

Aggregating metasearch engine results based on maximal entropy OWA operator

Sang Xiuzhi Liu Xinwang

(School of Economics and Management, Southeast University, Nanjing 211189, China)

Abstract: The maximal entropy ordered weighted averaging (ME-OWA) operator is used to aggregate metasearch engine results, and its newly analytical solution is also applied. Within the current context of the OWA operator, the methods for aggregating metasearch engine results are divided into two kinds. One has a unique solution, and the other has multiple solutions. The proposed method not only has crisp weights, but also provides multiple aggregation results for decision makers to choose from. In order to prove the application of the ME-OWA operator method, under the context of aggregating metasearch engine results, an example is given, which shows the results obtained by the ME-OWA operator method and the minimax linear programming (minimax-LP) method. Comparison between these two methods are also made. The results show that the ME-OWA operator has nearly the same aggregation results as those of the minimax-LP method.

Key words: maximal entropy ordered weighted averaging operator; minimax linear programming; metasearch engine; information aggregation

doi: 10.3969/j.issn.1003-7985.2013.02.006

In recent years, the world wide web (WWW) has been a main place to find information about any area. It has become the most frequently used tool for locating specific information. However, several attempts have been reported that the results being searched for with the same query in different search engines differentiate significantly at the same time^[1-2]. But what users want is to find accurate information through employing a general-purpose search engine in a timely and cost-effective way, which seems to be unrealistic^[1,3]. Therefore, aggregating the retrieval results from different search engines can significantly improve the quality of final metasearch engine results^[4]. A metasearch engine is a system that supports unified access to multiple existing search engines^[5-6].

Diaz et al.^[1] demonstrated that the metasearch engine provides a significant improvement in coverage of search

effectiveness. Dreilinger and Howe^[7] adopted a pragmatic approach to the problem of information retrieval on the web. A comprehensive survey for building an efficient metasearch engine was provided by Meng et al.^[8]. Keyhanipour et al.^[4] used the optimistic OWA operator to aggregate metasearch engine results. Emrouznejad^[2] proposed the most preferred OWA operator to aggregate search engine results. Herrera-Viedma et al.^[9] analyzed the role of aggregation operators to access information on the web. Amin and Emrouznejad^[5] proposed the minimax linear programming model for aggregating multiple search engine results. Diaz et al.^[10] proposed a fuzzy analytic network process (ANP) and weighted fuzzy ANP methods for merging metasearch engine results. To the best of the author's knowledge, no single research is reported to find the metasearch engine results of a specific query using the crisp OWA operator weights.

There are some other kinds of OWA operator weights determining methods^[11-15]. This paper proposes a new application of the ME-OWA operator method for aggregating metasearch engine results obtained from a specific query. Three reasons contribute to our decision for using the ME-OWA operator. First, it is simple, comprehensible and convenient for users to know. Secondly, the weights obtained by the ME-OWA operator change monotonically and smoothly with the position index, which can well reflect the importance of the retrieval documents in different positions^[12]. Thirdly, if we use other OWA operator methods, the frameworks of finding metasearch engine results are the same.

This paper first gives a general approach to finding metasearch engine results. Then it divides the existing OWA operator methods into two kinds: one is to acquire unique aggregation results, the other is to acquire multiple aggregation results. But the existing solutions to the second kind are fuzzy search engine weighted approaches. Whereas, aggregating metasearch engine results is a crisp problem, we should use crisp weights to aggregate the metasearch engine results. The ME-OWA operator not only provides crisp weights, but also has multiple aggregation results for decision makers to choose from. Meanwhile, we elaborate on the advantages of the proposed method. In the end, an example is provided to illustrate its application, which extends the aggregation results of the minimax-LP method^[5].

Received 2012-11-21.

Biographies: Sang Xiuzhi (1983—), female, graduate; Liu Xinwang (corresponding author), male, doctor, professor, xwliu@seu.edu.cn.

Foundation item: The National Natural Science Foundation of China (No. 71171048).

Citation: Sang Xiuzhi, Liu Xinwang. Aggregating metasearch engine results based on maximal entropy OWA operator[J]. Journal of Southeast University (English Edition), 2013, 29(2): 139 – 144. [doi: 10.3969/j.issn.1003-7985.2013.02.006]

1 Existing Methods for Aggregating Metasearch Engine Results

As is stated above, we just discuss the problem of aggregating metasearch engine results in the OWA operator context.

1.1 A general approach to aggregate metasearch engines results

The OWA operator aggregation process can be divided into two phases:

1) Aggregating phase. Submit queries to the meta-search engine system, which dispatches them to the selected search engines; then each search engine extracts a retrieval list. Count the frequency λ_{ji} of document D_i appearing in the j -th place and compute the weighting vectors \mathbf{W} with the relevant OWA operator method; then obtain aggregation values for metasearch engine results.

2) Ranking phase. Order the aggregation results according to a given criterion (i. e., the degree of orness associated with the OWA operator) to obtain the optimal rank.

Let us consider an application of the OWA operator in the metasearch system^[2]. Suppose that there are m search engines ($m \geq 2$) denoted by SE_1, \dots, SE_m . Without loss of generality, we only consider the first l -th ranked list of documents retrieved from each search engine. Let D_1, \dots, D_r denote the documents appearing in the retrieval lists. The corresponding data can be recorded in Tab. 1.

Tab. 1 The appearing place of documents

Documents	The first place	...	The i -th place	...	The l -th place
D_1	λ_{11}	...	λ_{1i}	...	λ_{1l}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
D_j	λ_{j1}	...	λ_{ji}	...	λ_{jl}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
D_r	λ_{r1}	...	λ_{ri}	...	λ_{rl}

In Tab. 1, λ_{ji} ($j = 1, 2, \dots, r; i = 1, 2, \dots, l$) denotes the frequency of document D_j appearing in the i -th ranked place.

The relevancy index of document D_k is defined as

$$z_k = \sum_{i=1}^l \lambda_{ki} w_i \quad (1)$$

where w_i ($i = 1, 2, \dots, l$) is the unknown weight defined for the i -th place. It is noted that the weight should satisfy $w_i \geq w_{i+1}$ because the greater the frequency number, the more relevant to the query, and we should give them more weights.

In Eq. (1), it is obvious that the aggregation value z_k is dependent upon the weighting vector. So, the problem of aggregating metasearch engine results can be changed into obtaining a proper weighting vector under the OWA operator.

Yager^[16] was the first introducing the concept of the OWA operator which has the ability to derive optimal re-

sults based on the weighting vectors after an aggregation process and it is defined as follows.

Definition 1 An OWA operator of dimension n is a mapping $F_W: \mathbf{R}^n \rightarrow \mathbf{R}$ that has weights $\mathbf{W} = \{w_1, w_2, \dots, w_n\}^T$ with properties as

$$w_1 + w_2 + \dots + w_n = 1 \quad 0 \leq w_i \leq 1; i = 1, 2, \dots, n \quad (2)$$

Such that

$$F_W(X) = F_W(x_1, x_2, \dots, x_n) = \sum_{i=1}^n w_i y_i \quad (3)$$

where y_i is the i -th largest value in x_i .

Yager^[16] also introduced two important measures with respect to the weighting vector \mathbf{W} of problem (3), which are defined as follows.

Definition 2^[16] Assume that $\mathbf{W} = \{w_1, w_2, \dots, w_n\}^T$ is the weighting vector of problem (3), the orness level is defined as

$$\text{orness}(\mathbf{W}) = \sum_{i=1}^n \frac{n-i}{n-1} w_i \quad (4)$$

where $\text{orness}(\mathbf{W}) = \alpha$ and $\alpha \in [0, 1]$.

Definition 3^[16] Assume that $\mathbf{W} = \{w_1, w_2, \dots, w_n\}^T$ is the weighting vector of problem (3), the dispersion measure is defined as

$$\text{Dispersion}(\mathbf{W}) = - \sum_{i=1}^n w_i \ln w_i \quad (5)$$

1.2 Existing methods to aggregate metasearch engine results

1.2.1 Method of acquiring unique solution

It means that we can obtain an unique optimal weighting vector by solving a certain method without artificially adjusting any parameters, related researches can be referred to Refs. [2, 5]. Here, we just introduce the minimax-LP method^[5], because we will compare the ranking with that of the ME-OWA operator.

The minimax-LP method minimizes the maximum deviation from relevancy indices corresponding to the aggregated elements, which can be expressed as

$$\begin{aligned} & \min M \\ & \text{s. t. } M - d_j \geq 0 \quad j = 1, 2, \dots, r \\ & \sum_{i=1}^l \lambda_{ji} w_i + d_j = 1 \quad j = 1, 2, \dots, r \\ & w_i - w_{i+1} \geq \varepsilon^* \quad i = 1, 2, \dots, l-1 \\ & w_l \geq \varepsilon^* \\ & d_j \geq 0 \end{aligned} \quad (6)$$

where $M = \max\{d_j: j = 1, 2, \dots, r\}$; d_j is a deviation variable; w_i is the unknown weight defined for the i -th place; λ_{ji} ($j = 1, 2, \dots, r; i = 1, 2, \dots, l$) is the frequency document D_j appeared in the i -th ranked place; ε^* is the maximum discriminating parameter, and the value is obtained as follows:

$$\varepsilon^* = \min \left\{ \frac{1}{\sum_{i=1}^r \lambda_{ji}(1-i+1)}, j = 1, 2, \dots, r \right\}$$

1.2.2 Method of obtaining multiple solutions

It means that we obtain multiple aggregation results through solving a model with certain predefined parameters to acquire different weighting vectors.

The linguistic quantifier was originally introduced by Zadeh^[17]. Later, Yager^[18] proposed a fuzzy linguistic regular increasing monotone (RIM) quantifier guided aggregation with the OWA operator, which is defined as follows.

Definition 4^[18] A fuzzy subset Q is called an RIM quantifier if $Q(0) = 0$, $Q(1) = 1$, and $Q(x) \geq Q(y)$ if $x \geq y$.

With an RIM quantifier Q , the weighting vector of the OWA operator can be obtained by

$$w_i = Q\left(\frac{i}{n}\right) - Q\left(\frac{i-1}{n}\right) \quad i = 1, 2, \dots, n \quad (7)$$

The quantifier guided aggregation with the OWA operator is

$$F_w(X) = \sum_{i=1}^n \left(Q\left(\frac{i}{n}\right) - Q\left(\frac{i-1}{n}\right) \right) x_i \quad (8)$$

where x_i is in a decreasing order.

The fuzzy linguistic quantifier Q generally represents the concept of a fuzzy majority in the arguments aggregation^[19]. Various applications are discussed in Refs. [1, 10].

It is so flexible that decision makers can obtain an optimal solution through multiple aggregation results. But fuzzy linguistic quantifier guided aggregation is used to solve fuzzy decision making problems; it is not suitable for solving crisp decision making problems. In this paper, we introduce the ME-OWA operator method to the application of aggregating metasearch engine retrieval results, which not only has crisp weights, but also has multiple aggregation results for decision makers to choose from.

2 Application of ME-OWA Operator Method in Metasearch Environment

2.1 ME-OWA operator and its computation

O'Hagan^[11] combined the principle of the maximum entropy with the OWA operator and proposed the following mathematical programming problem:

$$\begin{aligned} \max \quad & - \sum_{i=1}^n w_i \ln w_i \\ \text{s. t.} \quad & \sum_{i=1}^n \frac{n-i}{n-1} w_i = \alpha \quad \alpha \in [0, 1] \\ & \sum_{i=1}^n w_i = 1 \\ & w_i \geq 0 \quad i = 1, 2, \dots, n \end{aligned} \quad (9)$$

Various optimal solutions were discussed in Refs. [11 – 12, 20]. There is an important property of problem (9), which is expressed as follows.

Theorem 1^[12] Since α is the orness level of problem (9), when $0.5 \leq \alpha \leq 1$, the weights form a nonincreasing sequence, and $w_i \geq w_j$ for $i \leq j$. When $0 \leq \alpha \leq 0.5$, the weights form a nondecreasing sequence, and $w_i \leq w_j$ for $i \geq j$.

From Theorem 1, when $0.5 \leq \alpha \leq 1$, the greater the aggregated element is, the more emphasis will be given; when $0 \leq \alpha < 0.5$, the smaller the aggregated element is, the more emphasis will be given.

In the application of aggregating metasearch engine results, the more frequent the document is, the more relevant the query will be. So, we just consider the weights of problem (9) under orness level $\alpha \in [0.5, 1]$.

Using the method of Lagrange multipliers, Liu^[21] and Fullér et al.^[12] transformed problem (9) into a polynomial equation to determine the optimal weighting vectors. In this paper, we use the analytical solution of Liu's method^[21], which has a geometric form with $w_{i+1}/w_i = q$, and the weights w_i can be expressed as^[21]

$$w_i = \frac{q^{i-1}}{\sum_{j=0}^{n-1} q^j} \quad i = 1, 2, \dots, n \quad (10)$$

where q is the root solution to problem (9) that can be solved by

$$(n-1)\alpha q^{n-1} + \sum_{i=2}^n ((n-1)\alpha - i + 1) q^{n-i} = 0 \quad (11)$$

Various applications of problem (9) are discussed in Refs. [11 – 12, 20].

Next, we will introduce the analytical process of aggregating metasearch engine results.

2.2 Procedures of aggregating metasearch engine results

We divide the procedure into several steps as follows:

Step 1 Choose search engine (SE_1, \dots, SE_m) in the metasearch environment, and submit a specific query to these search engines.

Step 2 Determine the first l -th ranked list.

Step 3 To each document D_j , count the frequency λ_{ji} ($j = 1, 2, \dots, r; i = 1, 2, \dots, l$) appearing in the i -th place.

Step 4 Take n and α into Eq. (11), and compute the value of q .

Step 5 Take q into Eq. (10), and compute weighting vector W with given orness α .

Step 6 According to Eq. (1), calculate the aggregation results of each document D_j with the given weighting vector.

Step 7 With respect to the aggregation results, rank

the retrieval document D_j .

In the above process, Steps 1 to 6 are aggregating phases, and Step 7 is a ranking phase. Combining the conclusions in Theorem 1 with the application context, it is noted that the orness level α should be between 0.5 and 1.

2.3 Main advantages of ME-OWA operator method

The major advantages of the proposed method are presented as follows:

- 1) It is a well-developed method, which attracts much attention from the very beginning both in theory and application. It is simple, comprehensible, and convenient in application, and it has been used to aggregate information in certain areas.
 - 2) The polynomial equation of the Liu method^[21] of problem (9) is first used to determine the weighting vectors of the ME-OWA operator method, which is more simple.
 - 3) It overcomes the fundamental shortcomings of the fuzzy RIM linguistic quantifier OWA operator method applied in aggregating metasearch engine results.
- The fuzzy RIM linguistic quantifier OWA operator is more suitable for solving fuzzy aggregation problems. However, aggregating metasearch engine results is a crisp problem, and the weights obtained by the ME-OWA operator method are crisp numbers. The proposed method not only provides right multiple aggregation results for decision makers, but also has nearly the same aggregation results as those of the minimax-LP method.
- 4) It determines the optimal weighting vector under the maximal entropy, which has the ability to ascertain the optimal aggregation results after the aggregating processes.

When the orness level $\alpha \in [0.5, 1]$, the weights obtained by problem (9) monotonically decrease with the position index, which just satisfies the requirement of aggregating the retrieval document. $\alpha = 1$ is used to present the situation when the decision maker is maximally optimistic, and $\alpha = 0.5$ is used to represent the situation when the decision maker faces a moderate assessment.

3 Case Study of ME-OWA Operator Method

We now use the ME-OWA operator method to aggregate metasearch engine results, which is given by Amin and Emrouznejad^[5]. The factors involve the number of search engines, the number of the first ranked list considered, the frequency of the document in every ranked place and the ME-OWA weighting vectors under different orness levels.

3.1 Computing process

Step 1 Four search engines $SE_1 = \text{Google}$, $SE_2 = \text{Yahoo}$, $SE_3 = \text{MSN}$ and $SE_4 = \text{Altavista}$ are selected. Que-

ries $Q = \{\text{metasearch engine, aggregation}\}$ are submitted to the four search engines.

Step 2 Suppose that we just rank the first ten documents retrieved from the four search engines, and there are 25 documents appearing in the first ten places, which is denoted as D_1 to D_{25} .

Step 3 Count the frequency of each document appearing in each place within the four search engines, which is listed in Tab. 2.

Tab. 2 Frequency of the first ten results retrieved by four search engines

Documents	Place									
	1	2	3	4	5	6	7	8	9	10
D_1	1	2	1	0	0	0	0	0	0	0
D_2	0	1	0	0	0	0	0	0	0	0
D_3	0	0	1	0	0	0	0	0	1	0
D_4	0	0	0	1	0	1	0	0	1	0
D_5	0	0	0	0	1	0	0	0	0	0
D_6	0	0	0	0	0	1	0	0	0	0
D_7	0	0	0	0	0	0	1	0	0	0
D_8	0	0	0	0	1	0	0	1	0	0
D_9	1	0	1	0	0	0	0	0	1	0
D_{10}	0	0	0	0	0	0	0	0	0	1
D_{11}	1	1	1	0	0	0	0	0	0	0
D_{12}	1	0	0	1	0	0	0	0	0	0
D_{13}	0	0	0	0	1	0	0	0	0	0
D_{14}	0	0	0	0	0	1	0	1	0	0
D_{15}	0	0	0	1	0	0	1	0	0	0
D_{16}	0	0	0	0	1	0	0	1	0	0
D_{17}	0	0	0	1	0	0	0	0	0	1
D_{18}	0	0	0	0	0	1	0	0	0	0
D_{19}	0	0	0	0	0	0	1	0	0	0
D_{20}	0	0	0	0	0	0	0	0	1	0
D_{21}	0	0	0	0	0	0	0	0	0	1
D_{22}	0	0	0	0	0	0	1	0	0	0
D_{23}	0	0	0	0	0	0	0	1	0	0
D_{24}	0	0	0	0	0	0	0	0	1	0
D_{25}	0	0	0	0	0	0	0	0	0	1

Step 4 Take $n = 10$ and $\alpha = 0.5$ into Eq. (12), and obtain $q = 1$.

Step 5 Take $q = 1$ into Eq. (11), and obtain $w_1 = w_2 = \dots = w_{10} = 0.1$.

Step 6 According to Eq. (1), the aggregation results of all the documents in Tab. 2 are calculated, and we obtain

$$z_1 = 0.4, z_4 = z_9 = z_{11} = 0.3$$
$$z_2 = z_8 = z_{12} = z_{14} = z_{15} = z_{16} = z_{17} = 0.2$$
$$z_3 = z_5 = z_6 = z_7 = z_{10} = z_{13} = z_{18} = z_{19} = \dots = z_{25} = 0.1$$

Step 7 By the same method, compute the weighting vectors and the corresponding aggregation results under the orness level $\alpha = 0.6, 0.7, 0.8, 0.9, 1.0$, which are listed in Tab. 3 and Tab. 4, respectively.

Tab. 3 Weighting vectors in ME-OWA method

α	Weight									
	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}
0.5	0.100 0	0.100 0	0.100 0	0.100 0	0.100 0	0.100 0	0.100 0	0.100 0	0.100 0	0.100 0
0.6	0.156 9	0.140 4	0.125 6	0.112 3	0.100 5	0.089 9	0.080 4	0.072 0	0.064 4	0.057 6
0.7	0.233 6	0.184 0	0.145 0	0.114 3	0.090 1	0.071 0	0.056 0	0.044 1	0.034 7	0.027 4
0.8	0.342 7	0.227 2	0.150 6	0.099 8	0.066 2	0.043 9	0.029 1	0.019 3	0.012 8	0.008 5
0.9	0.525 0	0.249 5	0.118 6	0.056 4	0.026 8	0.012 7	0	0	0	0
1.0	1.000 0	0	0	0	0	0	0	0	0	0

Tab. 4 Aggregation results obtained by ME-OWA method

Document	α					
	0.5	0.6	0.7	0.8	0.9	1.0
D_1	0.4	0.563 3	0.746 6	0.947 7	1.142 6	1
D_2	0.1	0.140 4	0.184 0	0.227 2	0.249 5	0
D_3	0.2	0.190 0	0.179 7	0.163 4	0.118 6	0
D_4	0.3	0.266 6	0.220 0	0.156 5	0.069 1	0
D_5	0.1	0.100 5	0.090 1	0.066 2	0.026 8	0
D_6	0.1	0.089 9	0.071 0	0.043 9	0.012 7	0
D_7	0.1	0.080 4	0.056 0	0.029 1	0	0
D_8	0.2	0.172 5	0.134 2	0.085 5	0.026 8	0
D_9	0.3	0.346 9	0.413 3	0.506 1	0.643 6	1
D_{10}	0.1	0.057 6	0.027 4	0.008 5	0	0
D_{11}	0.3	0.422 9	0.562 6	0.720 5	0.893 1	1
D_{12}	0.2	0.269 2	0.347 9	0.442 5	0.581 4	1
D_{13}	0.1	0.100 5	0.090 1	0.066 2	0.026 8	0
D_{14}	0.2	0.161 9	0.115 1	0.063 2	0.012 7	0
D_{15}	0.2	0.192 7	0.170 3	0.128 9	0.056 4	0
D_{16}	0.2	0.172 5	0.134 2	0.085 5	0.026 8	0
D_{17}	0.2	0.169 9	0.141 7	0.108 3	0.056 4	0
D_{18}	0.1	0.089 9	0.071 0	0.043 9	0.012 7	0
D_{19}	0.1	0.080 4	0.056 0	0.029 1	0	0
D_{20}	0.1	0.064 4	0.034 7	0.012 8	0	0
D_{21}	0.1	0.057 6	0.027 4	0.008 5	0	0
D_{22}	0.1	0.080 4	0.056 0	0.029 1	0	0
D_{23}	0.1	0.072 0	0.044 1	0.019 3	0	0
D_{24}	0.1	0.064 4	0.034 7	0.012 8	0	0
D_{25}	0.1	0.057 6	0.027 4	0.008 5	0	0

Step 8 Rank the first ten retrieval documents and compare them with those of the minimax-LP method, which is shown in Tab. 5. The last column is calculated by the minimax-LP method^[5].

Tab. 5 Rank obtained by ME-OWA operator and minimax-LP method

Rank	α						Minimax-LP method
	0.5	0.6	0.7	0.8	0.9	1.0	
1	D_1	D_1	D_1	D_1	D_1	D_1	D_1
2	D_4	D_{11}	D_{11}	D_{11}	D_{11}	D_{11}	D_{11}
3	D_9	D_9	D_9	D_9	D_9	D_9	D_9
4	D_{11}	D_{12}	D_{12}	D_{12}	D_{12}	D_{12}	D_{12}
5	D_3	D_4	D_4	D_2	D_2	D_2	D_4
6	D_8	D_{15}	D_2	D_3	D_3	D_3	D_{15}
7	D_{12}	D_3	D_3	D_4	D_4	D_4	D_2
8	D_{14}	D_8	D_{15}	D_{15}	D_{15}	D_{15}	D_8
9	D_{15}	D_{16}	D_{17}	D_{17}	D_{17}	D_{17}	D_{16}
10	D_{16}	D_{17}	D_8	D_8	D_8	D_8	D_3

3.2 Discussion

The characteristics of the proposed method are shown as follows:

1) It is noted that when the documents have the same aggregation results, the order in which the aggregation results are arranged is determined by the document position index. If there are more than ten qualified documents appearing in the final list, we just keep the first ten documents, and others will be omitted. For example, in Tab. 4, when $\alpha = 0.5$, documents D_3 , D_8 , D_{12} , D_{14} , D_{15} , D_{16} and D_{17} have the same aggregation result 0.2, and we order them according to their position index. In this case, document D_{17} is ordered in the 11-th place, so we omit it accordingly.

2) Compared with the minimax-LP method, from Tab. 5, it is shown that some documents, e. g. , D_1 , D_{11} , D_9 , D_{12} , constantly have the same rank almost under all weighting vectors. Interestingly, when $\alpha = 0.6$, the ranking is almost the same as that of the minimax-LP method.

Meanwhile, the proposed method fully considers people’s decision attitudes. Different orness levels will directly affect people’s decision results. However, the minimax-LP method cannot deal with them.

3) Compared with the existing methods based on the OWA operators^[1,5], it not only deals with multiple crisp weighting vectors, but also has nearly the same aggregation results as those of the minimax-LP method.

4 Conclusion

We present a new decision making process based on the maximal entropy OWA operator and apply it in aggregating metasearch engine results problems. We study the main advantages of the new aggregation method and analyze the applicability of the proposed method. It is found that the proposed method is a more general formulation of the OWA operator. In this paper, we focus on an application under a metasearch engine context and see that depending on the aggregation operator, the results may lead to a different ranking, one of which is almost the same as that of the minimax-LP method.

References

[1] Diaz E D, De A, Raghavan V. A comprehensive OWA-

- based framework for result merging in metasearch[C]//*Lecture Notes in Computer Science*. Springer, 2005, **3542**: 193 – 201.
- [2] Emrouznejad A. MP-OWA: the most preferred OWA operator[J]. *Knowledge-Based Systems*, 2008, **21**(8): 847 – 851.
- [3] Sugiura A, Etzioni O. Query routing for web search engines: architecture and experiments [J]. *Computer Networks*, 2000, **33**(1/2/3/4/5/6): 417 – 429.
- [4] Keyhanipour A H, Moshiri B, Kazemian M, et al. Aggregation of web search engines based on users' preferences in WebFusion [J]. *Knowledge-Based Systems*, 2007, **20**(4): 321 – 328.
- [5] Amin G R, Emrouznejad A. Finding relevant search engines results a minimax linear programming approach[J]. *Journal of Operational Research Society*, 2010, **61**(3): 1144 – 1150.
- [6] Amin G R, Emrouznejad A. Optimizing search engines results using linear programming[J]. *Expert Systems with Applications*, 2011, **38**(9): 11534 – 11537.
- [7] Dreilinger D, Howe A E. Experiences with selecting search engines using metasearch[J]. *ACM Trans Inf Syst*, 1997, **15**(3): 195 – 222.
- [8] Meng W Y, Yu C, Liu K L. Building efficient and effective metasearch engines[J]. *ACM Comput Surv*, 2002, **34**(1): 48 – 89.
- [9] Herrera-Viedma E, Gijón J L, Alonso S, et al. Applying aggregation operators for information access systems: an application in digital libraries[J]. *International Journal of Intelligent Systems*, 2008, **23**(12): 1235 – 1250.
- [10] De A, Diaz E D. Fuzzy analytical network models for metasearch [C]//*Studies in Computational Intelligence*. Springer, 2012, **399**: 197 – 210.
- [11] O'Hagan M. Aggregating template or rule antecedents in real-time expert systems with fuzzy set[C]//*Proceedings of the 22nd Asilomar Conference on Signals, Systems, Computers*. Piscataway, NJ: IEEE, 1988: 681 – 689.
- [12] Fullér R, Majlender P. An analytic approach for obtaining maximal entropy OWA operator weights[J]. *Fuzzy Sets and Systems*, 2001, **124**(1): 53 – 57.
- [13] Majlender P. OWA operators with maximal Rényi entropy [J]. *Fuzzy Sets and Systems*, 2005, **155**(3): 340 – 360.
- [14] Xu Z S. An overview of methods for determining OWA weights[J]. *International Journal of Intelligent Systems*, 2005, **20**(8): 843 – 865.
- [15] Liu X W. Models to determine parameterized ordered weighted averaging operators using optimization criteria [J]. *Information Sciences*, 2012, **190**(1): 27 – 55.
- [16] Yager R R. On ordered weighted averaging aggregation operators in multicriteria decision making [J]. *IEEE Transactions on Systems, Man and Cybernetics*, 1988, **18**(1): 183 – 190.
- [17] Zadeh L A. A computational approach to fuzzy quantifiers in natural languages[J]. *Computing and Mathematics with Applications*, 1983, **9**(1): 149 – 184.
- [18] Yager R R. Quantifier guided aggregation using OWA operators[J]. *International Journal of Intelligent Systems*, 1996, **11**(1): 49 – 73.
- [19] Liu X W, Han S L. Orness and parameterized RIM quantifier aggregation with OWA operators: a summary [J]. *International Journal of Approximate Reasoning*, 2008, **48**(1): 77 – 97.
- [20] Filev D, Yager R R. Analytic properties of maximum entropy OWA operators[J]. *Information Sciences*, 1995, **85**(1/2/3): 11 – 27.
- [21] Liu X W, Chen L H. On the properties of parametric geometric OWA operator [J]. *International Journal of Approximate Reasoning*, 2004, **35**(2): 163 – 178.

基于极大熵 OWA 算子的元搜索引擎搜索结果集成

桑秀芝 刘新旺

(东南大学经济管理学院, 南京 211189)

摘要:将极大熵有序加权平均 (ME-OWA) 算子应用于元搜索引擎搜索结果的集成, 并使用最新的极大熵有序加权平均算子的解析算法求解集成结果. 在现有的加权有序集成算子 (OWA) 环境下, 把用于集成元搜索引擎搜索结果的方法分为 2 类: 有唯一解和多解信息集成方法. 所提出的极大熵有序加权平均算子不仅可以得到精确的权重值, 还能为决策者提供多种集成结果供选择. 为了证明此集成算子的实用性, 在集成元搜索引擎搜索结果的应用背景下, 用极大熵有序加权平均算子和极小极大线性规划算法进行求解, 并对 2 种方法的运算结果加以比较和讨论. 结果显示, 极大熵有序加权平均算子具有和极小极大线性规划算法几乎一样的结果.

关键词:极大熵有序加权平均算子; 极小极大线性规划; 元搜索引擎; 信息集成

中图分类号: C934