AN OVERVIEW OF EXTRACTIVE BASED AUTOMATIC TEXT SUMMARIZATION SYSTEMS

Kanitha.D.K [1] And D. Muhammad Noorul Mubarak[2]

[1]Compuational Linguistics, Department of Linguistics, University of Kerala, Kariavattom, Thiruvananthapuram.
[2]Department of Computer Science, University of Kerala, Kariavattom, Thiruvananthapuram,

ABSTRACT:

The availability of online information shows a need of efficient text summarization system. The text summarization system follows extractive and abstractive methods. In extractive summarization, the important sentences are selected from the original text on the basis of sentence ranking methods. The Abstractive summarization system understands the main concept of texts and predicts the overall idea about the topic. This paper mainly concentrated the survey of existing extractive text summarization models. Numerous algorithms are studied and their evaluations are explained. The main purpose is to observe the peculiarities of existing extractive summarization models and to find a good approach that helps to build a new text summarization system.

KEYWORDS:

Text summarization, Abstractive summarization, Extractive summarization, Statistical methods, Latent semantic analysis.

1. INTRODUCTION

Large number of text materials is available on internet in any topic. The user searches a number of web pages to find out the relevant information. It takes time and effort to the user. An efficient summarizer generate summary of document within a limited time. Mani and Maybury (1999) defined an automatic text summarization as the process of distilling the most important information from a source (or sources) to produce an abridged version for a particular user (or users) and task (or tasks) [26].

Text summarization methods can be classified into abstractive and extractive summarization (Hahn.U, and Mani.I. 2000) [15]. In abstractive summarization Natural language generation techniques are used for summarization. It understands the original document and retells it in few words same as human summarization. The extractive summarization method select the important sentences, paragraphs etc from the original document and concatenate into shorter form. The sentences are extracted on the basis of statistical, heuristic and linguistic methods. Most of the text summarization systems used extractive summarization method based on statistical and algebraic methods which generate an accurate summary in large datasets and give overall opinion about the document. Abstractive summarization approaches are more complex than extractive summarization.

This paper primarily aims to examine the efficiency of summarization methods. This paper is

DOI:10.5121/ijcsit.2016.8503

organized as follows. Section 1 describes a brief introduction about text summarization techniques. Section 2 describes the existing models that focusing on extractive techniques. Section 3 discusses the advantages and disadvantages of each method. Section 4 describes some of the standards for evaluating summaries automatically and Section 5 concludes the paper.

2. EARLY WORK ON TEXT SUMMARIZATION

The text summarization systems started in early 1950's. Most of the early work on text summarization focused on single document summarization in technical articles. Due to lack of powerful computers and technological developments, summarization systems consider some simple surface level features of sentences like word frequency, position, length of the sentence etc. In 1970's Artificial Intelligence technology was developed and most of the summarization systems depend on AI technology. The AI technology based summarization systems are domain dependent systems.

In 1980 some summarization systems are developed on the basis of cognitive science theory. In 1990 Information retrieval methods are used for domain independent summarization. The IR technique doesn't consider synonymy, and polysemy. In 1995 Machine learning techniques are developed and it is highly used in summarization systems. The machine learning algorithms are bayesian classifier, hidden Markov model, long linear model, neural network etc. Now the statistical and mathematical techniques are widely used for extractive text summarization. The technological developments and its advantages and disadvantages are explained in Table1.

Table 1: Technological Developments in Text Summarization

Year	Methods	Advantages	Disadvantages	Models
1958	Simple surface level features of sentences.	The sentences which include most frequent words are selected as summary sentences.	Duplication in summary sentences.	Luhn,1958[25]; Edmundson,1969[11] etc.
1970	Artificial Intelligence.	Frames or templates are used to identify the conceptual relation of entities and extract the relation between entities by an assumption.	Only limited frames or templates may lead to incomplete analysis of conceptual entities.	Azzam, Humphreys, and Gaizauskas, 1999[2]; DeJong, 1979[10]; Graesser, 1981; McKeown and Radev, 1995[27]; Schank and Abelson, 1977[32]; Young and Hayes, 1985[35] etc.
1980	Cognitive science theories	The system can overcome the redundancy in some extent. Extract the representative sentences from the source text.	Complex task and limited to specified area.	Rinehart, S. D., Stahl, S. A., Erikson, L.G. 1986[31]; Jones, R. C. 2006[19]; Johnston, P. H. 1983[18]etc.
1990	Information retrieval techniques	Generate significant sentence from source text same as information retrieval techniques.	Doesn't consider the semantic aspects such as synonymy and polysemy	Aone, Okurowski, Gorlinsky, & Larsen, 1997[1]; Goldstein, Kantrowitz, Mittal, & Carbonell, 1999[13]; Hovy & Lin, 1997 [16]etc.
1995	Machine	Different machine	Computationally	Kupiec. J, Pedersen. J

	learning techniques	learning algorithms are used and provides more generalized summary.	complex and lack of semantic analysis of source text.	and Chen. F. (1995)[21]; Conroy, J. M. & O'Leary, D. P. 2001[9]; Osborne, M. 2002[28] etc.
1997	Statistical and Algebraic methods	Depended on some heuristics, linguistics and mathematical techniques. Easy to implement.	Without any syntactic analysis of the source text.	Gong and Liu, 2001[14]; Steinberger, J. and Jezek, K. 2004[29] etc.

2.1 Some Models In Extractive Text Summarization

2.1.1 Luhn Method (1958)[25]

Luhn created the first automatic text summarizer for summarize technical articles. The author ranked each sentence in the document on the basis of word frequency and phrase frequency. After performing the stemming and stop word removal, then calculates the word frequency. He stated that the word frequency shows a useful measure for significant factor of a sentence. All sentences are ranked on the basis of significant factor and get top rank sentences. The top ranked sentences are selected as summary sentences.

2.1.2 Boxendale Model (1958)[6]

Boxendale proposed a position method for sentence extraction. He argued that some significant sentences are placed in some fixed positions. The author checked 200 paragraphs in newspaper articles and 85% of the paragraphs, the topic sentence come first and 7% come last. So he stated that in newspaper articles the first sentence in each paragraph got high chance to include in summary. In 1997 Lin and Hovy claimed that Baxendale position method is not a suitable method for sentence extraction in different domains. Because the discourse structure of a sentence varies from different domains.

2.1.3 Edmundson Method (1969)[11]

Edmundson developed a new method in automatic summarization. This method computes the candidate sentence by adding some features of sentences such as keywords, cue phrases, title plus heading and sub heading words and sentence location. This sentence scoring parameters are used to extract the top ranked sentences. The stop words are removed from the source document. The sentences include cue words like conclusion, according to the study etc gets high score. This method also gives high score to title word, heading and sub-heading words which are included in the sentences. Through location feature, conclusion sentences in technical documents and the first and last sentences in the newspaper articles gets high score. The score of each sentence is computed as follows:

$$S_i = w_1*C_i+w_2*K_i+w_3*T_i+w_4*L_i \ldots\ldots\ldots\ldots\ldots \quad (1)$$

Where *Si* is the score of *sentence i. Ci, Ki and Ti* are the scores of sentence *i* based on the number of cue phrases, keywords and title words. *Li* is the score of location in the document. w*1, w2, w3 and w4* are the weights for linear combination of the four scores.

2.1.4 TRAINABLE DOCUMENT SUMMARIZER (KUPIEC. J, PEDERSEN. J AND CHEN. F. ,1995) [21].

Trainable Document Summarizer executes sentence extraction on the basis of some sentence weighting methods. The important methods used in this summarizer are:

- Sentence length cutoff feature - sentences containing less than a pre-specified number of words are excluded by sentence length cutoff feature.
- Cue words and phrases related sentences are included
- The first sentence in each paragraph is included
- Thematic words -The most frequent words are included.

Thus the sentences are ranked on the basis of the above features and high scored sentences are selected as summary sentences.

2.1.5 ANES (BRANDOW, MITZE AND RAU 1995)[8]

ANES text extraction system is a domain-independent summary system for summarize news articles. The process of summary generation has four major elements such as:

a. Calculation of the *tf*idf* weights for all terms.
b. Terms with a high *tf*idf* weight plus headline-words.
c. Summing over all signature word weights plus the relative location score.
d. Select the high scored sentences as summary sentences.

2.1.6 BARZILAY & ELAHADAD SYSTEM, 1997[4]

Barzilay & Elahadad**,** develop a summarizer based on lexical chain method. The sentences are extracted by the collection of the similar words which form a lexical chain. The concept of lexical chain was introduced in Morris and Hirst, 1991. The lexical chain links the semantically related terms with the different parts of source document. Barzilay and Elhadad used a wordnet to construct the lexical chains.

2.1.7 BOGURAEV, BRANIMIR & KENNEDY (BOGURAEV, BRANIMIR AND CHRISTOPHER KENNEDY, 1997)[7]

The authors develop a single document and domain independent system. The linguistic techniques are used to identify the main topic. The sentences are selected on the basis of noun phrases, title word and topic related sentences.

2.1.8 FOCISUM (KAN, MIN-YEN AND KATHLEEN MCKEOWN (1999))[20]

The summarization system follows a question answering approach. It is a two stage system, first takes a question then summarizes the source text then gives answer to the question. The system first uses a named entity extractor to find the important term of the document. The system also follows existing information extraction features of sentence like word frequency and type of terms. The result is a concatenation of sentence fragments and phrases found in the original document.

2.1.9 SUMMARIST (HOVY AND LIN 1999)[16]

Lin and Hovy, 1997 studied the importance of sentence position method proposed by Baxendale, 1958. In 1999, Lin and Hovy develop a machine learning model for summarization using decision trees instead of a naive Bayes classifier. Summarist system produces summaries of the web documents. The system provides abstractive and extractive based summaries. Summarist first identifies the main topics of the document using the chain of lexically connected sentences. Wordnet and dictionaries are used for identify the lexically connected sentences. The statistical techniques such as position, cue phrases, numerical data, proper name, word frequency etc are used for extractive summary.

2.1.10 MULTIGEN (BARZILAY, MCKEOWN AND ELHADAD, 1999)[4]

MultiGen is a multi document summarization system. The system identified similarities and differences across the documents by applying the statistical techniques. It extracted high weight sentences that represent key portion of information in the set of related documents. This is done by apply the machine learning algorithm to group paragraph sized chunks of text in related topics. Sentences from these clusters are parsed and the resultant trees are merged together to form the logical representations of the commonly occurring concepts. Matching concepts are selected on the basis of the linguistic knowledge such as stemming, part-of-speech, synonymy and verb classes.

2.1.11 CUT AND PASTE SYSTEM (JING, HONGYAN AND KATHLEEN MCKEOWN. 2000)[20]

The Cut and Paste system designed to understand the key concepts of the sentences. These key concepts are then combined to form new sentences. The system first copies the surface form of these key concepts and pasted them into the summary sentences. The key concepts are achieved by probabilities learnt from a training corpus and lexical links.

2.1.12 CONROY ET AL. (CONROY, J. M. & O'LEARY, D. P., 2001)[9]

The work presented by Conroy, J. M. & O'Leary, D. P., considered the probability of inclusion of a sentence in summary depends on whether the previous sentence is related next sentence based on HMM (Hidden Markov Model).The sentences are classified into two states such as summary sentences and non summary sentences. The lexically connected sentences are selected into summary sentences.

2.1.13 SWESUM(HERCULES DALIANIS., 2000)[34]

SweSum create summaries from Swedish or English texts either the newspaper or academic domains. Sentences are extracted according to weighted word level features of sentences. It uses statistical, linguistic and heuristic methods to generate summary. The methods are Baseline, First sentence, Title, Word frequency, Position score, Sentence length, Proper names and Numerical data etc. The processed text is newspaper articles so the first sentence in the paragraphs got high score. The formula is, $1/n$, where n is the line number, this method is called Baseline. It built a combination of function on above parameters and extracts the required summary sentences.

2.1.14 MEAD (RADEV, H. Y. JING, M. STYS AND D. TAM, 2001)[30]

MEAD computed the score of a sentence on the basis of a centroid score. The centroid score is formed on the basis of *tf-idf* values, similarity to the first sentence of the document, position of the sentence in the document, sentence length etc. The highest ranked sentences are selected as summary sentences. This summarizer produced single and multi document summaries.

2.1.15 WEBINESSENCE (RADEV, 2001)[30]

This system is an improved version MEAD summarizer. It is a web based summarizer for web pages. The architecture of the system includes two stages. The first stage the system collects URLs from the different web pages and extracts the news articles in same event. The second stage clusters the data from different documents. A centroid algorithm is used for find the representative sentences. Avoid repetition and generate a final summary.

2.1.16 TEXT SUMMARIZATION USING TERM WEIGHTS (R.C.BALABANTARY, D.K.SAHOO, B.SAHOO, M.SWAIN. 2012)[5]

The authors developed a statistical approach to summarize the source text. The sentences are split into tokens and remove the stop words. After remove the stop words then a weight value is assigned to each individual term. The weight is calculated on the basis of frequency of a term in the sentence divided by frequency of term in the document. Then add a additional score to the weight of terms which are appear in bold, italic, underlined or any combination of these. Then rank the individual sentence according to their weight value that is calculated as weight of individual term divided by total number of terms in that sentence. Finally, extract the higher ranked sentences include the first sentence of the first paragraph of the input text to generate summary.

2.1.17 LSA FOR DOCUMENT SUMMARIZATION [22]

LSA is a technique for extracting the hidden semantic representation of terms, sentences, or documents (Landauer & Dumais, 1997). It is an unsupervised method for extract the semantics of terms by examines the co-occurrence of words. The first step of this approach is the representation of input documents as a word by sentence matrix A. Each row represents a word from the document and each column represents a sentence in the document. So A=mXn matrix that means ‘m’ words and ‘n’ sentences. The Singular Value Decomposition (SVD) from linear algebra is applied to matrix A. The SVD of mXn matrix is defined as A=U∑VT. Matrix U is an mXn matrix of real numbers. Matrix ∑ is diagonal nXn matrix. The V^T matrix is nXn matrix each row represented as sentences. Gong and Liu (2001) [14]proposed a method of LSA for document summarization to recognize the important topics in the document without the use of wordnet. They consider each rows of matrix V^T and select the sentences with the highest value. Steinberger and Jezek (2004)[33] proposed an improved method for document summarization. Murray, Renals and Carletta (2005) proposed an approach for summarizing meeting recordings using LSA. Text summarization using a trainable summarizer and latent semantic analysis are proposed by Yeh, Ke, Yang and Meng (2005). This approach sentence ranking depends on graph based method and LSA based method.

2.1.18 POURVALI AND ABADEH MOHAMMAD (2012) [29]

The authors approach was based on lexical chains method and the exact meaning of each word in the text is determined by using WordNet and Wikipedia. The score of sentence is determined by the number and type of relation in the chains. The sentences that got highest chains are selected as final summary sentences.

2.1.19 S.T. KHUSHBOO, R.V.D.DHARASKAR AND M.B.CHANDAK (2010)[13]

They proposed a method based on graph based algorithms for text summarization. This method constructs a graph from the source text. The nodes are represented as sentences and the edges are represents the semantic relation between sentences. The weight of each node is calculated and the highest ranking sentences are selected for final summary.

2.1.20 DISCOURSE BASED SUMMARIZER (LI CHENGCHENG, 2010)[24]

The author proposed a summarizer depend on rhetorical structure theory. This technique based on analysis of discourse structure of sentence. The sentence score is calculated on its relevance factor. The relevant sentences got the highest weight and irrelevant sentences got low weight.

3. COMPARISON OF EXTRACTIVE TEXT SUMMARIZATION MODELS

No.	Models	Criteria for sentence selection	Type of document	Level of processing	Corpus	Advantages	Disadvantages
1.	Luhn method (1958)	Word frequency and phrase frequency.	Single	Surface	Technical articles.	The highest word frequency sentences are selected to summary sentences.	Duplication in summary.
2.	Baxendale method (1958)	Position method.	Single	Surface	Technical documents.	It is used in the system where machine learning systems are complex.	It is related to the discourse structure of sentence. The discourse structure of sentence varies from different domain.
3.	Edmunson method (1969)	Word frequency, cue phrases, title and heading words, sentence location.	Single	Surface	Technical documents.	Foundation for many existing extractive summarization method.	Redundancy in the summary and computationally complex.
4.	Trainable Document Summarizer (1995)	Machine learning techniques.	Single	Surface	Technical documents.	It provides a universal summary.	Machine learning techniques are computationally complex.

5.	ANES (1995)	Tf*idf	Single	Surface	Domain independent	The main topic related sentences are included in summary sentences.	The summarizer doesn't handle various sub topics.
6.	Barzilay & Elahadad (1997)	Lexical chain method	Single	Entity		Consider the semantic relationship among sentences and provides representative summary.	It requires deep syntactic and semantic structure of a sentence.
7.	Bogurаev & Kennedy (1997)	Noun phrases, Title related terms	Single	Entity		Extract the sentences in same context.	Requires linguistic knowledge.
8.	Focisum (1998)	Named entity recognition and information extraction techniques.	Single	Entity	News articles.	Extract the information same way as a question answering system. The number of word co-occur in the questions are extracted as summary.	Requires a question generator for information extraction. The summary is the result of question.
9.	Summarist (1999)	Statistical and linguistical	Single and multi document	Surface	Web documents.	Extract the representative sentences as summary sentences.	Computationally complex method.
10.	MultiGen (1999)	Syntactic analysis	Multi document	Entity	News articles from different web pages.	Generate multi document summaries.	Require the language processing tools.
11.	Cut and Paste System (1999)	Statistical	Single	Surface		Generate cohesive summary.	Complex method
12.	SweSum (Hercules Dalianis, 2000)	Statistical, linguistic methods	Single	Surface	News article	Generate the representative summary.	Restricted to some specific domain.
13.	Conroy, J. M. & O'Leary, D. P. 2001	HMM	Single	Surface	News article	Lexically related sentences.	Difficult to compute
14.	MEAD (Radev, H. Y.	Cluster based	Single and multi	Surface	News article	Summary from single and multiple	Duplication in summary.

	Jing, M. Stys and D. Tam, 2001)		document			documents.	
15.	WebInEssence (Radev, 2001)	Cluster based	Single and multi document	Surface	News articles	Summarize news articles in different web pages.	
16.	Discourse based, 2010	Statistical and linguistic method.		Discourse	News articles	Linguistic analysis of source text.	Compute all the rhetorical relation between sentences is difficult.
17.	Graph based,2010	Statistical	Single and multi document	Surface		Graph based method to form final summary.	Computationally complex.
18.	Text Summarization using Term Weights (R.C.Balabantary, D.K.Sahoo, B.Sahoo, M.Swain. 2012)	Statistical based		Surface		Extract more relevant sentences.	It generally depends on format of the text.
19.	Pourvali, 2012	Statistical		Surface		Generate semantic based summary.	Requires language processing tools.
20.	LSA based summarization	Statistical and algebraic method	Single and multiple	Surface/ Entity	News articles, technical documents, books etc.	Semantically related sentences and easy to implement.	Non availability of syntactic analysis and world knowledge.

4. EVALUATION OF SUMMARIZATION SYSTEMS

Evaluation of summaries is an important aspect of text summarization. A general policy to evaluate the quality of a summarization system is absent in existing models. The authors provide different approaches for summary evaluation. In some systems the quality of a summary is determined by grammatically and its relevancy to the user. If the summary is satisfactory then the system summary meets the needs of a user.

Mainly the evaluation method can be classified as intrinsic and extrinsic methods. The intrinsic methods evaluate the quality of summary on the basis of manual summary. The extrinsic evaluation evaluates how the summary affects the other task. Most of the summarization system follows combination of methods to evaluate the quality of summary. Precision and Recall measures are used by the most of the extractive based summarization systems. Most of the systems evaluate the quality of summary on the basis of manual summary. Comparing manual summaries with system summaries are not appropriate. Because the human select the different sentence in different times same way the different authors choose different sentences as summary sentences. Recently some system follows SEE, ROUGE, BE methods for summary evaluation (Lin,C.Y., Hovy, E. 2003)[23].

5. CONCLUSION

This paper examines the efficiency and accuracy of existing summarization systems. The summarizer systems in earlier stage mainly concentrate some simple statistical features of sentences and summarize only the technical articles. For a generic summarization these systems are not produce the satisfactory result. The above extractive summarization systems follows statistical, linguistic and heuristics methods. The statistical methods are tf method, tf-idf method, graph based, machine learning, lexical based, discourse based, cluster based, vector based, LSA based etc. The statistical methods follow supervised and unsupervised learning algorithms. The machine learning algorithm related models are generates coherent and cohesive summary but the algorithms are computationally complex and needs large storage capacity. These algorithms are overcome the redundancy in some extent and the systems are domain independent. Some systems follow the statistical and linguistic based methods. It also generate good summary but the linguistic analysis of source document required heavy machinery for language processing. The lexical based method requires semantic dictionaries and thesaurus. The discoursed based methods analyze the rhetorical structure of documents. Complete analysis of source document is very difficult. At the same time the statistical and algebraic method LSA extract the semantically related sentences without the use of wordnet and online dictionaries. The systems provide a domain independent generic summary rather than a query based summary. The LSA based systems summarize the large datasets within the limited time and produce satisfactory result.

REFERENCES

[1] Aone, C., Okurowski, M. E., Gorlinsky, J., & Larsen, B. (1997). A scalable summarization system using robust NLP. In Proceedings of the ACL’97/EACL’97 workshop on intelligent scalable text summarization (pp. 10–17), Madrid, Spain.

[2] Azzam, S., Humphreys, K. R., Gaizauskas. (1999). Using co- reference chains for text summarization. In Proceedings of the ACL’99 workshop on co-reference and its applications (pp. 77–84), College Park, MD, USA.

[3] Baldwin., Breck., & Thomas.S.Morton. (1998). Dynamic coreference-based summarization. In Proceedings of the Third Conference on Empirical Methods in Natural Language Processing, Granada, Spain,June.

[4] Barzilay, R., & Elhadad, M. (1997). Using lexical chains for text summarization. In Proceedings ISTS'97.

[5] Balabantary.R.C., Sahoo.D.K., Sahoo.B., & Swain.M (2012). “Text summarization using term weights”, International Journal of computer applications, Volume 38-No.1

[6] Baxendale, P. B. (1958). ‘Machine-made index for technical literature: an experiment’. IBM Journal, 354–361.

[7] Boguraev., Branimir., & Christopher Kennedy. (1997). Sailence-based content characterisation of text documents. In proceedings of ACL'97 Workshop on Intelligent, Scalable Text Summarization,Pages 2-9,Madrid, Spain.

[8] Brandow., Ronald., Karl Mitz., & Lisa. F. Rau. (1995). Automatic condensation of electronic publications by sentence selection. Information Processing and Management, 31(5):657-688.

[9] Conroy, J. M., & O'Leary, D. P. (2001). Text summarization via hidden Markov models and pivoted QR matrix decomposition. Tech. Rep., University of Maryland, College Park.

[10] DeJong, G.F. (1979). Skimming stories in real time: an experiment in integrated understanding. Doctoral Dissertation. Computer Science Department, Yale University.

[11] Edmundson, H.P. (1969). New Methods in Automatic Extracting, Journal of the ACM, 16(2):264-285.

[12] Eduard Hovy., (2003) The Oxford Handbook of Computational Linguistics, Oxford University Press, Oxford, chapter 32.

[13] Goldstein, J., Kantrowitz, M., Mittal, V., & Carbonell, J. (1999). Summarizing text documents: sentence selection and evaluation metrics. In Proceedings of the 22nd annual international ACM SIGIR conference on research and development in information retrieval (SIGIR'99) (pp. 121–128), Berkeley, CA, USA.

[14] Gong.Y., & Liu. X (2001) Generic text summarization using relevance measure and latent semantic analysis. In Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pages 19–25.Hahn, U., & Mani, I. (2000). The challenges of automatic summarization.

[15] Hahn, U., & Mani. I. (2000). The challenges of automatic summarization.

[16] Hovy, E., & Lin, C.Y. (1997). Automatic text summarization in SUMMARIST. In Proceedings of the ACL'97/EACL'97 workshop on intelligent scalable text summarization (pp. 18–24), Madrid, Spain.

[17] Jen-Yuan Yeh, Hao-Ren Ke, Wei-Pang Yang, & I-Heng Meng. (2005). Text summarization using a trainable summarizer and Latent Semantic Analysis.

[18] Johnston, P. H. (1983). Reading Comprehension Assessment: A Cognitive Basis. Newark, Delaware: IRA.

[19] Jones, R. C. (2006). Reading quest: Summarizing strategies for reading comprehension. TESOL Quarterly, 8, 48-69.

[20] Kan., Min-Yen., & Kathleen McKeown. (1999). Information extraction and summarization: Domain independence through focus types. Technical report, Computer Science Department, Columbia University, New York.

[21] Kupiec, J., Pedersen, J., & Chen, F. (1995). A trainable document summarizer. In Proceedings of the 18th annual international ACMSIGIR conference on research and development in information retrieval (SIGIR'95) (pp. 68–73), Seattle, WA, USA.

[22] Landauer, T., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge, Psychological Review, 104: 211-240.

[23] Lin.C.Y, & Hovy.E, "Automatic Evaluation of Summaries Using N-gram Concurrence Statistics", in 2003 Language Technology Conference, Edmonton, Canada, 2003.

[24] Li Chengcheng, .(2010)."Automatic text summarization based on rhetorical structure theory", Computer application system modeling, 2010 International Conference.

[25] Luhn, H.P. (1958). The automatic creation of literature abstracts. IBM Journal of Research and Development, 2:159–165.

[26] Mani, I., & Maybury, M.T. (Eds.) (1999). Advances in automated text summarization. Cambridge, US: The MIT Press.

[27] McKeown, K., & Radev, D. R. (1995). Generating summaries of multiple news articles. In Proceedings of the 18th annual international ACM SIGIR conference on research and development in information retrieval (SIGIR'95) (pp. 74–82), Seattle, WA, USA.

[28] Osborne, M., (2002). Using maximum entropy for sentence extraction. In Proceedings of the Acl-02, Workshop on Automatic Summarization, Volume 4 (Philadelphia Pennsylvania), Annual Meeting of the ACL, Association for Computational Linguistics, Morristown.

[29] Pourvali,M., & Abadeh Mohammad,S.(2012). "Automated text summarization base on lexical chain and graph using of word net and wikipedia knowledge base", IJCSI International Journal of Computer Science, Issues No.3, Vol.9.

[30] Radev, H. Y. Jing, M. Stys & D. Tam. (2004) Centroid-based summarization of multiple documents. Information Processing and Management, 40: 919-938.

[31] Rinehart, S. D., Stahl, S. A., & Erikson, L.G. (1986). Some effects of summarization training on reading. Reading Research Quarterly, 21(4), 422-435.

[32] Schank, R., & Abelson, R. (1977). Scripts, Plans, Goals, and Understanding. Hillsdale, NJ: Lawrence Erlbaum Associates.

[33] Steinberger, J., and Jezek, K. (2004). Using latent semantic analysis in text summarization and summary evaluation. Proceedings of ISIM'04, pages 93-100.
[34] SweSum - A Text Summarizer for Swedish Hercules Dalianis, NADA-KTH, SE-100 44 Stockholm, Sweden.
[35] Young, S. R., & Hayes, P. J. (1985). Automatic classification and summarization of banking telexes. In Proceedings of the 2ndConference on Artificial Intelligence Applications (pp. 402–408).