

COMPARISON OF DIFFERENT AUTOMATIC TEXT SUMMARIZATION
SYSTEMS USING STANDARD PERFORMANCE EVALUATIONS

NUR HAFIZAH BINTI ABD MUNIR

A project report submitted in partial fulfillment of the
requirements for the award of the degree of
Master of Science (Computer Science)

Faculty of Computer Science and Information Systems
Universiti Teknologi Malaysia

ARPIL 2009

*“To my dearest beloved husband ~ Nazrie,
My dearest beloved father ~ Hj. Abd Munir and mother ~ Hjh. Salmah,
My dearest younger brother ~Hafiz and younger sister ~ Hidayah,
My dearest beloved mother-in-law, Hjh. Rohijah,
sincere thanks for all your love, care, support and believe in me.
Special thanks to all my dear lecturers and friends,
sweet memory remains in our hearts forever.
Thanks to all for being there, throughout this journey”*

ACKNOWLEDGEMENT

In the name of Allah, Most Merciful, Most Compassionate. It is God's willing; make me able to complete this project within period given. I would like to express the deepest gratitude to my supervisor, Associate Professor Dr. Naomie Salim for her advice, guidance, support, tolerance and attention toward the accomplishment of this study. My sincere appreciation also goes to my examiners, Associate Professor Dr. Harihodin Selamat and Dr. Mohd. Shahizan Othman on their helpful comments and suggestions in evaluating this project.

Not forgetting, I am also obliged to express my greatest appreciation to my lovely husband and family members, who have been fully giving me their commitments and supports whenever I need any helps in whatever sources. Sincere thanks for the everlasting love, care and supports along my journey of life.

I am very grateful towards all my fellow friends as they have been very supportive and always giving me assistance in various occasions. And, all the staffs and lecturers of Faculty of Computer System and Information System, Universiti Teknologi Malaysia who have been directly or indirectly influential and supportive to this project are also entitled for an appreciation on their knowledge, high motivation and self esteem, made my experience of learning here worthwhile. The sweet memory with all of you will cherish and never been forgotten forever.

ABSTRACT

There are many automatic summarization systems can be used to produce a summary from a single text documents. From the different automatic summarization system, it can be found that the system will produce a different content of summary results although the percentage of sentences out of whole single text document is setting to the same value. Therefore, in this study, three automatic summarization systems are used to produce the summary results; Microsoft Word Automatic Summarization, Shvoong Summarization and Simple Text Summarization in PHP. The performance of those results are investigated and measured using standard performance evaluation such recall, precision and f-measure. The dataset collection used in this study is collected from The New Straits Time and The Stars online and it is about Iskandar Region Development Authority (IRDA). Two automatic summarization system are already existed which is Microsoft Word Automatic Summarization and Shvoong Summarization and only one summarization system is coded in PHP language, there is Simple Text Summarization in PHP. Many operations have been applied in this coded system such as removing stop word, stemming, normalizing, creating weighted term-frequency and applying the technique. The results from those systems are stored into the database. In this study, about 50 articles are used. The comparison between different automatic summarization systems was made using standard performance evaluation. The performance evaluation is fully analyzed without depending on human evaluator. One program of analyzing the performance is coded in PERL language to produce a statistic of all summary results from those three automatic summarization systems. From the experimental results, it can be concluded that the Shvoong Summarization is the most effective automatic summarization system for single text document.

ABSTRAK

Terdapat banyak sistem rumusan automatik (SRA) yang boleh digunakan bagi menghasilkan sesuatu rumusan daripada satu petikan. Daripada SRA yang berlainan, didapati bahawa rumusan yang dihasilkan juga adalah berbeza walaupun peratusan ayat yang dikeluarkan dari satu petikan disetkan pada nilai yang sama. Oleh itu, di dalam kajian ini, tiga SRA digunakan bagi menghasilkan rumusan di mana hasil rumusan bagi SRA ini diukur dan diselidiki dengan menggunakan penilaian pelaksanaan seperti pemanggilan balik (*recall*), ketepatan (*precision*) dan pengukuran-f (*f-measure*). Set data terkumpul yang digunakan di dalam kajian ini diperolehi daripada akhbar atas talian seperti *The New Strait Time* and *The Stars* dan ianya berkisar tentang Wilayah Pembangunan Iskandar (WPI). Dua daripada SRA adalah terdiri daripada sistem sedia ada iaitu Rumusan Automatik *Microsoft Word* dan Rumusan Shvoong dan hanya satu SRA yang dikodkan iaitu Rumusan Petikan Ringkas di dalam PHP. Banyak operasi yang digunakan secara praktikal di dalam program ini seperti membuang kata henti (*stop word*), mendapatkan kata dasar (*stemming*), pernormalan, mencipta pemberat kekerapan setiap perkataan dan penggunaan teknik rumusan. Hasil daripada semua sistem rumusan disimpan di dalam pangkalan data. Di dalam kajian ini, sebanyak 50 petikan akhbar atas talian digunakan. Perbandingan diantara SRA yang berlainan ini dibuat dengan menggunakan pengukuran penilaian pelaksanaan. Penilaian pelaksanaan ini secara keseluruhannya dianalisa dengan menggunakan sebuah program yang dikodkan di dalam bahasa PERL. Proses penganalisaan yang dijalankan tidak melibatkan hasil rumusan penilai manusia. Daripada keputusan kajian yang diperolehi, boleh disimpulkan bahawa Rumusan Shvoong adalah merupakan rumusan yang paling berkesan bagi satu petikan.

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE
	DECLARATION	ii
	DEDICATION	iii
	ACKNOWLEDMENTS	iv
	ABSTRACT	v
	ABSTRAK	vi
	TABLE OF CONTENTS	vii -x
	LIST OF TABLES	xi
	LIST OF FIGURES	xii
	LIST OF ABBREVIATIONS	xiii
	LIST OF SYMBOLS	xiv
	LIST OF APPENDICES	xv
1	INTRODUCTION	
1.1	Introduction	1
1.2	Problem Background	3
1.3	Problem Statements	7
1.4	Aim of the Study	7
1.5	Objectives of the Project	8
1.6	Scopes of the Project	8
1.7	Organization of Thesis	9

1.8	Summary	9
-----	---------	---

2 LITERITURE REVIEW

2.1	Introduction	10
2.2	Summarization System	11
2.3	Extraction versus Abstraction Summarization	14
2.4	Types of Summarizations	15
2.4.1	Single Document Summarization	15
2.4.2	Multi-Document Summarization	17
2.5	Stop Words	20
2.6	Types of Stemmer	21
2.6.1	The Lovins	21
2.6.2	The Porter	22
2.6.3	The Dawson	22
2.6.4	The Paice/Husk	23
2.6.5	The Krovetz	24
2.7	Weighting Schemes	25
2.7.1	Term Frequency (<i>tf</i>)	28
2.7.2	Inverse Document Frequency (<i>idf</i>)	30
2.8	Summarization Techniques	31
2.8.1	Luhn's Keyword Cluster	31
2.8.2	Full Coverage (FC)	33
2.8.3	Title Term Frequency	35
2.8.4	Singular Vector Decomposition (SVD)-Based	36
2.8.5	Text Segmentation	38
2.8.6	Sentence Scoring	39
2.8.7	Modified Term Weighting	40
2.8.8	Segment Ranking	42
2.8.9	Simple Text Summarization in PHP	43

2.9	Performance Evaluations	44
2.9.1	Responsiveness	44
2.9.2	Linguistic Quality	45
2.9.3	Rouge	46
2.9.4	Pyramid	47
2.9.5	Readability	49
2.9.6	Recall, Precision and F-Measure	50
2.10	Discussion	52
2.11	Summary	54

3 METHODOLOGY

3.1	Introduction	55
3.2	Project Framework	56
3.2.1	First Stage: Preparing Collection	56
3.2.2	Second Stage: Parsing the Document into Sentences	57
3.2.3	Third Stage: Getting Summary from Summarization Systems.	57
3.2.3.1	Removing Stop Words.	58
3.2.3.2	Stemming Process.	59
3.2.3.3	Normalizing Process.	59
3.2.3.4	Creating Weighted Term-Frequency.	59
3.2.3.5	Applying Technique.	60
3.2.3.6	Getting Summary from Simple Text Summarization in PHP	61
3.2.3.7	Getting Summary from Microsoft Word Automatic Summarization.	62
3.2.3.8	Getting Summary from Shvoong Summarization.	63

3.2.4	Fourth Stage: Measuring Performance Evaluation.	64
3.3	System Requirements	65
3.3.1	Software Justification	65
3.3.2	Hardware Specification	66
3.4	Summary	66
4	EXPERIMENTAL RESULTS AND ANALYSIS	
4.1	Introduction.	67
4.2	First Stage: Preparing Collection.	67
4.3	Second Stage: Parsing the Document into Sentences.	68
4.4	Third Stage: Getting Summary from Summarization Systems.	68
4.5	Fourth Stage: Measuring Performance Evaluations.	69
4.6	Comparison of Summarization Systems	74
4.7	Discussion	78
4.8	Summary	79
5	CONCLUSION AND FUTURE WORK	
5.1	Introduction	80
5.2	Finding	81
5.3	Contribution	81
5.4	Conclusion	82
5.5	Suggestion for Future Work	82
	REFERENCES	83 – 85
	Appendices A - M	86 - 120

LIST OF TABLES

TABLE NO.	TITLE	PAGE
2.1	Stopwords.	20
2.2	Example Weighting Schemes by Chisholm and Kolda (1999).	26
2.3	Weighting Schemes in Summarization Performance.	27
2.4	Term Frequency Factors and Its Description.	28
3.1	Recall and Precision Formulation.	64
3.2	Hardware Specifications.	66
4.1	Recall Measurement Results.	69
4.2	Precision Measurement Results.	72
4.3	F-measure Results.	75

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
2.1	Basis Overview of Summarization System.	13
2.2	Example of a Weighted Graph.	17
2.3	Summarization Process for Multi-Document.	18
2.4	Application of SVD-Based.	37
3.1	Flow of Project Framework.	56
3.2	Simple Text Summarization in PHP Interface.	62
4.1	Recall Graph.	71
4.2	Precision Graph.	74
4.3	F-measure Graph.	77

LIST OF ABBREVIATIONS

AMD	-	Advanced Micro Devices
ASCII	-	American Standard Code for Information Interchange.
DUC	-	Document Understanding Conference.
FC	-	Full Coverage.
GB	-	Gigabyte
HTML	-	Hyper-Text Markup Language.
IDF	-	Inverse Document Frequency.
IR	-	Information Retrieval.
IRDA	-	Iskandar Region Development Authority.
LLR	-	Log-Likelihood Ratio.
MEAD	-	Multi-document Summarizer.
MS-DOS	-	Microsoft Disk Operating System
MySQL	-	My Structure Query Language.
PERL	-	Practical Extraction and Reporting Language.
PHP	-	PHP Hypertext Preprocessor.
SCU	-	Summarization Content Units.
SRA	-	Sistem Rumusan Automatik.
SVD	-	Singular Vector Decomposition.
TF	-	Term Frequency.
TIME	-	Technology Information Multimedia And Entertainment.
TREC	-	Text Retrieval Conference.
UNICODE	-	Unique, Universal, and Uniform Character Encoding.
WPI	-	Wilayah Pembangunan Iskandar.

LIST OF SYMBOLS

A_k	-	Vector of sentence k
C	-	Vector normalization
Q	-	Quadratic
Σ	-	Singular Value Matrix / Summation
σ	-	Diagonal Elements (Sigma)
V	-	Right Singular Vector Matrix
U	-	Left Singular Vector Matrix
A	-	Target Matrix
i	-	Term
j	-	Document
$\sqrt{}$	-	Square Root
\log	-	Logarithm
\in	-	Element of
$ T_i $	-	Length of Summarization Content Unit i

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
A	The Project Gantt Chart.	86 – 90
B	The Stopword List.	91 – 96
C	The Porter Stemmers' Flow.	97
D	The Paice/Husk Stemmers' Flow.	98
E	The Stopword Processes.	99
F	The Calculation for Highest Rating Sentence.	100
G	The Steps in Getting Summary Results for Microsoft Word Automatic Summarization.	101
H	The Steps in Getting Summary Results for Online Shvoong Summarization.	102 – 103
I	The Diagram on How to Get the Number of Title Words in the Document and Summary.	104
J	The Dataset Collection Directory and Its Documents.	105
K	The Samples of Parsed Sentence File and Its ID in MySQL Database.	106 – 107
L	The Information About MySQL Database	108
M	The Samples of Summary Results	109 – 120

CHAPTER 1

INTRODUCTION

1.1 Introduction

The growing amounts of information available electronically require tools for fast assessing the content of the information resources. A text summarization system may be thought of as such a tool. Summarization is one of the most common acts of language behavior. Text summarization system can be defined as a process of condensing a source text while preserving its information content and maintaining readability. The goal of the text summarization system is to produce a concise representation with a minimal loss of information of a document or set of documents.

Summaries have been made in order to gain access to and control the flood of information. What is a worth reading and what is useful for particular purpose should be known because nobody want to waste time by reading useless information. By giving an overview of content, summaries will save readers' times. Dagstuhl, (1993) has made a

brief explanation about the importance of text summarization, with access to computers capable of dealing with large textual database.

Radev *et al.* (2002) have provided a sketch of the current state of the art of summarization including single-documents summarization through extraction which is the beginning of abstractive approach to single-documents summarization and a variety of approaches to multi-documents summarization. The major approaches will be explained in detail in chapter 2.

Summary generation systems seek to identify document contents that convey the most “important” information within the document. Where, importance may depend on the use to which the summary is to be put. There are two basic approaches to summarization that are information extraction with subsequent text generation and summaries composed of extracted sentences or phrases. Sentence extracted summaries have been formed by scoring the sentences in the document using some criteria, ranking the sentences and then taking a number of the top ranking sentences as the summary. Various studies have led to the proposal of the following criteria of measuring sentence significance for effective summary generation like sentence position within the document, word frequency within the full-text, the presence or absence of certain words or phrases in the sentence and a sentence’s relation to other sentences, words or paragraphs within the source document. Each sentence score is computed as the sum of its constituent words and other scores (Adesina and Jones, 2001).

Algorithms for extractive summarization are typically based on techniques for sentence extraction and attempt to identify the set of sentences that are most important for the overall understanding of a given document. Some of the most successful approaches consist of supervised algorithms that attempt to learn what makes a good summary by training on collections of summaries built for a relatively large number of

training documents. However, the price paid for the high performance of such supervised algorithms is their inability to easily adapt to new languages or domains as new training data are required for each new data type. The technique for extractive summarization relying on iterative graph-based algorithm had been applied to the summarization of documents in different language without any requirement for additional data. Additionally, it shows that a layered application of the single-documents summarization technique can result into an efficient multi-document summarization tool (Mihalcea and Tarau, 2004).

1.2 Problem Background

As the amount of online information increases, systems that can automatically summarize one or more documents become increasingly desirable. Recent research has investigated types of summaries, techniques to create them and performances evaluation for the summarization. Several evaluation competitions in the style of the National Institute of Standards and Technology (NIST) Text Retrieval Conference (TREC) have helped determine baseline performance levels and provide a limited set of training material (Radev *et al.*, 2002). The Document Understanding Conferences (DUC) also involved in providing the appropriate framework for system independent evaluation of text summarization system.

Knowingly, the main problem in achieving an effective text summarization system is to create a summary with a wider coverage of the document contents and determine less redundancy. Consequentially, an investigation of the most appropriate

techniques must be done to select sentences that are highly ranked and different from each other.

The performance of the text summarization system can be affected by text summarization techniques, weighting schemes and summary evaluation. But the most important task in this system is its' performances evaluation. There are many experiments that have been done to achieve the most appropriate performances for text summarization system for single and multiple documents. For example, Gong and Liu (2001) proposed two generic text summarization techniques that create text summaries by ranking and extracting sentences from original documents. The first techniques used standard information retrieval technique (relevance measure) to rank sentences relevance, while the second technique used the latent semantic analysis technique (SVD-based). Both techniques had been experimented with nine weighting schemes and the standard evaluation method (Recall, Precision and F-measure) to identify semantically important sentences for summary creation. As the result, the two different techniques produced very similar output.

Daniel *et al.* (2004) have proposed Full-Coverage summarizer (FC) to leverage existing information retrieval technology by extracting key-sentences on the premise that the relevance of a sentence is proportional to its similarity to the whole documents. The operational flow of FC summarizer is approximately similar with relevance measure which is proposed by Gong and Liu (2001). By using TIME and DUC as a dataset, their techniques can produce sentences-based summaries up to 78% smaller than the original text with only 3% loss in retrieval performance.

Mihalcea and Ceylan (2007) have explored the problem of book summarization. About 50 books together with its summary had been used as a dataset for evaluation and each of them have two manually created summaries. The average length of book

collection is about 92,000 words with summary length between 6,500 (Cliff's Notes) and 7,500 (Grade Save) words. In this research, there have two stages namely initial experiment and specific experiment. In initial stage, book summarization has been done using a re-implementation of an existing state-of-the-art summarization system like centroid-based technique. This technique has implemented in MEAD by Radev *et al.* (2004) which can be optimized and made very efficient summarization for very long documents such as books. Specific experiment for the dataset had been done in the second stage. The specific experiment has decided to be done because the dataset consist of very large documents and correspondingly the summarization of such document required techniques that count for the length. Several have been selected to test the dataset such as sentence position (positional score), test segmentation, modified term weighting, segment ranking and the combination of some existing techniques. For performance evaluation, all techniques in this specific stage have been evaluated by Rouge evaluation toolkit, recall, precision and f-measure. As a conclusion, the research has made two important combinations. First, a new summarization benchmark, specifically targeting the evaluation of systems for book summarization had been introduced. Second, the system that developed for the summarization of short documents do not fare well when applied to very long documents such as books. Instead, a better performance can be achieved with a system that accounts for the length of the documents.

Using different weighting schemes on summarization system can effects the performance evaluation in producing short and accurate summaries for the document. Weighting schemes can be defined by local and global weighting and also normalization factor. For example, Gupta *et al.* (2007) have examined the focused-based summary by using four weighting schemes such raw frequency (word probability), $R(w)$ and Log-Likelihood Ratio (LLR). The variant of Log-Likelihood, LLR with cut-off, LLR (C) and LLR (CQ) also examined. As a result, the focused summarizer LLR (CQ) is the best and it significantly outperforms the focused summarizer based on frequency. Also, LLR (assign weights to all words) performs significantly worse than LLR (C). Both LLR and

LLR (C) are sensitive to the introduction of topic relevance in producing somewhat better summaries in the focused scenario compared to generic scenario. In other experimentation, Gong and Liu (2001) have studied nine common weighting schemes for two generic summarization which are summarization by relevance measure (represented by summarizer 1) and summarization by latent semantic analysis (represented by summarizer 2). By adding the global weighting and/or vector normalization, the performance of summarization could be changed. So, from both experimentation, can be said that, applying different weighting schemes on various summarization techniques will produce the different result for the performance of the summary.

The most important task in summarization is its performance evaluation. Summaries can be evaluated from the point of view of coverage (the extent to which a system summary bears on the context of the sources text) and quality (consistency and chronological coherence estimation) (Biryukov, 2004). Usually, performance evaluations could be evaluated using the standard precision, recall and f-measure within human evaluator or only by system evaluator itself. Besides, performance evaluation also can be evaluated by human evaluation (pyramid method) and automatic evaluation (Rouge method). In literature review, the detail of performance evaluations will be discussed.

This project is focused on generic summarization systems which it provided the author's points of view of the input text, giving equal important to all major themes in it. Three summarization systems are investigated in such Microsoft Office Word 2003 Automatic Summarization, online Shvoong Summarization and Simple Text Summarization in PHP. Standard performance evaluation methods like recall, precision and F-measure are used for analyzing a good summary for the dataset collection.

1.3 Problem Statement

This project aims to provide a comprehensive comparison of different summarization systems based on performance evaluation for finding out which one is better in finding a good summary to dataset collection.

The purpose of the project is to make a comparison of different automatic text summarization systems by using recall, precision and f-measure to analyze the performance of those systems for single-documents. The research questions to be answered in this project is *which is the most effective automatic text summarization system can be used in performing a good summary for single-document?*

1.4 Aim of the Study

The aim of the study is to investigate and compare the performance of Microsoft Office Word 2003 Automatic Summarization, online Shvoong Summarization and Simple Text Summarization in PHP in producing a summary for the single-document in the dataset collection.

1.5 Objectives of the Project

In order to achieve the aim of the project, several objectives are identified:

- (i) To produce summary results of different automatic text summarization systems for single-documents.
- (ii) To analyze effects of performance evaluation on different automatic text summarization systems using recall, precision and f-measure.
- (iii) To recommend the most effective automatic text summarization systems based on the result from performance evaluation.

1.6 Scope of the Project

- (i) About 50 articles related to Iskandar Region Development Authority (IRDA) are collected and used as dataset in this project. The dataset is obtained from The New Strait Times (NST) and The Star Newspaper Online.
- (ii) Only single-documents are investigated in this project.
- (iii) This project used three automatic text summarization systems which are:
 - Microsoft Office Word 2003 Automatic Summarization.
 - Online Shvoong Summarization.
 - Simple Text Summarization in PHP.

- (iv) A standard performance evaluation such precision, recall and f-measure are used to evaluate the performance on a summary result from each automatic text summarization systems.

1.7 Organization of Thesis

There are five chapters in this thesis like introduction for the project is included in chapter 1, the discussion of literature review is in chapter 2, methodology of the project are explained in chapter 3, the experimental results and analysis discussed in chapter 4 and the last chapter 5 presented the conclusion and suggestion for future work.

1.8 Summary

In this chapter, the introduction of the project such the definition of text summarization system, problem background, problem statements, aim of the study, objectives, scopes and organization of this project are included and explained. Project I and Project II planning for this study also done and illustrated in Gantt chart in Appendix A.

REFERENCES

- Adesina, M.L. and Jones, J.F. (2001). Applying Summarization Technique for Term Selection in Relevance Feedback. In *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '01*. September 9-12. New Orleans, Louisiana, USA: ACM, pages 1 – 9
- Biryukov, M. *et al.* (2005). Multidocument Question Answering Text Summarization Using Topic Signature. *5th Dutch-Belgian Information Retrieval Workshop (DIR'5)*. March. Belgium: Digital Information Management (JDIM), Volume 3 (Issue 1).
- Chali, Y. and Kolla, M. (2004). Summarization Techniques at DUC 2004. In *Proceedings of the Document Understanding Conference 2004*. May 6-7. USA: pages, 1 – 7.
- Chisholm, E. and Kolda, T.G. (1999). *New Term Weighting Formulas for The Vector Space Method in Information Retrieval*. Oak Ridge National Laboratory: Technical Report.
- Dagstuhl, S. (1993). Introduction to “Text Summarization” workshop. *John Hutchins*, (University of East Anglia, Norwich, UK).
- Dawson, J.L. (1974). “Suffix Removal for Word Conflation”. *Bulletin of the Association for Literary & Linguistic Computing*. Volume 2 (Issue 3): pages 33 - 46.
- Gong Y. and Liu, X. (2001). Generic Text Summarization Using Relevance Measure and Latent Semantic Analysis. In *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*

- SIGIR '01*. September 9-12. New Orleans, Louisiana, USA: ACM, pages 19 - 25.
- Gupta, S. *et al.* (2007). Measuring Important and Query Relevance in Topic-focused Multi-document Summarization, (*Stanford University, Stanford*).
- Hirohata, M. *et al.* (2005). Sentence Extractor-Based Presentation Summarization Techniques And Evaluation Metrics. In *Proceedings (ICASSP '05) IEEE International Conference 2005*. March 81-23. Japan: IEEE, Volume 1: pages, 1065-1068.
- Kiani A and Akbarzadeh M.R (2006). Automatic Text Summarization Using: Hybrid Fuzzy GA-GP. In *International Conference on Fuzzy Systems*. July 16 – 21. Vancouver, Canada: IEEE, page 977 – 983.
- Krovetz, R. (1993). “Viewing Morphology as an Inference Process”. In R. Korfhage *et al.*, *Proc. 16th ACM SIGIR Conference*. June, 27 - July, 1. Pittsburgh: ACM, pages 191 - 202.
- Kruengkrai, C. and Jaruskulchai, C. (2003). Generic Text Summarization Using Local and Global Properties of Sentences. In *Proceedings of the IEEE/WIC International Conference on Web Intelligence (WI'03)*. October 13 - 17. Bangkok, Thailand: IEEE, pages 201 – 206.
- Lovins, J.B. (1968). “Development of a Stemming Algorithm”. *Mechanical Translation and Computational Linguistics* Volume 11: pages 22 – 31.
- Mallett, D. *et al.* (2004). Information-Content Based Sentence Extraction for Text Summarization. In *Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'04)*. April 5-7. Alberta, Canada: IEEE, pages 214 - 418.
- Mihalcea, R. and Ceylan, H. (2007). Explorations in Automatic Book Summarization. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. June. Prague: Association for Computer Linguistics, pages 380 - 389
- Mihalcea, R. and Tarau, P. (2004). A Language Independent Algorithm for Single and Multiple Document Summarization. *Department of Computer Science and Engineering, (University of North Taxes)*.

- Nekova, A. (2006). Summarization Evaluation for Text and Speech: Issues and Approaches. In *INTERSPEECH 2006 – ICSLP Ninth International Conference on Spoken Language Processing*. September 17-21. Pittsburg, USA: pages 2079 - 2082.
- Radev, D.R. *et al.* (2002). Introduction to the Special Issue on Summarization. *Computational Linguistics*. Volume 28: pages 399 - 408.
- Robertson, S. (2004). Understanding Inverse Document Frequency: On Theoretical Arguments for IDF. *Journal of Documentation*. Volume 60: pages 503 - 520.
- Saggion, H. (2005), Topic-based Summarization DUC (2005), In *Proceedings of the Document Understanding Conference 2005*. October 9-10. Canada: NIST, pages 1 – 6.