence similarity to create the edge between two sentences [2], [12], [7], [4], [13], [14].

Rada Mihalcea and Paul Tarau [2] proposed an undirected weighted graphs model that rely on sentence similarity to create edges. Here, the similarity is measured as function of their content overlap. In other words, a sentence that addresses certain concepts in a text, gives the reader a "recommendation" to refer to other sentences in the text that address the same concepts, thus a link can be drawn between any two sentences that share common content. References [7] and [4] apply the same concept of similarity.

In *Complex Networks and Extractive Summarization* [13] the authors use common meaningful nouns of a sentence to create the edges. In other words, they exclude all other words from sentences before running the similarity algorithm.

In [12] the authors use a bag-of-words model to represent each sentence as an N-dimensional vector, where *N* is the number of all possible words in the target language. The similarity between two sentences is then defined by the cosine between their two corresponding vectors.

Shuzhi Sam Ge, *et al.* [14] also propose the cosine similarity metric together with some discourse relationships in a vector representation of sentences. Every sentence is modeled by a weighted term-frequency vector. In order to measure the discourse relationships (called discourse relationship weight) among them, the connectives proposed in [11]are used. The edges between the sentences are created based on sentence similarity and the discourse relationship weight.

Benbrahim and Ahmad [15] model the edges of a graph with cohesion links (such as repetition, synonymy, antonym and hyponymy) between words in sentences.

The article entitled *Discourse indicators for content selection in summarization* [11] presents an approach that relies on discourse information to build edges of the graph. An edge between two nodes representing a relation is directed in case asymmetric relations such as *Cause* and *Condition* are found while they are not considered to have a direction for symmetric relations like *Similarity* and *Contrast*. This information are obtained from the Penn Discourse Treebank [16]. The authors also take into account structural and semantic information to generate the summary. The semantic information is not encapsulated in the graph, however.

Another important aspect not to be neglected is the level of linguistic representation considered in the model. This may vary, the approaches usually considers only superficial characteristics (such as repetition of words, etc).

In the next section, the proposed four-dimension graph model is introduced.

III. The Four-Dimension Graph Model

The new graph model to improve text summarization combines the most successful approaches presented in the literature. The main characteristics of the proposal are:

- The related work referenced in the last section mainly considers superficial characteristics, such as repetition of words, to build the graph. The new approach presented here enhances those features by performing a syntactic analysis of the text discourse together with semantic and linguistic aspects. Besides that, the new algorithm finds relationships between sentences, such as coreferences, that other approaches do not take into account.
- It relies on coreference resolution for identifying entities in the documents that are referenced by either pronouns or nominal descriptions. The experiments performed allow saying that coreference resolution can play a very important role in the task at hand once some terms, such as pronouns, have no explicit connection with other sentences in a document if only traditional and semantic similarity measures are at hand. None

of the aforementioned approaches applies coreference resolution.
- Similarity measures (content overlap, cosine), semantic metrics, coreference resolution, and discourse connectives are taken into account to build a graph that corresponds to the text to be summarized. New dimensions are being considered, since usually the competing approaches are restricted to one or two dimensions only. The results presented in Section IV show that the combination strategy proposed here substantially improves the performance of the TextRank algorithm.

A new graph-based representation to improve extractive text summarization is proposed. The sentences from the document are seen as nodes and four different kinds of edges are used:

- Similarity;
- Semantic Similarity;
- Coreference Resolution;
- Discourse Relations.

The motivation and details on the creation of each dimension is explained in the following sections. In addition, the prototype based on this proposal executes some preprocessing services in order to improve the edge creation methods.

A. Similarity

The similarity method to create edges is the simplest and most widely used, as introduced in section II. This method measures the overlap content between pairs of sentences. If it exceeds a threshold score, selected by the user, the edge between the sentence pair is created. The traditional way to perform this method is, for example, content overlap and similarity using cosine metric. Four similarity methods were implemented:

1) Centrality: Whenever the vocabulary of a sentence overlaps with the other sentences in a document [17] one talks about “sentence centrality”. No semantic “value” is considered. Centrality could be calculated as follows:

$$Score = \frac{Ks \cap KOs}{Ks \cup KOs} \tag{1}$$

where,
Ks = Keywords in s, and
KOs = Keywords in other sentences.

2) Cosine Similarity: It creates a bag-of-words model to represent each sentence as an N-dimensional vector. The vector encapsulates the word and its frequency in the text. The similarity between two sentences is then defined by the cosine between the two corresponding vectors. The formula 2 calculates the cosine similarity.

$$Wsim(Si, Sj) = \frac{\overrightarrow{Si}.\overrightarrow{Sj}}{\left\|\overrightarrow{Si}\right\| x \left\|\overrightarrow{Sj}\right\|} \tag{2}$$

where,
$\overrightarrow{Si}$ and $\overrightarrow{Sj}$ are weighted term-frequency vectors of sentence Si and Sj.

3) Entropy Measure: Entropy measures the degree of uncertainty of a given probabilistic event. For the sentence similarity problem, “maximum uncertainty” means “desired relatedness”. Among sentences, the low values for 0 and 1 means precisely “undesirable relatedness” - sentences almost equal (p = 1) or sentences too dissimilar (p = 0) [18].

4) Word Co-occurrence: The chance of two terms from a text appear alongside each other in a certain order is called “word co-occurrence”. N-grams [19], which are contiguous sequence of n items from a given sequences of text or speech, were implemented here as measures of word co-occurrence. The more often terms appear in sequence in sentences, the higher the words co-occurrence score.

B. Semantic Similarity

Some techniques to find the similarity between sentences rely only on the words from the sentences (as presented in the previous section). Such approaches are not able to find the existing semantic relations between them, as “meaning” of the words ane not considered. On the other hand, the semantic similarity resort to relations such as synonyms, hyponym, and hypernym [9]. The following steps to make for semantic annotation of the sentences are used here:

- Sentences are represented as a word vector, in which only nouns are retained;
- the semantic similarity scores for each pair of word between two sentences is calculated;
- The results are combined by summing up the scores;
- The final scores are normalized yielding a value between 0 and 1.

It is relevant to remark that semantic similarity is only calculated if instances of both words appear in the WordNet, otherwise the score value for the pair is zero. In other words, this measure seeks the degree of similarity between two terms or words. Semantic similarity measures based on WordNet have been widely used in NLP applications [20]. These measures make use of the structure of WordNet to produce a numerical score that quantifies the degree of similarity between two concepts. In this work, the five different measures described below were used.

1) Resnik Measure: Resnik [21] introduced a measure that is based on information content, which is a value that indicates the specificity of a concept, augmenting the concepts in the *is-a* WordNet hierarchy. The measure of relatedness between two concepts is the information content of the most specific concept that both concepts have in common, i.e., their lowest common subsumer in the *is-a* hierarchy. Equations 3 and 4 shows how Resnik measures is calculated:

$$I(c) = -log(\frac{Pr(c)}{Pr(root)}) \quad (3)$$

$$Resnik(c1, c2) = I(LSO(c1, c2)) \quad (4)$$

where,
I(c) = Informativeness of the word c; [1].
Pr(c) = The probability that a concept appears in a large scale corpus. More abstract concepts have a higher value of probability than more specific ones;
Root = The root of the hierarchy to which *c* belongs; and
LSO(c1,c2) = The most specific ancestor in common between c1 and c2.

2) Lin Metric: The Lin metric extends Resnik, by taking the ratio of the shared information content to that of the individual concepts. The measure is calculated using Equation 5.

$$LIN(c1, c2) = (\frac{2 * Resnik(c1, c2)}{I(c1) - I(c2)}) \quad (5)$$

3) Wu and Palmer Metric: This Metric of similarity [22] relies on finding the most specific WordNet [2] concept that subsumes both the concepts being measured. The path length from this shared concept to the root of the ontology is scaled by the sum of the distances of the concepts to the subsuming concept.

4) Path Metric: This metric [23] computes the semantic relatedness of word senses by counting the number of nodes along the shortest path between the senses in the *is-a* hierarchies of WordNet. The path lengths include the end nodes. In other words, the longer the path length indicates less relatedness. The relatedness value returned is the multiplicative inverse of the path length (distance) between the two concepts: relatedness = 1 / distance. If the value of those two concepts are identical, then the distance between them is one; therefore, their relatedness is also 1.

5) Leacock and Chodorow Metric: This metric [24] counts up the number of edges between the senses in the *is-a* hierarchy of WordNet. The value is then scaled by the maximum depth of the WordNet *is-a* hierarchy. A relatedness value is obtained by taking the negative log of this scaled value. The formula 6 shows how this metric is calculated.

$$sim(c1, c2) = -log(dist(c1, c2)2 * maxDepth) \quad (6)$$

[1]The informativeness principle is an implicature in which one is licensed to apply additional knowledge of the world to infer an implicature that is informationally stronger than the actual utterance. For instance, the sentence "*He turned on the key and the engine started*" generally implicates the stronger utterance "*He turned on the key and this caused the engine to start*"

[2]http://wordnet.princeton.edu/

where,
dist(c1,c2) = Distance between c1 and c2 in WordNet hierarchy; and
maxDepth = maximum depth of the WordNet *is-a* hierarchy.

C. Coreference Resolution

Coreference resolution is the process by which we identify the noun that is referring to a same entity. There are three basic forms of coreference: *named*, *nominal* or *pronominal* [25]. For example, table I illustrates an example of mentions of the entity "Joe Smith" [26].

Table I
Coreferente Types

Name Mention:	Joe Smith, Mr. Smith
Nominal Mention:	the guy wearing a blue shirt
Pronoun Mentions:	he, him

The Stanford CoreNLP[3] tool was used here for indicating coreference in the experimentation corpus. When found, an edge in the graph with the relation is built.

D. Discourse Relations

Relations between sentences and parts in a text are represented by *discourse relations*. Reference [11] describes 16 classes of discourse relations in a text: Cause, Comparison, Condition, Contrast, Attribution, Background, Elaboration, Enablement, Evaluation, Explanation, Joint, Manner-Means, Topic-Comment, Summary, Temporal and Topic-Change.

The approach followed in this work was based on the results presented in reference [27], which presents a set of discourse structure relations based on contentful conjunctions. Some examples are presented in table II, which were used to generate discourse relation edges.

Table II
Discourse Relations [27]

Cause-effect	because; and so
Violated expectation	although; but; while
Condition	if; as long as; while
Similarity	and; (and) similarly
Contrast	by contrast; but
Temporal sequence	then; before; after; while
Attribution	according to; claim that; maintain that
Example	for example; for instance
Generalization	in general

The approach used here is interested in the relationship among sentences. Thus, whenever a discourse relation occurs in the beginning of a sentence, an edge to the previous sentence in the text is created.

[3]http://nlp.stanford.edu/software/corenlp.shtml

E. Preprocessing

The preprocessing services are divided into two different aspects, *structural* and *text analysis*. The first aspect consists of text splitters, whereas text analysis provides stop words removal, POS tagging, and lemmatization. In summary, the services provided are:

- **Paragraph Splitter:** It divides the text into paragraphs. It is implemented using regular expressions;
- **Sentence Splitter:** It divides the paragraphs into sentences. It is implemented using the Stanford CoreNLP;
- **Tokenization:** It divides the text into words. It is also implemented using regular expressions;
- **Stop Word Removal:** It removes words with little representative value for the document. Examples of such non representative words are articles and pronouns. It is implemented using the Lucene API[4];
- **POS tagging:** It takes as input a text in English and assigns to each token its morphological classification, such as noun, verb, adjective, etc.
- **Lemmatization:** This natural language subtask makes the mapping of verb forms such as the infinite tense and nouns into the singular form. Thus, the form of the word must be known. Also implemented using Stanford CoreNLP.

IV. Evaluation

This section reports on an experiment performed using the TextRank algorithm with the graph model presented in order to evaluate its performance in text summarization. The CNN Dataset was used and the result of the quantitative assessment is made using precision, recall, and f-measure.

The rest of this section describes:

- The TextRank algorithm;
- The dataset used;
- The methodology followed in the experiments to assess the quality of summaries;
- The results obtained.

A. TextRank

The TextRank graph [28] based ranking model for graphs extracted from texts written in natural language. It works in analogy to the PageRank algorithm [5] in ordering sentences in a document or in a collection of documents. It measures the importance of a sentence by counting the number of links to it.

The novelty in the work presented here is not in sentence ranking, but in proposing a model that creates links between sentences. The TextRank algorithm was employed to evaluate the proposed model.

[4]http://lucene.apache.org/core/

B. The CNN Dataset

The CNN corpus encompasses a dataset of news extracted from the news of CNN website[5]. It covers 400 texts with news articles assigned to 11 categories: Africa, Asia, business, Europe, Latin America, middle-east, sports, tech, travel, US, and world.

One important advantage of this dataset is its stylistic uniformity and cleanness, with well-written texts of general interest. Besides that, each text has its "highlights", a good quality and concise summary written by the original author of 3 or 4 sentences, that may be taken as gold standard. In addition, to the highlights a new evaluation set was created. Four of the authors of this paper selected for each text a number of sentences that were considered representative of the original text. Such number was equal to the number of sentences in the highlight plus one or two depending upon the size of the original text. Besides that, each of the authors in the selection experiment also made his "best mapping" between the sentences of the highlights and the original text itself. The chosen sentences were consensual in almost all the cases, with the exceptions being decided by the majority voting of the sentences in the set. Such set was also taken as a second gold-standard.

C. Evaluation Methodology

This section describes the methodology followed by the experiments to assess the quality of summaries.

1) Quantitative Analysis: The quantitative evaluation of the summaries generated by using the different scoring methods presented was performed with the ROUGE metrics. ROUGE (Recall-Oriented Understudy for Gisting Evaluation) [29] is a fully automated evaluator essentially widely used for such a purpose. It measures the content similarity between system-developed summaries and their corresponding gold summaries.

2) Qualitative Assessment: Four of the authors of this paper analyzed each original text in the CNN corpus and selected the sentences they considered ought to be included in its summary. The qualitative evaluation is done by counting the numbers of sentences selected by the system that match the human selected gold standard.

D. Results & Discussion

The experiments here follow the same model of the one presented in section IV. Table III presents the results to the selection of sentences in relation to each dimension and its many combinations.

Where,

- Similarity = Sim;
- Semantic Similarity = Sem;
- Coreference = Cor;
- Discourse Relation = Dis.

[5]www.cnn.com

Table III
Results of ROUGE having the sentences that match the highlights as gold standard

Method	Average_R	Average_P	Average_F
Sim	0,51(0,21)	0,38(0,14)	0,42(0,15)
Sem	0,36(0,22)	0,37(0,13)	0,35(0,16)
Sim + Cor	0,58(0,19)	0,39(0,13)	0,46(0,14)
Sem + Cor	0,54(0,23)	0,40(0,14)	0,44(0,16)
Sim + Dis	0,53(0,22)	0,40(0,14)	0,44(0,16)
Sem + Dis	0,36(0,22)	0,36(0,13)	0,34(0,16)
Sim + Cor + Dis	0,59(0,19)	0,39(0,13)	0,46(0,15)
Sem + Cor + Dis	0,53(0,23)	0,40(0,15)	0,44(0,17)
Sim + Sem + Cor + Dis	0,60(0,22)	0,40(0,13)	0,48(0,16)

We implement four and five services to similarity and semantic similarity metrics, respectively. However, in the experiments we use only one of them (that archive best result), there are: Entropy Measure to Similarity and Resnik Measure to Semantic Similarity.Figure 1 shows the results for the qualitative evaluation of the dimensions analyzed.

These results allow the following conclusions to be drawn:

- The use of traditional approach (similarity) achieved good results, but it could be further improved.
- Semantic similarity on its own does not yield good results.
- Coreference resolution improves the performance in all cases (with only syntactical similarity or adding semantic similarity).
- The discourse relation improves only the quantitative results of similarity.
- Discourse relation information usually reaches good results when it is used with coreference, once these metrics are related to text structure.
- The combination of similarity, coreference resolution and discourse relations generates the second best results, probably because similarity archives goods results when it is combined with structural aspects.
- The four dimensional model improves the TextRank algorithm both quantitatively (better precision, recall and f-measure) and qualitatively (improving 34,87% in relation to similarity model).

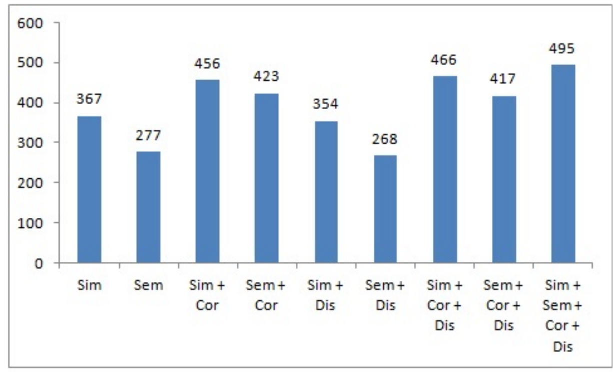

Figure 1. Number of Correct Sentences x Dimensions

V. Conclusions and Further Work

This paper proposes a new graph model which is general enough to be applicable in many text processing applications. Graph models show the relation between the sentences in a text, being particularly valuable for tasks of extractive and abstractive summarization, question-answering systems, information retrieval systems, among other text-related solutions.

The presented model uses the following four dimensions to create the graph:

- **Similarity**, which measures the sentences content overlaps.
- **Semantic similarity**, which employs ontologic conceptual relations such as synonyms, hyponym, and hypernym. To perform this step, sentences must first be represented as vectors with words and the semantic similarity scores for each pair of words using WordNet must be calculated.
- **Coreference resolution** links sentences that refer to the same subject. The experimental prototype developed here provides named, nominal, and pronominal coreference resolution.
- **Discourse relations** highlight the relevant relationships in the text.

Besides that, six preprocessing services are analyzed and implemented in the prototype in order to improve summarization performance: paragraph splitter, sentence splitter, word splitter, stop word removal, POS tagging, and lemmatization.

An experiment using the TextRank algorithm to evaluate the performance of the proposed model in context of extractive summarization as made. The CNN news dataset was used to perform both qualitative and quantitative evaluation. The results show that the proposed model outperforms related work both quantitatively and qualitatively. The TextRank algorithm using the four dimensional graph model achieves better precision, recall and f-measure (quantitative), and also displays better qualitative results (improving the traditional method in 34,87%).

As future work, the authors intend to: (i) Use other graph algorithms (for instance, Bushy Path and Aggregate Similarity [4]) to evaluate the proposed model; (ii) Apply the model presented to make abstractive summarizations; and (iii) Create clusters of sentences using the graph structure and study the possibility of selecting different perspectives of the summarized text.

References

[1] E. Lloret and M. Palomar, "Text summarisation in progress: a literature review," *Artif. Intell. Rev.*, vol. 37,

no. 1, pp. 1–41, Jan. 2012. [Online]. Available: http://dx.doi.org/10.1007/s10462-011-9216-z

[2] R. Mihalcea and P. Tarau, "TextRank: Bringing Order into Texts," in *Conference on Empirical Methods in Natural Language Processing*, Barcelona, Spain, 2004. [Online]. Available: http://acl.ldc.upenn.edu/acl2004/emnlp/pdf/Mihalcea.pdf

[3] O. Sornil and K. Gree-ut, "An Automatic Text Summarization Approach using Content-Based and Graph-Based Characteristics," in *Cybernetics and Intelligent Systems, 2006 IEEE Conference on*, 2006, pp. 1–6. [Online]. Available: http://dx.doi.org/10.1109/ICCIS.2006.252361

[4] M. A. Fattah and F. Ren, "Ga, mr, ffnn, pnn and gmm based models for automatic text summarization," *Comput. Speech Lang.*, vol. 23, no. 1, pp. 126–144, Jan. 2009. [Online]. Available: http://dx.doi.org/10.1016/j.csl.2008.04.002

[5] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web," in *Proceedings of the 7th International World Wide Web Conference*, Brisbane, Australia, 1998, pp. 161–172. [Online]. Available: citeseer.nj.nec.com/page98pagerank.html

[6] R. Mihalcea, "Graph-based ranking algorithms for sentence extraction applied to text summarization," in *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics, companion volume (ACL 2004)*, 2004. [Online]. Available: http://www.cs.unt.edu/~rada/papers.html

[7] H. Lin, J. Bilmes, and S. Xie, "Graph-based submodular selection for extractive summarization," in *Automatic Speech Recognition Understanding, 2009. ASRU 2009. IEEE Workshop on*, Dec. 2009, pp. 381–386. [Online]. Available: http://dx.doi.org/10.1109/ASRU.2009.5373486

[8] A. Hotho, A. Nurnberger, and G. Paas, "A brief survey of text mining," *LDV Forum - GLDV Journal for Computational Linguistics and Language Technology*, vol. 20, no. 1, pp. 19–62, May 2005.

[9] G. Varelas, E. Voutsakis, P. Raftopoulou, E. G. Petrakis, and E. E. Milios, "Semantic similarity methods in wordnet and their application to information retrieval on the web," in *Proceedings of the 7th annual ACM international workshop on Web information and data management*, ser. WIDM '05. New York, NY, USA: ACM, 2005, pp. 10–16. [Online]. Available: http://doi.acm.org/10.1145/1097047.1097051

[10] K. Raghunathan, H. Lee, S. Rangarajan, N. Chambers, M. Surdeanu, D. Jurafsky, and C. Manning, "A multi-pass sieve for coreference resolution," in *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, ser. EMNLP '10. Stroudsburg, PA, USA: Association for Computational Linguistics, 2010, pp. 492–501. [Online]. Available: http://dl.acm.org/citation.cfm?id=1870658.1870706

[11] A. Louis, A. Joshi, and A. Nenkova, "Discourse indicators for content selection in summarization," in *Proceedings of the 11th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, ser. SIGDIAL '10. Stroudsburg, PA, USA: Association for Computational Linguistics, 2010, pp. 147–156. [Online]. Available: http://dl.acm.org/citation.cfm?id=1944506.1944533

[12] G. Erkan and D. R. Radev, "Lexrank: graph-based lexical centrality as salience in text summarization," *J. Artif. Int. Res.*, vol. 22, no. 1, pp. 457–479, Dec. 2004. [Online]. Available: http://dl.acm.org/citation.cfm?id=1622487.1622501

[13] L. Antiqueira and M. G. V. Nunes, "Complex networks and extractive summarization," in *MSc Dissertation Contest of PROPOR'2010*, 2010. [Online]. Available: https://sites.google.com/site/lantiq/Antiqueira2010.pdf?attredirects=0&d=1

[14] S. S. Ge, Z. Zhang, and H. He, "Weighted graph model based sentence clustering and ranking for document summarization," in *Interaction Sciences (ICIS), 2011 4th International Conference on*, Aug. 2011, pp. 90–95.

[15] M. Benbrahim and K. Ahmad, "Computer-aided lexical cohesion analysis and text abridgment," University of Surrey, Computing Sciences Report CS-94-11, 1994, 60 p.

[16] A. J. Lee, R. Prasad and B. Webber., "Departures from tree structures in discourse: shared arguments in the penn discourse treebank," in *Proceedings of the Constraints in Discourse Workshop*, 2008.

[17] A. Abuobieda, N. Salim, A. Albaham, A. Osman, and Y. Kumar, "Text summarization features selection method using pseudo genetic-based model," in *Information Retrieval Knowledge Management (CAMP), 2012 International Conference on*, march 2012, pp. 193 –197.

[18] J. Cordeiro, G. Dias, and P. Brazdil, "New functions for unsupervised asymmetrical paraphrase detection," *JSW*, vol. 2, no. 4, pp. 12–23, 2007.

[19] J. B. Mariòo, R. E.Banchs, J. M. Crego, A. Gispert, P. Lambert, J. A. R. Fonollosa, and M. R. Costa-jussà, "N-gram-based machine translation," *Comput. Linguist.*, vol. 32, no. 4, pp. 527–549, Dec. 2006. [Online]. Available: http://dx.doi.org/10.1162/coli.2006.32.4.527

[20] T. Pedersen, "Information content measures of semantic similarity perform better without sense-tagged text," in *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, ser. HLT '10. Stroudsburg, PA, USA: Association for Computational Linguistics, 2010, pp. 329–332. [Online]. Available: http://dl.acm.org/citation.cfm?id=1857999.1858046

[21] P. Resnik, "Using information content to evaluate semantic similarity in a taxonomy," in *Proceedings of the 14th international joint conference on Artificial intelligence - Volume 1*, ser. IJCAI'95. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1995, pp. 448–453. [Online]. Available: http://dl.acm.org/citation.cfm?id=1625855.1625914

[22] Z. Wu and M. Palmer, "Verbs semantics and lexical selection," in *Proceedings of the 32nd annual meeting on Association for Computational Linguistics*, ser. ACL '94. Stroudsburg, PA, USA: Association for Computational Linguistics, 1994, pp. 133–138. [Online]. Available: http://dx.doi.org/10.3115/981732.981751

[23] S. Wubben and A. van den Bosch, "A semantic relatedness metric based on free link structure," in *Proceedings of the Eighth International Conference on Computational Semantics*, ser. IWCS-8 '09. Stroudsburg, PA, USA: Association for Computational Linguistics, 2009, pp. 355–358. [Online]. Available: http://dl.acm.org/citation.cfm?id=1693756.1693803

[24] C. Leacock, G. A. Miller, and M. Chodorow, "Using corpus statistics and wordnet relations for sense identification," *Comput. Linguist.*, vol. 24, no. 1, pp. 147–165, Mar. 1998. [Online]. Available: http://dl.acm.org/citation.cfm?id=972719.972726

[25] X. Luo, "Coreference or not: A twin model for coreference resolution," in *HLT-NAACL*, 2007, pp. 73–80.

[26] J. H. Clark and J. P. González-brenes, "Coreference resolution: Current trends and future directions," 2008.

[27] F. Wolf and E. Gibson, "Representing discourse coherence: A corpus-based study," *Comput. Linguist.*, vol. 31, no. 2, pp. 249–288, Jun. 2005. [Online]. Available: http://dx.doi.org/10.1162/0891201054223977

[28] R. Mihalcea and P. Tarau, "TextRank: Bringing order into texts," in *Proceedings of EMNLP-04and the 2004 Conference on Empirical Methods in Natural Language Processing*, July 2004.

[29] C.-Y. Lin, "Rouge: A package for automatic evaluation of summaries," in *Text Summarization Branches Out: Proceedings of the ACL-04 Workshop*, M.-F. Moens and S. Szpakowicz, Eds. Barcelona, Spain: Association for Computational Linguistics, July 2004, pp. 74–81.