Automatic Text Summarization and it's Methods - A Review

Neelima Bhatia
Amity School of Engineering and Technology
Amity University Noida, India
nbneelima@gmail.com

Arunima Jaiswal
Amity School of Engineering and Technology
Amity University Noida, India
ajaiswal34@amity.edu

Abstract— Text summarization is an incipient practice for verdict out the summary of the text article. Text summarization has grew so uses such as Due to the enormous aggregate of information getting augmented on internet; it is challenging for the user to verve through altogether the information accessible on web. The large availability of internet content partakes constrained a broad research area in the extent of automatic text summarization contained by the Natural Language Processing (NLP), especially statistical machine learning communal. Terminated the bygone half a century, the defaulting has been addressed from numerous diverse standpoints, in erratic domains and using innumerable archetypes. In this survey paper we investigate the popular and important work done in the field of single and multiple document summarizations, generous distinctive prominence towards pragmatic approaches and extractive techniques. Particular auspicious slants that quintessence on unambiguous minutiae of the summarization are also deliberated. Exceptional consideration is ardent to involuntary assessment of summarization classifications, as forthcoming investigation on summarization is sturdily reliant over evolvement in this problem space.

Keywords: Document summarization, Automatic text summarization, Systematic review, Survey, Natural Language Processing.

I. INTRODUCTION

Automatic Text Summarization, is the procedure in which, a text is reduced to create a summary from an original document. Reduction of text is a very complex problem which, in spite of the progress in the area thus far, poses many challenges to the scientific community. It is also a relevant application in today's information society given the exponential growth of textual information online and the need to promptly assess the contents of text collections. It has long been assumed that summarization presupposes to understand the input text, which means for identifying the important point of the document, explicit (semantic) representation of the text must be calculated therefore, text summarization became an interesting application to test the understanding capabilities of artificial systems.

There was a resurge of interest in summarization in the nineties with the organization of a number of relevant scientific events [1] and a peak of interest from the year 2000 with the development of evaluation programs such as the DUC (document understanding conferences) [2] & the TAC (text Analysis conferences) [3] in the United States (US). Two fundamental questions in text summarization are; **(i)** how to select the essential content of a document, and **(ii)** how to express the selected content in a condensed manner [4, 5].

Text summarization research has many times concentrated more on the product: the summary and less on the cognitive basis of text understanding and production that underlay human summarization. Some of the limitations of current systems would benefit from a better understanding of the cognitive basis of the task. Nevertheless, formalizing the content of open domain documents is still a research issue, so most systems are only based on collecting the summarized text only from the original document (not creating new definitions), Where the conversion of the original text into a summary is concerned, there are two main approaches of defining a summary in summarizing a text document: ***an extractive summary***, selecting and extracting sentences from the original document & ***an abstractive summary*** (i.e., an intellectual section of document), a short summary in which some of its substantial is absent within the input document [5]. The output summaries can correspondingly classified into indicative or informative, depending on their particular purpose intended, to alert or to inform respectively.

II. AUTOMATIC TEXT SUMMARIZATION

It is the method of filtering in which significant facts from a document to create a shortened variety for specific customers & tasks Maybury 1995 et at, [6].Since synopsis is essential, most significant factor to summarization is the level of compression (which is the ratio of length of the summary to length of the source) anticipated.

Borko & *Bernier* 1975 *et al,* [7] provide a dissimilarity between indicative summaries, in which it deliver to notifying the purpose for choosing the text for more reading, & informative summaries, in which it can take position instead of the source. These Summaries can correspondingly provide the evaluative function, Lancaster 1991 et al, [8], Sparck-Jones 1999 et al, [9]. Summaries of the text can be delivering to a reader's profits & skill, & gives a subject related summaries, otherwise they can be projected at a complete distribution, as this case to be called general abstracts. It's likewise worthwhile to differentiate among such summaries that are eliminates of original input source, & those which are summaries comprising an extra text created by the summarizer system. Summaries can extent to more than one document and called multi-document summarization. *Radev* in [10], *M. Bloedorn* in [11] here we are anxious mostly with single-document summaries. Lastly, here author create a English summaries of English text but according to requirement & demand it can translate or summarized.

The summaries that can be created from the original document, which are divided into dissimilar categories & this categories depending on their using schedule, for e.g.: movie

summaries, headlines, outlines, chronologies abridgments minutes, biographies, sound bites, etc. [12]. Consequently, the summarization scheme falls into at slightest one, often more than one, opening in every group that are shown above, & thus must too be evaluated along numerous magnitudes using the dissimilar events.

Types of Automatic Text Summarization Systems
There are 3 kinds of text summarization systems listed below:
- Abstract vs. Extract summary
- Single document vs. multi-document
- Generic vs. query-based

III. METHOD-WISE APPROACHES FOR AUTOMATIC TEXT SUMMARIZATION

A. Term Frequency Based Method

TF-IDF or term-frequency inverse-document-frequency model introduced by Salton [13] in 1989, in this method to find the ratio of the number of terms to the frequency of the quantity of the documents containing the term which defines the score of a term in the document. By calculating the related terms in the sentence, successively sentences can be recorded for illustration.

$$TF(t) = \frac{\text{Number of times term t appears in a document}}{\text{Total number of terms in the document}}$$

$$IDF(t) = \log_e \frac{\text{Total number of documents}}{\text{Number of documents with term t in it}}$$

Fig. 1. In Figure, weight of a term n in a document d is the product of its *frequency in the document* as term frequency (TF) & its *inverse document frequency* (IDF) as reflected in the figure.

Author $J. Fukumoto$ in [14] presented a multi-document text summarization method. In this method, the information are extracted from the multiple text using TF/IDF based sentebnce & also use the single document summarization for multi-document to create a summary of the text. His proposed scheme adaptively summarizes a document set into three types: **(a)** Single topic type, **(b)** multiple-topic type, & **(c)** other topic type. In first one, the additional document represent more information or following event of the first document, & so on for the succeeding documents. In the second type, documents define the similar occasion type such as a set of traffic coincidences. In the final step proposed in [14], documents is connected each other but not categorized into the first two types. For summarization, at first sentences are extracted from each document based on TF/IDF, sentence position and weighing using intention type, such as "*request*", "*obligation*" & "*necessary*" etc., of a sentence. In step two, for creating the reduced form of the single document, first remove the unrelated portion of the sentence & then arranged the important portion of the document in original order & send them for document set type arrangement.

Author $Y. Ouyang$ in [15] proposed a new approach, in it multi-document abstraction is capable to combine a variety of purposes. Human summarization concepts in which the man can search the important topic in a document set for representing the topic and also write something about the topic and man may search the sub topic and continue the search. For obtaining the hierarchical method, relate the all activities of human summarizer.

B. Graph Based Method

In a graph, $I. Mani$ et al in [16] represented the sentences through the set of document nodes that are created by taking the node as vertices & to the semantic relations amid the sentences as edge weights. Spreading activation technique is used here to find nodes that are related to the node for applying a certain algorithm. The Weight of nodes of activating node is an exponentially decaying function & is also the distance among nodes & the Weight of a neighbouring node is designed as the function of an associated weight & exciting the node weight. Successively the techniques for getting threshold number of output nodes are used to determine the neighbour of opening nodes & accumulate the activating nodes to the output until.

In graph-based ranking algorithms has been also shown to be active in TS. The text representation as a graph, the graph topology relates the information on the basis of texts which are the attachment of the numerous components. In a multi document summarization, all the applicant text which are shown in a graph or in the system which are involved in the condensed form of the text , above which a predefined threshold value of the two sentences are linked if the similarity occur in text representation in a graph which is proposed in LexRank by $Erkan$ in [17]. Different graph-based algorithms which is represented by $Mihalcea$ in [18] are analyzed with calculating their use of the automatically removal of the document in the situation of the TS. Additionally, for both generic & query-focused multi-document summarization. $Wan\ et\ al.$ in [19], proposed a technique which is based on the similarity of the graphs, in this techniques the more information are extracted and this can be done by taking the combined form of the text & then comparing the each pair of document and find the similarity of the document , and also distinguish the related information and eliminate the unrelated information.

Giannakopoulos et al. in 2008 proposed a method in which the related information is extracted from the original document using the character & word graph. [20] Whereas $Plaza\ et\ al.$ in [21] construct the graph using the ideas that are known with Word net [22] and is-a relationships, for each sentence in a document which are then used to create a graph representation. This method has been proven effectively in

different areas, such as image captions, biomedical documents, newswire etc.
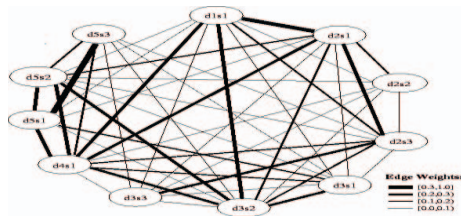

Fig. 2. Figure shows the graph based similarity index generated by Weighted cosine similarity diagram for the cluster [16] for automatic text summarization.

Using graph based method proposed by *Rada Mihalcea et al* in [23] projected an algorithm named Text Rank in the ground of natural language processing. In each sentence of the text the vertex is added for creating the graph. Different vertices are combined using to find the similarity in the related sentence & it is based on the text coinciding due to this coinciding a score is generated for each vertex. And this can be done by the arrangement of vertices according to their score for applying the iteration method & then highest scorer text are selected to create an abstract.

C. Time Based Method

On comparable occurrence using empirical analysis *McKeown et al* in [24] presented a scheme in which condensed form of the document is the series of summary objects. Summaries length alters that depend upon on the basis of the accessible possessions of text. As under the ARPA human language technology program author projected a method called *SUMMONS* that use the templates to précis the overall text input that are designed by the message understanding schemes developed [25]. All the research is based on the method that how the procedure change time to time using sequences of different event or using the numerous way in the same event. In *SUMMONS* method, which is defined as the group of templates where templates show the extracted information from the source with the help of message understanding scheme   It first groups messages together , categorises commonalities among them, & notes how the discourse effects expressing by setting realization flags.
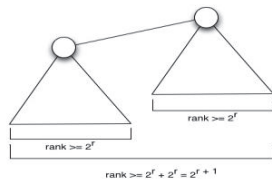

Fig. 3. Primarily, when a node is a root of its own tree, it is trivially true.

For multi document summarization, *X. Wan* et al in [26] projected a timed text rank algorithm which is an improvement of the graph based method called text rank & combined with a new sequential dimension. Because of novel

information is added in the document for evolving, due to this the current information is more important as comparison to the previous document

Fig. 4. Text Rank generation for multi-document text summarization and assimilated a different temporal dimension.

D. Clustering (separation & merging) Based Method

J. Goldstein et al in [29] offered a method for the text extraction method. In this method the multi-document summarization creates on the single-document text summarization approaches using the extra information in the relation between the documents. For multi document summarization four minimum requirements are recognized. **(a)** *a clustering-* to find associated information which is able to cluster alike documents & the passages, **(b)** *coverage the capacity* to determine & extract the main points across the texts, **(c)** *anti redundancy-* among passages the ability to reduce redundancy in the shortened form of the texts, **(d)** *a summary cohesion criteria-* for reader comfort, capability to mix the text. The proposed method underlined on "relevant novelty" which is a metric for reducing redundancy & exploiting both significance & assortment. The technique works as follows: **(i)** Cluster the documents into the passages, & the index them, **(ii)** Classify passages appropriate to question using cosine likeness with the threshold below which passages are forbidden, **(iii)** Using "*relevant novelty*" metric, dependent on the wanted size of summary, elect a amount of passages, **(iv)** Bear a resemblance to the selected passages into a summary text.

In numerous topics, automatic summarization systems use clusters to produce an important summary. The documents are signified using term-frequency inverse-document-frequency method. Within $TF - IDF$ method The term frequency is the average amount of occurrence in the cluster. If the $TF - IDF$ is higher then the topic is represented by words in the cluster & related sentences is selected on the basis of similar sentences with the topic of cluster. Author *Zhang* in [30] according to their resemblance of the words among two sentences & semantic similarity of words due to this the measurement of similarity among the sentences is designed & then, the K-means method is used to gather the sentences of the document in the clusters.
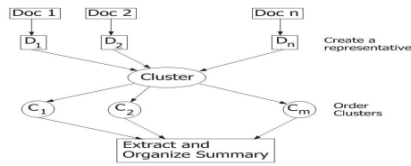
Fig. 5. *Text Miner Cluster*: The key sentences are extracted from clusters in ascending order of the size of clusters.

Yulia, et al. [31] propose a method based on the following steps: *terms selection*, *terms weighting* and *sentences selection*. In the primary step has been a three models & one out of three model of the text is take out: n-grams model, Maximal frequent Sequence model and Bag-of-words model. In the secondary stage, the terms are prejudiced by using the Boolean technique, TF, IDF etc. Ultimately, to acquire a sentence representing segments, the *Expectation-maximization algorithm* (EM) is used to form comparable collections of judgments that included in the summary.

E. Semantic Dependency Based Method

S. Hariharan et al in [27] projected an enhancement to be based on 2 graphical approaches specifically- the LexRank (which means threshold) & the LexRank (which means continuous) accessible by *Erkan and Radev* [28]. Both method are established for the web link analysis, which is based on the alteration of most general page ranking algorithms proposed. A relation between the sentences is measured as the vote company from 1 sentence to another sentence. Discounting technique predicts that once a sentence is selected then next sentence is designated based on contributions made by remaining $n-1$ sentences only. Subsequently, the probability of the repetition of information in following sentences is reduced, & the summaries that are created will be interrelated & meaningful. In graph based method, the significance to position of sentence can be given in the way by giving inclination to the sentences that happens earlier out of other documents well-thought-out.
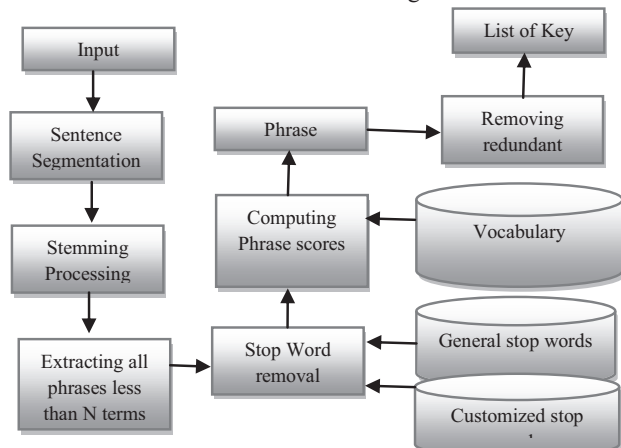

Fig. 6. Illustration of key phrase extraction in LexRank and Continuous LexRank methods.

F. Topic-Based Approaches

Edmundson In [32] for cue word identification summaries are produced for determining the related sentence in the form of the axioms or words that it contains in a document. The Sentences which contain phrases like "*in conclusion*" or "*the aim of this paper*" may be noble pointers of appropriate facts. In other approaches, take revenue of the benefits of joining topics' identification & segmentation projected by *Boguraev* in [33] *Neto* in [34], *Angheluta* in [35], or *Harabagiu* in [36]. But, in [36], for the representations of events ,the topic erection is measured in terms of topics themes, that are repeated throughout the document gathering, & therefore signify the repetitive information. There are Five dissimilar methods of representing topics are examined: **(1)** via topic signatures. Author *Lin* in [37], proposed the information using a set of terms where it is supposed that can be represented the topic of a document; **(2)** via enhanced topic signatures. The aim of this method to find the related relation among two topic concepts and this is differ from first method; **(3)** via thematic signatures, using the TextTiling algorithm proposed by *Hearst* in [38] which is known by segmenting documents, & then assigning labels to themes & further provide the rank; **(4)** via modelling the content arrangement of documents. Particular topic is created by single content model in which all texts describing a given topic. Lastly, the previous technique is used to signify topics within text **(5)** templates, subsequent knowledge of the field of classifying exact articles, info extraction, or facts.

Additionally, in a single document summarization, topic identification are combined with term frequency is proposed by *Teng* in [39]. Firstly the projected procedure calculates the relation of the sentence, & then due to sentence clustering it achieves the topic identification. In the second step, in accordance to the term frequency value, sentences from local topics are selected. Moreover, document topic words are used to detect significant text (*Kuo and Chen* 2008) [40]. The vector space representation is used to identify the multiple theme which is combined with the collection of document is addressed in *Ando et al.* in [41]. And for suitable building space models for verbal items Iterative Residual Rescaling is used.

Fig. 7. Figure shows one by one concept layers subjugated to summarize text from documented data in topic based approach [42].

G. Discourse Based Approaches

A linguistic point of view, all the previous stated methods faces the summarization difficulty, for example developing discourse relations. Rhetorical Structure Theory projected by *Mann et al*; in [45] worked on the basis for the summarization method that are established by *Marcu* in [46], ranging the rhetorical relations, & using this type of discourse depiction to regulate the most significant written elements in a document.

Khan et al. In [47] add the language information to summarization procedure due to this the Rhetorical Structure Theory is shared within a common summarizer. Even though the outcome found by different technique do not improved the ones achieved by the common summarizer, it was requested that the difficulty of this technique depend on the parser which could not notice all the Rhetorical Structure Theory relations, otherwise language info could have improve the whole summarization routine. Additionally, *Cristea et al.* in[48] a method comparable to Rhetorical Structure Theory is defined, differing from the earlier ones, in the lack of relative names & the use of binary trees. This summarization method is planned to exploit the coherence & cohesion of a document.

There are two leading thought-provoking subjects for TS are **Cohesion** & **coherence**. Author *Cunha et al.* in [49] combines statistical & language methods to show that improve consequences with respect to use only one type of methods.

Gonçalves et al. in [50], in the extractive summarization method co-reference chains are used to deal with common referential cohesion problems. A post processing system is recognized in order to change referential terms in the most possible coherent way, & it is useful after the summary is made, finding considerable improvements in assessment to the single summaries. However, the use of co-reference chains is not original in TS. The earliest methodologies can be originate in Baldwin & Morton (1998) [51], & Azzam et al. (1999) [52]. The main announcement is that the longest co-reference chain requires the main topic of the document, & shorter chains indicate subtopics. Therefore, one possible method for building précises is to choose only those sentences associated in the longest chain. This method helps to reserve the coherence of the document. A comparable knowledge is to use verbal chains, which comprises of defining groupings of semantic linked words. Some author also used lexical chains which is also useful in summarizing text documents. This method has also widely used in summarization, & approaches like defined by *Barzilay* in [53], *Medelyan* in [54] or *Ercan* in [55], apply lexical chains to yield summaries. It is important to mention here that, to categorise all the articles that are related within a document or across documents, evade summaries from the common dangling anaphora phenomenon, thus producing more coherent subsequent summaries (*Elsner* in [56]), this only consist of noun part of the text in it.

H. Latent Semantic Analysis (LSA) Based Approaches

Gong and Liu [57] Proposed an automatic summarization system of news text with the use of LSA as a way to identify the significant topics in the documents without using lexical possessions like WordNet. In this way, the SVD is applied to matrix A to decompose into three novel matrices as follows: $A = UWV^T$.

Fig. 8. Shows Propagating, reversing, intensifying, and weakening.

The proposed that the row of the matrix V^T can be considered as various topics covered in the original text, while all support signifies a sentence in the document. And finally, in order to create an extractive summary, they reproduce each row of matrix V^T successively, & extract a sentence from it which has maximum values. *Yeh, et al.* In [58] additional technique using LSA was projected. It is a mixed method among graphs based technique & LSA based technique. After using the SVD on a matrix of words per sentence & saving of these magnitudes, the consistent matrix is constructed.

Each support represents the sentence semantic (logic) symbol which is used, in its place of an existence frequency vector of keyword, in order to signify a document as a graph of relatives among sentences. A ranking algorithm is then useful to the subsequent graph.

The technique which is explained in LSA is Cross & Topic which realizes sentence selection [59].

I. Approaches Based on Lexical Chains (Cohesion lexies)

The automatic text summarization by lexical chains was presented by *Barzilay* & *Elhadad* in [43]. This method uses the word Net database knowledge to determine the relations of cohesion between terms then composes chains based on these terms. Scores are given which is created on the amount & type of relatives in the chains.

A similar method with graphs using the knowledge bases of word Net and Wikipedia was presented by author *Pourvali* in [44]. This method consists initially in finding the exact meaning of each word in the text using WorldNet, then builds the lexical chains and removes those which have a weak score compared with the others.

J. Approach Based on Fuzzy Logic

Farshad, et al. in [60], another method to automatic summarization has been projected, this time it is based on a

fuzzy logic. This technique takes into account every characteristics of the text such as word frequency, similarity to keywords, similarity to the title words, sentences location, statistics of co-occurrence of lexical chain, indicative expression etc. After removing these characteristics & depending on the consequences, a value of 0-1 is allocated to all sentence of the text according to the features of sentences & rules obtainable in base of knowledge. The consequence of sentence in the last summary's value acquired at the output defines degree. *Esther Hannah, et al.* in [61], dissimilar features of each sentence were taken into account, like sentence length, Sentence to sentence resemblance, title words, term weight etc. the values of these characteristics are used by the inference engine to make the score of each sentence of the text. *L. Suanmali* in [62] has proposed to remove significant sentences as a summary the text summarization based on fuzzy logic technique.
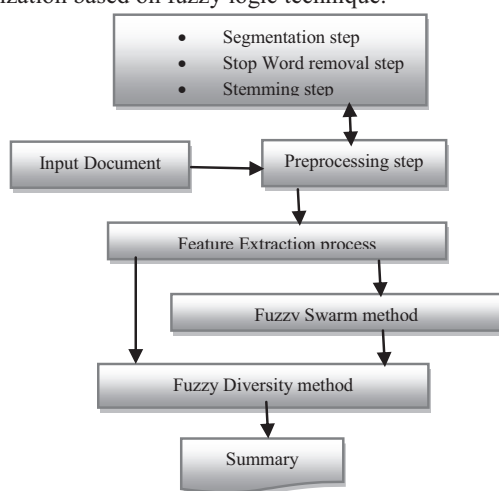

Fig. 9. The output of the fuzzy logic controller is resolute by a procedure called defuzzification. As can be seen in the fuzzy control block diagram, this process converts the output text of the related rules into a single arithmetical value that is denormalized & then used to control the wrought variable.

In paper [63], Author has used the hybrid method-Statistical and Linguistic methods. The evaluation is carried out by using precision and recall method. The ranking or score for every sentence are derived using fuzzy logic technique, and also by means of all 8 feature scores. In paper [64] some state of methods for text summarization author has analyzed. Author proposed a new method of text summarization using fuzzy logic & discussed these methods' main disadvantages. In [65], focus is on the automatic text summarization. To extract the sentences author used fuzzy logic.

In [66], author proposed an automatic text summarization method using genetic algorithm, fuzzy logic, semantic role classification & their combinations based on sentence extraction to produce the high quality summaries. GA used in text summarization. In [67], author has applied the fuzzy logic approach for important features to extract the

sentences. In this paper the result is compared with the Microsoft Word 2007 & baseline summarizer summarizers. Author in [68] proposed a method to extract key sentences based fuzzy sets, by estimating the relevance of sentences as a summary of a document in [68]. To avoid the alike semantic meaning this method does not uses raw words, its only uses senses.

IV. CONCLUSION

The quantity of text evolution outstanding to the WWW (World Wide Web) has called for a prerequisite to advance effectual & exact summarization schemes. Despite the fact, exploration on summarization approaches underway nearby 50 years formerly, but there is a lot of optimum work is required in this field. In previous attempts in this field, consideration has coasted from summarizing scientific term to newscast articles, automated mail letters, announcements, & blogs. Self-possessed extractive methodologies and abstractive methodologies partake remained endeavoured, contingent preceding the solicitation on arrow. Typically, abstractive type summarization necessitates extensive equipment for language cohort & is difficult to repeat and/or take it to comprehensive domains. The topical admiration of an operative newswire (A service of transmitting the latest news stories via Internet) summarization structures endorses this type of prerogative.

This detailed assessment accentuates extractive methodologies to summarization using numerous approaches. A discrepancy has through amongst single/solitary document and multi-document summarization. Subsequently, some motivating efforts are done up to now within earlier research, This work preferred towards embrace a transitory conversation for few good approaches that we found to be of good scope for future research, also, these works are emphasis only on trivial minutiae associated with a wide-ranging summarization progression and not on proposing an all-inclusive summarization scheme.

As a final point, particular contemporary inclinations in automatic appraisal of summarization schemes have been fathomed. The truncated inter-annotator preparation figures pragmatic throughout manual estimations put forward upcoming experiments of this problem area profoundly be contingent in capacity to find competent approach for inevitably weighing such kind of text summarization system, and also on enlargement of procedures which impartial adequate to ordinarily putative by the peoples who are doing researches on this area.

V. REFERENCES

[1] Sparck Jones, K., Endres-Niggemeyer, B.: Automatic summarizing. Inf. Process. Manag. 31(5), 625–630 (1995).

[2] Over, P., Dang, H., Harman, D.: DUC in context. Inf. Process. Manag. 43, 1506–1520 (2007). DOI 10.1016/j.ipm.2007.01.019.

[3] Owczarzak, K., Dang, and H.: Overview of the tac 2010 summarization track. In: Proceedings of TAC 2010, NIST, Gaithersburg, MD, USA (2010).

[4] Sparck Jones, K.: What might be in a summary? In: K. Knorz, Womser-Hacker (eds.) Information Retrieval 93: (1993).

[5] Mani, I.: Automatic Text Summarization. John Benjamins, Amsterdam/Philadelphia (2001).

[6] Maybury, M. (1995) Generating summaries from event data. Iizfor. Process. Manage. 31(5): 735-751.

[7] Borko, H. and Bernier, C. (1975) Abstracting Concepts and Methods. Academic Press.

[8] Lancaster, F. W. (1991) Indexing and Abstracting in Theory and Practice. University of Illinois Graduate School of Library and Information Science.

[9] Sparck-Jones, K. (1999) Automatic summarizing: factors and directions. In: Mani, I. and Maybury, M., editors, Advances in Automatic Text Summarization, pp. 1-12. MIT Press.

[10] Radev, D. and McKeown, K. (1998) Generating natural language summaries from multiple on-line sources. Computational Linguistics.

[11] Mani, I. and Bloedorn, E. (1997) Multi-document Summarization by graph search and merging. Proceedings 14th National Conference on Artificial Intelligence (AAAI-97), Providence, RI, pp. 622-628.

[12] Inderjeet Mani and Mark T. Maybury, editors. 1999. Advances in Automatic Text Summarization. MIT Press, Cambridge, MA.

[13] G. Salton, "Automatic Text Processing: the transformation, analysis, and retrieval of information by computer," Addison Wesley Publishing Company, USA, 1989.

[14] Jun'ichi Fukumoto, "Multi-Document Summarization Using Document Set Type Classification," Proceedings of NTCIR- 4, Tokyo, 2004.

[15] You Ouyang, Wenji Li and Qin Lu, "An Integrated Multi-document Summarization Approach based on Word Hierarchical Representation," Proceedings of the ACL-IJCNLP Conference Short Papers, Suntec, Singapore, pp. 113–116, 2009.

[16] Inderjeet Mani and Eric Bloedorn, "Multi-document summarization by graph search and matching," AAAI/IAAI, vol. cmplg/9712004, pp. 622-628, 1997.

[17] Erkan G, Radev DR (2004) LexRank: Graph-based Lexical Centrality as Salience in Text Summarization. J Artif Intell Res (JAIR) 22:457–479/

[18] Mihalcea R (2004) Graph-based ranking algorithms for sentence extraction, applied to text summarization. In: Proceedings of the ACL 2004 on interactive poster and demonstration sessions. p 20/

[19] Wan X, Yang J, Xiao J (2007) towards a unified approach based on affinity graph to various multi-document summarizations. In: Proceedings of the 11th European conference. pp 297–308/

[20] Giannakopoulos G, Karkaletsis V, Vouros G (2008a) Testing the use of n-gram graphs in summarization sub-tasks. In: Proceedings of the text analysis conference (TAC)/

[21] Plaza L, Díaz A, Gervás P (2008) Concept-graph based biomedical automatic Summarization Using Ontologies. In: Coling 2008: Proceedings of the 3rd text graphs workshop on graph-based algorithms for natural language processing. pp 53–56/

[22] Fellbaum C (1998) word Net: an electronical lexical database. The MIT Press, Cambridge.

[23] Rada Mihalcea and Paul Tarau, "Text-rank: Bringing Order into Texts," Proceeding of the Conference on Empirical Methods in Natural Language Processing, Barcelona, Spain, 2004.

[24] Kathleen McKeown and Dragomir R. Radev, "Generating Summaries of Multiple News Articles," Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval, Seattle, Washington, pp. 74–82, 1995.

[25] Message Understanding Conference (MUC). Proceedings of the Fourth Message Understanding Conference (MUC- 4). DARPA Software and Intelligent Systems Technology Office, 1992.

[26] Xiaojun Wan, "Timed Text Rank: Adding the Temporal Dimension to Multi-Document Summarization," Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Amsterdam, pp. 867- 868, 2007.

[27] Shanmugasundaram Hariharan, Thirunavukarasu Ramkumar and Rengaramanujam Srinivasan, "Enhanced Graph Based Approach for Multi Document Summarization," The International Arab Journal of Information Technology, 2012.

[28] G. Erkan and D. Radev, "LexRank: Graph-based Lexical Centrality as Salience in Text Summarization," Journal of Artificial Intelligence Research, vol. 22, pp. 457-479, 2004.

[29] Jade Goldstein, Vibhu Mittal, Mark Kantrowitz and Jaime Carbonell, "Multi-Document Summarization by Sentence Extraction," ANLP/NAACL Workshops. Association for Computational Linguistics, Morristown, New Jersey, pp. 40-48, 2000.

[30] P.-Y. Zhang and L. Cun-He, "Automatic text summarization based on sentences clustering and extraction," Computer Science and Information Technology, 2009. ICCSIT 2009. 2nd IEEE International Conference. 2009.

[31] L. Yulia, H. René García, S. Romyna Montiel, R. Rafael Cruz, and G. Alexander, "EM clustering algorithm for automatic text summarization," in Proceedings of the 10th Mexican international Conference on Advances in Artificial Intelligence - Volume Part I (MICAI'11), Vol. Part I. Springer Verlag, Berlin, Heidelberg, 2011, pp. 305-315.

[32] Edmundson HP (1969), "New methods in automatic extracting." In: Mani I, Maybury M (eds) Advances in automatic text summarization.

[33] Boguraev BK, Neff MS (2000) Discourse segmentation in aid of document summarization. In: Proceedings of the 33rd Hawaii international conference on system sciences, vol 3. p 3004.

[34] Neto JL, Santos A, Kaestner CAA, Freitas AA (2000) Generating text summaries through the relative importance of topics. In: IBERAMIA-SBIA '00: proceedings of the international joint conference, 7th Ibero-American conference on AI. pp 300–309.

[35] Angheluta R, Busser RD, Francine Moens M (2002) the use of topic segmentation for automatic summarization. In: Proceedings of the ACL-2002 post-conference workshop on automatic summarization. Pp 66–70.

[36] Harabagiu S, Lacatusu F (2005) Topic themes for multi-document summarization. In: SIGIR '05: proceedings of the 28th annual international ACM SIGIR conference on research and development in information retrieval.

[37] Lin CY, Hovy E (2000), the automated acquisition of topic signatures for text summarization. In: Proceedings of the 18th conference on computational linguistics. pp 495–501.

[38] Hearst MA (1997) Text Tiling: segmenting text into multi-paragraph subtopic passages. Comput Linguist 23(1):33–64.

[39] Teng Z, Liu Y, Ren F, Tsuchiya S, Ren F (2008) Single document summarization based on local topic identification and word frequency. In: MICAI '08: proceedings of the 2008 seventh Mexican international conference on artificial intelligence. pp 37–41. http://dx.doi.org/10.1109/MICAI.2008.12.

[40] Kuo JJ, Chen HH (2008) Multi-document Summary Generation: Using Informative and Event Words. ACM Trans Asian Lang INF Process (TALIP) 7(1):1–23.

[41] Ando R, Boguraev B, Byrd R, Neff M (2005) Visualization-enabled multi-document summarization by Iterative Residual Rescaling. Nat Lang Eng.

[42] Z. Teng, Y. Liu, F. Ren, S. Tsuchiya, and F. Ren, "Single document summarization based on local topic identification and word frequency," in MICAI '08: Proceedings of the 2008 Seventh Mexican International Conference on Artificial Intelligence, 2008, pp. 37–41.

[43] R. Barzilay and M. Elhadad, "Using lexical chains for text summarization," in Proceedings of the ACL/EACL'97 Workshop on Intelligent Scalable Text Summarization, Madrid, Spain, 1997, pp. 10– 17.

[44] M. Pourvali and S. Abadeh Mohammad, "Automated text summarization base on lexicales chain and graph using of word net and Wikipedia knowledge base," IJCSI International Journal of Computer Science Issues, No. 3, vol. 9, 2012

[45] Mann WC, Thompson SA (1988) Rhetorical structure theory: Toward a functional theory of text organization. Text 8(3):243–281

[46] Marcu D (1999) Discourse trees are good indicators of importance in text. In: Advances in automatic text summarization. pp 123–136

[47] Khan AU, Khan S, Mahmood W (2005) MRST: a new technique for information summarization. In: The second world enformatika conference, WEC'05. pp 249–252

[48] Cristea D, Postolache O, Pistol I (2005) Summarization through discourse structure. In: Proceedings of the computational linguistics and intelligent text processing, 6th International conference (CICLing 2005).

[49] Cunha ID, Fernandez S, Velázquez-Morales P, Vivaldi J, SanJuan E, Moreno JMT (2007) a new hybrid summarizer based on vector space model, statistical physics and linguistics. In: MICAI 2007: advances in artificial intelligence. pp 872–882.

[50] Gonçalves PN, Rino L, Vieira R (2008) Summarizing and referring: towards cohesive extracts. In: DocEng '08: proceeding of the eighth ACM symposium on document engineering. pp 253–256

[51] Baldwin B, Morton TS (1998) Dynamic coreference-based summarization. In: Proceedings of the third conference on empirical methods in natural language processing (EMNLP-3)

[52] Azzam S, Humphreys K, Gaizauskas R (1999) Using coreference chains for text summarization. In: Proceedings of the ACL'99 workshop on coreference and its applications

[53] Barzilay R, Elhadad M (1999) Using lexical chains for text summarization. In: Advances in automatic text summarization. pp 111–122

[54] Medelyan O (2007) Computing lexical chains with graph clustering. In: Proceedings of the ACL 2007 student reasearch workshop. pp 85–90

[55] Ercan G, Cicekli I (2008) Lexical cohesion based topic modeling for summarization. In: Proceedings of the 9th international conference in computational linguistics and intelligent text processing. pp 582–592

[56] Elsner M, Charniak E (2008) Co-reference-inspired coherence modeling. In: Proceedings of ACL-08: HLT, short papers. pp 41–44

[57] Y. Gong and X. Liu, "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, 2001, pp. 19–25.

[58] J.-Y. Yeh, H.-R. Ke, W.-P. Yang, and I.-H. Meng, "Text summarization using a trainable summarizer and latent semantic analysis," In Special Issue of Information Processing and Management on An Asian Digital Libraries Perspective, vol. 41, pp. 75–95, 2005.

[59] Ozsoy, Cicekli, "Text Summarization of Turkish Texts using Latent semantic Analysis", In proceeding of 23rd International conference Computational Linguistics, Beijing, pp. 869-876, 2010.

[60] K. Farshad, K. Hamid, E. Esfandiar, and K. D. Pooya, "Optimizing text summarization based on fuzzy logic," in Proceedings of Seventh IEEE/ACIS International Conference on Computer and Information Science, IEEE, University of Shahid Bahonar Kerman, UK, 2008, pp. 347-352.

[61] M. Esther Hannah, T. V. Geetha, and M. Saswati, "Automatic extractive text summarization based on fuzzy logic: A sentence oriented approach," in Proceedings of the Second international conference on Swarm, Evolutionary, and Memetic Computing - Volume Part I (SEMCCO'11), Bijaya Ketan Panigrahi, Ponnuthurai Nagaratnam Suganthan, Swagatam Das, and Suresh Chandra Satapathy (Eds.), Vol. Part I. 2011, pp. 530-538..

[62] L. Suanmali, N. Salim and M.S. Binwahlan, "Fuzzy Logic Based Method for Improving Text Summarization", International Journal of Computer Science and information Security , 2009, vol.2, No. 1, pp. 4- 10.

[63] Mrs.A.R.Kulkarni, Dr.Mrs.S.S.Apte "A DOMAIN-SPECIFIC AUTOMATIC TEXT SUMMARIZATION USING FUZZY LOGIC ", International Journal of Computer Engineering and Technology (IJCET), ISSN 0976- 6367(Print), ISSN 0976 - 6375(Online) Volume 4, Issue 4, (2013).

[64] Farshad Kyoomarsi ,Hamid Khosravi ,Esfandiar Eslami ,Pooya Khosravyan Dehkordy; "Optimizing Text Summarization Based on Fuzzy Logic ",Seventh IEEE/ACIS International Conference on Computer and Information science, 978-0-7695-3131-1, 2008.

[65] Ladda Suanmali, Naomie Salim and Mohammed Salem Binwahlan, "Feature-Based Sentence Extraction Using Fuzzy Inference rules ", 2009 IEEE International Conference on Signal Processing Systems, 2009.

[66] Ladda Suanmali, Naomie Salim and Mohammed Salem Binwahlan "Fuzzy Genetic Semantic Based Text Summarization ", 2011 Ninth International Conference on Dependable, Autonomic and Secure Computing ,978-0-7695-4612-4 ,2011 IEEE .

[67] Ladda Suanmali, Mohammed Salem Binwahlan and Naomie Salim "Sentence Features Fusion for Text Summarization Using Fuzzy Logic ", 2 IEEE Ninth International Conference on Hybrid Intelligent Systems, 978-0-7695-3745-0, 2009.

[68] Hsun-Hui Huang, Yau-Hwang Kuo, Horng-Chang Yang, "Fuzzy-Rough Set Aided Sentence Extraction Summarization", IEEE Proceedings of the First International Conference on Innovative Computing, Information and Control (ICICIC'06), 0-7695-2616-0/0