RAIRO-Oper. Res. 46 (2012) 289–303
DOI: 10.1051/ro/2012019

**RAIRO Operations Research**
www.rairo-ro.org

# METASEARCH INFORMATION FUSION USING LINEAR PROGRAMMING

Gholam R. Amin[1], Ali Emrouznejad[2] and Hamid Sadeghi[3]

**Abstract.** For a specific query merging the returned results from multiple search engines, in the form of a metasearch aggregation, can provide significant improvement in the quality of relevant documents. This paper suggests a minimax linear programming (LP) formulation for fusion of multiple search engines results. The paper proposes a weighting method to include the importance weights of the underlying search engines. This is a two-phase approach which in the first phase a new method for computing the importance weights of the search engines is introduced and in the second stage a minimax LP model for finding relevant search engines results is formulated. To evaluate the retrieval effectiveness of the suggested method, the 50 queries of the 2002 TREC Web track were utilized and submitted to three popular Web search engines called Ask, Bing and Google. The returned results were aggregated using two exiting approaches, three high-performance commercial Web metasearch engines and our proposed technique. The efficiency of the generated lists was measured using TREC-Style Average Precision (TSAP). The new findings demonstrate that the suggested model improved the quality of merging considerably.

**Keywords.** Linear programming, search engine, metasearch, information fusion, information retrieval.

**Mathematics Subject Classification.** 90C05, 90C90, 68P20.

---

Received February 5, 2011. Accepted September 10, 2012.

[1] Postgraduate Engineering Centre, Islamic Azad University, South Tehran Branch, Tehran, Iran. g_amin@azad.ac.ir

[2] Aston Business School, Aston University, Birmingham, UK.

[3] Department of Computer Engineering, Hashtgerd Branch, Islamic Azad University, Alborz, Iran.

Article published by EDP Sciences

## 1. Introduction

A metasearch aggregation deals with the problem of fusion multiple search engines results in order to retrieve the most relevant resources for a submitted query. Many research studies have investigated that using the results of different web search engines can significantly improve the aggregated ranked list of documents [3, 8, 9, 19]. Meng *et al.* [14] proposed a survey for building efficient metasearch engine. Also, Yao *et al.* [25] reviewed the state of art methods regarding the web data fusion. Wu *et al.* [24] discussed on the experiments of linear combination of data merging and Wu [21] used the statistical principles in a metasearch aggregation. Several aggregation methods are proposed in the literature, for example Amin and Emrouznejad [2] introduced an improved ordered weighted averaging (OWA) model for aggregation of different results under uncertainty. Emrouznejad [10] used the OWA operator for aggregation of Web search engines. Also Herrera-Viedma *et al.* [12] analyzed the role of aggregation operators in the development of new technologies to access information on the Web. In information retrieval data fusion has been investigated by many authors. Wu and Crestani [23] developed an approach to estimate the performance of every input retrieval system using a subjective based method. Farah and Vanderpooten [11] proposed a rank aggregation method within a multiple criteria framework. Amin and Sadeghi [4] successfully applied prioritized aggregation operators for preference aggregation of returned results of Web search engines. Recent years have seen many attempts to enhance the quality of the aggregated documents [1, 5, 18].

Recently, Zhou *et al.* [28] introduced relevance features and a new ranking framework for content-based multimedia information retrieval (CBMIR). Also, Wu [22] suggested a multiple linear regression technique to obtain suitable weights for linear combination method.

Also, Amin and Emrouznejad [3] originated a linear programming (LP) model for finding relevant results of multiple search engines. The proposed LP model [3] considers all search engines equally important however this is not a true assumption in real world.

The aim of this paper is to delineate the importance weights of different search engines in the process of metasearch information fusion. The paper proposes a minimax LP model for aggregating of a metasearch engine results for a specific query. The main idea in the proposed model is to take into account the importance weights of different search engines and to compute more relevant aggregated documents. This is a two-phase method which in the first phase we suggest a new measure for computing the importance weights of underlying search engines and in the second stage we develop a minimax LP model for finding relevant search engines results.

The paper discusses on the property of the developed model. To test the quality of the proposed rank aggregation method we considered ten different queries related to "Operations Research" in the real world environment using two well web search engines, Google and Yahoo. We also used the experience of some OR experts to judge on the relevancy of the retrieved documents. The experimental result shows that our LP based rank aggregation can improve the quality of the result comparing with Amin and Emrouznejad [3], AE10, and Borda count methods.

The remainder of this paper is organized as follows. Section two, gives a brief explanation of the original LP formulation for a metasearch aggregation. The section also suggests a more general minimax LP model by considering the importance weights of underlying search engines. Section 3 develops a weighting method for computing the importance weights of different search engines. This is followed by investigating some properties of the suggested model in Section 4. Section 5 provides an experimental evaluation to test the quality of the proposed LP method. Section 6 gives the conclusion of the paper.

## 2. Metasearch information fusion

Let us consider a metasearch engine containing $m$ different search engines $SE_1, \ldots, SE_m$, where $m \geqslant 2$. Assume a specific query is issued to the metasearch engine. Then it passes the query to each of the search engines and receives $m$ ranked lists of resources or documents. Without loss of generality we consider only the first $l$th ranked results of documents retrieved from each search engine. Let $L_k$ denote the list of ranked documents obtained from the $k$th search engine, $k = 1, \ldots, m$. Therefore Table 1 is prepared for the metasearch engine.

Where, $D_{kj}$ is the retrieved document from the $k$th search engine in the $j$th ranked place, $k = 1, \ldots, m$, $j = 1, \ldots, l$. Now the problem of metasearch documents fusion is finding the first $l$th ranked relevant documents among the above table corresponding to the issued query. Let us $D_1, \ldots, D_r$ denote the distinct documents given in Table 1. To obtain the most relevant documents Amin and Emrouznejad [3], hereafter AE10, proposed the following minimax linear programming formulation.

$$\begin{array}{l} \min\ M \\ s.t. \\ \quad M - d_i \geqslant 0 \qquad\quad i = 1, \ldots, r \\ \quad \sum\limits_{j=1}^{l} \lambda_{ij} w_j + d_i = 1 \quad i = 1, \ldots, r \\ \quad w_j - w_{j+1} \geqslant \varepsilon^* \qquad j = 1, \ldots, l-1 \\ \quad w_l \geqslant \varepsilon^* \\ \quad d_i \geqslant 0 \quad i = 1, \ldots, r \end{array} \tag{2.1}$$

where, $\lambda_{ij}$ denotes the number of search engines (or lists) voted to the $i$th document in the $j$th ranked place, $w_j$ is the unknown weight assigned to the $j$th ranked place (for each $i = 1, \ldots, r,\ j = 1, \ldots, l$), $d_i$ is the deviation from relevancy index of the $i$th document, *i.e.* $d_i = 1 - z_i = 1 - \sum_{j=1}^{l} \lambda_{ij} w_j i = 1, \ldots, r$.

Also,

$$M = \max\{d_i : \ i = 1, \ldots, r\}$$

TABLE 1. The retrieved lists of documents.

| Lists \ places | The first place | ... | The $j$th place | ... | The $l$th place |
|---|---|---|---|---|---|
| $L_1$ | $D_{11}$ | $\cdots$ | $D_{1j}$ | $\cdots$ | $D_{1l}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $L_k$ | $D_{k1}$ | $\cdots$ | $D_{kj}$ | $\cdots$ | $D_{kl}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $L_r$ | $D_{r1}$ | $\cdots$ | $D_{rj}$ | $\cdots$ | $D_{rl}$ |

denotes the maximum deviation among the relevancy indices, and $\varepsilon^*$ is a feasible discrimination parameter as used by AE10 [3]. According to the minimax model (2.1) document $p$ has higher ranking place than document $q$ if and only if $d_p^* < d_q^*$, where $(w_1^*, \ldots, w_l^*, d_1^*, \ldots, d_r^*, M^*)$ is an optimal solution of the model. The aim in this paper is to improve model (2.1) for which the results of the metasearch aggregation can provide more relevant resources. The main disadvantage in model (2.1) is that the search engines are not equally important in practice, hence a weighting system is needed to differentiate the importance of the search engines as the current minimax model (2.1) considers the returned lists of the search engines be equally important.

Assume $v_k$ $(v_k > 0)$ denotes the importance weight assigned to the $k$th search engine (or the $k$th list), where $\sum_{k=1}^{m} v_k = 1$. For each search engine $k$ $(k = 1, \ldots, m)$, document $i$ $(i = 1, \ldots, r)$, and every place $j (j = 1, \ldots, l)$ we define

$$\delta_{ij}^k = \begin{cases} 1 & \text{if } SE_k \text{ gives the } i^{th} \text{ document in the } j^{\text{th}} \text{ position} \\ 0 & \text{otherwise} \end{cases}$$

So the relevancy index of the $i$th document can be restated as follows

$$\hat{z}_i = \sum_{j=1}^{l} \sum_{k=1}^{m} \delta_{ij}^k \, v_k w_j \quad i = 1, \ldots, r$$

Therefore the minimax LP model (2.1) can be written as in the following form.

$$\begin{aligned} \min \ & M \\ s.t. \ & \\ & M - d_i \geqslant 0 \quad i = 1, \ldots, r \\ & \sum_{j=1}^{l} \sum_{k=1}^{m} \delta_{ij}^k \, v_k w_j + d_i = 1 \quad i = 1, \ldots, r \\ & w_j - w_{j+1} \geqslant \hat{\varepsilon}^* \quad j = 1, \ldots, l-1 \\ & w_l \geqslant \hat{\varepsilon}^* \\ & d_i \geqslant 0 \quad i = 1, \ldots, r \end{aligned} \tag{2.2}$$

where, $\hat{\varepsilon}^*$ is a feasible discrimination parameter satisfying $\hat{\varepsilon}^* \in (0, \hat{\beta}]$, and

$$\hat{\beta} = \hat{\varepsilon}^*_{\max} = \min \left\{ \frac{1}{\hat{\beta}_i} : i = 1, \ldots, r \right\} \quad \hat{\beta}_i = \sum_{j=1}^{l} \sum_{k=1}^{m} (l - j + 1) \, \delta_{ij}^k \, v_k. \tag{2.3}$$

Clearly if we assume the entire search engines to be equally important, *i.e.* $v_k = \frac{1}{m} \forall k$, then

$$\hat{z}_i = \sum_{j=1}^{l} \sum_{k=1}^{m} \delta_{ij}^{k} \, v_k w_j = \sum_{k=1}^{m} \sum_{j=1}^{l} \delta_{ij}^{k} \, v_k w_j = \frac{1}{m} \sum_{k=1}^{m} \sum_{j=1}^{l} \lambda_{ij} w_j$$
$$= \sum_{j=1}^{l} \lambda_{ij} \, w_j = z_i \quad i = 1, \ldots, r.$$

Therefore model (2.2) generalizes the existing minimax model (2.1) proposed by AE10 [3]. In the next section we propose a technique for measuring the importance weights corresponding to the search engines.

## 3. Search engines weights

This section suggests an empirical method for computing the importance weights of the search engines. First let $v_k = \frac{1}{m}$ for all search engines $k = 1, \ldots, m$. We obtain the relevancy score of each document, for a specified query, using the minimax model (2.1) as follows.

$$z_i^* = \sum_{j=1}^{l} \lambda_{ij} w_j^* = 1 - d_i^* \quad i = 1, \ldots, r$$

where, $(w_1^*, \ldots, w_l^*, d_1^*, \ldots, d_r^*, M^*)$ is an optimal solution of model (2.1). From the above scores we rank the first $l$th relevant documents.

Assume $L_0 : D_{i_1} \succ D_{i_2} \succ \ldots \succ D_{i_l}$ denotes the initial aggregated list of documents. Now we measure the distance between the initial aggregated list $L_0$ and the list corresponding to the $k$th search engine, $L_k$ for each $k = 1, \ldots, m$, as follows.

$$d(L_0, L_k) = \sum_{j=1}^{l} \phi_j^k \qquad k = 1, \ldots, m \tag{3.1}$$

where,

$$\phi_j^k = \begin{cases} \frac{\left| j - \alpha_j^k \right|}{j} & if \; D_{i_j} \in L_k \\ \frac{l+1}{j} & if \; D_{i_j} \notin L_k \end{cases}$$

where, $\alpha_j^k$ is the position of $D_{i_j}$ in the list $L_k$, for each $k = 1, \ldots, m$ and $j = 1, \ldots, l$. In the case of $D_{i_j} \notin L_k$, that is $D_{i_j}$ is missed by the $k$th search engine, we define the related distance term as $l + 1$ divided by the corresponding position, as the longest possible distance term. Note that if $L_k = L_0$, for some $k = 1, \ldots, m$, then the above distance is zero. Without loss of generality, we assume $d(L_0, L_k) > 0$ for each search engine. Now we define the importance weight as below.

$$v_k = \frac{\theta_k}{\theta_1 + \ldots + \theta_m} \tag{3.2}$$

where, $\theta_i = (d(L_0, L_i))^{-1}$ for $i = 1, \ldots, m$. As $\theta_k$ is the inverse of the distance of the $k$th list from the initial aggregated list then the importance weight of the $k$th search engine defined in (3.2) can be interpreted as the inverse of the distance.

## 4. Property of the developed model

Now we investigate the relationship between the developed minimax model (2.2) and the LP formulation (2.1). Let $\hat{\lambda}_{ij} = \sum_{k=1}^{m} \delta_{ij}^k v_k$ for each $i = 1, \ldots, r$, $j = 1, \ldots, l$ and consider the following model.

$$\begin{aligned} \hat{z}_p^* = max \ & \sum_{j=1}^{l} \hat{\lambda}_{pj} w_j \\ s.t. \ & \\ & \sum_{j=1}^{l} \hat{\lambda}_{ij} w_j \leqslant 1 \quad i = 1, \ldots, r \\ & w_j - w_{j+1} \geqslant \hat{\varepsilon}^* \ \ j = 1, \ldots, l-1 \\ & w_l \geqslant \hat{\varepsilon}^* \end{aligned} \tag{4.1}$$

where, $\hat{z}_p^*$ is the score of the $p$th document, $p = 1, \ldots, r$, obtained by considering the importance weights of the search engines. Similar to the result obtained by AE10 [3], the following theorem holds when the search engines are not equally important.

**Theorem 4.1.** *Models (2.2) and (4.1) give the same relevancy score for the pth document ($p = 1, \ldots, r$).*

*Proof.* The proof is similar to the proof of Theorem 4 shown in AE10 [3]. □

*According to the above theorem to investigate the relationship between the optimal solutions of models (2.1) and (2.2) it is sufficient to obtain the relationship between model (4.1) and the following model.*

$$\begin{aligned} z_p^* = max \ & \sum_{j=1}^{l} \lambda_{pj} w_j \\ s.t. \ & \\ & \sum_{j=1}^{l} \lambda_{ij} w_j \leqslant 1 \quad i = 1, \ldots, r \\ & w_j - w_{j+1} \geqslant \varepsilon^* \quad j = 1, \ldots, l-1 \\ & w_l \geqslant \varepsilon^* \end{aligned} \tag{4.2}$$

*where, $\varepsilon^* \in (0, \beta]$ as defined in AE10 [3]. Let us suppose $\varepsilon^* = \varepsilon^*_{\max} = \beta$. Clearly model (4.2) is equivalent to the following model.*

$$\begin{aligned} z_p^* = max \ & \sum_{j=1}^{l} \lambda_{pj} w_j \\ s.t. & \\ & \sum_{j=1}^{l} \lambda_{ij} w_j \leqslant 1 \quad i = 1, \ldots, r \\ & w_j \geqslant (l - j + 1)\,\varepsilon^*_{\max} \quad j = 1, \ldots, l. \end{aligned} \tag{4.3}$$

*Denote $X$ as the feasible region of model (4.3). First, note that $\tilde{\mathbf{w}} = (\tilde{w}_1, \ldots, \tilde{w}_l)$ with $\tilde{w}_j = (l - j + 1)\,\varepsilon^*_{\max}$, $j = 1, \ldots, l$ is an extreme point of $X$, as $\tilde{\mathbf{w}} \in X$ and at least $l + 1$ linear independent hyperplanes of $X$ are binding at $\tilde{\mathbf{w}}$. Now we show that $\tilde{\mathbf{w}}$ is the only extreme point of $X$. On the contrary, assume that $\bar{\mathbf{w}} = (\bar{w}_1, \ldots, \bar{w}_l) \in X$ denotes another extreme point of $X$, that is $\bar{\mathbf{w}} \neq \tilde{\mathbf{w}}$. Therefore*

$$\bar{w}_j = \begin{cases} (l - j + 1)\,\varepsilon^*_{\max} & if \ j \in T_1 \\ \pi_j \varepsilon^*_{\max} & if \ j \in T_2 \end{cases}$$

*where, $\pi_j > (l - j + 1)$ for each $j \in T_2$ and $T_2 \neq \phi$, $T_1 \cup T_2 = \{1, \ldots, r\}$. The other constraints of $X$ imply that*

$$\sum_{j=1}^{l} \lambda_{ij} \bar{w}_j = \varepsilon^*_{\max} \left( \sum_{j \in T_1} (l - j + 1)\,\lambda_{ij} + \sum_{j \in T_2} \pi_j \lambda_{ij} \right) \leqslant 1 \quad i = 1, \ldots, r$$

*So*

$$\varepsilon^*_{\max} \leqslant \frac{1}{\sum_{j \in T_1} (l - j + 1)\,\lambda_{ij} + \sum_{j \in T_2} \pi_j \lambda_{ij}} \quad i = 1, \ldots, r.$$

*According to the assumption the above inequalities yield*

$$\varepsilon^*_{\max} < \frac{1}{\sum_{j \in T_1} (l - j + 1)\,\lambda_{ij} + \sum_{j \in T_2} (l - j + 1)\lambda_{ij}} \quad i = 1, \ldots, r.$$

*So*

$$\varepsilon^*_{\max} < \min \left\{ \frac{1}{\sum_{j=1}^{l} (l - j + 1)\,\lambda_{ij}} \ : \ i = 1, \ldots, r \right\} = \varepsilon^*_{\max}.$$

*Which is a contradiction. Therefore we have the following theorem.*

**Theorem 4.2.** *If $\varepsilon^* = \varepsilon^*_{\max}$ then $\mathbf{w}^* = (w_1^*, \ldots, w_l^*) = (l, l - 1, \ldots, 1)\,\varepsilon^*_{\max}$ is an optimal solution of model (4.2).*

A same result holds for model (4.1). That is

**Theorem 4.3.** *If $\hat{\varepsilon}^* = \hat{\varepsilon}^*_{\max}$ then $\hat{\mathbf{w}}^* = (\hat{w}_1^*, \ldots, \hat{w}_l^*) = (l, l - 1, \ldots, 1)\,\hat{\varepsilon}^*_{\max}$ is an optimal solution of model (4.1).*

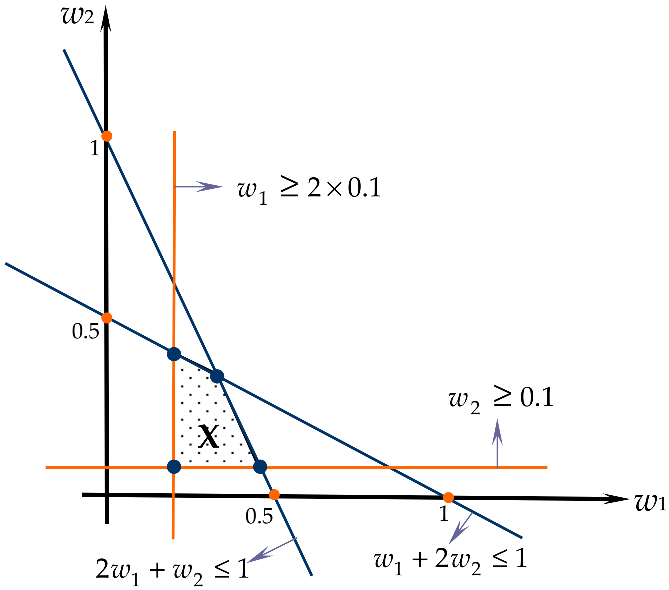


FIGURE 1. Feasible region X when $\varepsilon^* = 0.1 < \varepsilon^*_{\max}$.

Figure 1 presents the feasible region of model (4.3) corresponding for two documents.

As the figure shows we have $\lambda_{11} = 2$, $\lambda_{12} = 1$, $\lambda_{21} = 1$, $\lambda_{22} = 2$ and $\varepsilon^* = 0.1$. Also it illustrates if we impose the maximum discrimination among the positions, $i.e.\varepsilon^* = \varepsilon^*_{\max} = 0.2$, then the corresponding region becomes singleton.

The results of Theorems 2 and 3 conclude the optimal solutions of minimax models (2.1) and (2.2) in the case of imposing the maximum discrimination among the places.

In the next section we give experimental results for considering the importance weights of different search engines in a metasearch aggregation process.

## 5. Experimental results

In the following section we evaluate the proposed metasearch method. To see how the proposed method works we first give a simple example consisting of three search engines with a single query. This is followed by a more general illustration including ten queries related to "Operations Research". We use IR metrics to show the quality of aggregated documents using the importance of different search engines.

TABLE 2. The first tenth results of search engines.

| Search engines \ results | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $L_1$ = Google | $D_1$ | $D_2$ | $D_3$ | $D_4$ | $D_5$ |
| $L_2$ = Bing | $D_1$ | $D_2$ | $D_6$ | $D_7$ | $D_8$ |
| $L_3$ = Ask | $D_2$ | $D_1$ | $D_4$ | $D_9$ | $D_7$ |

### 5.1. A numerical illustration

Now, we illustrate the consideration of unequal importance weights for different search engines in a metasearch aggregation. We use the following three well known search engines, $m = 3$,

$$SE_1 = Google,\ \ SE_2 = Bing,\ \ SE_3 = Ask$$

Let us submit a query $q$, *"Operational Research"*, to the above search engines. Without loss of generality we only consider the first five ranked results (documents) returned from the search engines, that is $l = 5$. Table 2 shows the nine distinct retrieved resources, $r = 9$.

To extract the initial aggregated list, $L_0$, we consider an equal importance weights for all search engines. The corresponding minimax model (2.1) is asfollows.

$$\begin{aligned}
&\min\ M\\
&s.t.\\
&\ M - d_i \geqslant 0 \quad i = 1, \ldots, 9\\
&\ 2w_1 + w_2 + d_1 = 1, \quad w_1 + 2w_2 + d_2 = 1\\
&\ w_3 + d_3 = 1,\ \ w_3 + w_4 + d_4 = 1\\
&\ w_5 + d_5 = 1,\ \ w_3 + d_6 = 1\\
&\ w_4 + w_5 + d_7 = 1,\ \ w_5 + d_8 = 1,\ w_4 + d_9 = 1\\
&\ w_1 - w_2 \geqslant \varepsilon^*_{\max} = 0.0714,\ \ w_2 - w_3 \geqslant \varepsilon^*_{\max} = 0.0714\\
&\ w_3 - w_4 \geqslant \varepsilon^*_{\max} = 0.0714,\ \ w_4 - w_5 \geqslant \varepsilon^*_{\max} = 0.0714\\
&\ w_5 \geqslant \varepsilon^*_{\max} = 0.0714,\ \ d_i \geqslant 0 \quad i = 1, \ldots, 9.
\end{aligned}$$

where, $\varepsilon^*_{\max}$ is the maximum discrimination parameter among the places. According to Theorem 2 the optimal solution of the above model gives the initial aggregated list as shown below.

$$L_0 : \mathrm{D}_1 \succ \mathrm{D}_2 \succ \mathrm{D}_4 \succ \mathrm{D}_7 \succ \mathrm{D}_3.$$

Now, we compute the distance between $L_0$ and $L_k$, the retrieved list corresponding to the $k$th search engine $k = 1, 2, 3,$. According to the proposed formulation given in (3.1) we have

$$\begin{aligned}
d(L_0, L_1) &= \phi^1_1 + \ldots + \phi^1_5 = 0 + 0 + \tfrac{1}{3} + \tfrac{6}{4} + \tfrac{2}{5} = 2.23\\
d(L_0, L_2) &= \phi^2_1 + \ldots + \phi^2_5 = 0 + 0 + \tfrac{6}{3} + 0 + \tfrac{6}{5} = 3.2\\
d(L_0, L_3) &= \phi^3_1 + \ldots + \phi^3_5 = 1 + \tfrac{1}{2} + 0 + \tfrac{6}{4} + \tfrac{1}{5} = 3.2.
\end{aligned}$$

Notice that document $D_7$ which is the fourth result in the initial aggregated list $L_0$ is a missed document by the first search engine, that is $D_7 \notin L_1$. This impact is shown by the impact of the term $\phi_4^1 = \frac{6}{4}$ in the computing of $d(L_0, L_1)$. So, we have

$$\theta_1 = 0.4484,\ \theta_2 = 0.3125,\ \theta_3 = 0.3125$$

and therefore according to equation (3.2) we normalize the importance weights of the search engines as follows.

$$v_1 = \frac{0.4484}{1.0734} = 0.4178,\ v_2 = \frac{0.3125}{1.0734} = 0.2911,\ v_3 = \frac{0.3125}{1.0734} = 0.2911.$$

Hence, the corresponding minimax LP model (2.2) can be written as below

$$\begin{aligned}
&\min M\\
&s.t.\\
&M - d_i \geqslant 0 \quad i = 1, \ldots, 9\\
&(0.4178 + 0.2911)\, w_1 + 0.2911 w_2 + d_1 = 1,\ \ 0.2911\, w_1 + (0.4178 + 0.2911) w_2 + d_2 = 1\\
&0.4178 w_3 + d_3 = 1,\ \ 0.2911 w_3 + 0.4178 w_4 + d_4 = 1,\ 0.4178 w_5 + d_5 = 1\\
&0.2911 w_3 + d_6 = 1,\ 0.2911 w_4 + 0.2911 w_5 + d_7 = 1,\ \ 0.2911 w_5 + d_8 = 1\\
&0.2911 w_4 + d_9 = 1,\ \ w_1 - w_2 \geqslant \hat{\varepsilon}^*_{\max} = 0.2123,\ \ w_2 - w_3 \geqslant \hat{\varepsilon}^*_{\max} = 0.2123\\
&w_3 - w_4 \geqslant \hat{\varepsilon}^*_{\max} = 0.2123,\ \ w_4 - w_5 \geqslant \hat{\varepsilon}^*_{\max} = 0.2123,\ \ w_5 \geqslant \hat{\varepsilon}^*_{\max} = 0.2123\\
&d_i \geqslant 0 \quad i = 1, \ldots, 9.
\end{aligned}$$

where, the maximum discrimination parameter $\hat{\varepsilon}^*_{\max}$, is derived according to the compact form (2.3). Now Theorem 3 concludes that

$$\begin{aligned}
&d_1^* = 1 - (0.4178 + 0.2911)\, \hat{w}_1^* + 0.2911 \hat{w}_2^* = 0\\
&d_2^* = 0.0887,\ d_3^* = 0.7338,\ d_4^* = 0.637,\ d_5^* = 0.9112\\
&d_6^* = 0.8145,\ d_7^* = 0.8144,\ d_8^* = 0.9381,\ d_9^* = 0.8763.
\end{aligned}$$

So, the relevancy scores of documents are as follows

$$\hat{z}_1^* = 1,\ \hat{z}_2^* = 0.9113,\ \hat{z}_3^* = 0.2662,\ \hat{z}_4^* = 0.363,\ \hat{z}_5^* = 0.0888,\ \hat{z}_6^* = 0.1855,$$
$$\hat{z}_7^* = 0.1856,\ \hat{z}_8^* = 0.0619,\ \hat{z}_9^* = 0.1237.$$

Therefore, we have the following aggregated list.

$$\mathrm{D}_1 \succ \mathrm{D}_2 \succ \mathrm{D}_4 \succ \mathrm{D}_3 \succ \mathrm{D}_7.$$

In the coming section we evaluate the quality of the aggregated documents by weighting method.

### 5.2. Evaluation

In order to investigate the effectiveness of the proposed LP-based method an experiment using TREC dataset was conducted. The Text REtrieval Conference

(TREC) is a workshop series sponsored by the US National Institute of Standards and Technology (NIST) which sets standards to appraise the retrieval efficacy of an IR system. There are concrete TREC tracks for each type of problem in the IR context. Some are: blog track, robust retrieval track, Web track, spam track, million query track and so on. Among them, the Web track is the most pertinent dataset to our result merging approach, because it aims to explore retrieval behavior when the collection to be searched is a large hyperlinked structure such as the World Wide Web [20]. The 2002 TREC Web track offers 50 various queries. Each one comprises an index number, title, description and narrative. The title field consists of few words which is related to a query. The description field is one or two sentences which describes the topic and the intent of user in more detail. The narrative provides detailed explanation regarding the topic and describes what documents should be considered relevant to the corresponding topic.

In the present experiment, three popular Web search engines namely, Ask, Bing (Microsoft's search engine, formerly MSN Search, Windows Live Search and Live Search) and Google as the basis retrieval systems were chosen. It is interesting to note that Yahoo was excluded from the list because since July 2009, Yahoo Search has been powered by Bing. Next, each of the 50 queries was submitted to all the mentioned search engines. For each query, the top 10 results were retrieved. The obtained results were aggregated through our suggested LP-based merging technique. Moreover, its performance was compared with the two existing approaches (AE10 and Borda Count) and three commercial Web metasearch engines called Dogpile, MetaCrawler and WebFetch. According to [15] the three mentioned metasearch engines are high-quality ones which employ Ask, Bing and Google as their underlying resources.

The efficacy of the generated ranked result lists were measured via a renowned performance indicator in the IR field named TREC-Style Average Precision (TSAP). It has been widely used in the literature [7, 13, 16, 17, 26, 27]. TSAP is a human-based evaluation criterion that quantifies the relevance of each generated result list considering an issued query. TSAP at cutoff $n$, indicated as $TSAP@n$, is defined as:

$$TSAP@n = \frac{\sum_{i=1}^{n} r_i}{n}$$

where $r_i = 1/i$ if the $i^{th}$ ranked item is relevant and $r_i = 0$ otherwise. It is obvious that $TSAP@n$ takes into account both the number and ranks of the relevant documents in the top $n$ results. In the other words, $TSAP@n$ tends to yield a larger value when more relevant documents appear in the top $n$ results and also when they are ranked higher [13]. In order to compare the performance of each merging approach, the mean of TSAP@5 and TSAP@10 over all 50 queries were computed. Note that TSAP@5 ranges from 0 to 2.283, while TSAP@10 ranges from 0 to 2.928. The relevance of the documents in the aggregated result lists were judged by 4 experts. The results are shown in Table 3.

TABLE 3. Performance comparison of the proposed LP-based merging method with the other existing approaches using TSAP@5 and TSAP@10 measures.

| | AE10 | Borda Count | Dogpile | MetaCrawler | WebFetch | LP method |
|---|---|---|---|---|---|---|
| TSAP@5 | 1.534 | 1.386 | 1.501 | 1.454 | 1.436 | 1.680 |
| TSAP@10 | 1.915 | 1.643 | 1.832 | 1.769 | 1.757 | 2.092 |

As we can see from Table 3, our proposed LP-based method gives the best performance among all the approaches, followed by AE10 and then Dogpile. The next two positions were occupied by MetaCrawler and WebFetch, which almost shared similar performance. However, MetaCrawler performed slightly better than WebFetch on average. Moreover, Borda Count had the worst performance. Obviously, in comparison to the others, the TSAP@5 and TSAP@10 of the weighted LP model is significantly high which indicates more score to a document that appeared in more lists and more top places.

In order to check the closeness degree between the ranked result lists generated by our introduced method and those obtained from the other approaches a well-known statistical test called "extended version of Spearman's Footrule distance" [6] was conducted. Let $\sigma_1$ and $\sigma_2$ be two various ranked lists. Also, for each element (in our case URL) $i \in \sigma_j$ $(j \in \{1, 2\})$, let $\sigma_j(i)$ denote the position or rank of element $i$ in list $\sigma_j$. The extended version of Spearman's Footrule between top $K$ results of $\sigma_1$ and $\sigma_2$ is defined as:

$$CD^K(\sigma_1, \sigma_2) = \sum_Z \left| \frac{1}{\sigma_1(i)} - \frac{1}{\sigma_2(i)} \right| + \sum_S \left( \frac{1}{\sigma_1(j)} - \frac{1}{(K+1)} \right) + \sum_T \left( \frac{1}{\sigma_2(j)} - \frac{1}{(K+1)} \right)$$

where $Z$ is the set of overlapping documents, $S$ denotes the set of items that are only in the first list ($\sigma_1$) and $T$ implies the set of documents that appear only in the second list ($\sigma_2$). This measure has to be normalized as well, thus

$$NCD^K = 1 - \frac{CD^K}{\max CD^K}$$

where

$$\max\ CD^K = 2 \sum_{i=1}^{K+1} \left( \frac{1}{i} - \frac{1}{K+1} \right).$$

Hence, $NCD^K$ implies the normalized closeness degree between top $K$ results of two ranked lists in response to a specific query [6]. Moreover, the experiment can be run with various queries to get more stable result. The average of $NCD^K$

TABLE 4. Comparison of the ranked result lists generated by the proposed LP-based merging method with those obtained from the other existing approaches using $ANCD_{50}^{5}$ and $ANCD_{50}^{10}$ measures.

| | AE10 | Borda Count | Dogpile | MetaCrawler | WebFetch |
|---|---|---|---|---|---|
| $ANCD_{50}^{5}$ | 0.909 | 0.783 | 0.885 | 0.847 | 0.831 |
| $ANCD_{50}^{10}$ | 0.882 | 0.726 | 0.849 | 0.820 | 0.794 |

over $R$ different queries (in our case $R = 50$) is defined as:

$$ANCD_R^K = \frac{\sum_{i=1}^{R} NCD_i^K}{R}.$$

The measure assigns more weight to identical or near-identical rankings among the top-ranking documents. It attempts to capture the intuition that identical or near-identical rankings among the top documents are more valuable to the user than such similarities in ranking among the lower-placed documents. $ANCD$ ranges from 0 to 1. Clearly, the higher the $ANCD$, the stronger the correlation between the lists. Table 4 reports $ANCD_{50}^{5}$ and $ANCD_{50}^{10}$ between our suggested technique and the other methods.

As it is shown in Table 4, the strongest correlation exists between AE10 and our proposed method, which is significant. This value implies the result lists generated by these two algorithms are similar but not the same. Findings also reveal that Dogpile is highly correlated with our introduced technique. Also, note that even though in these two mentioned cases the reported values show high degrees of correlation, but they prove the methods create various lists of ranked results.

## 6. Conclusion

This paper suggested a minimax linear programming formulation for fusion of multiple search engines results retrieved for a specific user query. The paper has investigated that taking into account different levels of importance weights for the underlying search engines can provide more relevant aggregated results. Also, we developed a new method for obtaining the importance weights corresponding to the underlying search engines. Furthermore, an experimental investigation was used to show the quality of aggregated documents by the proposed mathematical model. The findings disclosed that the new model outperformed two existing approaches and three popular commercial Web metasearch engines.

*Acknowledgements.* This work was supported by Department of Research, Islamic Azad University, South Tehran Branch, under project 16/509, 88/9/8. The authors are grateful to anonymous reviewers and the editor of ROR for their constructive comments as result the paper has been improved substantially. The authors also thank to Dr Victoria Uren at Aston University for her useful comment.

## References

[1] L. Akritidis, D. Katsaros and P. Bozanis, Effective rank aggregation for metasearching. *J. Syst. Soft.* **84** (2011) 130–143.

[2] G.R. Amin and A. Emrouznejad, An extended minimax disparity to determine the OWA operator weights. *Comput. Ind. Eng.* **50** (2006) 312–316.

[3] G.R. Amin and A. Emrouznejad, Finding relevant search engines results: a minimax linear programming approach. *J. Oper. Res. Soc.* **61** (2010) 1144–1150.

[4] G.R. Amin and H. Sadeghi, Application of Prioritized Aggregation Operators in Preference Voting. *Int. J. Intell. Syst.* **25** (2010) 1027–1034.

[5] R.A. Baeza-Yates and B. Ribeiro-Neto, Modern information retrieval: the concepts and technology behind search, 2nd edition. ACM Press Books (2010).

[6] J. Bar-Ilan, M. Mat-Hassan and M. Levene, Methods for comparing rankings of search engine results. *Comput. Netwo.* **50** (2006) 1448–63.

[7] S.K. Deka and N. Lahkar, Performance evaluation and comparison of the five most used search engines in retrieving web resources. *Online Inf. Rev.* **34** (2010) 757–771.

[8] E.D. Diaz, A. De and V. Raghavan, A comprehensive OWA-based framework for result merging in metasearch. *Lect. Notes Comput Sci.* **3642** (2005) 193–201.

[9] A. De, E.D. Diaz and V. Raghavan, A fuzzy search engine weighted approach to result merging for metasearch. *Lect. Notes Comput Sci.* **4482** (2007) 95–102.

[10] A. Emrouznejad, MP-OWA: The Most Preferred OWA Operator. *Knowl-Based Syst.* **21** (2008) 847–851.

[11] M. Farah and D. Vanderpooten, An outranking approach for rank aggregation in information retrieval. *Proceedings of the 30th ACM SIGIR conference on Research and development in information retrieval.* Amsterdam, The Netherlands (2007) 591–598.

[12] E. Herrera-Viedma, J. Lopez Gijon, S. Alonso, J. Vilchez, C. Garcia and L. Villen, Applying Aggregation Operators for Information Access Systems: An Application in Digital Libraries. *Int. J. Intell. Syst.* **23** (2008) 1235–1250.

[13] Y. Lu, W. Meng, L. Shu, C. Yu and K.L. Liu, Evaluation of Result Merging Strategies for Metasearch Engines. *Lect. Notes Comput. Sci.* **3806** (2005) 53–66.

[14] W. Meng, C. Yu and K.L. Liu, Building efficient and effective metasearch engines. *ACM Comput. Surv.* **34** (2002) 48–89.

[15] H. Sadeghi, Assessing metasearch engine performance. *Online Inf. Rev.* **33** (2009) 1058–1065.

[16] H. Sadeghi, Empirical challenges and solutions in construction of a high-performance metasearch engine. *Online Inf. Rev.* **36** (2012) 713–723.

[17] S. Shekhar, K.V. Arya, R. Agarwal and R. Kumar, A WEBIR Crawling Framework for Retrieving Highly Relevant Web Documents: Evaluation Based on Rank Aggregation and Result Merging Algorithms, *International Conference on Computational Intelligence and Communication Networks (CICN)* (2011) 83–88.

[18] M. Shokouhi and L. Si, Federated Search. *Found. Trends Inf. Retr. FTIR* **5** (2011) 1–102.

[19] T.P.C. Silva, E.S. De Moura, J.M.B. Cavalcanti, A.S. Da Silva, M.G. De Carvalho and M.A. Gonçalves, An evolutionary approach for combining different sources of evidence in search engines. *Inform. Syst.* **34** (2009) 276–289.

[20] E. Voorhees, Overview of TREC 2002, in *Proceedings of the 11th Text REtrieval Conference (TREC)*, Gaithersburg, MD, USA (2002) 1–15.

[21] S. Wu, Applying statistical principles to data fusion in information retrieval. *Expert Syst. Appl.* **36** (2009) 2997–3006.

[22] S. Wu, Linear combination of component results in information retrieval. *Data Knowl. Eng.* **71** (2012) 114–126.

[23] S. Wu and F. Crestani, Data fusion with estimated weights, CIKM '02: *Proceedings of the eleventh international conference on Information and knowledge management* (2002) 648–651.

[24] S. Wu, Y. Bi, X. Zeng and L. Han, The Experiments with the Linear Combination Data Fusion Method in Information Retrieval. *Lect. Notes Comput. Sci.* **4976** (2008) 432–437.

[25] J.T. Yao, V. Raghavan and Z. Wu, Web information fusion: A review of the state of the art. *Inform Fusion* **9** (2008) 446–449.

[26] X.S. Xie and G. Zhang, Study of Optimizing the Merging Results of Multiple Resource Retrieval Systems by a Particle Swarm Algorithm. *International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)* (2011) 39–42.

[27] S. Zhou, M. Xu and J. Guan, LESSON: A system for lecture notes searching and sharing over Internet. *J. Syst. Soft.* **83** (2010) 1851–1863.

[28] G.T. Zhou, K.M. Ting, F.T. Liu and Y. Yin, Relevance feature mapping for content-based multimedia information retrieval. *Pattern Rec.* **45** (2012) 1707–1720