IJE TRANSACTIONS A: Basics Vol. 36 No. 07, (July 2023) 1219-1229

International Journal of Engineering

Journal Homepage: www.ije.ir

A Fuzzy Fusion Framework for Generating Purpose-oriented Texts

M. Tayefeh Mahmoudi*[a], K. Badie[b]

[a] *Data Analysis & Processing Research Group, IT Research Faculty, ICT Research Institute, Iran*
[b] *E-Content & E-Services Research Group, IT Research Faculty, ICT Research Institute, Iran*

PAPER INFO

Paper history:
Received 07 February 2023
Received in revised form 11 April 2023
Accepted 16 April 2023

Keywords:
Fuzzy Framework
Text Fusion
Informative Text
Class Prototype
Key Segment

ABSTRACT

Fusing textual information, as type of information fusion, has been of great significance to those interested in making informative texts out of the existing ones. The main idea behind text fusion, like any other type of information fusion, is to merge the partial texts from different sources in such a way that the outcome can hold a reasonably high relevance with regard to certain objectives. In this paper, a fuzzy framework is proposed for text generation, according to which a range of relevant texts are merged to yield producing a new text that can help the users fulfill a certain functionality in plausible manner. The focal point in our approach with regard to fusion is the distance between the class prototype of a text on the one side and the feature vectors belonging to different subsets of the existing texts on the other side. Results of experiments, show that the suggested framework can be a suitable alternatives for performing fusion in the cases that the identity of the existing texts from the viewpoint of the texts considered is unclear. This would turn into an effective utilization of the existing texts for the purpose of generating new texts.

doi: *10.5829/ije.2023.36.07a.03*

1. INTRODUCTION

In recent years, text fusion has been of great significance to those interested in making informative texts out of the existing ones [1]. Texts fusion, in this way, has got numerous applications with regard to issues such as text summarization and text mining as well [2, 3]. The main idea behind text fusion, like any type of information fusion, is to merge partial texts from different sources in such a way that the outcome can hold a reasonably high relevance with regard to certain objectives.

Provided that organizational texts can help their users perform their functionalities in a reasonable way, one interesting application of text fusion is to merge the existing texts in such a manner that the outcome can hold such characteristic. For instance, if the functionality under consideration is "planning", the purpose of text fusion would be to provide a new text through merging the existing one, that can help planners do their functionality of planning in a plausible manner. Since the functionalities of a text may occasionally share similar aspects that could reflect in the way their essential segments show up in the corresponding texts, we therefore need to make use of those approaches to fusion which can handle the ambiguity or uncertainty which may arise due to such sort of similarity in functionalities. In this sense, fuzzy logic can be considered as a helpful means for handling this uncertainty or ambiguity. Taking this point into account, a fuzzy framework is proposed in this paper for text fusion according to which a variety of texts are to be selected whose merging can yield a high expectation for helping the user fulfill a certain functionality.

The main concern in our approach is the distance between the prototypical classes of text considered for a variety of pre-defined functionalities on the one side, and the feature vectors belonging to different subsets of the existing texts, on the other side. Feature vectors in our approach are represented in terms of a number of membership degrees, each standing for the status of affiliation toward a certain feature in the desired class of text. In this paper, to evaluate the performance of our proposed framework we decided to compare it with the performance of OWA operators which are well- known

*Corresponding Author Email: mahmodi@itrc.ac.ir (M. T. Mahmoudi)

Please cite this article as: M. Tayefeh Mahmoudi, K. Badie, N. Rostamzadeh, A Fuzzy Fusion Framework for Generating Purpose-oriented Texts, ***International Journal of Engineering, Transactions A: Basics***, Vol. 36, No. 07, (2023), 1219-1229

for information fusion purposes. Therefore, beside the proposed framework as a novel approach to text fusion, a part of the novelty in our work lies in examining the capability of OWA operators in fusing the textual information from the viewpoint of synthesizing a new text that can meet certain goals. Results of experimentation show that the suggested framework can be regarded a suitable alternative for fusion in the case that the identity of the existing texts from the viewpoint of the considered functionalities is not clear.

2. RELATED WORKS

By text integration or fusion, we mean to blend or fuse some pre-experienced texts to finally form a new text which can meet certain goals and /or certain criteria. A classical approach to integrating text can lie in multi-document text summarization according to which a final summery is to be made based on the information encapsulated in each document [4]. The main assumption in case of such an approach is that, each document may include its own peculiar information that can contribute to the final summery. Many methods have been proposed for multi-document text summarization out of which the recent works based on using deep neural networks are of particular significance [5]. The significance of deep neural net based text summarization mostly goes back to the fact that past experiences of summarization through joining different parts (from different documents) would have the ability to realize a new summary in a promising way. Another perspective in text integration can be met in integrating and blending concepts for creative generation of stories [6, 7] and automatic story telling [8] specially for generating game scenarios. Within this context, those who design games may generate and also explore some scenarios of game with high complexity through three main stages of domain implementation, solution generation and story board generation based on a kind of composition between the related items [9]. It seems that the main advantage of the proposed framework, compared to deep neural network-based document summarization, is no need for training patterns due to its dependence on the presence of linguistically-significant notions (such as What, Where, When, …) while considering some membership functions. In fact the effect of the very training which is essential to operationalizing deep neural network is somewhat encapsulated in the corresponding membership functions.

Fusion of dynamic web contents/ services through an extreme personalization as well as heterogeneous devices together with a interaction channels using content caching and adaptive aggregation algorithms as well as fuzzy utility based patterns mining are typical examples for fusion [10, 11].

Semantic data/information integration using fuzzy rules may also be widely used for text fusion [12], datasets integration [13] and information fusion [14] as well. Here, the relevant information is fused on the basis of a kind of logical reasoning to enhance the power of current semantic web systems [15] and search engines [16]. In this respect, the process of fuzzy conceptual matching can help an efficient retrieval of intelligent information and knowledge, and would be able to be integrated into the other commercial search engines [16]. Information fusion under fuzzy environment can also be applied for retrieving fuzzy information, whose fuzzy numbers are to represent the strength degrees according to which criteria of priority are met for the documents [17]. On the basis of the concept of fusing enterprise information, a kind of neural network has been presented which functions on the ground of choquet fuzzy integral [18].

Beside the above approaches, compositional adaptation techniques used in case-based reasoning have also the ability to combine the experienced sources (of information), like book chapters or texts to be used in learning in the form of a new configuration that can serve the current problems of teaching or research/ development support in a plausible manner [19, 20]. Although the propositions used for presenting the resources have crisp nature, it can however be possible to make use of fuzzy logic in these approaches both at representation and inference levels to provide the final solution in a favorable way.

Concept composition can also be considered as an approach to creating textual contents [21, 22].

In the meantime, due to the achievements of issues like semantic web, many researchers have tended to developing digital libraries based on ontological structures to share exchange and retrieval information efficiency [23]. Moreover, fuzzy queries have also been taken into account to build their queries in a more precise way in order to give assistance to readers for searching information [24]. In this regard, fuzzy ontology is utilized to present uncertain information in the case of digital libraries whereas fuzzy queries are used to retrieve information from this ontology. It is obvious that, using such a system can give more precise integrational results [24, 25]. It should also be mentioned that, collecting large amount of documents, in the way delivered by the search engines in Internet, is difficult and time-consuming for users to read and analyze. Here, the burden of managing information can be greatly eased through detecting common and distinctive topics within a set of documents as well as generating multi-document summaries. To achieve this goal, we first apply the well-known ordered weighted averaging (OWA) fusion operator to text fusion while examining the results, and then propose a new fuzzy framework according to which the input text is transformed into a set of membership degrees with regard to a set of nominal values, and an appropriate distance function is then considered to measure the distance between the feature vector comprising these membership degrees and the ideal prototype vectors already defined for the prototype classes of a text. In both cases (applying OWA operators and proposed framework) possible types of texts (research, development, analysis, …) are represented in terms of a variety of key segments whose nominal values (L, M, H) take part in characterizing the types of the texts. Here, linguistically significant notions like What, Where, When , Who, … are used to determine the status of each nominal value. We found out through experimentation that the results obtained through the new framework is quite close to those obtained through applying AVG OWA. Moreover it was perceived that the proposed framework functions better than i) MIN OWA in the sense of avoiding the items in the resulted text which are not necessarily compatible with the requirements of the desired key segment, and ii) MAX OWA in the sense of observing some crucial items in the resulted text which are usually neglected by this operator mostly due to the point that text fusion through MAX OWA calls for severe constraints. The benefit of such an approach goes back to the fact that the input texts are treated in a generic way thus providing a chance for the system to merge them in a way as generic as possible.

3. THE PROPOSED APPROACH

3. 1. Basic Concepts The basic idea behind our approach is based on the point that there exists a number of key segments in a text whose status of emphasis gives sense to the type of functionality which is expected to be performed by it when a text is exposed to its user. A partial ontology of these key segments that ought to be observed by texts authors in some way is shown in Figure 1 [26]. This ontology has in reality been designed to describe each key segment in depths; in such a way that the entire key segments can be differentiated properly in the input text.

Taking this ontology into account, each class of text functionality can be represented in terms of the related key segments as features and their values as the feature's values. The feature's values in our approach are Low (L),

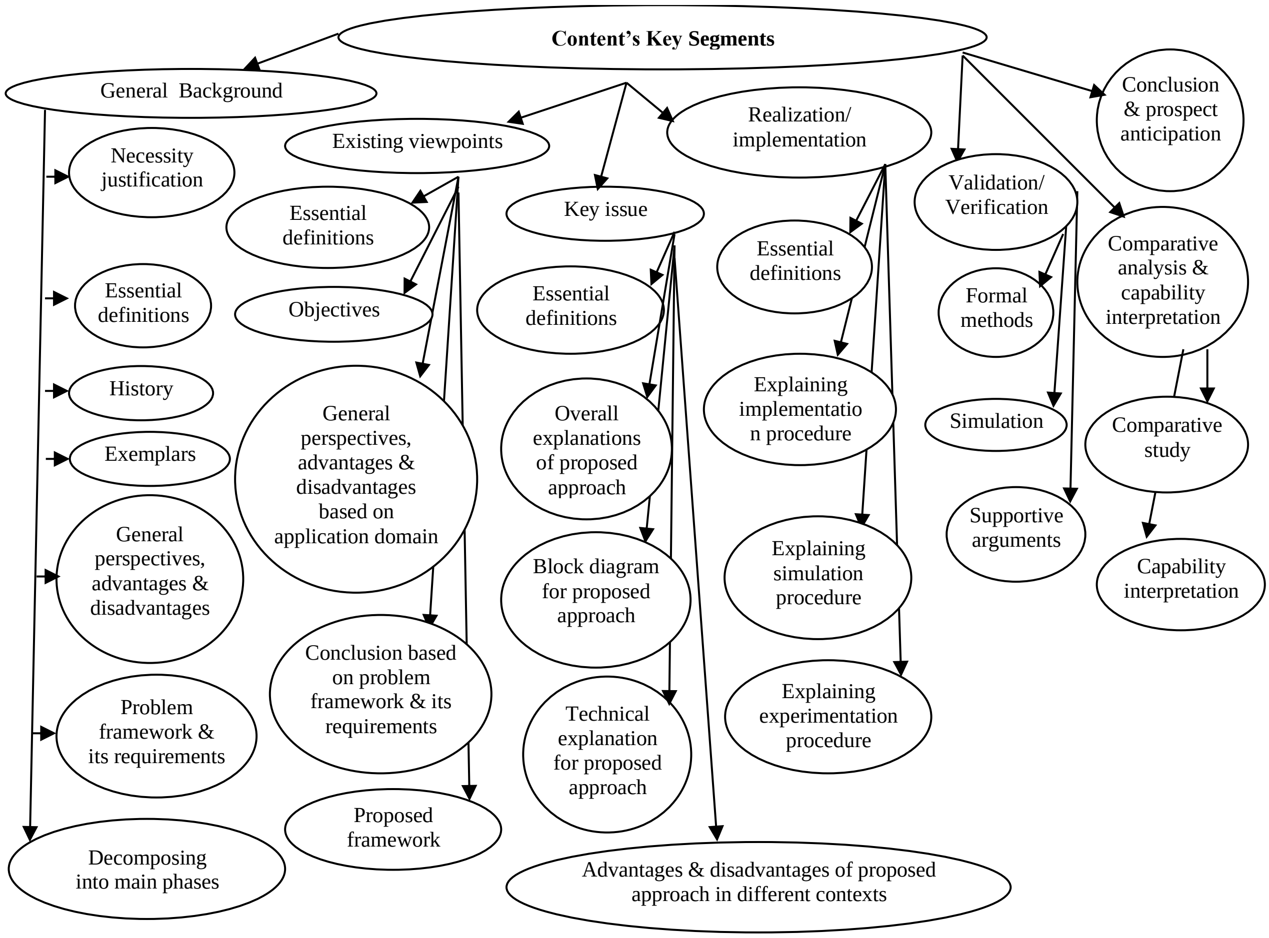

Figure 1. The Ontology of Text's Key Segments

Medium (M) and High (H), which are decided based on the status of the explanations regarding linguistically-significant notions such as What, Which, Who, Whom, Where, When, How and Why that are the items to be tackled in generating texts. Having represented classes of text functionality in the above manner, our main objective would be to find a certain combination, out of the existing texts, whose similarity degree regarding the corresponding features' values in the desired class of functionality as the purpose of fusion can be sufficiently high. To show the similarity degree, in our approach, membership degrees of the existing features' values with regard to the prototypical values for the same features in the desired class of text functionality are taken into account. These prototypical values are in reality the quantitative versions of the linguistic variable values "L", "M" and "H" which have been considered to indicate the status of the existing key segments (General Background, Existing Viewpoints, …) for different text functionalities. In our approach we decided to consider "1" for "L", "2" for "M" and "3" for "H" as the simplest version. The ground for calculating such a measure is decided to be the total number of the predicates as well as the arguments, which tackle the afore-mentioned linguistically- significant notions for each key segment as an feature. This is because, provided that these notions are significant from the view-point of a text's status of richness, the status of the related predicates/arguments can naturally be a good indicator to show how far it can be regarded close to the corresponding feature value in the desired class of text functionality. To show that a certain combination of texts (with certain functionalities), belongs adequately to a certain class of text functionality, in our approach we make use of a number of criteria that tackle respectively (i) the fact that the similarity of this combination to the desired class ought to be more than the corresponding similarities for the other combinations, and (ii) the fact that this similarity ought to be more than a certain threshold.

3. 2. The Proposed Approach Based on Text Fusion The main point in our framework is merging a range of relevant texts to yield producing a new text that can help the users fulfill a certain functionality. It would thus be important to first select those texts as the input texts whose topics are somewhat similar in some aspects. Having done so, the next step would be to identify the significant key segments in each input text through detecting the entities which have been realized to be essential for each key segment as illustrated in the ontology of Figure 1. To perform this task, linguistically significant notions such as What, Where, Who, … are searched for, to determine the status of membership degrees with regard to the considered linguistic variable values "L", "M" and "H" for the corresponding key segments. As the next step, the possible combinations of the feature vectors comprising these membership degrees are constituted, and out of these combinations those which hold the minimum distance toward the prototypical feature vectors predefined for each text functionality, are selected as the most appropriate alternatives for fusion. Below, the above mentioned steps are discussed in detail. Figure 2 illustrates the essential steps to the proposed framework for fusion.

3. 2. 1. Text Functionalities and Status of Linguistic Variable Values for the Selected Features With respect to fusion issues, identifying the features of those texts that are to be fused, is of great importance.

To facilitate the fusion process, the major key segments (Figure 1) are used as the features: General Background, Existing viewpoints, Key issue, Proposed approach realization/ implementation, Validation/ Verification, Comparative analysis & capability interpretation, Conclusion & prospect anticipation. We have realized that such features are consistent for a variety of texts whose functionality is to help users with their tasks in the corresponding organizations. Interestingly the key segments discussed above are also popular among many knowledge workers like those involved in research, innovation, development, planning and analysis issues with the final goal of disseminating their works' results in terms of suitable texts. Examples for the important functionalities in an organization, are Planning/ Scheduling, Research, Innovation, Development/ Optimization/ Improvement. Education/ Promotion, Analysis/ Assessment/ Assurance, Guidance, and Justification.

As far as linguistically significant notions are concerned, "L" (Low), "M" (Medium) and "H" (High)

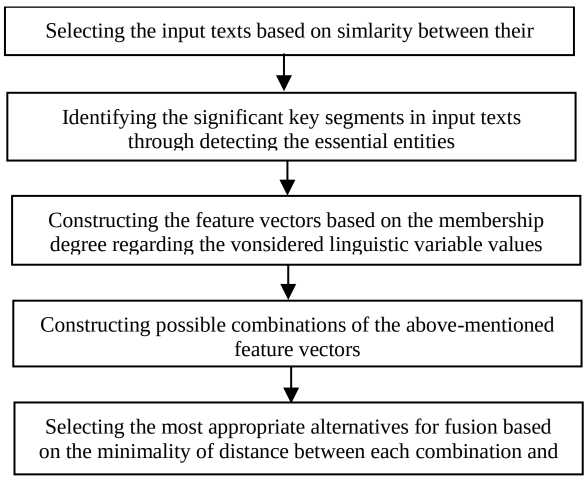

Figure 2. The steps essential to the proposed framework for fusion

stand for the extent that notions like "What", "Who", "Whom", "Where", "When", "How", and "Why", are to be addressed to create a partial text for each of the afore-mentioned key segments. For example with regard to the key segment "General Background", the items with regard to "Necessity", "Essential Definitions", "History", "Exemplars", "General Perspectives", "Advantages & Disadvantages", "Requirements of Problem Framework", and "Main Phases" are considered as "Whats", the agents who have taken part in the "History" of an issue are considered as "Whose", the space and time where a "framework" can function well are considered as "Wheres" and "Whens", the way a problem framework works as well as the way a problem is "decomposed into main phases" are considered as “Hows", and finally the reason that the "necessity" of an issue is justified as well as the reasons a "framework" may or may not function properly with regard to certain situations are considered as "Whys". Referring to Figure 1, the interpretation of "Whats", "Whose", "Whoms", "Wheres", "Whens", "Hows" and "Whys" can be equally performed for the other types of key segment taking into account their hyponyms. It is to be noted that the nominal values already agreed for each functionality are responsible for doing this task to indicate how far the above-mentioned notions ought to be addressed. Table 1 illustrates the status of the corresponding nominal values with regard to these notions.

We may see from the table that the depth of a notion depends on the amount of a nominal value. Let us say, the notion "How" is also called for when we have a transition from the nominal value "L" to the nominal value "M", and in the same manner, by transition from "M" to "H", the notion "Why", which organically stands for an explanation with more depth, would be worth being added.

Taking this point into account, for each feature, the fuzzy membership functions can be considered based on the nominal values already discussed. Table 2 shows the corresponding nominal values for the selected text features. The membership functions for these values are also shown in Figure 3.

TABLE 1. Status of the nominal values

Nominal Value	Status of Linguistically Significant notions
L (Low)	Addresses What, Where, When, Who , and Whom
M (Medium)	Addresses What, Where, When, Who, Whom, and How
H (High)	Addresses What, Where, When, Who, Whom, How, and Why

TABLE 2. Status of linguistic variables values for the selected features taking into account the existing functionalities

Input Text Features	Research	Development / Planning	General Learning	Justification	Innovation	Analysis/ Assessment
General Background	H	M	L	L	M	L
Existing Viewpoints	H	M	H	L	M	L
Key Issue	H	M	M	M	M	M
Proposed Approach Realization/ Implementation	H	M	L	M	L	M
Validation/ Verification	H	M	L	M	L	H
Comparative Analysis & Capability Interpretation	H	M	L	L	L	L
Conclusion & Prospect Anticipation	H	H	L	L	L	L

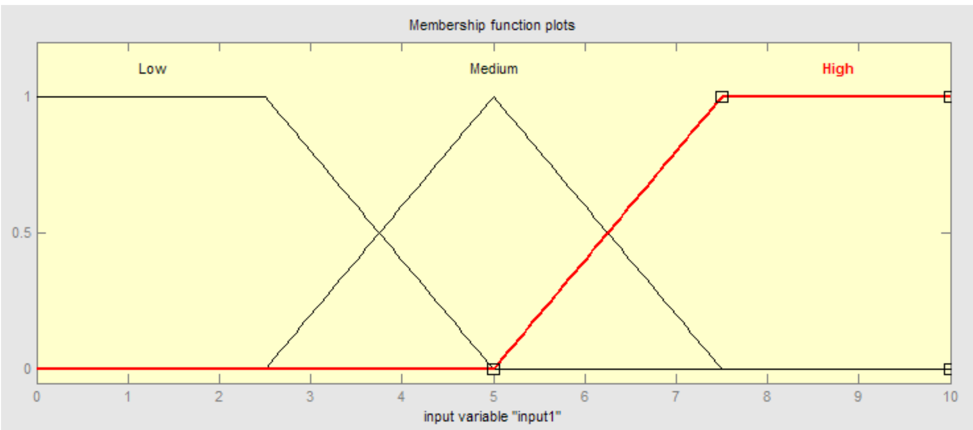

Figure 3. Membership functions for each of the linguistic variable values for the features in the text

3. 2. 2. Process of Fusion Using Weights In order to perform fusion, we first need to make a membership degree matrix for each input text document within which each row represents the status of membership degrees for the related key segment. Figure 4(a) illustrates such a matrix for the i-th and jth text documents.

To fuse these texts documents, all their combination forms have to be considered. For instance, if there are three texts documents to be fused, the possible combinations would be as follows:

(doc i) , (doc j), (doc k), (doc i , doc j), (doc I, doc k), (doc j, doc k), (doc I, doc j, doc k)

Let us consider the combination of the ith and jth text documents. Taking into account the matrix representation for each text (as shown in Figure 4(a)), the result of fusion can be represented as the matrix of Figure 4(b).

To obtain a text suitable for a certain functionality, matrix of each texts' functionalities, which is represented in Table 2, has to be compared with the equivalent membership degrees from the matrix of document.

Suppose that our purpose is to select a combination of texts which can best fit the functionality of "research".

To figure out the best combination for fusion, the following conditions should be satisfied as the necessary criteria.

Condition 1: distance (Research, Xi) < distance (Development, Xi), distance (Learning, Xi),…..

Condition 2: (distance (Development, Xi)- distance (Research, Xi)) + (distance (learning, Xi) – distance (Research, Xi))... → Max

Based on the type of the functionality expected from fusion, the coefficients called "importance coefficient" is to be considered for different functionalities, in order to improve the process of selecting the best combination. To achieve this, the conditions are changed to the following forms:

$$d(R, X_i) = (d_1^{i_R}, d_2^{i_R}, \ldots\ldots, d_7^{i_R})$$

$$\sum_{j=1}^{7} w\, d_j^{i^2{}_R, X}$$

$$\sum_{j=1}^{7} w\, d_j^{i^2{}_R} < \sum_{j=1}^{7} w\, d_j^{i^2{}_D}, \sum_{j=1}^{7} w\, d_j^{i^2{}_P}, \ldots\ldots$$

$$\left(\sum_{j=1}^{7} w\, d_j^{i^2{}_D} - \sum_{j=1}^{7} w\, d_j^{i^2{}_R}\right) + \left(\sum_{j=1}^{7} w\, d_j^{i^2{}_P} - \sum_{j=1}^{7} w\, d_j^{i^2{}_R}\right) + \ldots\ldots --> \max$$

To determine the importance coefficient (w), suppose that WL=1, WM=2, WH=3, for each function, W should be calculated per each feature in a way that ∑w =1 can be satisfied.

For instance, weights for each feature can be calculated as follows:

W_H: 3 3 3 3 3 3 3
Research : [H H H H H H H]
$W_{Research}$: [3/21 3/21 3/21 3/21 3/21 3/21 3/21]

While for Learning the same weights would better be calculated as follows:

$W_{L,M,H}$: 1 3 2 1 1 1 1
Learning:[L H M L L L L]
$W_{Learning}$: [1/10 3/10 2/10 1/10 1/10 1/10 1/10]

It is to be noted that Ws can also be determined by OWA. For this purpose, three forms of OWA for the discussed importance coefficients are considered as follows [27, 28]:

- Max: or oring aggregation operator is given by: W' = [l, 0,..., 0]
- Min: or anding aggregation operator is given by: W' = [0, 0,..., 1]
- Average: W' = [1/n, 1/n,..., 1/n]

In this approach, the best combinations for the existing functions are calculated and upon the requested objective, the combination results are represented and compared with the results achieved for applying OWA.

$$\begin{array}{cc} \text{Text i} & \text{Text j} \\ \begin{pmatrix} \mu_L^{j,1} & \mu_M^{j,1} & \mu_H^{j,1} \\ \vdots & & \\ \mu_L^{j,q} & \mu_M^{j,q} & \mu_H^{j,q} \end{pmatrix} & \begin{pmatrix} \mu_L^{i,1} & \mu_M^{i,1} & \mu_H^{i,1} \\ \vdots & & \\ \mu_L^{i,q} & \mu_M^{i,q} & \mu_H^{i,q} \end{pmatrix} \end{array}$$

(a)

$$\begin{pmatrix} Max(\mu_L^{i,1}, \mu_L^{j,1}) & Max(\mu_M^{i,1}, \mu_M^{j,1}) & Max(\mu_H^{i,1}, \mu_H^{j,1}) \\ \vdots & & \\ Max(\mu_L^{i,q}, \mu_L^{j,q}) & Max(\mu_M^{i,q}, \mu_M^{j,q}) & Max(\mu_H^{i,q}, \mu_H^{j,q}) \end{pmatrix}$$

(b)

Figure 4. (a) Matrix for the i-th and jth text documents (b) Fusion of matrix i-th and jth text documents

4. EXPERIMENTAL RESULTS

To validate the performance of the proposed framework, a number of texts with unclear nature from the view-point of the desired class of functionality were selected as the input texts. These texts are the research papers addressed in the literature [29, 30] as $text_1$, $text_2$ and $text_3$. Let us say, $text_1$ is on “Particle Swarm Optimization” (as mentioned in the literature [29]), $text_2$ is on “Evolving the Structure of the Particle Swarm Optimization Algorithms” (as mentioned in the literature [30]), and $text_3$ is on “Optimization of Nonlinear Constrained Particle Swarm” (as mentioned in the literature). We selected these research papers as input texts since they address the same topic of "particle swarm", so as to conduct fusion in a way that can make sense with regard to the considered text's key labels (General Background, Existing Viewpoint, Key Issue, …). Experimentation was done in such a manner that, each input text was investigated from the perspective of the key segments as the features discussed in the paper, taking into account the status of the nominal values belonging to them. As discussed in section 3.2.1, these nominal values stand for the extent that notions like "What", "Who", "Whom", "Where", "When", "How" and "Why" are addressed in the way already discussed with regard to the ontology of the key segments in Figure 1.

Result of experimentation for each feature has been shown in Table 3 by a separate color for the three input texts. The number of predicates/ arguments detected for each feature in the input texts is also mentioned in Table 3.

Having normalized the given numbers in Table 4 into the range of [1 10] using

X= Lower band+ (Upper band – Lower band) α /Max

process is performed based on the weights that have been explained in the previous section.

In this example, as the presupposed range is [1 10] and the Max value is 200, the normalizing formula would be:

$x=1+(10-1)\alpha /200)$

Simulation results are presented in Table 4. As it is seen, each number defines the status of the texts that are appropriate to be combined for a certain functionality.

The feature values for the selected texts for fusion are as follows:

General Background, Existing viewpoints, Key issue, Proposed approach realization/ Implementation, Validation/ Verification, Comparative analysis & capability interpretation, Conclusion & prospect anticipation.

$Text_1$: [50, 10, 35, 26, 200, 15, 4] → Normalized $Text_1$: [3.25, 1.45, 2.57, 2.17, 10, 1.67, 1.18]
$Text_2$: [4, 21, 37, 70, 0, 14, 5] → Normalized $Text_2$: [1.18, 1.94, 2.66, 4.15, 1, 1.63, 1.22]
$Text_3$: [7, 11, 18, 45, 0, 29, 8] → Normalized $Text_3$: [1.31, 1.49, 1.81, 3.02, 1, 2.30, 1.36]

The orders of certain functionalities which have been considered as the purpose of fusion are respectively: research, development, learning, justification, guidance, innovation and analysis.

The results of fusion with the mentioned feature values are as follows. As it is illustrated, to generate text for "research", combination of $Text_2$ and $Text_3$ are appropriate according to the proposed approach, while to generate a text for "development/planning", the combination of $Text_1$ & $Text_3$ would be suitable. In the same way, we may have different combinations of texts based on the type of the desired functionality.

Results of applying Max OWA reveal that, if the objective is to have a text for "research", $Text_2$ is more suitable, while for having a text for "development/planning", $Text_1$ is appropriate.

Applying Min OWA shows that the combination of $Text_1$ and $Text_2$ leads to a proper result once "research" has been considered as the desired text functionality. Finally, applying AVG OWA, illustrates the same results as those obtained by applying the proposed framework. Figure 5 shows the results obtained.

For the moment, to combine the texts, the parts corresponding to the related features are just added together, paying no attention to the fact that they may have some common parts. This may lead to a possible presence of somewhat similar parts in the text obtained as the result of fusion. To circumvent this problem, a variety of techniques such as semantic information

TABLE 3. Number of predicates for each feature in selected Texts

	Text Functionalities		
Text's Key Labels	**$Text_1$ [29]**	**$Text_2$ [30]**	**$Text_3$ [31]**
General Background	50	4	7
Existing Viewpoints	10	21	11
Key Issue	35	37	18
Proposed Approach Realization/ Implementation	26	70	45
Validation/ Verification	200	0	0
Comparative Analysis & Capability Interpretation	15	14	29
Conclusion & Prospect Anticipation	4	5	8

TABLE 4. Definition of simulation results

Numerical Result	**Suitable Text**
0	None of the combinations are suitable for the desired functionality
1	$Text_1$ is suitable for the desired functionality
2	$Text_2$ is suitable for the desired functionality
3	$Text_3$ is suitable for the desired functionality
4	The combination of $Text_1$ and $Text_2$ are suitable for the desired functionality
5	The combination of $Text_1$ and $Text_3$ are suitable for the desired functionality
6	The combination of $Text_2$ and $Text_3$ are suitable for the desired functionality
7	The combination of $Text_1$ and $Text_2$ and $Text_3$ are suitable for the desired functionality

Each Number Shows Which Combination of Texts Are Appropriate for Purpose-Oriented Fusion

0: None of the Texts is good for fusion
1: Text1 is good for fusion
2: Text2 is good for fusion
4: Text1 & Text2
5: Text1 & Text3
6: Text2 & Text3
7: Text1 & Text2 & Text3
The results of Proposed Approach for Fusion :
6 5 0 0 0 0 0

The results of Max OWA:
2 1 0 0 0 0 0

The results of Min OWA:
4 0 0 0 0 0 0

The results of AVG OWA:
6 5 0 0 0 0 0

Figure 5. Implementation Results for Fusion

processing as well as natural language processing may come effective to resolve this kind of overlap or redundancy. Approaching this problem can be considered as a continuation of the present work in future.

Figure 6 shows a part of the result of text fusion that has been obtained through the above mentioned fusion methods. It should be noticed that our proposed framework for fusion in some sense can be regarded as an approach to summarization since its main concern is to pick out those parts in the input text whose significance is in the way they comply with some desired functionalities. Let us say, it can be regarded as a kind of multi-document summarization with respect to certain functionalities for a text (research, planning, assessment, …) taking into account text's key labels (such as "General Background", "Existing Viewpoints", "Proposed Approach Realization Implementation", …) as the important aspects in a summery.

Research Text

Proposed Approach: Fusion of $Text_2$ [30] and $Text_3$

General Background:

The partial texts obtained through applying our framework to $Text_2$ [30]

Particle Swarm Optimization (PSO) is a population based stochastic optimization technique well developed in literature [8]. Standard PSO algorithm randomly initializes a group of particles (solutions) and then searches for optima by updating all particles along a number of generations. In any iteration, each particle is updated by following some rules [16]. Standard model implies that particles are updated synchronously [16]. This means that the current position and speed for a particle is computed taking into account only information from the previous generation of particles.

The partial texts obtained through applying our framework to $Text_3$

The particle swarm optimization algorithm (PSOA) was firstly proposed in literature [4, 5] and has deserved some attention during the last years in the global optimization field. PSOA is based on the population of agents or particles and tries to simulate its social behavior in optimal exploration of problem space. During time (iterations in the optimization context) each agent possesses a velocity vector that is a stochastic combination of its previous velocity and the distances of its current position to its own best ever position and to the best ever swarm position. The weights of the last two directions are controlled by two parameters called cognitive and social parameters [6].
PSOA belongs to a class of stochastic algorithms for global optimization and its main advantages are the easily parallelization and simplicity. PSOA seems to outperform the genetic algorithm for some difficult programming classes, namely the unconstrained global optimization problems [6]. In spite of the referred advantages, PSOA possesses some drawbacks, namely its parameters dependency and the slow convergence rate in the vicinity of the global minimum.

MAX OWA: $Text_2$ [30]

General Background:

The partial texts obtained through applying MAX OWA to $Text_2$ [30]

Particle Swarm Optimization (PSO) is a population based stochastic optimization technique well developed [8]. Standard PSO algorithm randomly initializes a group of particles (solutions) and then searches for optima by updating all particles along a number of generations. In any iteration, each particle is updated by following some rules [16]. Standard model implies that particles are updated synchronously [16]. This means that the current position and speed for a particle is computed taking into account only information from the previous generation of particles.

MIN OWA: Fusion of $Text_1$ [29] & $Text_2$ [30]

General Background:

The partial texts obtained through applying MIN OWA to fuse $Text_1$ [29] & $Text_2$ [30]

Although the principle of Particle Swarm Optimization is a quite new approach, its ancestors reach back into history as it emerged from biological research and simulation on swarming animals. The first computer-related work in this area was provided by Craig Reynolds which published a paper about simulating bird swarms in literature [6]. The idea was to simulate realistic swarms mainly for computer graphics and movies. The result was some simulated swarm of whose the individuals called \Boids\ [7]. These were directed by three simple rules, which were implemented and caused a near-realistic swarming behaviour:

- Separation: Do not run into flockmates
- Alignment: Align the own heading to the average of the neighbours
- Cohesion: Move toward the average position of neighbours.

For making this possible a new software was written named MASSIVE which controls this mass of agent technology-equipped computer actors (CGIs) and their states.

A quite funny anecdote about this battle sequence is that in the early testing-runs it was working way too good. The directors noticed some group of orcs which fled the battle because they were too scared. This was adjusted later on, as orcs are said to fear nothing at all...

The concepts embedded into the Boids were refined and later led into some new area of computer graphics which is called behavioral animation. The most impressive usage are probably the immense battle sequences in the trilogy Lord of the Rings where about 250,000 individual fighters.

AVG OWA: Fusion of $Text_2$ [30] and $Text_3$
General Background:

The partial texts obtained through applying AVG OWA to fuse $Text_2$ [30] & $Text_3$
From $Text_2$ [30]

Particle Swarm Optimization (PSO) is a population based stochastic optimization technique developed in literature [8]. Standard PSO algorithm randomly initializes a group of particles (solutions) and then searches for optima by updating all particles along a number of generations. In any iteration, each particle is updated by following some rules [16]. Standard model implies that particles are updated synchronously [16]. This means that the current position and speed for a particle is computed taking into account only information from the previous generation of particles.

From $Text_3$

The particle swarm optimization algorithm (PSOA) was firstly proposed in literature [4, 5] and has deserved some attention during the last years in the global optimization field. PSOA is based on the population of agents or particles and tries to simulate its social behavior in optimal exploration of problem space. During time (iterations in the optimization context) each agent possesses a velocity vector that is a stochastic combination of its previous velocity and the distances of its current position to its own best ever position and to the best ever swarm position. The weights of the last two directions are controlled by two parameters called cognitive and social parameters [6].

PSOA belongs to a class of stochastic algorithms for global optimization and its main advantages are the easily parallelization and simplicity. PSOA seems to outperform the genetic algorithm for some difficult programming classes, namely the unconstrained global optimization problems [6]. In spite of the referred advantages, PSOA possesses some drawbacks, namely its parameters dependency and the slow convergence rate in the vicinity of the global minimum.

Figure 6. The results of Texts Fusion

5. CONCLUDING REMARKS

A fuzzy framework was presented in this paper, which is capable of fusing the input texts in such a manner that the outcome can have the claim of belonging to certain text functionalities by considering a number of text's key labels. As discussed in the paper, each class of text functionality is represented in terms of some key segments as features and the related values, which are decided based on the status of the explanations regarding some linguistically- significant notions. It was discussed that the membership degrees of the features' values in a combination of texts is calculated based on the total number of the predicates/ arguments that tackle them for each key segment as a feature. To show that a certain combination of texts belongs sufficiently to a certain class of functionality, we made use of a variety of criteria that can assure this belongingness in a reasonable manner. Results of an experimentation on some texts (picked out from some research papers) show the fact that the proposed framework can be a suitable alternative for performing fusion in the cases that the class identity of the input text documents is unclear. Moreover, It should be noticed that the proposed framework for fusion can be regarded as an approach to multi-document summarization as well since its main concern is to pick out those parts in the input texts whose significance is in the way they comply with some desired functionalities.

As the final point, it is observed that, from the viewpoint of fusion methodology, the proposed approach to text fusion can be regarded similar to the fusion operator AVG OWA. This seems to be mainly due to the smooth distribution of weights in both AVG OWA –based fusion and our proposed framework as well.

6. REFERENCES

1. Ying, L., Zhe, W., Jie, F., Tingge, Z., Linna, L. and Jiming, L., "Multi-modal public opinion analysis based on image and text fusion", ***Journal of Frontiers of Computer Science & Technology***, Vol. 16, No. 6, (2022), 1260. doi: 10.3778/j.issn.1673-9418.2110056.
2. Ernst, O., Caciularu, A., Shapira, O., Pasunuru, R., Bansal, M., Goldberger, J. and Dagan, I., "Proposition-level clustering for multi-document summarization", in Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. (2022), 1765-1779.
3. Srividya, K., Mirayababu, K. and Mary Sowjanya, A., "Mining interesting aspects of a product using aspect-based opinion mining from product reviews (research note)", ***International Journal of Engineering***, Vol. 30, No. 11, (2017), 1707-1713. doi: 10.5829/ije.2017.30.11b.11.
4. Nazari, N. and Mahdavi, M., "A survey on automatic text summarization", ***Journal of AI and Data Mining***, Vol. 7, No. 1, (2019), 121-135. doi: 10.22044/JADM.2018.6139.1726.
5. Ma, C., Zhang, W.E., Guo, M., Wang, H. and Sheng, Q.Z., "Multi-document summarization via deep learning techniques: A survey", ***ACM Computing Surveys***, Vol. 55, No. 5, (2022), 1-37. doi: https://doi.org/10.1145/3529754.
6. Fauconnier, G. and Turner, M., "Conceptual integration networks", ***Cognitive science***, Vol. 22, No. 2, (1998), 133-187.
7. Pereira, F.C. and Cardoso, A., "Conceptual blending and the quest for the holy creative process", in Second Workshop on Creative Systems, Approaches to Creativity in Artificial Intelligence and Cognitive Science, European Conference on Artificial Intelligence (ECAI 2002), Lyon, France, July. (2002).
8. Yu, M.-H., Li, J., Chan, Z., Yan, R. and Zhao, D., "Content learning with structure-aware writing: A graph-infused dual conditional variational autoencoder for automatic storytelling", in Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 35, (2021), 6021-6029.
9. Zubek, R., "Elements of game design, MIT Press, (2020).
10. Ravi, J., Yu, Z. and Shi, W., "A survey on dynamic web content generation and delivery techniques", ***Journal of Network and Computer Applications***, Vol. 32, No. 5, (2009), 943-960. https://doi.org/10.1016/j.jnca.2009.03.005
11. Mohbey, K., "High fuzzy utility based frequent patterns mining approach for mobile web services sequences", ***International Journal of Engineering***, ***Transactions B: Applications,*** Vol. 30, No. 2, (2017), 182-191. doi: 10.5829/idosi.ije.2017.30.02b.04.
12. Nikolov, N. and Stoehr, P., "Integrating biomedical publications with existing metadata", in 2008 21st IEEE International Symposium on Computer-Based Medical Systems, IEEE. (2008), 653-655.
13. Kieler, B., "Semantic data integration across different scales: Automatic learning generalization rules", ***International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences***, Vol. 37, No., (2008), 685-690.
14. Hong-Sheng, S., "Multi-source fuzzy information fusion method based on bayesian optimal classifier", ***Acta Automatica Sinica***, Vol. 34, No. 3, (2008), 282-287.
15. Nikravesh, M., "Beyond the semantic web: Fuzzy logic-based web intelligence", ***Soft Computing in Ontologies and Semantic Web***, (2006), 149-209. doi: 10.1007/978-3-540-33473-6_7.
16. Nikravesh, M., Takagi, T., Tajima, M., Shinmura, A., Ohgaya, R., Taniguchi, K., Kawahara, K., Fukano, K. and Aizawa, A., "Enhancing the power of search engines and navigations based on conceptual model: Web intelligence", ***Soft Computing for Information Processing and Analysis***, (2005), 35-92. https://doi.org/10.1007/3-540-32365-1_3
17. Lee, H.-S., Chou, M.-T., Tseng, W.-K., Fang, H.-H. and Yeh, C.-H., "A new information fusion method for fuzzy information retrieval", in Knowledge-Based Intelligent Information and Engineering Systems: 11th International Conference, KES 2007, XVII Italian Workshop on Neural Networks, Vietri sul Mare, Italy, September 12-14, 2007. Proceedings, Part II 11, Springer., (2007), 1293-1298.
18. Wang, Z., "A new approach to enterprise information fusion: Choquet-fuzzy-integral-based neural network", in Third International Conference on Natural Computation (ICNC 2007), IEEE. Vol. 5, (2007), 212-216.
19. Arshadi, N. and Badie, K., "A compositional approach to solution adaptation in case-based reasoning and its application to tutoring library", in Proceedings of 8th German Workshop on Case-Based Reasoning. Lammerbuckel. Vol. 11, (2000).
20. Tayefeh Mahmoudi, M., Badie, K., Moosaee, M. and Souri, A., "A compositional adaptation-based approach for recommending learning resources in software development", ***International Journal of Engineering***, ***Transactions A: Basics,*** Vol. 35, No. 7, (2022), 1317-1329. doi: 10.5829/IJE.2022.35.07A.11.
21. Wang, P., Zamora, J., Liu, J., Ilievski, F., Chen, M. and Ren, X., "Contextualized scene imagination for generative commonsense reasoning", arXiv preprint arXiv:2112.06318, (2021).
22. Mahmoudi, M.T., Badie, K., Kharrat, M., Khamaneh, S.B. and Khales, M.Y., "Content personalization in organizations via composing a source content model with user model", in IC-AI, (2008), 943-949.
23. Petty, M.D. and Weisel, E.W., Model composition and reuse, in Model engineering for simulation. 2019, Elsevier.57-85.
24. Lin, B.Y., Zhou, W., Shen, M., Zhou, P., Bhagavatula, C., Choi, Y. and Ren, X., "Commongen: A constrained text generation challenge for generative commonsense reasoning", arXiv preprint arXiv:1911.03705, (2019).
25. Sharma, A. and Kumar, S., "Machine learning and ontology-based novel semantic document indexing for information retrieval", ***Computers & Industrial Engineering***, Vol. 176, (2023), 108940. https://doi.org/10.1016/j.cie.2022.108940
26. Wang, H., Du, J., Li, M. and Li, W., An ontology-based query system for university domain, in Advances in natural computation, fuzzy systems and knowledge discovery. 2021, Springer.632-643.
27. Quan, T.T., Hui, S.C. and Cao, T.H., "Ontology-based fuzzy retrieval for digital library", in Asian Digital Libraries. Looking

Back 10 Years and Forging New Frontiers: 10th International Conference on Asian Digital Libraries, ICADL 2007, Hanoi, Vietnam, December 10-13, 2007. Proceedings 10, Springer. (2007), 95-98.

28. Badie, K., Kharrat, M., Mahmoudi, M.T., Mirian, M.S., Ghazi, T.M. and Babazadeh, S., Ontology-driven creation of contents: Making efficient interaction between organizational users and their surrounding tasks, in User interface design for virtual environments: Challenges and advances. 2012, IGI Global.156-170.

29. Yager, R.R., "On ordered weighted averaging aggregation operators in multicriteria decisionmaking", ***IEEE Transactions on Systems, Man, and Cybernetics***, Vol. 18, No. 1, (1988), 183-190.

30. Alcantud, J.C.R., Santos-García, G. and Akram, M., "Owa aggregation operators and multi-agent decisions with n-soft sets", ***Expert Systems with Applications***, Vol. 203, (2022), 117430. https://doi.org/10.1016/j.eswa.2022.117430

Persian Abstract

چکیده

همجوشی اطلاعات متنی، به عنوان نوعی از همجوشی اطلاعات، برای آنان که علاقمند به استخراج و تولید محتواهای حاوی اطلاعات از میان متون موجود هستند، از اهمیت بسزایی برخوردار گردیده است. ایده اصلی که در پس همجوشی متن همانند هرنوع دیگری ازهمجوشی اطلاعات وجود دارد،ادغام بخشهایی از متون برگرفته از منابع مختلف است به قسمی که نتیجه حاصل از انسجام منطقی ای در راستای اهداف خاص مورد انتظار از متن برخوردار باشد.در این مقاله یک چارچوب فازی جهت تولید متن پیشنهاد می شود که برمبنای آن بخشهایی از متون مرتبط طوری با هم ادغام گردند که نیاز کاربر رابصورت قابل قبولی برآورده سازند. نکته اصلی در چارچوب پیشنهادی ما در خصوص همجوشی، فاصله میان طبقه نمونه متن از یکسو و بردار ویژگی های زیرمجموعه های مختلف از متون موجوداز سوی دیگر است. نتایج حاصل از آزمایش نشان می دهد که چارچوب پیشنهادی می تواند جایگزین مناسبی برای انجام همجوشی در مواردی باشد که هویت متون موجوداز وضوح کافی برخوردار نباشد. این امر منجر به تولید کارآمد متون جدید بر مبنای متون موجود می گردد.