# INFORMATION
# TECHNOLOGY JOURNAL

# Intelligent Model for Automatic Text Summarization

[1]M.S. Binwahlan, [1]N. Salim and [2]L. Suanmali
[1]Faculty of Computer Science and Information Systems, University of Technology Malaysia,
81310, Skudai, Johor, Malaysia
[2]Faculty of Science and Technology, Suan Dusit Rajabhat University, Dusit, Bangkok, 10300, Thailand

**Abstract:** The navigation through hundreds of the documents in order to find the interesting information is a tough job and waste of the time and effort. Automatic text summarization is a technique concerning the creation of a compressed form for single document or multi-documents for tackling such problem. In this study, we introduced an intelligent model for automatic text summarization problem; we tried to exploit different resources advantages in building of our model like advantage of diversity based method which can filter the similar sentences and select the most diverse ones and advantage of the non diversity method used in this study which is the adaptation of intelligent techniques like fuzzy logic and swarm intelligence for building that method which gave it a good ability for picking up the most important sentences in the text. The experimental results showed that our model got the best performance over all methods used in this study.

**Key words:** Binary tree, diversity, fuzzy, summarization, swarm, summary

## INTRODUCTION

The navigation through hundreds of the documents in order to find the interesting information is a tough job and waste of the time and effort. Automatic text summarization is a technique concerning the creation of a compressed form for single document or multi-documents for tackling such problem. The benefits of automatic text summarization system's availability increase the need for existence of such systems; the most important benefits of using a summary is its reduced reading time and providing quick guide to the interesting information. The aim of automatic text summarization techniques is to find the most important text units and present them as summary of the original document. Each technique differs from another in the way of discovering such text units. The automatic text summarization system which is built based on exploiting of the advantages of different resources in form of combined model could produce a good summary for the original document. Many techniques have been proposed for automatic text summarization problem based on different methodologies, (Luhn, 1958; Baxendale, 1958; Edmundson, 1969) used the shallow features to score the text units and selecting the highest score text units as summary (Kupiec et al., 1995; Lin and Hovy, 1997; Lin, 1999; Conroy and Oleary, 2001; Osborne, 2002; Svore et al., 2007; Fattah and Ren, 2009)

could add advantage to the mechanism of the text unit scoring which is the exploitation of the data to create a criteria or weights to be used in the scoring coefficient through the applying of the machine learning for automatic text summarization problem. The discourse structure based techniques (Ono et al., 1994; Barzilay and Elhadad, 1997; Marcu, 1998) employed the discourse structure for the sentence scoring.

All techniques mentioned above did not pay attention to the problem of the redundancy which causes the low quality of the created summary. The method which was built for dealing with the problem of redundancy is MMR (Carbonell and Goldstein, 1998), many methods made use of MMR either directly or after modifying it (Kraaij et al., 2001; Mori et al., 2005; Ye et al., 2005; Liu et al., 2006; Zajic et al., 2006; Filippova et al., 2007; Lin et al., 2007).

Applying the fuzzy logic for text summarization still needs more investigation; a few studies were done in this direction (Kiani and Akbarzadeh, 2006; Kyoomarsi et al., 2008). In our previous work (Binwahlan et al., 2009c), we found the integration of the fuzzy logic and swarm intelligence could get a good performance.

The improvement of the summary quality remains the key research problem and needs much work like incorporate more than one good technique. Aretoulaki (1994) proposed a hybrid system, the system was built

**Corresponding Author:** Mohammad Salem Binwahlan, Faculty of Computer Science and Information Systems,
University of Technology Malaysia, 81310, Skudai, Johor, Malaysia
Tel: +60177582859, +6075532208   Fax: +6075532210

based on four modules, where each module tries to look for specific features and information in the input text, then the outputs of those modules are passed to Artificial Neural Network (ANN) to score the text units as important and unimportant based on those outputs of the four modules. Alemany and Fort (2003) presented a summarize based on lexical chains, in which, the cohesive properties of the text were combined with coherence relations to produce good summaries. a different hybrid model was introduced by Cunha *et al.* (2007), which combines mainly three systems, each system produces its own extract, then an algorithm creates the final summary by selecting the highest score sentences from the three extracts after scoring of those extract sentences. The summary creation under the condition of the redundancy and the summary length limitation is a challenge problem. Therefore in this study, we introduce an intelligent model for automatic text summarization problem; we try to exploit different resources advantages in building of our model like advantage of diversity based method (Binwahlan *et al.*, 2009d) which can filter the similar sentences and select the most diverse ones and advantage of the non diversity method used in this study which is the adaptation of intelligent techniques like fuzzy logic and swarm intelligence for building that method which gave it a good ability for picking up the most important sentences in the text (Binwahlan *et al.*, 2009c).

## MATERIALS AND METHODS

**Swarm diversity based text summarization:** Swarm diversity based method (Binwahlan *et al.*, 2009d) is an integration of the two methods (MMI diversity based text summarization (Binwahlan *et al.*, 2009a) and swarm based text summarization (Binwahlan *et al.*, 2009b). In MMI method, the score of the sentence in the binary tree is calculated using:

$$Score_{BT}(s_i) = impr(s_i) + (1 - (1 - impr(s_i) \times friendsNo(s_i)))$$  (1)

where, $Score_{BT}(s_i)$ is the score of the sentence $s_i$ in the binary tree building process, $impr(s_i)$ is the importance of the sentence si calculated using normal features Eq. 6 and friends No $(s_i)$ is the number of sentences which are similar to sentence $s_i$.

In this equation the importance of the sentence appears in two positions, both two importances of the sentence are calculated using a simple combination of the text features score. The sentence importance in the second position was replaced by the sentence importance which is calculated using swarm weights based adjusted features scores. The new formula of scoring of the sentence in the binary tree became as the following:

$$Score_{BT}(s_i) = impr(s_i) + (1 - (1 - swarm\_impr(s_i) \times friends No(s_i)))$$  (2)

where: swarm_impr($s_i$) is the importance of the sentence si calculated using Eq. 7.

The reason of making the features are the centre of the integrating of the two methods is because the features are the cornerstone in the generation process of the text summary. The summary quality is sensitive for those features in terms of how the sentences are scored based on the used features.

**Fuzzy swarm based text summarization:** Fuzzy swarm based text summarization (Binwahlan *et al.*, 2009c) was implemented using Matlab fuzzy logic toolbox which contains built-in Mamdani's fuzzy inference method (Mamdani and Assilian, 1975). The trapezoidal membership function Eq. 3 was used for fuzzifing of the crisp numerical values of the five text features; those values are limited to the universe of discourse in the range (0, 1). The features values are adjusted using the weights resulting in the training of the Particle Swarm Optimization (PSO) (Kennedy and Eberhart, 1997); this forms the central point of merging of the fuzzy logic with swarm intelligence. To determine the degree to which the input values belong to each of the appropriate fuzzy sets, the trapezoidal membership function was used due to its simplicity and widely use. Three fuzzy sets were used: low, medium and high.

$$A_{ij}(x_j; a_{ij}, b_{ij}, c_{ij}, d_{ij}) = max\left(min\left(\frac{x_j - a_{ij}}{b_{ij} - a_{ij}}, 1, \frac{d_{ij} - x_j}{d_{ij} - c_{ij}}\right), 0\right)$$  (3)

where, $a_{ij} \le b_{ij} \le c_{ij} \le d_{ij}$ must hold.

The parameters a and d locate the feet of the trapezoid and the parameters b and c locate the shoulders.

The output of the trapezoidal membership function is a fuzzy degree of membership (in the range (0, 1)) in the fuzzy set. Figure 1 shows the membership functions of fuzzification of the input value of the sentence centrality feature (SC).
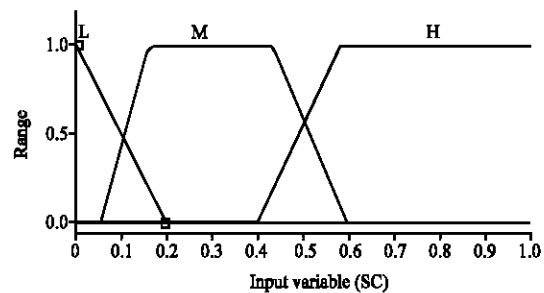


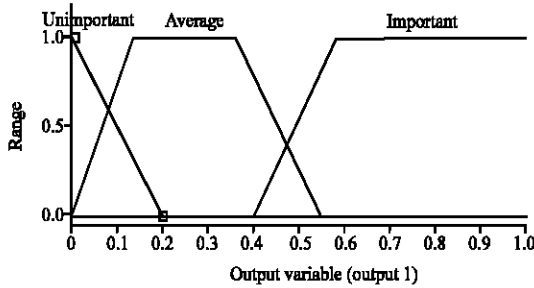Fig. 1: The trapezoid membership functions of the sentence centrality feature (SC)

Fig. 2: The trapezoid membership function of the output

For Inference process, the facts resulted in the fuzzification step need to be merged with a series of the production rules (IF-THEN rules) to perform the fuzzy reasoning process; around 200 IF-THEN rules were defined for that purpose. The following is an example for those rules.

[If WSS is H and SC is H and S_FD is M and SS_NG
is H and KWRD is H then output is important]

The output fuzzy membership function was the trapezoid membership function as shown in the Fig. 2.

For defuzzification which is to convert the fuzzy results of the inference into a crisp output which represents the final score of the sentence, we used the centroid method (Sivanandam *et al.*, 2006) Eq. 4, which returns the center (one crisp number) of the area under the curve of the output fuzzy set.

$$z = \frac{\sum_{j=1}^{q} Z_j u_c(Z_j)}{\sum_{j=1}^{q} u_c(Z_j)} \qquad (4)$$

where, z is the center of mass and $u_c$ is the membership in class c at value $Z_j$.

After getting the scores of all sentences produced by the fuzzy inferences system, the sentences are reranked based on those scores in descending order, then the top n sentences are selected as summary, where n is equal to the compression rate which is 20% of the total number of the document sentences.

**The intelligent model:** To propose our intelligent model, we try to investigate the performance of combining of two methods: swarm diversity based method (Binwahlan *et al.*, 2009d) and fuzzy swarm based method (Binwahlan *et al.*, 2009c).

The text features used in this method are: sentence centrality, Title-Help Sentence (THS) and Title-Help Sentence Relevance Sentence (THSRS), Word Sentence
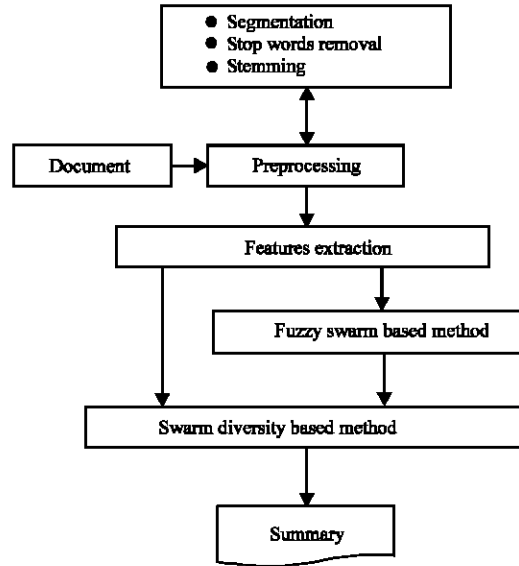


Fig. 3: Fuzzy swarm diversity model

Score (WSS), key word feature and the similarity to first sentence (Binwahlan *et al.*, 2009c).

**Our model consists of two modules:** Fuzzy swarm module which is in charge of selecting of the initial centroids which required for clustering of the sentences in the second module and swarm diversity module which is in charge of generating the summary after scoring the sentences, filtering the similar sentences and selecting the most diverse one. The task of the second module depends on clustering of the similar sentences; the clustering process uses the initial centroids determined by the fuzzy swarm module. In conclusion, there are two summaries created by our model, first summary generated by fuzzy swarm module and used as input for second module swarm diversity which uses its sentences as initial centroids for clustering process. Then the second module generates the final summary.

To summarize a document using this model, it is required first to cluster the document sentences into clusters (using k-means clustering algorithm) where each cluster contains the most similar sentences. The clusters number is determined automatically by the summary length (number of sentences in the final summary). The selection of the initial centroids is very important step in the clustering process, therefore we employed fuzzy swarm based method (Binwahlan *et al.*, 2009c) to pick up the most important sentences and these sentences are used as initial centroids. This is the central point of combining of the two methods (swarm diversity based method (Binwahlan *et al.*, 2009d) and fuzzy swarm based method (Binwahlan *et al.*, 2009c) (Fig. 3). Each sentences

cluster is represented as one binary tree or more. The first sentence which is presented in the binary tree is that sentence with higher number of friends (higher number of similar sentences), then the sentences which are most similar to already presented sentence are selected and presented in the same binary tree. The sentences in the binary tree are ordered based on their scores. The score of the sentence in the binary tree building process is calculated based on the importance of the sentence and the number of its friends using Eq. 5.

$$Score_{BT}(s_i) = impr(s_i) + (1-(1-swarm\_impr(s_i) \times friends \ No(s_i))) \quad (5)$$

$$impr(S_i) = avg(WSS(S_i) + SC(S_i) + SS\_NG(S_i) + sim\_fsd(S_i) + kwrd(S_i)) \quad (6)$$

where, WSS is word sentence score, SC is sentence centrality, SS_NG is average of THS and THSRS features and Sim_fsd is the similarity of the sentence $s_i$ with the first document sentence calculated using cosine similarity measure and $kwrd(s_i)$ is the key word feature.

$$Score(s) = \sum_{i=1}^{5} w_i \times score\_f_i(s) \quad (7)$$

where, Score (s) is the score of the sentence s, $w_i$ is the weighted of the feature i produced by swarm, i = 1 to 5 and score_$f_i$(s) is the score of the feature i.

Each level in the binary tree contains $2^{ln}$ of the higher score sentences, where ln is the level number, ln = 0,1,2,..,n, the top level contains one sentence which is a sentence with highest score. In case, there are sentences remaining in the same cluster, a new binary tree is built for them by the same procedure.

Swarm diversity method is used to select one sentence from the binary tree of each sentence cluster to be included in the final summary. In the binary tree, a level penalty is imposed on each level of sentences which is 0.01 times the level number. The purpose of the level penalty is to reduce the noisy sentences score. The summary sentence is selected from the binary tree by traversing all levels and applying swarm diversity Eq. 8 on each level sentences.

$$MMI(S_i) = Arg \max_{S_i \in CS \backslash SS} \left[ (Score_{BT}(S_i) - \beta(S_i)) - \max_{S_j \in SS} (Rel(S_i, S_j)) \right] \quad (8)$$

Where:

$$REL(s_i, s_j) = avg(nfriends(s_i, s_j) + ngrams(s_i, s_j) + sim(s_i, s_j)) \quad (9)$$

where, $Rel(s_i, s_j)$ is the relevance between the two competitive sentences, $s_i$ is the unselected sentence in the current binary tree, $s_j$ is the already selected sentence, ss is the list of already selected sentences, cs is the competitive sentences of the current binary tree, $\beta$ is the penalty level and $Nfriends(s_i, s_j)$ is the shared friends (the group of sentences which are similar to both sentences $s_i$ and $s_j$), $ngrams(s_i, s_j)$ the shared ngrams (the group of ngrams which are contained in both sentences $s_i$ and $s_j$) and the similarity between those two sentences.

## RESULTS

The DUC 2002 document sets (D061j, D062j, D063j, D064j, D065j, D066j, D067f, D068f, D069f, D070f, D071f, D072f, D073b and D077b) comprising 100 documents were used for testing the proposed model. ROUGE (Recall-Oriented Understudy for Gisting Evaluation) toolkit (Lin, 2004) is used for evaluation, where ROUGE compares a system generated summary against a human generated summary to measure the quality. ROUGE is the main metric in the DUC text summarization evaluations. It has different variants, in our experiment, we use ROUGE-N (N = 1 and 2) and ROUGE-L. In DUC 2002 document sets, each document set contains two model or human generated summaries for each document. We gave the names H1 and H2 for those two model summaries. The human summary H2 is used as benchmark to measure the quality of our proposed model summary, while the human summary H1 is used as reference summary. Beside the human with human benchmark (H2-H1) (H2 against H1); we also use another benchmark which is MS word summarize (Msword).

The evaluation is based on the generalizing of the results via confidence limits where the aim of generalization is to get a value which can express all values in the population of the results. For each summary, evaluation values (recall, precision and f-measure) are created using the evaluation measure ROUGE (Lin, 2004) which is standard measure. To measuring the performance of the proposed method, we check each evaluation value separately which is computationally expensive. Doing so is tough job and a waste of resources. The solution is to use the sample of results (summaries evaluation values) to calculate a range within which value in the evaluation is likely 95% to fall, which is the 95% confidence interval. The minimum and maximum values in that range are called the confidence limits, where the evaluation values are between the confidence limits. The ROUGE (Lin, 2004) generalizes the evaluation results using bootstrapping by resampling method.

We ran three experiments using the same data set, first is MMI diversity based method (Binwahlan *et al.*,

Table 1: MMI, swarm diversity, the proposed model, Msword summarize and H2-H1 comparison: average recall, average precision and average f-measure using rouge-1 at the 95%-confidence interval

| Methods | AVG-R | AVG-P | AVG-F |
|---|---|---|---|
| Msword | 0.39306 | 0.48487 | 0.42477 |
| MMI | 0.42288 | 0.48837 | 0.44420 |
| Swarm diversity | 0.42678 | 0.49368 | 0.44848 |
| Proposed model | 0.43141 | 0.48795 | 0.45060 |
| H2-H1 | 0.49657 | 0.49613 | 0.49605 |

Table 2: MMI, swarm diversity, the proposed model, Msword summarize and H2-H1 comparison: average recall, average precision and average f-measure using rouge-2 at the 95%-confidence interval

| Methods | AVG-R | AVG-P | AVG-F |
|---|---|---|---|
| Msword | 0.16325 | 0.21066 | 0.17947 |
| MMI | 0.18213 | 0.22242 | 0.19504 |
| Swarm diversity | 0.18629 | 0.22817 | 0.19971 |
| Proposed model | 0.18934 | 0.22305 | 0.20062 |
| H2-H1 | 0.20957 | 0.20940 | 0.20938 |

Table 3: MMI, swarm diversity, the proposed model, Msword summarize and H2-H1 comparison: average recall, average precision and average f-measure using rouge-L at the 95%-confidence interval

| Methods | AVG-R | AVG-P | AVG-F |
|---|---|---|---|
| Msword | 0.36605 | 0.45272 | 0.39604 |
| MMI | 0.38865 | 0.45131 | 0.40889 |
| Swarm diversity | 0.39261 | 0.45692 | 0.41330 |
| Proposed model | 0.39596 | 0.44987 | 0.41414 |
| H2-H1 | 0.46524 | 0.46490 | 0.46479 |

2009a), second is swarm diversity based method (Binwahlan *et al.*, 2009d) and third is the proposed intelligent model experiment. We present the three evaluation measures ROUGE-1, ROUGE-2 and ROUGE-L with the metrics recall, precision and f-measure.

The results are shown in the Table 1-3 using ROUGE-1, ROUGE-2 and ROUGE-L, respectively. We can see in table 1 for example there is a big difference between recall and precision; therefore, we will compare all methods based on the average f-measure evaluation which is a balance between recall and precision.

Based on the average f-measure evaluation for ROUGE-1, Table 1, the swarm diversity based method is better than MMI diversity based method. The proposed intelligent model outperforms the swarm diversity based method and MMI diversity based method. The benchmark Msword got less performance than the individual methods and the proposed intelligent model. The same thing can be said on the performance of the methods for ROUGE-2 in Table 2 and for ROUGE-L in Table 3.

## DISCUSSION

In this experiment, we have tested the hypothesis of exploiting of the advantages of different resources in form of a combined intelligent model could produce a good summary, it was not rejected. The obtained results of the proposed model supported and show the importance of

such intelligent model for solving the automatic text summarization problem. For comparison purposes, the experimental results of the proposed model testing have shown good performance. Based on f-measure evaluation, the proposed model outperforms MMI diversity based method (Binwahlan *et al.*, 2009a), swarm diversity based method (Binwahlan *et al.*, 2009d) and ms word summarize by: 0.0064, 0.00212 and 0.02583, respectively for rouge-1, 0.00558, 0.00091 and 0.02115, respectively for rouge-2 and 0.00525, 0.00084 and 0.0181 respectively for rouge-L. the main reason which made the proposed model performs better than the previous techniques is in the proposed model, much attention paid to the clustering phase by employing the fuzzy swarm based method to select the initial centroids, which is most important step in the clustering process. For future work, we plan to introduce a different hybrid model of the two techniques combined in the current model based on an idea of leaving each technique to create its own summary then the final summary will be generated by an algorithm which can pick up the important sentences from each summary and include them in the final summary. This requires evaluation mechanism which can rescore the sentences of each individual summary.

## ACKNOWLEDGMENT

## REFERENCES

Alemany, A.L. and M.F. Fort, 2003. Integrating cohesion and coherence for automatic summarization. ACL, Budapest, 2: 1-8.

Aretoulaki, M., 1994. Towards a hybrid abstract generation system. Proceedings of the International Conference on New Methods in Language Processing, (ICNMLP'1997), Manchester, pp: 220-227.

Barzilay, R. and M. Elhadad, 1997. Using lexical chains for text summarization. Proceedings of the Intelligent Scalable Text Summarization Workshop, August 1997, Madrid, Spain: ACL, pp: 10-17.

Baxendale, P., 1958. Machine-made index for technical literature - an experiment. IBM J. Res. Dev., 2: 354-361.

Binwahlan, M.S., N. Salim and L. Suanmali, 2009a. MMI diversity based text summarization. IJCSS Int. J. Comput. Sci. Security, 3: 23-33.

Binwahlan, M.S., N. Salim and L. Suanmali, 2009b. Swarm based text summarization. Proceedings of the International Conference on IACSIT Spring Conference, Apr. 17-20, Singapore, pp: 145-150.

Binwahlan, M.S., N. Salim and L. Suanmali, 2009c. Fuzzy swarm based text summarization. J. Comput. Sci., 5: 338-346.

Binwahlan, M.S., N. Salim and L. Suanmali, 2009d. Integrating of the diversity and swarm based methods for text summarization. Proceedings of the 5th Postgraduate Annual Research Seminar, Jun. 17-19, Johor, Malaysia, pp: 523-527.

Carbonell, J. and J. Goldstein, 1998. The use of MMR, diversity-based reranking for reordering documents and producing summaries. Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Aug. 24-28, Melbourne, Australia, pp: 335-336.

Conroy, J.M. and D.P. Oleary, 2001. Text summarization via hidden markov models. Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Sept. 9-12, New Orleans, Louisiana, United States, pp: 406-407.

Cunha, I., S. Fernández, P.V. Morales, J. Vivaldi, E. SanJuan and J.M. Torres-Moreno, 2007. A New Hybrid Summarize Based on Vector Space Model, Statistical Physics and Linguistics. In: MICAI 2007, LNAI 4827, 2007, A. Gelbukh and A.F. Kuri Morales (Eds.). Springer-Verlag, Berlin Heidelberg, pp: 872.

Edmundson, H.P., 1969. New methods in automatic extracting. J. Assoc. Comput. Machinery, 16: 264-285.

Fattah, M.A. and F. Ren, 2009. GA, MR, FFNN, PNN and GMM based models for automatic text summarization. Comput. Speech Language, 23: 126-144.

Filippova, K., M. Mieskes, V. Nastase, S.P. Ponzetto and M. Strube, 2007. Cascaded filtering for topic-driven multi-document summarization. Proceedings of the Document Understanding Conference, Apr. 26-27, Rochester, New York, pp: 30-35.

Kennedy, J. and R.C. Eberhart, 1997. A discrete binary version of the particle swarm algorithm. Proceedings of the IEEE International Conference on Computational Cybernetics and Simulation, (ICCCS'97), New York, pp: 4104-4108.

Kiani, B.A. and T.M.R. Akbarzadeh, 2006. Automatic text summarization using: Hybrid Fuzzy GA-GP. Proceedings of the IEEE International Conference on Fuzzy Systems, Jul. 16-21, IEEE Xplore Press, Vancouver, BC, Canada, pp: 977-983.

Kraaij, W., M. Spitter and M.V.D. Heijden, 2001. Combining a mixture language model and naive bayes for multi-document summarization. Proceedings of Document Understanding Conference, Sept. 13-14, New Orleans, LA, pp: 109-116.

Kupiec, J., J. Pedersen and F. Chen, 1995. A trainable document summarize. Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Jul. 9-13, New York, USA., pp: 68-73.

Kyoomarsi, F., H. Khosravi, E. Eslami, P.K. Dehkordy and A. Tajoddin, 2008. Optimizing text summarization based on fuzzy logic. Proceedings of the 7th IEEE/ACIS international Conference on Computer and information Science, May 14-16, IEEE Computer Society, Washington, DC. USA., pp: 347-352.

Lin, C.Y. and E. Hovy, 1997. Identifying topics by position. Proceedings of the 5th Conference on Applied Natural Language Processing, Mar. 31-Apr. 03, San Francisco, CA, USA., pp: 283-290.

Lin, C.Y., 1999. Training a selection function for extraction. Proceedings of the 18th Annual International ACM Conference on Information and Knowledge Management, Nov. 2-6, Kansas City, Kansas, pp: 55-62.

Lin, C., 2004. Rouge: A package for automatic evaluation of summaries. Proceedings of the Workshop on Text Summarization Branches Out, 42nd Annual Meeting of the Association for Computational Linguistics, Jul. 25-26, Barcelona, Spain, pp: 74-81.

Lin, Z., T. Chua, M. Kan, W. Lee, Q. L. Sun and S. Ye, 2007. NUS at DUC 2007: Using evolutionary models of text. Proceedings of the Document Understanding Conference, Apr. 26-27, Rochester, New York, pp: 1-8.

Liu, D., Y. Wang, C. Liu and Z. Wang, 2006. Multiple Documents Summarization Based on Genetic Algorithm. In: Fuzzy Systems and Knowledge Discovery, Wang, L. (Ed.). Springer-Verlag, Berlin, Heidelberg, pp: 355–364.

Luhn, H.P., 1958. The automatic creation of literature abstracts. IBM J. Res. Dev., 2: 159-165.

Mamdani, E.H. and S. Assilian, 1975. An experiment in linguistic synthesis with a fuzzy logic controller. Int. J. Man. Mach. Stud., 7: 1-15.

Marcu, D., 1998. Improving summarization through rhetorical parsing tuning. Proceedings of the 6th Workshop on Very Large Corpora, August 1998, Montreal, Canada, pp: 206-215.

Mori, T., M. Nozawa and Y. Asada, 2005. Multi-answer-focused multi-document summarization using a question-answering engine. ACM Trans. Asian Language Inform. Process., 4: 305-320.

Ono, K., K. Sumita and S. Miike, 1994. Abstract generation based on rhetorical structure extraction. Proceedings of 15th International Conference on Computational Linguistics, Kyoto, Japan, Aug. 5-9, Association for Computational Linguistics Morristown, New Jersey, USA., pp: 344-348.

Osborne, M., 2002. Using maximum entropy for sentence extraction. Proceedings of the ACL02 Workshop on Automatic Summarization, Jul. 11-12, Morristown, New Jersey, pp: 1-8.

Sivanandam, S.N., S. Sumathi and S.N. Deepa, 2006. Introduction to Fuzzy Logic using MATLAB. 1st Edn., Springer-Verlag, New York, ISBN-10: 3540357807, pp: 430.

Svore, K., L. Vanderwende and C. Burges, 2007. Enhancing single-document summarization by combining RankNet and third-party sources. Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, June 2007, Association for Computational Linguistics, Prague, pp: 448-457.

Ye, S., L. Qiu, T. Chua and M. Kan, 2005. NUS at DUC 2005: Understanding documents via concept links. Proceedings of Document Understanding Conference, Oct. 9-10, Vancouver, Canada, pp: 1-6.

Zajic, D.M., B.J. Dorr, R. Schwartz and J. Lin, 2006. Sentence compression as a component of a multi-document summarization system. Proceedings of the 2006 Document Understanding Workshop, Jun. 8-9, New York, pp: 1-7.