(1) Name of the Workshop: Workshop on Artificial Intelligence

(2) Title of the Paper: A New Method for Fuzzy Information Retrieval Based on Geometric-Mean Averaging Operators

(3) Abstract: In this paper, we present a new method for fuzzy information retrieval based on geometric-mean averaging (GMA) operators. We use some examples to compare the proposed GMA operators with the existing averaging operators. We also prove some properties of the proposed GMA operators. The proposed GMA operators can deal with fuzzy information retrieval in a more flexible and more intelligent manner.

(4) Authors: Shi-Jay Chen (Ph.D. Student),
Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, R. O. C.
E-mail: D8815006@mail.ntust.edu.tw
Tel: (02) 27333141 ext. 7212

Professor Shyi-Ming Chen,
Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, R. O. C.
E-mail: smchen@et.ntust.edu.tw
Tel: (02) 27376417

(5) Contact Author: Professor Shyi-Ming Chen,
Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, R. O. C.
E-mail: smchen@et.ntust.edu.tw
Tel: (02) 27376417

(6) Keywords: Fuzzy Information Retrieval, Fuzzy Query, Geometric Mean, GMA Operators, T-Operators.

A New Method for Fuzzy Information Retrieval Based on Geometric-Mean Averaging Operators

Shi-Jay Chen and Shyi-Ming Chen

Department of Computer Science and Information Engineering
National Taiwan University of Science and Technology,
Taipei, Taiwan, R. O. C.

Abstract

In this paper, we present a new method for fuzzy information retrieval based on geometric-mean averaging (GMA) operators. We use some examples to compare the proposed GMA operators with the existing averaging operators. We also prove some properties of the proposed GMA operators. The proposed GMA operators can deal with fuzzy information retrieval in a more flexible and more intelligent manner.

Keywords: Fuzzy Information Retrieval, Fuzzy Query, Geometric Mean, GMA Operators, T-Operators.

1. Introduction

From [17], we can see that fuzzy sets [25] are very useful in information retrieval (IR). In [3]-[5], [8], [10], [11], [13], [15]-[20] and [24], they all used the T-operators [6], namely T-norms and T-conorms, to deal with the AND and OR operations for fuzzy information retrieval, respectively. However, in [9], Kim et al. pointed out that the existing T-operators are not well model human's behavior for document ranking.

In [9], [12] and [14], Lee et al. pointed out that there are three averaging operators (i.e., P-Norm operators [21], Infinite-One operators [23], and Waller-Kraft operators [24]), which are suitable for achieving high retrieval effectiveness in information retrieval systems. According to [12], the three averaging operators have the following common characteristics: (1) The resulting values of the three averaging operators are controlled by an associated parameter, respectively. For instance, the resulting values of the P-Norm operators are controlled by a parameter p, where $1 \leq p \leq \infty$; the resulting values of the Infinite-One operators are controlled by a parameter γ, where $0 \leq \gamma \leq 1$; the resulting values of the Waller-Kraft operators are controlled by a parameter γ, where $0 \leq \gamma \leq 1$. (2) The resulting values of the three averaging operators are always in the range between "Min" and "Max". However, the three averaging operators still have some drawbacks when we use them to deal with fuzzy information retrieval. That is, sometime they will get unreasonable retrieval results in fuzzy information retrieval. Thus, it is important to develop new averaging operators to overcome the drawbacks of the existing averaging operators for fuzzy information retrieval.

In this paper, we present a new method for fuzzy information retrieval based on geometric-mean averaging (GMA) operators. We use some examples to compare the proposed GMA operators with the existing averaging operators. We also prove some properties of the proposed GMA operators. The proposed GMA operators can deal with fuzzy information retrieval in a more flexible and more intelligent manner.

The rest of this paper is organized as follows. In Section 2, we briefly review the definitions of information retrieval based on the conventional fuzzy set model from [9]. We also briefly review T-operators [6], [12], three existing averaging operators [21], [23], [24], and some analytic results of the T-operators and the averaging operators [9], [12], [14]. In Section 3, we point out the drawbacks of the three existing averaging operators (i.e., the P-Norm operators, the Infinite-One operators, and the Waller-Kraft operators). In Section 4, we present new averaging operators, called the GMA operators, based on the geometric mean for handling the AND and OR operations in fuzzy information retrieval. We use some examples to compare the proposed GMA operators with the existing averaging operators. Furthermore, we also prove some properties of the proposed GMA operators. In Section 5, we extend the proposed GMA operators to deal with weighted fuzzy queries in fuzzy information retrieval. The conclusions are discussed in Section 6.

2. Preliminary

In this section, we briefly review the definitions of information retrieval based on the conventional fuzzy set model [10]. We also briefly review T-operators [6], [12], three existing averaging operators [21], [23], [24], and some analytic results of the T-operators and the averaging operators [9], [12], [14].

2.1 Information Retrieval Based on the Conventional Fuzzy Set Model

In [9], Kim et al. pointed out that an information retrieval system based on the conventional fuzzy set model is defined by a quadruple <T, Q, D, F>, where

(1) T is a set of index terms, T=$\{t_1, t_2, \ldots, t_m\}$. The index terms are used to represent queries and documents.

(2) Q is a set of queries. Each query $q \in$ Q is a Boolean expression composed of index terms t_j, where $1 \leq j \leq$ m, and logical operators "AND", "OR" and "NOT".

(3) D is a set of documents, D $=\{d_1, d_2, \ldots, d_n\}$. Each document $d_i \in$ D is represented by ((t_1, e_{i1}), (t_2, e_{i2}), …, (t_m, e_{im})), where e_{ij} denote the weight (i.e., the degree of strength) of term t_j in document d_i, $e_{ij} \in [0, 1]$, $1 \leq i \leq$ n, and $1 \leq j \leq$ m.

(4) F is a retrieval function,

$$F\text{: D} \times \text{Q} \rightarrow [0, 1], \tag{1}$$

where $F(d_i, q)$ denotes the degree of satisfaction of document d_i with respect to the query q, $F(d_i, q) \in [0, 1]$, and $1 \leq i \leq$ n.

2.2 A Review of T-Operators

The T-operators [6] (i.e., T-norms and T-conorms) are very useful for handling the decision-making problems and they usually used the AND and OR operations to deal with fuzzy information retrieval [3]-[5], [8], [10], [11], [13], [15]-[20], [24]. In [1] and [7], Alsina and Höhle et al. introduced the operators of the T-norms ($\wedge$) and the T-conorms ($\vee$) of fuzzy sets. Let T be a T-norm and let S be a T-conorm, where T: $[0, 1] \times [0, 1] \rightarrow [0, 1]$ and S: $[0, 1] \times [0, 1] \rightarrow [0, 1]$. In [12], Lee et al. summarized some T-norms and T-conorms as shown in Table 1.

Table 1. Some T-norms and T-conorms [12]

T-Norms		**T-Conorms**	
Min(x, y),	Logical Product	Max(x, y),	Logical Sum
$x \times y$,	Algebraic Product	$x + y - x \times y$,	Algebraic Sum
$\frac{xy}{x+y-xy}$,	Hamacher Product	$\frac{x+y-2xy}{1-xy}$,	Hamacher Sum
$\begin{cases} x & if\ y=1 \\ y & if\ x=1 \\ 0 & otherwise \end{cases}$	Drastic Product	$\begin{cases} x & if\ y=0 \\ y & if\ x=0 \\ 1 & otherwise \end{cases}$	Drastic Sum
Max($x + y$ - 1, 0),	Bounded Product	Min($x + y$, 1),	Bounded Sum

Based on [6], we can see that the T-operators can be used in conventional Boolean retrieval systems when the evaluating values e_{ij} of index terms t_j in documents d_i are either 0 or 1, where $1 \leq i \leq$ n and $1 \leq j \leq$ m.

2.3 A Review of Averaging Operators

In the following, we briefly review three averaging operators from [21], [23] and [24] shown as follows:

(1) P-Norm Operators [21]:

$$F(d_i, q_{\text{AND}}) = F(d_i, t_1 \text{ AND } t_2 \text{ AND} \cdots \text{ AND } t_m)$$
$$= 1 - \left(\frac{\sum_{j=1}^{m} (1 - e_{ij})^p}{m} \right)^{\frac{1}{p}}, \tag{2}$$

$$F(d_i, q_{\text{OR}}) = F(d_i, t_1 \text{ OR } t_2 \text{ OR} \cdots \text{ OR } t_m)$$
$$= \left(\frac{\sum_{j=1}^{m} e_{ij}{}^p}{m} \right)^{\frac{1}{p}}, \tag{3}$$

where $1 \le p \le \infty$ and $1 \le i \le$ n. If $p = 1$, then $F(d_i, q_{\text{AND}}) = F(d_i, q_{\text{OR}})$ and they are the same as the arithmetic mean. If $p = \infty$, then formula (2) became

$$\begin{aligned} F(d_i, q_{\text{AND}}) &= F(d_i, t_1 \text{ AND } t_2 \text{ AND} \cdots \text{ AND } t_m) \\ &= 1 - \text{Max}[(1 - e_{i1}), (1 - e_{i2}), \cdots, (1 - e_{im})] \\ &= \text{Min}[e_{i1}, e_{i2}, \cdots, e_{im}], \end{aligned} \tag{4}$$

and formula (3) became

$$\begin{aligned} F(d_i, q_{\text{OR}}) &= F(d_i, t_1 \text{ OR } t_2 \ OR \cdots \text{ OR } t_m) \\ &= \text{Max}(e_{i1}, e_{i2}, \cdots, e_{im}). \end{aligned} \tag{5}$$

(2) Infinite-One Operators [23]:

$$F(d_i, q_{\text{AND}}) = F(d_i, t_1 \text{ AND } t_2 \text{ AND} \cdots \text{ AND } t_m)$$
$$= \gamma \times \text{Min}(e_{i1}, e_{i2}, \cdots, e_{im}) + (1 - \gamma) \times \frac{\sum_{j=1}^{m} e_{ij}}{m}, \tag{6}$$

$$F(d_i, q_{\text{OR}}) = F(d_i, t_1 \ OR \ t_2 \ OR \cdots OR \ t_m)$$
$$= \gamma \times \text{Max}(e_{i1}, e_{i2}, \cdots, e_{im}) + (1 - \gamma) \times \frac{\sum_{j=1}^{m} e_{ij}}{m}, \tag{7}$$

where $0 \le \gamma \le 1$ and $1 \le i \le$ n. If $\gamma = 0$, then $F(d_i, q_{\text{AND}}) = F(d_i, q_{\text{OR}})$ and they are the same as the arithmetic mean. If $\gamma = 1$, then formula (6) became the operator of logical product, and formula (3) became the operator of logical sum.

(3) Waller-Kraft Operators [24]:

$$F(d_i, q_{\text{AND}}) = F(d_i, t_1 \text{ AND } t_2 \text{ AND} \cdots \text{AND } t_m) \\ = (1-\gamma) \times \text{Min}(e_{i1}, e_{i2}, \cdots, e_{im}) + \gamma \times \text{Max}(e_{i1}, e_{i2}, \cdots, e_{im}), \quad (8)$$

where $0 \le \gamma \le 0.5$ and $1 \le i \le$ n,

$$F(d_i, q_{\text{OR}}) = F(d_i, t_1 \text{ } OR \text{ } t_2 \text{ } OR \cdots OR \text{ } t_m) \\ = (1-\gamma) \times \text{Min}(e_{i1}, e_{i2}, \cdots, e_{im}) + \gamma \times \text{Max}(e_{i1}, e_{i2}, \cdots, e_{im}), \quad (9)$$

where $0.5 \le \gamma \le 1$ and $1 \le i \le$ n. If $\gamma = 0.5$, then $F(d_i, q_{\text{AND}}) = F(d_i, q_{\text{OR}}) = \frac{\text{Min}(e_{i1}, e_{i2}, \cdots, e_{im}) + \text{Max}(e_{i1}, e_{i2}, \cdots, e_{im})}{2}$. If $\gamma = 0$, then formula (8) became the operator of logical product. If $\gamma = 1$, then formula (9) became the operator of logical sum.

2.4. Some Analytic Results of the T-Operators and the Averaging Operators

In [9], Lee et al. defined the following three properties for evaluating the T-operators and the averaging operators:

Definition 2.1 An operator θ is *"single operand dependency"* if $\theta(x, y)$ is either x or y, where $x \in [0, 1]$ and $y \in [0, 1]$, and this type of operator is called the *"single operand dependency"* operator.

Definition 2.2 An operator θ is *"negatively compensatory"* if $\theta(x, y)$ is less than Min (x, y) or greater than Max (x, y) for all $x, y \in [0, 1]$, and this type of operator is called the *"negatively compensatory"* operator.

Definition 2.3 An operator θ is *"positively compensatory"* if $\theta(x, y)$ is greater than Min (x, y) and less than Max (x, y) for all $x, y \in [0, 1]$, and this type of operator is called the *"positively compensatory"* operator.

In [9], [12] and [14], Lee et al. pointed out that the operators of logical product (i.e., Min(x, y)) and logical sum (i.e., Max(x, y)) shown in Table 1 are inappropriate for handling fuzzy information retrieval because these two operators have the *"single operand dependency"* property. In the following, we use two examples to explain why these two operators are inappropriate to deal with fuzzy information retrieval.

Example 2.1: Assume that there are two documents d_1 and d_2, and assume that there is a query q_1 shown as follows:

d_1 = {(Information, 0.5), (System, 0.5)},

d_2 = {(Information, 0.9), (System, 0.4)},

q_1 = Information AND System.

If the operator of logical product is used for the AND operations, then the degrees of satisfaction of the documents d_1 and d_2 with respect to the query q_1 can be evaluated and are equal to 0.5 (i.e., Min(0.5, 0.5) = 0.5) and 0.4 (i.e., Min(0.9, 0.4) = 0.4), respectively, and the system will retrieve the document d_1. However, intuitively, the document d_2 is more suitable than the document d_1 with respect to the query q_1.

Example 2.2. Assume that there are two documents d_3 and d_4, and assume that there are two queries q_2 and q_3 shown as follows:

$d_3 = \{(t_1, 0), (t_2, 0.8), (t_3, 1), \ldots, (t_{99}, 1), (t_{100}, 1)\}$,

$d_4 = \{(t_1, 0), (t_2, 0.1), (t_3, 0.1), \ldots, (t_{99}, 0.1), (t_{100}, 1)\}$,

$q_2 = t_1$ AND t_2 AND t_3 AND … AND t_{99} AND t_{100},

$q_3 = t_2$ OR t_{100}.

If the operator of logical product is used for the AND operations, then the degrees of satisfaction of the documents d_3 and d_4 with respect to the query q_2 are the same (i.e., Min(0, 0.8, 1, …, 1, 1) = Min(0, 0.1, 0.1, …, 0.1, 1) = 0). However, intuitively, the document d_3 is more suitable than the document d_4 with respect to the query q_2. If the operator of logical sum is used for the OR operation, the degrees of satisfaction of the documents d_3 and d_4 with respect to the query q_3 are the same and are equal to 1, respectively (i.e., Max(0.8, 1) = Max(0.1, 1) = 1). However, intuitively, the document d_3 is more suitable than the document d_4 with respect to the query q_3.

From [9], [12] and [14], we can see that the remaining T-operators except the operators of logical product and logical sum shown in Table 1 are still inappropriate for handling fuzzy information retrieval due to the fact that they have the *"partially single operand dependency"* and *"negatively compensatory"* properties. In [9], Kim et al. pointed out that the *"partially single operand dependency"* operator can avoid the problem described in Example 2.1, but it still have the same problem described in Example 2.2. The *"negatively compensatory"* property can cause the problem illustrated in the following example.

Example 2.3. Assume that there is a document d_5 and assume that there are two queries q_4 and q_5 shown as follows:

d_5 = {(Information, 0.5), (System, 0.5), (Management, 0.5)},

q_4 = Information AND System,

q_5 = Management.

If we use the operator of algebraic product (i.e., $x \times y$) for the AND operations, then the degrees of satisfaction of the document d_5 with respect to the queries q_4 and q_5 can be evaluated and are equal to 0.25 (i.e., 0.5 × 0.5 = 0.5) and 0.5, respectively. However, it is unreasonable that the degree of satisfaction of the document d_5 with respect to the query q_4 is less than that of the document d_5 with respect to the query q_5.

From the above three examples, we can see that the Min and Max operators suffer from the problem of *"single operand dependency"*. Other T-operators (i.e., "Algebraic product and Algebraic sum", "Hamacher product and Hamacher sum", "Drastic product and Drastic sum", and "Bounded product and Bounded sum") shown in Table 1 have the problems of not only *"partially single operand dependency"*, but also *"negative compensatory"*. Because the T-operators have these problems, some averaging operators are proposed to overcome these problems [21], [23], [24]. In [9], Kim et al. pointed out that P-Norm operators [21] (i.e., formulas (2) and (3)), Infinite-One operators [23] (i.e., formulas (6) and (7)), and Waller-Kraft operators [24] (i.e., formulas (8) and (9)) have the *"positively compensatory"* property. From [9], we can see that *"positively compensatory"* operators are neither *"partially single operand dependency"* nor *"negatively compensatory"*, they can avoid all the problems described previously if the fuzzy information retrieval system using the *"positively compensatory"* operators as the evaluating formulas for the AND and OR operations. Thus, in [9], kim et al. pointed out that the three averaging operators (i.e., P-Norm operators, Infinite-One operators, and Waller-Kraft operators) are suitable to achieve high retrieval effectiveness for fuzzy information retrieval.

3. Analysis of the Existing Averaging Operators

From [9], we can see that P-Norm operators [21], Infinite-One operators [23], and Waller-Kraft operators [24] are appropriate to deal with the AND and OR operations for fuzzy information retrieval respectively. However, according to our research, the three averaging operators still have the following drawbacks:

(1) From [21], we can see that the resulting value of P-Norm operators is controlled by a parameter p, and the values of the parameter p is between 1 and ∞. From [23] and [24], we can see that the resulting values of the Infinite-One operators and Waller-Kraft operators are controlled by a parameter γ, and the values of the parameter γ is between 0 and 1. However, it is very subjective and very hard to determine an appropriate value of the parameter γ between 0 and 1 and to determine

the appropriate parameter p between 1 and ∞ for fuzzy information retrieval.

(2) According to the Infinite-One operators (i.e., formulas (6) and (7)), if $\gamma = 1$, then formula (6) became the operator of logical product (i.e., Min(x, y)) for the AND operations and formula (7) became the operator of logical sum (i.e., Max(x, y)) for the OR operations. However, from Section 2, we can see that the operators of logical product and logical sum have the *"single operand dependency"* property, and from Example 2.1 and Example 2.2, we can see that the Infinite-One operators are inappropriate for fuzzy information retrieval if the parameter $\gamma = 1$. In the same way, if we use P-Norm operators (i.e., formulas (2) and (3)) to deal with fuzzy information retrieval, it will has the same drawback describe above if the parameter $p = \infty$; if we use Waller-Kraft operators (i.e., formulas (8) and (9)) to deal with fuzzy information retrieval, it will has the same drawback describe above if the parameter $\gamma = 0$ in formula (8) and the parameter $\gamma = 1$ in formula (9).

(3) According to Infinite-One operators (i.e., formulas (6) and (7)), if $\gamma = 0$, then the two operators are the same as the arithmetic mean, and the resulting values of the AND and OR operations are the same. That is, the system can not distinguish the degrees of satisfaction of the documents with respect to the queries for the AND and OR operations.

Example 3.1. Assume that there is a document d_6, and assume that there are two queries q_6 and q_7 shown as follows:

$$d_6 = \{(\textit{Information}, 0.2), (\textit{System}, 0.6)\},$$

$$q_6 = \textit{Information} \text{ AND } \textit{System},$$

$$q_7 = \textit{Information} \text{ OR } \textit{System}.$$

If formula (6) is used for the AND operation and formula (7) is used for the OR operation and if $\gamma = 0$, then the degrees of satisfaction $F(d_6, q_6)$ and $F(d_6, q_7)$ of the document d_6 with respect to the queries q_6 and q_7, respectively, can be evaluated as follows:

$$\begin{aligned} F(d_6, q_6) &= F(d_6, \textit{Information} \text{ AND } \textit{Systems}) \\ &= \frac{0.2 + 0.6}{2} \\ &= 0.4, \end{aligned}$$

$$F(d_6, q_7) = F(d_6, \mathit{Information}\ \mathrm{O}R\ \mathit{Systems}) = \frac{0.2+0.6}{2} = 0.4.$$

In this situation, the system can not distinguish the degrees of satisfaction of the document d_6 with respect to the queries q_6 and q_7, respectively.

In the same way, if we use the P-Norm operators (i.e., formulas (2) and (3)) to deal with fuzzy information retrieval, it will have the same drawback if the parameter p = 1; if we use the Waller-Kraft operators (i.e., formulas (8) and (9)) to deal with fuzzy information retrieval, it will have the same drawback if the parameter $\gamma = 0.5$.

(4) According to the Infinite-One operators (i.e., formulas (6) and (7)), if $\gamma = 0.5$, then the system can distinguish the degrees of satisfaction of the documents with respect to the queries for the AND and OR operations [9]. However, it still has a drawback illustrated as follows.

Example 3.2. Assume that there are two documents d_7 and d_8, and assume that there is a query q_8 shown as follows:

d_7 = {(*Information*, 0.2), (*System*, 0.7), (*Management*, 0.9)},

d_8 = {(*Information*, 0.3), (*System*, 0.4), (*Management*, 0.8)},

q_8 = *Information* AND *System* AND *Management*.

If formula (6) is used for the AND operation and if $\gamma = 0.5$, then the degrees of satisfaction $F(d_7, q_8)$ and $F(d_8, q_8)$ of the documents d_7 and d_8 with respect to the query q_8, respectively, can be evaluated as follows:

$$F(d_7, q_8) = F(d_7, \mathit{Information}\ \mathrm{AND}\ \mathit{System}\ \mathrm{AND}\ \mathit{Management}) = 0.5 \times \mathrm{Min}(0.2, 0.7, 0.9) + 0.5 \times \frac{0.2+0.7+0.9}{3} = 0.5 \times 0.2 + 0.5 \times 0.6 = 0.4,$$

$$F(d_8, q_8) = F(d_8, \mathit{Information}\ \mathrm{AND}\ \mathit{System}\ \mathrm{AND}\ \mathit{Management}) = 0.5 \times \mathrm{Min}(0.3, 0.4, 0.8) + 0.5 \times \frac{0.3+0.4+0.8}{3} = 0.5 \times 0.3 + 0.5 \times 0.5 = 0.4.$$

In this situation, the system can not distinguish the degrees of satisfaction of the document d_7 and the document d_8 with respect to the query q_8, respectively. However, intuitively, the document d_7 is more suitable than the document d_8 with respect to the query q_8.

(5) If we use Waller-Kraft operators (i.e., formulas (8) and (9)) to deal with fuzzy information retrieval, the operators have a drawback illustrated in the following example.

Example 3.3. Assume that there are two documents d_9 and d_{10}, and assume that there is a query q_9 shown as follows:

d_9 = {(*Information*, 0.1), (*System*, 0.2), (*Management*, 0.9)},

d_{10} = {(*Information*, 0.1), (*System*, 0.8), (*Management*, 0.9)},

q_9 = *Information* AND *System* AND *Management*.

If formula (8) is used for the AND operations, then the degrees of satisfaction $F(d_9, q_9)$ and $F(d_{10}, q_9)$ of the documents d_9 and d_{10} with respect to the query q_9 can be evaluated as follows:

$$\begin{aligned} F(d_9, q_9) &= F(d_9, \textit{Information} \text{ AND } \textit{System} \text{ AND } \textit{Management}) \\ &= (1-\gamma)\times \text{Min}(0.1, 0.2, 0.9) + \gamma \times \text{Max}(0.1, 0.2, 0.9) \\ &= (1-\gamma)\times 0.1 + \gamma \times 0.9, \end{aligned}$$

$$\begin{aligned} F(d_{10}, q_9) &= F(d_{10}, \textit{Information} \text{ AND } \textit{System} \text{ AND } \textit{Management}) \\ &= (1-\gamma)\times \text{Min}(0.1, 0.8, 0.9) + \gamma \times \text{Max}(0.1, 0.8, 0.9) \\ &= (1-\gamma)\times 0.1 + \gamma \times 0.9, \end{aligned}$$

where $0 \le \gamma \le 0.5$. That is, $F(d_9, q_9) = F(d_{10}, q_9)$. However, intuitively, the document d_{10} is more suitable than the document d_9 with respect to the query q_9.

Thus, if we want to use the averaging operators for the AND and OR operations in fuzzy information retrieval, it is important to develop new averaging operators to overcome the drawbacks of the above three averaging operators.

4. Fuzzy Information Retrieval Based on the Proposed Geometric-Mean Averaging Operators

In the following, we present the new averaging operators, called the Geometric-Mean Averaging (GMA) operators, for fuzzy information retrieval shown as follows:

$$\begin{aligned} F(d_i, q_{\text{AND}}) &= F(d_i, t_1 \text{ AND } t_2 \text{ AND} \cdots \text{AND } t_m) \\ &= \left[\prod_{j=1}^{m}(\alpha + e_{ij})\right]^{\frac{1}{m}} - \alpha, \end{aligned} \tag{10}$$

$$F(d_i, q_{\mathrm{OR}}) = F(d_i, t_1 \text{ OR } t_2 \text{ OR} \cdots \text{OR } t_m)$$
$$= (\alpha+1) - \left[\prod_{j=1}^{m}(\alpha + 1 - e_{ij})\right]^{\frac{1}{m}}, \tag{11}$$

where $\alpha \in \{0, 1\}$, $1 \leq i \leq n$, $1 \leq j \leq m$, $F(d_i, q_{\mathrm{AND}}) \in [0, 1]$ and $F(d_i, q_{\mathrm{OR}}) \in [0, 1]$. The values of $F(d_i, q_{\mathrm{AND}})$ and $F(d_i, q_{\mathrm{OR}})$ of the proposed GMA operators are controlled by a parameter α, where α is either 0 or 1. If the evaluating values e_{ij} of terms t_j in documents d_i are either 0 or 1, where $1 \leq i \leq 4$ and $1 \leq j \leq 2$, then the example shown in Table 2 indicates that the proposed GMA operators are compatible with the traditional Boolean operators (i.e., Table 2(a)) if the parameter α is 0 and are compatible with the extended Boolean operators [21] (i.e., Table 2(b)) if the parameter α is 1. For example, let's consider the document d_2 shown in Table 2(a) and formulas (10) and (11). If $t_1 = 0$, $t_2 = 1$ and the parameter $\alpha = 0$, then we can get

$$F(d_2, q_{\mathrm{AND}}) = F(d_2, t_1 \text{ AND } t_2)$$
$$= [0 \times 1]^{\frac{1}{2}}$$
$$= 0,$$

$$F(d_2, q_{\mathrm{OR}}) = F(d_2, t_1 \text{ OR } t_2)$$
$$= 1 - [(1-0) \times (1-1)]^{\frac{1}{2}}$$
$$= 1.$$

Furthermore, let's consider the document d_2 shown in Table 2(b) and formulas (10) and (11). If $t_1 = 0$, $t_2 = 1$ and the parameter $\alpha = 1$, then we can get

$$F(d_2, q_{\mathrm{AND}}) = F(d_2, t_1 \text{ AND } t_2)$$
$$= [1 \times 2]^{\frac{1}{2}} - 1$$
$$= 0.4142,$$

$$F(d_2, q_{\mathrm{OR}}) = F(d_2, t_1 \text{ OR } t_2)$$
$$= 2 - [(2-0) \times (2-1)]^{\frac{1}{2}}$$
$$= 0.5858.$$

From Table 2(a), we can see that when $\alpha = 0$, the proposed GMA operators can be used in the traditional Boolean information retrieval environment.

Table 2. Applying the Proposed GMA Operators for Information Retrieval

(a) Parameter $\alpha = 0$

	Terms		Query	
	t_1	t_2	t_1 AND t_2	t_1 OR t_2
Documents d_1	0	0	0	0
Documents d_2	0	1	0	1
Documents d_3	1	0	0	1
Documents d_4	1	1	1	1

(b) Parameter $\alpha = 1$

	Terms		Query	
	t_1	t_2	t_1 AND t_2	t_1 OR t_2
Documents d_1	0	0	0	0
Documents d_2	0	1	0.4142	0.5858
Documents d_3	1	0	0.4142	0.5858
Documents d_4	1	1	1	1

In the following, we analyze the proposed GMA operators in different situations, i.e., the parameter α is 0 and the parameter α is 1.

Situation 1: If the parameter α is 0, then the proposed GMA operators (i.e., formulas (10) and (11)) became

$$F(d_i, q_{\text{AND}}) = F(d_i, t_1 \text{ AND} \cdots \text{AND } t_m) = \left[\prod_{j=1}^{m} e_{ij}\right]^{\frac{1}{m}}, \quad (12)$$

$$F(d_i, q_{\text{OR}}) = F(d_i, t_1 \text{ OR} \cdots \text{OR } t_m) = 1 - \left[\prod_{j=1}^{m} (1 - e_{ij})\right]^{\frac{1}{m}}. \quad (13)$$

In this situation, the proposed GMA operators (i.e., formulas (12) and (13)) have the *"partially single operand dependency"* property. From Section 2, we can evaluate that this property can overcome the problem of Example 2.1. Furthermore, formulas (12) and (13) do not have the *"negative compensatory"* property. Thus, it can avoid the problem of Example 2.3. In the following, we use formula (12) to deal with Example 2.1 and Example 2.3, respectively.

(1) If we use formula (12) to deal with Example 2.1, we can evaluate the degree of satisfaction $F(d_1, q_1)$ of the document d_1 with respect to the query q_1 shown as follows:

$$F(d_1, q_1) = F(d_1, \textit{Information} \text{ AND } \textit{System})$$
$$= [0.5 \times 0.5]^{\frac{1}{2}}$$
$$= 0.5.$$

In the same way, we can obtain the degree of satisfaction $F(d_2, q_1)$ of the document d_2 with respect to the query q_1, where $F(d_2, q_1) = 0.6$, and the system will retrieve the document d_2, and it coincides the viewpoint of the human's intuition.

(2) If we use formula (12) to deal with Example 2.3, we can evaluate the degree of satisfaction $F(d_5, q_4)$ of the document d_5 with respect to the query q_4 shown as follows:

$$F(d_5, q_4) = F(d_5, \textit{Information} \text{ AND } \textit{System} \text{ AND } \textit{Management})$$
$$= [0.5 \times 0.5 \times 0.5]^{\frac{1}{3}}$$
$$= 0.5.$$

In the same way, we can evaluate the degree of satisfaction $F(d_5, q_5)$ of the document d_5 with respect to query q_5, where $F(d_5, q_5) = 0.5$. In this situation, we can evaluate the same degrees of satisfaction of the document d_5 with respect to the queries q_4 and q_5, respectively, and it coincides the viewpoint of the human's intuition.

Furthermore, the proposed GMA operators can overcome the drawbacks of Example 3.1, Example 3.3 and Example 3.3. In the following, we use formulas (12) and (13) to deal with Example 3.1, Example 3.2 and Example 3.3, respectively.

(1) If we use formulas (12) and (13) to deal with Example 3.1, we can evaluate the degrees of satisfaction of the document d_6 with respect to the queries q_6 and q_7, respectively, shown as follows:

$$F(d_6, q_6) = F(d_6, \textit{Information} \text{ AND } \textit{System})$$
$$= [0.2 \times 0.6]^{\frac{1}{2}}$$
$$= 0.3464,$$

$$F(d_6, q_7) = F(d_6, \textit{Information} \text{ OR } \textit{System})$$
$$= 1 - [(1 - 0.2) \times (1 - 0.6)]^{\frac{1}{2}}$$
$$= 0.4343.$$

According to the values of $F(d_6, q_6)$ and $F(d_6, q_7)$, the system can distinguish the degrees of satisfaction of the document d_6 with respect to the queries q_6 and q_7, respectively.

(2) If we use formula (12) to deal with Example 3.2, we can evaluate the degree of satisfaction $F(d_7, q_8)$ of the document d_7 with respect to the query q_8 shown as follows:

$$F(d_7, q_8) = F(d_7, \textit{Information} \text{ AND } \textit{System} \text{ AND } \textit{Management})$$
$$= [0.2 \times 0.7 \times 0.9]^{\frac{1}{3}}$$
$$= 0.5013.$$

In the same way, we can evaluate the degree of satisfaction $F(d_8, q_8)$ of the document d_8 with respect to query q_8, where $F(d_8, q_8)$ = 0.4579. The system will retrieve the document d_7, and it coincides the viewpoint of the human's intuition.

(3) If we use formula (12) to deal with Example 3.3, we can evaluate the degree of satisfaction $F(d_9, q_9)$ of the document d_9 with respect to the query q_9 shown as follows:

$$F(d_9, q_9) = F(d_9, \textit{Information} \text{ AND } \textit{System} \text{ AND } \textit{Management})$$
$$= [0.1 \times 0.2 \times 0.9]^{\frac{1}{3}}$$
$$= 0.2621.$$

In the same way, we can evaluate the degree of satisfaction $F(d_{10}, q_9)$ of the document d_{10} with respect to query q_9, where $F(d_{10}, q_9)$ = 0.416. The system will retrieve the document d_{10}, and it coincides the viewpoint of the human's intuition.

From the previously discussions, we can see that when the parameter α is 0, the proposed GMA operators are very useful to deal with fuzzy information retrieval. However, it still have the same problem of Example 2.2, i.e., the degrees of satisfaction $F(d_3, q_2)$ and $F(d_4, q_2)$ are all evaluated as 0, and the degrees of satisfaction $F(d_3, q_3)$ and $F(d_4, q_3)$ are all evaluated as 1. In this situation, we can use the proposed GMA operators and set the parameter α to 1 to overcome this problem. Let us consider the following situation.

Situation 2: If the parameter α is 1, then the proposed GMA operators (i.e., formulas (10) and (11)) became

$$F(d_i, q_{\text{AND}}) = F(d_i, t_1 \text{ AND} \cdots \text{AND } t_m)$$
$$= \left[\prod_{j=1}^{m} (1 + e_{ij}) \right]^{\frac{1}{m}} - 1, \tag{14}$$

$$F(d_i, q_{\text{OR}}) = F(d_i, t_1 \text{ OR} \cdots \text{OR } t_m)$$
$$= 2 - \left[\prod_{j=1}^{m} (2 - e_{ij}) \right]^{\frac{1}{m}}. \tag{15}$$

In this situation, the proposed GMA operators have the *"positively compensatory"* property if the parameter α is 1. In [14], Lee pointed out that the *"positively compensatory"* operators are functions of the form

$$p: [0, 1]\times[0, 1]\rightarrow[0, 1].$$

They must satisfy the follow two properties:

Property 1: $p(x, x) = x$; i.e., p is idempotent.

Property 2: $\text{Min}(x, y) < p(x, y) < \text{Max}(x, y)$, where $x \neq y$.

In the following, we prove the properties of the proposed GMA operators when the parameter α is 1.

Property 1: $F(d, x \text{ AND } x) = x$ and $F(d, x \text{ OR } x) = x$; i.e., $F(d, x \text{ AND } x)$ and $F(d, x \text{ OR } x)$ are idempotent.

Proof:

$$\begin{aligned} F(d, x \text{ AND } x) &= [(1+x)\times(1+x)]^{\frac{1}{2}} - 1 \\ &= (1+x) - 1 \\ &= x, \end{aligned}$$

$$\begin{aligned} F(d, x \text{ OR } x) &= 2 - [(2-x)\times(2-x)]^{\frac{1}{2}} \\ &= 2 - (2-x) \\ &= x. \end{aligned}$$

Thus, we can see that the proposed GMA operators are idempotent when the parameter α is 1.

Property 2: $\text{Min}(x, y) < F(d, x \text{ AND} y) < F(d, x \text{ OR } y) < \text{Max}(x, y)$, where $x \neq y$.

Proof: If $x = 0$ and $y = 1$, then we can see that $\text{Min}(x, y) = 0$ and $\text{Max}(x, y) = 1$. If we use formulas (14) and (15), we can get

$$F(d, x \text{ AND } y) = [(1+x)\times(1+y)]^{\frac{1}{2}} - 1$$
$$= [(1+0)\times(1+1)]^{\frac{1}{2}} - 1$$
$$= 0.4142,$$

$$F(d, x \text{ OR } y) = 2 - [(2-0)\times(2-1)]^{\frac{1}{2}}$$
$$= 2 - (2)^{\frac{1}{2}}$$
$$= 0.5858.$$

In the same way, we can evaluate the same results if $x = 1$ and $y = 0$ shown as follows

$$F(d, x \text{ AND } y) = [(1+x)\times(1+y)]^{\frac{1}{2}} - 1$$
$$= [(1+1)\times(1+0)]^{\frac{1}{2}} - 1$$
$$= 0.4142,$$

$$F(d, x \text{ OR } y) = 2 - [(2-1)\times(2-0)]^{\frac{1}{2}}$$
$$= 2 - (2)^{\frac{1}{2}}$$
$$= 0.5858.$$

In summary, we can see that $\text{Min}(x, y) < F(d, x \text{ AND } y) < F(d, x \text{ OR } y) < \text{Max}(x, y)$, where $x \neq y$.

According to above properties, we can see that the proposed GMA operators have the *"positively compensatory"* property when the parameter α is 1. From [9], [12] and [14], we can see that the *"positively compensatory"* operators do not have either the *"single operand dependency"* property or the *"negatively compensatory"* property. Therefore, the fuzzy information retrieval using *"positively compensatory"* operators can avoid the *"single operand dependency"* and *"negative compensatory"* problems of Example 2.1 to Example 2.3. In the following, we use formulas (14) and (15) to deal with Example 2.1, Example 2.2 and Example 2.3, respectively.

(1) If we use formula (14) to deal with Example 2.1, we can evaluate the degree of satisfaction $F(d_1, q_1)$ of the document d_1 with respect to the query q_1 shown as follows:

$$\begin{aligned} F(d_1, q_1) &= F(d_1, \textit{Information} \text{ AND } \textit{System}) \\ &= [(1+0.5)\times(1+0.5)]^{\frac{1}{2}} - 1 \\ &= [1.5\times 1.5]^{\frac{1}{2}} - 1 \\ &= 0.5. \end{aligned}$$

In the same way, we can evaluate the degree of satisfaction $F(d_2, q_1)$ of the document d_2 with respect to the query q_1, where $F(d_2, q_1) = 0.631$. The system will retrieve the document d_2, and it coincides the viewpoint of the human's intuition.

(2) If we use formula (14) to deal with Example 2.2, we can evaluate the degree of satisfaction $F(d_3, q_2)$ of the document d_3 with respect to the query q_2 shown as follows:

$$\begin{aligned} F(d_3, q_2) &= F(d_3, t_1 \text{ AND } t_2 \text{ AND} \cdots \text{AND } t_{100}) \\ &= [(1+0)\times\cdots\times(1+1)]^{\frac{1}{100}} - 1 \\ &= 1.9841 - 1 \\ &= 0.9841. \end{aligned}$$

In the same way, we can evaluate the degree of satisfaction $F(d_4, q_2)$ of the document d_4 with respect to the query q_2, where $F(d_4, q_2) = 0.1055$. The system will retrieve the document d_3, and it coincides the viewpoint of the human's intuition.

Then, we use formula (15) to evaluate the degree of satisfaction $F(d_3, q_3)$ of the document d_3 with respect to the query q_3 shown as follows:

$$\begin{aligned} F(d_3, q_3) &= F(d_3, t_2 \text{ O}R\ t_{100}) \\ &= 2 - [(2-0.8)\times(2-1)]^{\frac{1}{2}} \\ &= 2 - 1.0954 \\ &= 0.9046. \end{aligned}$$

In the same way, we can evaluate the degree of satisfaction $F(d_4, q_3)$ of the document d_4 with respect to the query q_3, where $F(d_4, q_3) = 0.6216$. The system will retrieve the document d_3, and it coincides the viewpoint of the human's intuition.

(3) If we use formula (12) to deal with Example 2.3, we can evaluate the degree of satisfaction $F(d_5, q_4)$ of the document d_5 with respect to the query q_4 shown as follows:

$$
\begin{aligned}
F(d_5, q_4) &= F(d_5, \textit{Information} \text{ AND } \textit{System} \text{ AND } \textit{Management}) \\
&= [(1+0.5)\times(1+0.5)\times(1+0.5)]^{\frac{1}{3}} - 1 \\
&= [1.5\times1.5\times1.5]^{\frac{1}{3}} - 1 \\
&= 0.5.
\end{aligned}
$$

Then, we can evaluate the degree of satisfaction $F(d_5, q_5)$ of the document d_5 with respect to query q_5, $F(d_5, q_5) =$ 0.5. In this situation, we can obtain the same evaluating results of the document d_5 for the queries q_4 and q_5, respectively, and it coincides the viewpoint of the human's intuition.

Furthermore, the proposed GMA operators can overcome the drawbacks of the existing averaging operators of Example 3.1, Example 3.2 and Example 3.3. In the following, we use formulas (14) and (15) to deal with Example 3.1, Example 3.2 and Example 3.3, respectively.

(1) If we use formulas (14) and (15) to deal with Example 3.1, we can evaluate the degrees of satisfaction of the document d_6 with respect to the queries q_6 and q_7 shown as follows:

$$
\begin{aligned}
F(d_6, q_6) &= F(d_6, \textit{Information} \text{ AND } \textit{System}) \\
&= [(1+0.2)\times(1+0.6)]^{\frac{1}{2}} - 1 \\
&= [1.2\times1.6]^{\frac{1}{2}} - 1 \\
&= 0.3856,
\end{aligned}
$$

$$
\begin{aligned}
F(d_6, q_7) &= F(d_6, \textit{Information} \text{ OR } \textit{System}) \\
&= 2 - [(2-0.2)\times(2-0.6)]^{\frac{1}{2}} \\
&= 0.4125.
\end{aligned}
$$

According to the values of $F(d_6, q_6)$ and $F(d_6, q_7)$, the system can distinguish the degrees of satisfaction of the document d_6 with respect to the queries q_6 and q_7, respectively.

(2) If we use formula (14) to deal with Example 3.2, we can evaluate the degree of satisfaction $F(d_7, q_8)$ of the document d_7 with respect to the query q_8 shown as follows:

$$F(d_7, q_8) = F(d_7, \textit{Information} \text{ AND } \textit{System} \text{ AND } \textit{Management})$$
$$= [(1+0.2)\times(1+0.7)\times(1+0.9)]^{\frac{1}{3}} - 1$$
$$= [1.2\times1.7\times1.9]^{\frac{1}{3}} - 1$$
$$= 0.5708.$$

In the same way, we can evaluate the degree of satisfaction $F(d_8, q_8)$ of document d_8 with respect to query q_8, where $F(d_8, q_8) =$ 0.4852. The system will retrieve the document d_7, and it coincides the viewpoint of the human's intuition.

(3) If we use formula (14) to deal with Example 3.3, we can evaluate the degree of satisfaction $F(d_9, q_9)$ of the document d_9 with respect to the query q_9 shown as follows:

$$F(d_9, q_9) = F(d_9, \textit{Information} \text{ AND } \textit{System} \text{ AND } \textit{Management})$$
$$= [(1+0.1)\times(1+0.2)\times(1+0.9)]^{\frac{1}{3}} - 1$$
$$= [1.1\times1.2\times1.9]^{\frac{1}{3}} - 1$$
$$= 0.3587.$$

In the same way, we can evaluate the degree of satisfaction $F(d_{10}, q_9)$ of the document d_{10} with respect to query q_9, where $F(d_{10}, q_9) =$ 0.5553. The system will retrieve the document d_{10}, and it coincides the viewpoint of the human's intuition.

5. Weighted Fuzzy Query Based on the Extended Geometric-Mean Averaging Operators

In Section 4, we only considered non-weighted fuzzy queries for fuzzy information retrieval. In [14], Lee pointed out that the retrieval effectiveness could be improved by assigning importance factors or weights to the terms and clauses in the queries. From [2], [14], [21] and [22], we can see that weighted queries are very useful in fuzzy information retrieval. Let us consider an example of weighted Boolean query q shown as follows [14]:

$$q = (((t_1, w_{q_1}) \text{ OR } (t_2, w_{q_2})), w_{q_1 \text{ OR } q_2}) \text{ AND } (t_3, w_{q_3}),$$

where w_{q_j} denotes the weight of the term t_j in the query q, $1 \le j \le 2$, and $w_{q_1 \text{ OR } q_2}$ denotes the weight of the clause "(t_1, w_{q_1}) OR (t_2, w_{q_2})" in the query q, where $w_{q_j} \in [0, 1]$ and $w_{q_1 \text{ OR } q_2} \in [0, 1]$.

In the following, we extend the proposed GMA operators of formulas (10) and (11) to formulas shown as follows:

$$F(d_i, q_{\text{AND}}) = F(d_i, (t_1, w_{q_1}) \text{ AND } (t_2, w_{q_2}) \text{ AND} \cdots \text{ AND } (t_m, w_{q_m}))$$

$$= \prod_{j=1}^{m} (\alpha + e_{ij})^{\frac{w_{q_j}}{\sum_{j=1}^{m} w_{q_j}}} - \alpha, \quad (16)$$

$$F(d_i, q_{\text{OR}}) = F(d_i, (t_1, w_{q_1}) \text{ OR } (t_2, w_{q_2}) \text{ OR } \cdots \text{ OR } (t_m, w_{q_m}))$$

$$= (\alpha + 1) - \prod_{j=1}^{m} (\alpha + 1 - e_{ij})^{\frac{w_{q_j}}{\sum_{j=1}^{m} w_{q_j}}}, \quad (17)$$

where $F(d_i, q_{\text{AND}}) \in [0, 1]$, $F(d_i, q_{\text{OR}}) \in [0, 1]$, $w_{q_j} \in [0, 1]$, $\alpha \in \{0, 1\}$, $1 \leq i \leq$ n, and $1 \leq j \leq$ m. The weight w_{q_j} of the term t_j in the query q in formulas (16) and (17) is a relative weight. In [14], Lee pointed out that users submit a query with "relative query weights" will be easier than with "absolute query weights". In the following, we use an example to illustrate how to use formulas (16) and (17) to deal with weighted fuzzy queries for fuzzy information retrieval.

Example 5.1. Assume that there is a document d_{11} and assume that there are two queries q_{10} and q_{11} shown as follows:

d_{11} = {(*Information*, 0.2), (*System*, 0.6), (*Management*, 0.7)},

q_{10} = (*Information*, 0.7) AND (*System*, 1),

q_{11} = (q_{10}, 0.6) OR (*Management*, 0.9).

If we use formula (16) to deal with the AND operation and the parameter α is 1, then the degree of satisfaction $F(d_{11}, q_{10})$ of the document d_{11} with respect to the query q_{10} can be evaluated as follows:

$$F(d_{11}, q_{10}) = F(d_{11}, (\textit{Informtion}, 0.7) \text{ AND } (\textit{System}, 1))$$

$$= [(1 + 0.2)^{\frac{0.7}{0.7+1}} \times (1 + 0.6)^{\frac{1}{0.7+1}}] - 1$$

$$= [(1.2)^{0.4118} \times (1.6)^{0.5882}] - 1$$

$$= 0.4212.$$

That is, the degree of satisfaction $F(d_{11}, q_{10})$ of the document d_{11} with respect to the query q_{10} is 0.4212. In the same way, if we use formula (17) to deal with the OR operation and the parameter α is 1, then the degree of satisfaction $F(d_{11}, q_{11})$ of the document d_{11} with respect to the query q_{11} can be evaluated as follows:

$$
\begin{aligned}
F(d_{11}, q_{11}) &= F(d_{11}, (q_{10}, 0.6) \text{ OR } (\mathit{Management}, 0.9)) \\
&= 2 - [(2 - 0.4212)^{\frac{0.6}{0.6+0.9}} \times (2 - 0.7)^{\frac{0.9}{0.6+0.9}}] \\
&= 2 - [(1.5788)^{0.4} \times (1.3)^{0.6}] \\
&= 0.5949.
\end{aligned}
$$

That is, the degree of satisfaction $F(d_{11}, q_{11})$ of the document d_{11} with respect to the query q_{11} is 0.5949.

6. Conclusions

In this paper, we have presented a new method for fuzzy information retrieval based on geometric-mean averaging (GMA) operators. We use some examples to compare the proposed GMA operators with the existing averaging operators. We also prove some properties of the proposed GMA operators. The proposed GMA operators can deal with fuzzy information retrieval in a more flexible and more intelligent manner.

References

[1] C. Alsina, E. Trillas, and L. Valverde, "On some logical connectives for fuzzy set theory," *Journal of Mathematical Analysis and Application*, vol. 93, no. 1, pp. 15-26, 1983.

[2] A. Bookstein, "Fuzzy requests: an approach to weighted Boolean searches," *Journal of the American Society for Information Science*, vol. 31, no. 4, pp. 240-247, 1980.

[3] D. A. Buell and D. H. Kraft, "Threshold values and Boolean retrieval systems," *Information Processing and Management*, vol. 17, no. 3, pp. 127-136, 1981.

[4] D. A. Buell, "A problem in information retrieval with fuzzy sets," *Journal of the American Society for Information Science*, vol. 36, no. 6, pp. 398-401, 1985.

[5] R. Fagin, "Combining fuzzy information from multiple systems," *Journal of Computer and System Sciences*, vol. 58, no. 1, pp. 83-99, 1999.

[6] M. M. Gupta and J. Qi, "Theory of T-norms and fuzzy inference methods," *Fuzzy Sets and Systems*, vol. 40, no. 3, pp. 431-450, 1991.

[7] U. Höhle, "Probabilistic uniformization of fuzzy topologies," *Fuzzy Sets and Systems*, vol. 1, no. 1, pp. 311-332, 1978.

[8] C. M. Kim and Y. G. Kim, "An improvement of Bandler-Kohout fuzzy information retrieval model using reduced set," *Proceedings of the 1999 IEEE International Fuzzy Systems Conference*, Seoul, Korea, pp. 22-25, 1999.

[9] M. H. Kim, J. H. Lee, and Y. J. Lee, "Analysis of fuzzy operators for high quality information retrieval," *Information Processing Letters*, vol. 46, no. 5, pp. 251-256, 1993.

[10] D. H. Kraft and D. A. Buell, "Fuzzy sets and generalized Boolean retrieval systems," *International Journal of Man-Machine Studies*, vol. 19, no. 1, 45-56, 1983.

[11] D. H. Kraft and D. A. Buell, "An extended fuzzy linguistic approach to generalize Boolean information retrieval," *Information Sciences*, vol. 2, no. 3, 119-134, 1994.

[12] J. H. Lee, W. Y. Kim, M. H. Kim, and Y. J. Lee, "On the evaluation of Boolean operators in the extended Boolean retrieval framework," *Proceedings of the Sixteenth Annual ACM Conference on Research and Development in Information Retrieval*, pp. 291-297, Pittsburgh, Pennsylvania, U. S. A., 1993.

[13] J. H. Lee, M. H. Kim, and Y. J. Lee, "Ranking Documents in thesaurus-based Boolean retrieval systems," *Information Processing and Management*, vol. 30, no. 1, pp. 79-91, 1994.

[14] J. H. Lee, "Properties of extended Boolean models in information retrieval," *Proceedings of the Seventeenth Annual ACM Conference on Research and Development in Information Retrieval*, pp. 182-190, Dublin, Ireland, 1994.

[15] D. Lucarella and R. Morara, "FIRST: fuzzy information retrieval system," *Journal of Information Science*, vol. 17, no. 2, pp. 81-91, 1991.

[16] R. L. Mántaras, U. Cortés, J. Manero, and E. Plaza, "Knowledge engineering for a document retrieval system," *Fuzzy Sets and Systems*, vol. 38, no. 2, pp. 223-240, 1990.

[17] S. Miyamoto, *Fuzzy Sets in Information Retrieval and Cluster Analysis*. Kluwer, Dordrecht, 1990.

[18] G. Pasi and R. A. M. Pereira, "A decision making approach to relevance feedback in information retrieval: a model based on soft consensus dynamics," *International Journal of Intelligent Systems*, vol. 14, no. 1, pp. 105-122. 1999.

[19] S. E. Robertson, "On the nature of fuzz: a diatribe," *Journal of the American Society for Information Science*, vol. 29, no. 4, pp. 304-307, 1978.

[20] W. M. Sachs, "An approach to associative retrieval through the theory of fuzzy sets," *Journal of the American Society for Information Science*, vol. 27, no. 1, pp. 85-87, 1976.

[21] G. Salton, E. A. Fox, and H. Wu, "Extended Boolean information retrieval," *Communications of the ACM*, vol. 26, no. 12, pp. 1022-1036, 1983.

[22] G. Salton, "A simple blueprint for automatic Boolean query processing," *Information Processing and Management*, vol. 24, no. 3, pp. 269-280, 1988.

[23] M. E. Smith, *Aspects of the P-Norm Model of Information Retrieval: Syntactic Query Generation, Efficiency, and Theoretical Properties*, Ph.D. Thesis, Cornell University, 1990.

[24] W. G. Waller and D. H. Kraft, "A mathematical model of a weighted Boolean retrieval system," *Information Processing and Management*, vol. 15, no. 5, pp. 235-245, 1979.

[25] L. A. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, no. 3, pp. 338-353, 1965.

[26] H. J. Zimmermann, *Fuzzy Set Theory and Its Applications*. Boston: Kluwer Academic Publishers, 1991.

[27] H. J. Zimmermann and P. Zysno, "Latent connectives in human decision making," *Fuzzy Sets and Systems*, vol. 4, no. 1, pp. 37-51, 1980.