On data fusion in information retrieval using different aggregation operators

Julien Ah-Pine

Web Intelligence and Agent Systems: An International Journal (0) 1–0
IOS Press

On data fusion in information retrieval using different aggregation operators

Julien Ah-Pine [a],
[a] *Xerox Research Centre Europe, 6 chemin de Maupertuis, 38240 Meylan, France*
E-mail: julien.ah-pine@xrce.xerox.com

Abstract. This paper is concerned with the problem of unsupervised rank aggregation in the context of metasearch in information retrieval. In such tasks, we are given many partial ordered lists of retrieved items provided by many search engines and we want to define a way for aggregating those lists in order to find out a consensus. One classical approach consists in aggregating, for each retrieved item, the scores given by the different search engines. Then, we use the resulting aggregated scores distribution in order to infer a consensus ordered list. In this paper we investigate whether aggregation operators defined in the fields of multi-sensor fusion and multicriteria decision making are of interest for metasearch problems or not. Moreover, another purpose of this paper is to introduce a new aggregation operator, its foundations and its properties. We finally test all these aggregation operators for metasearch tasks using the Letor 2.0 dataset. Our results show that among the studied aggregation functions, the ones which are more compensatory outperform the baseline methods CombSUM and CombMNZ.

Keywords: Aggregation operator, fuzzy results merging, data fusion, metasearch, unsupervised rank aggregation

1. Introduction

Data fusion problems arise when, for a given task, there are several experts or information sources that give different opinions, or scores with respect to a set of alternatives or strategies. These opinions or scores can correlate or not and we are faced with the problem of combining them in order to figure out a consensus.

That kind of problem can be encountered in many different contexts. First, in social choice theory which encompasses voting theory and welfare economics, aggregating preference relations is a central problem. Indeed, in an election for example, we have to aggregate the votes or preferences of many voters in order to elect a consensual candidate.

Another example is the multicriteria decision making field. In that case, the decision maker is given a set of alternatives and a set of criteria. Each alternative is given a degree of satisfaction according to each criterion. But, the criteria are often conflicting. Thus, the decision maker needs to find a way to come up with a compromise.

Next, in the multi-sensor fusion field, the aim is to combine information coming from distinct sources in order to achieve a resulting information that is expected to be more accurate than the ones provided by one source only.

Lately, in information retrieval field, there have been several tasks for which aggregating information coming from different sources is beneficial. In [12] for example, the authors investigate the use of unsupervised rank aggregation to tackle the problem of web spam. In that case, combining the results obtained by several search engines allows to filter spam. Indeed, authors of web pages that attempt to enhance the ranking of their own pages by analyzing the drawbacks of one search engine cannot do it for all search engines. Apart from web spam filtering, there are many other applications in the information access domain where aggregation techniques play a crucial role [16]. In that context, tools such as multimedia search engines or recommender systems are examples that make use of data fusion techniques.

In this paper, we are concerned with the metasearch problems in information retrieval. In that case, the goal

is to combine the results given by different search engines into a single consensus list of items which is expected to be more relevant than any of the search engines results if they were considered individually.

Accordingly our first aim is to investigate if some aggregation operators defined in the fields of multisensor fusion and multicriteria decision making can be applied to solve metasearch problems. Next, our second goal is to introduce a new aggregation operator and to compare its results with other techniques in the context of metasearch problems. This paper is an expanded version of the conference paper [2].

The rest of the paper is organized as follows. In section 2, we recall some previous works that address metasearch problems and we also introduce some notations. In section 3, we explain the pre-processing step which is necessary before applying any aggregation operators. The latter consists in normalizing each scores distributions so that their values become comparable. Then in section 4, we recall some famous aggregation functions that are studied in different contexts namely, weighted quasi-arithmetic means, triangular norms and conorms and triangular norms ordered weighted averaging operator (TOWA). Next, in section 5, we introduce a new aggregation operator, its set-based foundations and its combinatorial properties. We tested all the aforementioned approaches for metasearch tasks using the Letor 2.0 dataset [21]. We report and we analyse the results we obtained in section 6. Finally we conclude and sketch some future work in section 7.

2. Related work and notations

Basically, we suppose there are M experts (the search engines) and N alternatives (the documents). Each expert gives an ordered list of those N objects according to their own preferences (or scores).

However, when dealing with information retrieval tasks there are some particular points that are worth noticing (see [14] for more details). In general, the number of objects of the collection N is much larger than the number of search engines M. In voting theory for example, this is indeed the contrary: there are more voters ("experts") than candidates ("alternatives"). Besides, given a query, each search engine attempts to give in its top-list the most relevant objects. Therefore, the goal of search engines is not to score every objects in the collection. Furthermore, we can also assume that after a certain position, let say 100, objects that are ranked beyond are not relevant. Consequently, each search engine gives a partial ranking of the objects.

The top-lists provided by the M search engines are not mutually disjoint: there are overlaps between them. Thus, one can assume that objects that appear in many top-lists are more likely to be relevant. Accordingly, the problem we address is to find a single consensus list which combines the M rankings in order to achieve a more consistent list.

In the literature, there are different kinds of method to address the unsupervised rank aggregation problem. According to [26,14], we can notice two broad families of approaches:

- *positional methods*: for each object, we consider the scores given by each expert (the object's profile), we then aggregate these scores using different techniques and we finally re-rank the objects using the aggregated scores. The first positional method was proposed by Borda [5] but linear combination such as the weighted arithmetic mean is for instance, one of these methods [15]. Example of more recent approaches are the following ones [19,10]. In [19], it is proposed to learn the weight assigned to any search engine in an unsupervised manner. In [10], the authors apply aggregation functions based on ordered weighted averaging functions[1] (OWA), originally designed in the domain of multicriteria decision making, to metasearch problems.
- *majoritarian methods*: these methods use pairwise comparisons matrices of objects and are mostly based upon order relations aggregation techniques using association criteria such as Condorcet's criterion [9,23,24] or distance criteria such as Kendall distance [22,13]. Other methods have also been proposed using Markov chain models [12]. Lastly, outranking methods which were introduced in multicriteria decision problems have also been adapted for information retrieval tasks [27,14].

In this paper we are exclusively concerned with approaches that belong to the positional family. Before introducing the most popular techniques used in this context, we first give some basic notations:

[1] We particularly recall the triangular norm ordered weighted averaging function in subsection 4.3.

- $S_j(i)$ is the normalized score expert j assigns to object i (see the next section for details about the normalization step).
- $\mathcal{L}(j)$ is the ordered list of expert j.
- $\mathcal{L}_k(j)$ is the first k objects of $\mathcal{L}(j)$ or the top-k list of expert j. We will also denote $\mathbb{L}_k = \{\mathcal{L}_k(j); j = 1, \ldots, M\}$ the set of all top-k lists.
- $\mathbb{U} = \bigcup_{j=1}^{M} \mathcal{L}_k(j)$ is the union set of the M top-k lists. We will also denote $K = \#\mathbb{U}$, the number of elements in $\mathbb{U}$.
- $\forall i \in \mathbb{U} : \mathrm{NZ}(i, \mathbb{L}_k)$ is the number of top-k lists in which object i appears.

Using these notations, we can define below the general linear combination which represents the baseline methods for unsupervised rank aggregation problems in information retrieval [15], $\forall i \in \mathbb{U}$:

$$\mathrm{Comb}_{v,w}(i) = \mathrm{NZ}(i, \mathbb{L}_k)^v \sum_{j=1}^{M} w_j S_j(i) \tag{1}$$

where v is an integer and w a weight vector.

In this kind of approach, the most usually employed cases are the following ones [15], $\forall i \in \mathbb{U}$:

$$\mathrm{CombSUM}(i) = \sum_{j=1}^{M} S_j(i) \tag{2}$$

$$\mathrm{CombMNZ}(i) = \mathrm{NZ}(i, \mathbb{L}_k) \sum_{j=1}^{M} S_j(i) \tag{3}$$

CombSUM is the aggregated score given by the sum over the experts scores. CombMNZ weight the latter with respect to the number of experts who have ranked the element in their top-k list. According to CombMNZ, the greater $\mathrm{NZ}(i, \mathbb{L}_k)$, the higher the aggregated score. The NZ factor amounts to reinforcing the aggregated score on the basis of the number of times an object reaches a top-k list among the search engines. Suppose that for two objects i and i', $\mathrm{CombSUM}(i) = \mathrm{CombSUM}(i')$ than the $\mathrm{CombMNZ}$ approach would rank higher the object that appears most often in the top-k lists.

As mentioned previously, some methods have been proposed for learning the weight vector w in an unsupervised manner. In [19] for example, an unsupervised learning algorithm is suggested. However, this approach is a linear combination of scores. In our case, we rather focus on non-linear aggregation operators.

3. Normalization step

The normalization step is inherent to any positional method. Indeed, each expert may have its own particular score distribution which belongs to a certain range of values. As a result, before applying any aggregation function, we have to normalize each distribution if we want them to be mutually comparable. There exist several normalization methods [25]. However, most of the aggregation operators we are going to introduce in sections 4 and 5 use different tools taken from fuzzy logic. In other words, they manipulate truth measures that belong to $[0, 1]$. Thus, we propose to normalize the experts scores as follows.

Let us denote by $V_j(i)$, the original score value the expert j assigns to object i. Then, its associated normalized score denoted by $S_j(i)$ is given by, $\forall j = 1, \ldots, M; \forall i \in \mathcal{L}_k(j)$:

$$S_j(i) = \frac{V_j(i) - \min_{i' \in \mathcal{L}_k(j)}\{V_j(i')\}}{\max_{i' \in \mathcal{L}_k(j)}\{V_j(i')\} - \min_{i' \in \mathcal{L}_k(j)}\{V_j(i')\}} \tag{4}$$

With respect to expert j, the previous normalization formula is given for all objects belonging to the expert's top-k list, $\mathcal{L}_k(j)$. We still need to assign normalized scores to objects that belong to $\mathbb{U} - \mathcal{L}_k(j)$. The only thing we can assume about these latter objects is that they are not preferred to any of the objects belonging to $\mathcal{L}_k(j)$. Consequently, $\forall j = 1, \ldots, M; \forall i \in \mathbb{U} - \mathcal{L}_k(j)$, we set: $S_j(i) = 0$.

4. Recalling some aggregation operators

Many aggregation functions have been defined in the field of information fusion and multicriteria decision making (see for example [4,6,11]). We recall some of them namely weighted quasi-arithmetic means, triangular norms, triangular conorms and triangular norm OWA operators.

4.1. Weighted quasi-arithmetic means

Weighted quasi-arithmetic means, also called weighted generalized f-means, are functions that generalize the classical weighted arithmetic mean by using a transformation function f. In our case, we consider $f : [0, 1] \to [-\infty, \infty]$. f must be continuous, strictly monotonic, and it must also respects the following conditions: $\{f(0), f(1)\} \neq \{-\infty, \infty\}$. The quasi-

arithmetic mean associated to f is denoted $\mathbf{M}_f$. Given a sequence of M values $(a_1, \ldots, a_M) \in [0,1]^M$ and a system of weights (or a weight vector) $(w_1, \ldots, w_M)$ such that $\sum_{j=1}^{M} w_j = 1$, the weighted quasi-arithmetic mean is defined by:

$$\mathbf{M}_f(a_1, \ldots, a_M) = f^{-1}\left(\sum_{j=1}^{M} w_j f(a_j)\right)$$

We will consider more particularly, the case of power functions $f(a) = a^p$ with $p \in \mathbb{R}$ which are strictly monotonic functions. In that case, these operators are also called weighted power means.

$$\mathbf{M}_p(a_1, \ldots, a_M) = \left(\sum_{j=1}^{M} w_j (a_j)^p\right)^{1/p} \quad (5)$$

We can notice that CombSUM is equivalent to the arithmetic mean given by $p = 1$ associated to a uniform weight distribution. Weighted power means are non-decreasing with respect to p, $\forall -\infty < p \leq p' < +\infty$:

$$\mathbf{M}_p(a_1, \ldots, a_M) \leq \mathbf{M}_{p'}(a_1, \ldots, a_M)$$

Furthermore, they are lower bounded by the minimum function $\min(a_1, \ldots, a_M)$ which corresponds to the limit case $p \to -\infty$. Likewise, they are upper bounded by $\max(a_1, \ldots, a_M)$ which corresponds to the dual limit case $p \to +\infty$.

Other famous particular cases are the weighted harmonic, geometric and quadratic means which are respectively given by $p = -1$, the limit case $p \to 0$ and $p = 2$.

The weighted power mean is compensatory in the sense that its values are bounded by the minimal and the maximal values of its arguments. Particularly, as p grows, the averaged value tends to the maximum and we say that the limit value of the power mean has a disjunctive behavior. On the contrary, when p tends to $-\infty$, we say that the limit value of the power mean reflects a conjunctive behavior. These considerations will become clearer with the following subsection.

4.2. Triangular norms and conorms

Triangular norms (t-norms) are binary operations $T : [0,1]^2 \to [0,1]$ which are associative, commutative, non decreasing in each argument and such that $T(a,1) = a$, $\forall a \in [0,1]$. They are special functions used in probabilistic metric spaces [28], and fuzzy logic [18].

In fuzzy logic particularly, these operations extend the conjunction from the binary case to the multi-valued case. Let $\mathbb{A}$ and $\mathbb{B}$ be two events and $\mu(\mathbb{A}) = a$ and $\mu(\mathbb{B}) = b$ their associated truth values in $[0,1]$. Typically, in our case, atomic events are: "object i is relevant for expert j". Then, $T(a,b)$ measures the truth value of the conjunction $\mathbb{A} \cap \mathbb{B}$[2]. As t-norms are associative and commutative, they can be extended easily for measuring the conjunction of more than two events. We have for example: $T(a,b,c) = T(T(a,b),c)$ and so on.

T-norms are aimed at measuring the conjunction of events provided by different sources. However, the disjunction measures of events are also of interest. Then, departing from a t-norm T, we can define its related triangular conorm T^* (t-conorm) as follows:

$$T^*(a,b) = 1 - T(1-a, 1-b) \quad (6)$$

Similarly to t-norms, t-conorms allow one to extend the disjunction to the multi-valued case.

There exist non-parametric and parametric t-norms and t-conorms. The four fundamental non-parametric ones are:

- the minimum t-norm and its related t-conorm:

$$T_M(a,b) = \min(a,b)$$
$$T^*_M(a,b) = \max(a,b)$$

- the product t-norm and its related t-conorm:

$$T_P(a,b) = ab$$
$$T^*_P(a,b) = a + b - ab$$

- the Lukasiewicz t-norm and t-conorm:

$$T_W(a,b) = \max(a + b - 1, 0)$$
$$T^*_W(a,b) = \min(a + b, 1)$$

- the drastic product and its related t-conorm:

$$T_D(a,b) = \begin{cases} 0 & \text{if } (a,b) \in [0,1[^2 \\ \min(a,b) & \text{otherwise} \end{cases}$$

$$T^*_D(a,b) = \begin{cases} 1 & \text{if } (a,b) \in]0,1]^2 \\ \max(a,b) & \text{otherwise} \end{cases}$$

T_M and T_D are respectively, upper and lower bounds for t-norms. Thus, if T is a t-norm, we have $\forall (a,b) \in [0,1]^2$:

$$T_D(a,b) \leq T(a,b) \leq T_M(a,b) \quad (7)$$

[2] With a set-based view of events.

T_M gives the highest value for the conjunction of two events: it assumes that one event is included in the other one and thus a strong relationship between them. T_P gives the conjunction value of two events assuming that the latter are independent. On the contrary, T_W which is lower than T_P, supposes that there is a rather low dependence between the events. Finally, T_D gives a zero value to the conjunction of two events unless one of the two has a truth value 1.

Similarly, T^*_M and T^*_D are respectively, lower and upper bounds for t-conorms. Thus, if T^* is a t-conorm, we have $\forall (a,b) \in [0,1]^2$:

$$T^*_M(a,b) \leq T(a,b) \leq T^*_D(a,b)$$

There also exist numerous parametric t-norms and t-conorms. However, we will introduce only one pair of them: the Schweizer-Sklar family for which the non-parametric t-norms and conorms are particular cases.

The Schweizer-Sklar t-norm is given by, $\forall \lambda \in [-\infty, \infty]$:

$$T^{\lambda}_{SS}(a,b) = \begin{cases} T_M(a,b) \ \lambda = -\infty \\ T_P(a,b) \ \ \lambda = 0 \\ T_D(a,b) \ \ \lambda = \infty \end{cases}$$

and for $\lambda \in]-\infty, 0[\cup]0, \infty[$, we have:

$$T^{\lambda}_{SS}(a,b) = \left(\max(a^{\lambda} + b^{\lambda} - 1, 0)\right)^{\frac{1}{\lambda}}$$

Its corresponding t-conorm is as follows:

$$T^{*\lambda}_{SS}(a,b) = \begin{cases} T^*_M(a,b) \ \lambda = -\infty \\ T^*_P(a,b) \ \ \lambda = 0 \\ T^*_D(a,b) \ \ \lambda = \infty \end{cases}$$

and for $\lambda \in]-\infty, 0[\cup]0, \infty[$, we have:

$$T^{*\lambda}_{SS}(a,b) = 1 - \left(\max((1-a)^{\lambda} + (1-b)^{\lambda} - 1, 0)\right)^{\frac{1}{\lambda}}$$

T-norms or t-conorms can be used as aggregation operators. In that case they represent two extreme point of views which are respectively the conjunction and the disjunction over all their arguments. In our case, this would be the truth value of "object i is relevant for all experts" and the truth measure of "object i is relevant for at least one expert".

4.3. Triangular norm ordered weighted averaging operators (TOWA)

Ordered weighted averaging operators (OWA) are well-known aggregation techniques that were introduced by R.R. Yager in the field of multicriteria decision making in [30].

In the basic case, OWA operators aggregate a set of values $(a_1, \ldots, a_M)$ with respect to a system of weights $(w_1, \ldots, w_M)$ such that $w_j \in [0,1]; j = 1, \ldots, M$ and $\sum_{j=1}^{M} w_j = 1$. Let denote $(\sigma(1), \ldots, \sigma(M))$ a permutation of $(1, \ldots, M)$ that satisfies:

$$a_{\sigma(1)} \geq a_{\sigma(2)} \geq \ldots \geq a_{\sigma(M)}$$

Then the OWA operator is defined as follows:

$$\text{OWA}(a_1, \ldots, a_M) = \sum_{j=1}^{M} w_j a_{\sigma(j)} \tag{8}$$

As far as the weighted arithmetic mean is concerned, we can see that the system of weights attributes different measures of importance to the different experts j. On the contrary, OWA functions rather focus on the position with respect to the non-increasing order of the values. In other words, independently of the expert, OWA operators assign the same weight w_j to the j^{th} greatest value, $a_{\sigma(j)}$, in the sequence of measures to combine.

It is easy to see that OWA functions generalize order statistics. It is a compensatory aggregation operator bounded by the minimum and the maximum of its arguments, unlike t-norms and t-conorms but such as weighted power means.

In [31,32], the authors consider a particular extension of OWA functions by mixing t-norms and OWA functions. The resulting triangular norm OWA operator (TOWA) is defined below:

$$\begin{aligned} &\mathbf{TOWA}_T(a_1, \ldots, a_M) \\ &= \sum_{j=1}^{M} w_j T(a_{\sigma(1)}, \ldots, a_{\sigma(j)}) \end{aligned} \tag{9}$$

where T is a t-norm as described in the previous subsection.

TOWA functions generalize OWA operators since the latter correspond to the former when $T = T_M$. Indeed, in Eq. (9), $T_M(a_{\sigma(1)}, \ldots, a_{\sigma(j)})$ reduces to $a_{\sigma(j)}$ which leads to the classical OWA function.

TOWA allows one to take into account two different aspects that are of interest when combining data. First, using different t-norms enables one to consider different kinds of conjunctive relationship between subset of values based upon their position with respect to the non-increasing order. Second, the weight vector allows one to model fuzzy linguistic concepts such as "for all" or "at least m". Regarding those two latter examples, the corresponding weight systems would be: on

the one hand, $w_j = 0; \forall j \neq M$ and $w_M = 1$; on the other hand, $w_j = 0; \forall j \neq m$ and $w_m = 1$.

It is easy to see that the case $w_j = 0; \forall j \neq M$ and $w_M = 1$ reduces to the conjunction over all values given by the t-norm since we obtain $\mathbf{TOWA}_T(a_1, \ldots, a_M) = T(a_1, \ldots, a_M)$. As a result, $\mathbf{TOWA}_T(a_1, \ldots, a_M)$ is lower bounded by $T(a_1, \ldots, a_M)$ unlike $\mathrm{OWA}(a_1, \ldots, a_M)$ whose lower bound is $\min(a_1, \ldots, a_M)$. However, both functions are upper bounded by $\max(a_1, \ldots, a_M)$.

The distribution of weights allows one to control the compensatory degree (also called orness or andness degrees) of the TOWA functions. A general approach for computing the weights $w_j; j = 1, \ldots, M$, consists in using fuzzy quantifiers. The most used method is based upon regular increasing monotone (RIM) functions denoted Q [30]. The latter are functions from $[0, 1]$ to $[0, 1]$ such that $Q(0) = 0$, $Q(1) = 1$ and $a < b \Rightarrow Q(a) \leq Q(b)$. $Q(a)$ gives the degree of satisfaction for the decision maker, when $a\%$ of the criteria is fulfilled. Assuming that Q is a RIM function is the same as assuming that the more criteria fulfilled, the greater the satisfaction.

Given Q, one can define a weight vector as follows, $\forall j = 1, \ldots, M$:

$$w_j = Q\left(\frac{j}{M}\right) - Q\left(\frac{j-1}{M}\right) \tag{10}$$

In that case, w_j represents the marginal satisfaction a decision maker gets from having $j - 1$ criteria to j criteria fulfilled.

Power functions, $Q(a) = a^q$ with $q > 0$ are examples of RIM functions. In that case, the greater q, the more the TOWA operator has a conjunctive behavior.

4.4. Some basic properties of aggregation operators

Let us now replace ourselves in the context of metasearch problems and consider the notations introduced in section 2,

Then, in section 3, we have presented the normalization step which is aimed at making the scores between the different search engines comparable. With respect to Eq. (4), we propose to normalize the score of elements i between 0 and 1, for each search engine j. The resulting normalized score, denoted $S_j(i)$, is a non-negative measure and we proposed to interpret the latter as a truth measure. As a result, all of the aggregation operators that we have recalled previously can be employed.

Let us denote $\mathbf{C}$ any of the aggregation functions we have introduced so far. They are all functions from $[0, 1]^M$ to $[0, 1]$. Furthermore, they all respect the following basic properties (see for example [7] for a list of basic conditions for aggregation operators):

- Identity when unary:
  $\mathbf{C}(S_1(i)) = S_1(i)$
- Boundary conditions:
  $\mathbf{C}(0, \ldots, 0) = 0$ and $\mathbf{C}(1, \ldots, 1) = 1$
- Non-decreasing:
  $\{\forall j = 1, \ldots, M : S_j(i) \leq S_j(i')\}$
  $\Rightarrow \mathbf{C}(S_1(i), \ldots, S_M(i)) \leq \mathbf{C}(S_1(i'), \ldots, S_M(i'))$
- Symmetry (or commutativity):
  $\mathbf{C}(S_1(i), \ldots, S_M(i)) = \mathbf{C}(S_{\sigma(1)}(i), \ldots, S_{\sigma(M)}(i))$
  where σ is a permutation of $(1, 2, \ldots, M)$.

Moreover, it is worth mentioning again the different kinds of behavior that an aggregation operator can have: conjunctive (t-norms), disjunctive (t-conorms) and compensatory (weighted power means and TOWA). This can be underlined following the relations below:

$$0 \leq T \leq \min \leq \mathbf{M}_p \leq \max \leq T^* \leq 1$$

$$0 \leq T \leq \mathbf{TOWA}_T \leq \max \leq T^* \leq 1$$

5. A new consensus aggregation operator

In addition to the previous aggregation functions, we introduce a new approach for data fusion. Unlike t-norms and t-conorms which are not compensatory functions and which represent complete conjunctive and disjunctive behaviors, the method we are going to introduce is more compensatory. Thus, it is comparable to weighted quasi-arithmetic means and TOWA functions. After having introduced our proposal, we will particularly underline the relationships it has with the TOWA method.

5.1. Basic concepts and theoretical framework

Let us denote by $\mathbb{S}_j(i)$ the event: "object i is relevant for expert j". Then for each object i, we consider the following set of atomic events $\mathbb{X}_i = \{\mathbb{S}_1(i), \ldots, \mathbb{S}_M(i)\}$ and $\sigma(\mathbb{X}_i)$ a $\sigma-$algebra over $\mathbb{X}_i$. In other words, $\sigma(\mathbb{X}_i)$ is the set of events generated by $\mathbb{X}_i$ that is closed under set operations.

We give in Fig. 1 a Venn diagram representing the case $M = 3$. Each circle corresponds to an atomic event. The more irregular regions corresponding to different numbers are mutually disjoint events given by:

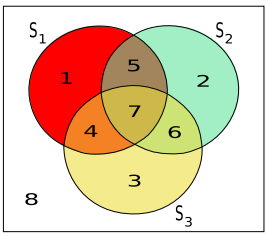

Fig. 1. Venn diagram and disjoint regions

1 $\mathbb{S}_1(i) \cap \overline{\mathbb{S}_2}(i) \cap \overline{\mathbb{S}_3}(i)$
2 $\mathbb{S}_2(i) \cap \overline{\mathbb{S}_1}(i) \cap \overline{\mathbb{S}_3}(i)$
3 $\mathbb{S}_3(i) \cap \overline{\mathbb{S}_1}(i) \cap \overline{\mathbb{S}_2}(i)$
4 $\mathbb{S}_1(i) \cap \mathbb{S}_3(i) \cap \overline{\mathbb{S}_2}(i)$
5 $\mathbb{S}_1(i) \cap \mathbb{S}_2(i) \cap \overline{\mathbb{S}_3}(i)$
6 $\mathbb{S}_2(i) \cap \mathbb{S}_3(i) \cap \overline{\mathbb{S}_1}(i)$
7 $\mathbb{S}_1(i) \cap \mathbb{S}_2(i) \cap \mathbb{S}_3(i)$
8 $\overline{\mathbb{S}_1}(i) \cap \overline{\mathbb{S}_2}(i) \cap \overline{\mathbb{S}_3}(i)$

Our assumption is that $(\mathbb{X}_i, \sigma(\mathbb{X}_i))$ is a measurable space[3] and for all atomic events $\mathbb{S}_j(i); j = 1, \ldots, M$, we are given a positive measure $\mu(\mathbb{S}_j(i)) \geq 0$. Since μ is a measure it then satisfies the σ-additivity property: if $\mathbb{A}$ and $\mathbb{B}$ are two incompatible events (disjoint sets) then $\mu(\mathbb{A} \cup \mathbb{B}) = \mu(\mathbb{A}) + \mu(\mathbb{B})$.

If we consider all M experts, we can define the following event: "object i is relevant for at least m experts among M". We will denote this event $\mathbb{E}_m^M(i)$. Formally, we can define the latter from a set viewpoint as follows, $\forall m = 1, \ldots, M; \forall i = 1, \ldots, K$:

$$\mathbb{E}_m^M(i) = \bigcup_{1 \leq j1 < \ldots < jm \leq M} \left(\bigcap (\mathbb{S}_{j1}(i), \ldots, \mathbb{S}_{jm}(i)) \right) \quad (11)$$

Given the measures, $\mu(\mathbb{S}_j(i)); j = 1, \ldots, M$, of event $\mathbb{S}_j(i); j = 1, \ldots, M$, we are interested in computing the measure of the events $\mathbb{E}_m^M(i)$ for $m = 1, \ldots, M$.

This is a combinatorial problem which generalizes Poincaré's formula[4] also known as the "inclusion-exclusion principle". Since μ is a measure, we can show [17,1] that $\mu(\mathbb{E}_m^M(i))$ can be expressed as a linear combination of the following quantities, $\forall i = 1, \ldots, K$:

– for $l = 1$:

$$\mathbf{S}_l^M(i) = \sum_{j=1}^{M} \mu(\mathbb{S}_j(i)) \quad (12)$$

– for $l = 2, \ldots, M$:

$$\mathbf{S}_l^M(i) = \sum_{1 \leq j1 < \ldots < jl \leq M} \mu\left(\mathbb{S}_{j1}(i) \cap \ldots \cap \mathbb{S}_{jl}(i)\right) \quad (13)$$

This linear combination is related to Jordan's combinatorial formulas [17,8], and we have, $\forall m = 1, \ldots, M; \forall i = 1, \ldots, K$:

$$\mu(\mathbb{E}_m^M(i)) = \sum_{l=m}^{M} (-1)^{l-m} \binom{l-1}{m-1} \mathbf{S}_l^M(i) \quad (14)$$

where $\binom{n}{p} = \frac{n!}{p!(n-p)!}$, are binomial coefficients.

We also have the following formulation [1]:

$$\mu(\mathbb{E}_m^M(i)) = \sum_{l=m}^{M} \binom{-m}{l-m} \mathbf{S}_l^M(i) \quad (15)$$

where $\binom{-n}{p} = \frac{(-n)(-n-1)\ldots(-n-p+1)}{p!}$, are the negative binomial coefficients.

5.2. Designing a consensus aggregation operator

Now that we have introduced the consensus measures $\mu(\mathbb{E}_m^M); m = 1, \ldots, M$, we make the following assumption on which our aggregation function is based: the greater the number of experts who find object i relevant, the higher the aggregated score corresponding to object i must be[5] [20].

In a literal formulation, our aggregation function could be defined as follows: we weight by m the measure that object i is relevant for at least m experts then we sum over m from 1 to M and we normalize this aggregated score according to $\sum_{m=1}^{M} m = M(M+1)/2$. As the proposed method gives increasing weights when m, the number of considered experts, grows, it is a method that satisfies the assumption we have mentioned beforehand.

We denote our consensus aggregation function $\mathbf{A}$. It is formally defined by the following equation, $\forall i =$

[3] It is not in general a probabilistic space since we cannot guarantee that the probabilities of the disjoint sets, depicted in Fig. 1 for example, always sum to 1.

[4] Poincaré's formula corresponds to the case $m = 1$.

[5] Notice that baseline methods given in Eq. (1) pursue the same goal by weighting by $\mathrm{NZ}(i, \mathbf{L}_k)^v$ the experts score sum. However, our proposal is different in spirit and corresponds to a non-linear and smoother way to reinforce the aggregated scores. See subsection 5.3 for more details.

$1, \ldots, K$:

$$\mathbf{A}(i) = \frac{1}{\frac{M(M+1)}{2}} \sum_{m=1}^{M} m\mu(\mathbb{E}_m^M(i)) \tag{16}$$

In its original formulation given by equations Eqs. (14) and (13), the computational cost of $\mathbf{A}(i)$ is very expensive. Indeed, in order to compute $\mathbf{S}_l^M(i)$ in Eq. (13), we need to enumerate $\binom{M}{l}$ combinations which grows exponentially. But, as we will show in the following paragraph we can *reduce the computational cost of* $\mathbf{A}(i)$ *to* $O(M^2)$ by using combinatorial properties.

5.3. The combinatorial "trick"

From Eq. (16) which has an exponential computational cost, we can prove the following statement which enables us to significantly reduce the computational cost of the proposed consensus aggregation operator, $\forall i = 1, \ldots, K$:

$$\mathbf{A}(i) = \frac{\sum_{j=1}^{M} \mu(\mathbb{S}_j(i)) + \sum_{1 \leq j < j' \leq M} \mu(\mathbb{S}_j(i) \cap \mathbb{S}_{j'}(i))}{\frac{M(M+1)}{2}} \tag{17}$$

Proof. In order to prove the property given in Eq. (17), we will use the two following combinatorial identities[6]:

$$\sum_{p=0}^{n} (-1)^p \binom{n}{p} = \mathbf{1}_{\{n=0\}} \tag{18}$$

$$\sum_{p=0}^{n} (-1)^{p+1} \binom{n}{p} p = \mathbf{1}_{\{n=1\}} \tag{19}$$

[6]For identities and properties referring to binomial coefficients and Pascal's triangle, a good reference is the following website `http://binomial.csueastbay.edu/`.

where $\mathbf{1}_{\{\mathbb{P}\}}$ is the indicator function which equals 1 if proposition $\mathbb{P}$ is true; 0 otherwise. We first have:

$$\begin{aligned}
&\frac{M(M+1)}{2}\mathbf{A}(i) \\
&= \sum_{m=1}^{M} m\mu(\mathbb{E}_m^M(i)) \\
&= \sum_{m=1}^{M} m \sum_{l=m}^{M} (-1)^{l-m} \binom{l-1}{m-1} \mathbf{S}_l^M(i) \\
&= \sum_{m=1}^{M} m \sum_{l=1}^{M} (-1)^{l-m} \binom{l-1}{m-1} \mathbf{S}_l^M(i) \mathbf{1}_{\{l \geq m\}} \\
&= \sum_{l=1}^{M} \sum_{m=1}^{M} m(-1)^{l-m} \binom{l-1}{m-1} \mathbf{S}_l^M(i) \mathbf{1}_{\{m \leq l\}} \\
&= \sum_{l=1}^{M} \sum_{m=1}^{l} m(-1)^{l-m} \binom{l-1}{m-1} \mathbf{S}_l^M(i)
\end{aligned}$$

Then we denote $m' = m - 1$; $l' = l - 1$ and $M' = M - 1$. Using these variables we obtain:

$$\begin{aligned}
&\frac{M(M+1)}{2}\mathbf{A}(i) \\
&= \sum_{l'=0}^{M'} \sum_{m'=0}^{l'} (m'+1)(-1)^{l'-m'} \binom{l'}{m'} \mathbf{S}_{l'+1}^{M'+1}(i) \\
&= \sum_{l'=0}^{M'} \mathbf{S}_{l'+1}^{M'+1}(i) \sum_{m'=0}^{l'} (m'+1)(-1)^{l'-m'} \binom{l'}{m'} \\
&= \sum_{l'=0}^{M'} \mathbf{S}_{l'+1}^{M'+1}(i) \left[\sum_{m'=0}^{l'} (-1)^{l'-m'} \binom{l'}{m'} \right. \\
&\quad \left. + \sum_{m'=0}^{l'} (-1)^{l'-m'} \binom{l'}{m'} m' \right]
\end{aligned}$$

Next, if we enumerate a bit this latter expression we can see that using the combinatorial identities given in Eqs. (18) and (19): we obtain $\mathbf{S}_1^M(i)$ for $l' = 0$; $\mathbf{S}_2^M(i)$ for $l' = 1$ and we obtain a null quantity, for $l' > 1$. This finally shows that $\frac{M(M+1)}{2}\mathbf{A}(i) = \mathbf{S}_1^M(i) + \mathbf{S}_2^M(i)$. □

We give in Fig. 2, an illustration of the property for $M = 3$.

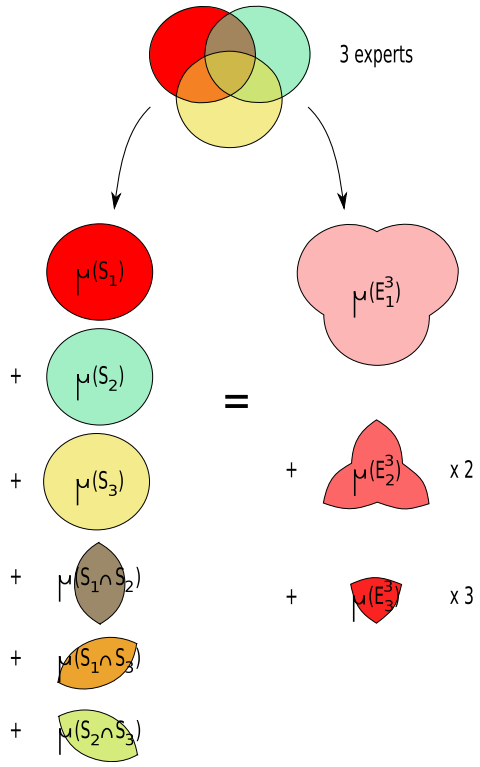

Fig. 2. Illustration of the combinatorial "trick" with 3 experts

Given Eq. (17), we can see that our proposal goes beyond a simple sum[7] since the introduced aggregation function allows us to take into account *pairwise conjunctions between experts*. Particularly, we can specify the type of relationship between pairs of experts by using conjunction operations. The following subsection focuses on this aspect.

5.4. *Conjunction measures and properties of the new aggregation operator*

Given an object i, the only data we are provided with are the normalized scores given by each search engine (or experts) $(S_1(i), \ldots, S_M(i))$. As $S_j(i)$ is a non-negative measure, we propose to set $\mu(\mathbb{S}_j(i)) = S_j(i); \forall j = 1, \ldots, M, \forall i = 1, \ldots, N$

From these inputs, we need to compute pairwise conjunction measures to apply our aggregation function $\mathbf{A}$ according to Eq. (17). Since we interpret $S_j(i)$ as a truth value, we propose to use t-norms to evaluate the conjunction measures between pairs of experts:

$$\begin{aligned} &\mu(\mathbb{S}_j(i) \cap \mathbb{S}_{j'}(i)) \\ &= T(\mu(\mathbb{S}_j(i)), \mu(\mathbb{S}_{j'}(i))) \\ &= T(S_j(i), S_{j'}(i)) \end{aligned} \tag{20}$$

where T is a t-norm such as described in subsection 4.2.

In the unsupervised case we use the same t-norm T for each pair of experts. Therefore we will use the notation $\mathbf{A}_T$ and we have:

$$\begin{aligned} &\mathbf{A}_T(i) \\ &= \frac{\sum_{j=1}^{M} S_j(i) + \sum_{1 \leq j < j' \leq M} T(S_j(i), S_{j'}(i))}{\frac{M(M+1)}{2}} \end{aligned} \tag{21}$$

Providing that we use t-norms as conjunction functions, we can precise the basic properties that our proposal satisfies. It is easy to see that $\mathbf{A}_T$ respect the following conditions (see subsection 4.4):

- Identity when unary
- Boundary conditions
- Non-decreasing
- Symmetry (or commutativity)

Besides, we have the following relation:

$$0 \leq \min_{i, 1 \leq j < j' \leq M} T(S_j(i), S_{j'}(i)) \leq \mathbf{A}_T \leq \max \leq 1$$

From a more general perspective, one could set for each pair of experts a specific kind of t-norm and even a particular conjunction operator. This is possible if we have any a priori information on the relationships between experts. However, since we focus on unsupervised rank aggregation tasks, we do not assume that we have such expertise and we rather use the approach given by Eq. (21).

5.5. *The relationship between* $\mathbf{A}$ *and TOWA*

In [31,32], the link between TOWA functions and multicriteria decision functions based upon previous works detailed in [3,29] is pointed out. In that framework, decision functions are based on events such as "alternative i fulfills at least m criteria among M". From a set viewpoint, the latter are basically the same as the consensus events we have introduced in Eq. (11).

More formally, if we replace the notations given in [32] with the ones used in this paper, a decision function $\mathbf{D}$ (an aggregation operator actually) that scores

[7] As the first part is similar to CombSUM or the arithmetic mean.

an alternative i with respect to a set of M criteria, is defined as follows:

$$\mathbf{D}(i) = \sum_{m=1}^{M} w_m \nu(\mathbb{E}_m^M(i)) \quad (22)$$

where $w_m \in [0,1]; \forall m = 1,\ldots,M; \sum_{m=1}^{M} w_m = 1$ and ν is a fuzzy measure.

This is very close to our proposal given in Eq. (16). In fact, there are two main differences between Yager's approach and ours:

1. In [31,32], the measure of the event $\mathbb{E}_m^M(i)$ relies on the possibility paradigm for which, the fuzzy measure of the disjunction between two events $\mathbb{A}$ and $\mathbb{B}$ is given by $\nu(\mathbb{A} \cup \mathbb{B}) = \max(\nu(\mathbb{A}), \nu(\mathbb{B}))$ whether $\mathbb{A}$ or $\mathbb{B}$ are disjoints or not. On the contrary, we have made in section 5.1, the assumption that the measure μ respects the classical measure paradigm that is to say the σ-additivity property. In that case, we rather have the following relation: $\mu(\mathbb{A} \cup \mathbb{B}) = \mu(\mathbb{A}) + \mu(\mathbb{B})$ if $\mathbb{A}$ and $\mathbb{B}$ are disjoints and; $\mu(\mathbb{A} \cup \mathbb{B}) = \mu(\mathbb{A}) + \mu(\mathbb{B}) - \mu(\mathbb{A} \cap \mathbb{B})$ otherwise.
2. There are two kinds of parameter for the TOWA: on the one hand, the t-norm(s)[8] for measuring the conjunction between nested subsets of experts ranked, for each object, according to the non-increasing order of their scores; on the other hand, the weight vectors $(w_1,\ldots,w_M)$ such as presented in subsection 4.3. In our case, we definitely set the weight vector to $(\frac{1}{M(M+1)/2}, \frac{2}{M(M+1)/2}, \ldots, \frac{M}{M(M+1)/2})$ since this is a reasonable choice to define a consensus aggregation operator. Thus, the weight vector is not a parameter to set anymore. The only type of parameter(s) to set is the t-norm(s)[9] for measuring the conjunction measures of pairs of experts.

Regarding the first point, we can notice that the possibility axiom has the advantage to avoid the complexity issue. Indeed, t-norms are monotonic functions such that, for all m-uple $(a_{j1},\ldots,a_{jm})$:

$$T(a_{\sigma(1)}, a_{\sigma(2)}, \ldots, a_{\sigma(m)}) \leq T(a_{j1}, a_{j2}, \ldots, a_{jm})$$

As a consequence we don not have to enumerate all combinations to compute $\nu(\mathbb{E}_m^M)$ since we have:

$$\begin{aligned} &\nu(\mathbb{E}_m^M) \\ &= \max_{1 \leq j1 < \ldots < jm \leq M} T(S_{j1}(i), \ldots, S_{jm}(i)) \\ &= T(S_{\sigma(1)}(i), S_{\sigma(2)}(i), \ldots, S_{\sigma(m)}(i)) \end{aligned} \quad (23)$$

One can see that if we replace the previous relation in Eq. (22) we obtain the definition of TOWA operator given in Eq. (9).

In comparison to Eq. (23) which relies on the possibility paradigm, the σ-additivity axiom leads to combinatorial issues as we can see by considering Jordan's combinatorial formula given in Eq. (14) and our definition of aggregation function given in Eq. (16). However, by assuming a reasonable condition concerning the weight vectors, as aforementioned in the second point, we have proved in section 5.3, that we can avoid the complexity problem of such an approach. More than a simple combinatorial "trick", the resulting form of our aggregation operator $\mathbf{A}$ given in Eq. (17) has an interesting semantic. It explicitly enables one to take into account the relationships between experts by considering conjunction measures between each pair of experts. Thus, despite similar basic concepts based on consensus events, our approach finally leads to a really different aggregation techniques in comparison to TOWA. In our method, we model the relationships between experts by considering pairs whereas for TOWA, the relationships between experts is rather based upon nested subsets that depend on the non-increasing order of their scores.

6. Experiments and discussion

In this section, using numerical examples, we show that the proposed aggregation operators can perform better than baseline approaches. We first describe the settings of the experiments we made. Then we report and discuss the results we obtained.

6.1. Experimental results

We tested the different methods introduced previously, on the three different datasets of the Letor 2.0 package [21]:

– OHSUMED dataset (subset of MEDLINE, a database of medical publications): 106 queries, 25 different ranking features (the experts)

[8] Possibly M different t-norms.
[9] Possibly $M(M-1)/2$ different t-norms.

– TREC 2003 web track (topic distillation task): 50 queries, 44 different ranking features (the experts)
– TREC 2004 web track (topic distillation task): 75 queries, 44 different ranking features (the experts)

For each dataset, we have different queries (topics) and for each ranking feature, we are given a distribution of scores among a set of objects (medical publications or html files). The ranking features are either low-level or high-level content features, between each judged query-document pairs. The reader can consult [21], for a more detailed description of the characteristics of these datasets and for the definition of the features that were extracted.

The Letor 2.0 package was initially constituted to benchmark "learning to rank" methods. In that context the goal is to learn, from a training set, how to combine the different experts. In our case, we are in an unsupervised context and we only use the relevance file for measuring a posteriori the performances of the proposed methods.

First, we normalized the data using the score normalization[10] given by Eq. (4) with $k = 1000$. Then, we computed the following aggregation functions:

– $\mathrm{CombSUM}$ (equivalent to $\mathbf{M}_p$ with $p = 1$ and with uniform weights), $\mathrm{CombMNZ}$ (with similar settings than $\mathrm{CombSUM}$).
– T-norms and t-conorms $T_M = \min$ and $T_M^* = \max$.
– the power-mean $\mathbf{M}_p$ with $p = 2$ and $p = 3$ and since we don not suppose any a priori information on the importance of each search engine, we took $w_j = 1/M; \forall j = 1, \ldots, M$.
– The $\mathbf{TOWA}_T$ approach with $T = T_M, T_P$ and T_{SS}^{λ} with $\lambda = 6$ and with $w_j = \left(\frac{j}{M}\right)^5 - \left(\frac{j-1}{M}\right)^5$. Regarding Eq. (10) this corresponds to take Q as a power function with $q = 5$. We chose this setting following the work presented in [10] that investigates the application of the TOWA technique (and other OWA based methods) to metasearch problems on similar datasets. In [10], $q = 5$ performed better than other settings. Besides, we recall that the case $T = T_M$ is the same as the classical OWA function given in Eq. (8).
– The new aggregation method $\mathbf{A}_T$ using either non-parametric t-norms T_M, T_P or the parametric t-norm $T = T_{SS}^{\lambda}$ with $\lambda = 6$.

[10] We found that it may happen that for some features, a unique value is assigned to all objects. In that case, we replaced this value by 0.

For measuring and comparing the performances of the different aggregation methods, we used the `trec_eval` tool[11] and we retained the mean average precision (MAP) and the precision at 10 (P@10) measures. The results are given in Tables 1, 2 and 3. In italic are the baseline results given by CombSUM. We put in bold the two best results among the aggregation operators we tested.

Table 1
Results for the OHSUMED dataset

Method	MAP	P@10
CombSUM	*39.52 %*	*44.81 %*
CombMNZ	38.68 %	44.34 %
Quadratic mean	40.37 %	47.36 %
Cubic mean	**40.65** %	47.45 %
T_M	0.82 %	0.38 %
T_M^*	39.71 %	49.25 %
$\mathbf{TOWA}_T$ with T_M	37.21 %	41.79 %
$\mathbf{TOWA}_T$ with T_P	**40.89** %	47.55 %
$\mathbf{TOWA}_T$ with T_{SS}^6	40.50 %	**48.68** %
$\mathbf{A}_T$ with T_M	38.49 %	43.96 %
$\mathbf{A}_T$ with T_P	39.44 %	44.91 %
$\mathbf{A}_T$ with T_{SS}^6	40.38 %	**47.93** %

Table 2
Results for the TREC 2003 dataset

Method	MAP	P@10
CombSUM	*16.83 %*	*13.20 %*
CombMNZ	15.42 %	12.40 %
Quadratic mean	16.57 %	13.60 %
Cubic mean	17.33 %	14.60 %
T_M	0.51 %	1.0 %
T_M^*	07.98 %	07.20 %
$\mathbf{TOWA}_T$ with T_M	13.03 %	11.40 %
$\mathbf{TOWA}_T$ with T_P	17.38 %	14.60 %
$\mathbf{TOWA}_T$ with T_{SS}^6	**18.18** %	**14.60** %
$\mathbf{A}_T$ with T_M	15.86 %	11.80 %
$\mathbf{A}_T$ with T_P	16.81 %	13.20 %
$\mathbf{A}_T$ with T_{SS}^6	**18.52** %	**15.40** %

6.2. Discussion

The first conclusion we can draw out from Tables 1, 2 and 3, is that the t-norm T_M and the t-

[11] cf. http://trec.nist.gov/

Table 3
Results for the TREC 2004 dataset

Method	MAP	P@10
CombSUM	*26.81 %*	*16.53 %*
CombMNZ	23.09 %	14.53 %
Quadratic mean	29.41 %	16.93 %
Cubic mean	29.48 %	**17.33** %
T_M	0.03 %	0.0 %
T_M^*	07.93 %	04.80 %
$\mathbf{TOWA}_T$ with T_M	18.59 %	12.93 %
$\mathbf{TOWA}_T$ with T_P	**30.76** %	**17.73**%
$\mathbf{TOWA}_T$ with T_{SS}^6	28.33 %	16.40 %
$\mathbf{A}_T$ with T_M	22.63 %	15.33 %
$\mathbf{A}_T$ with T_P	26.78 %	16.40 %
$\mathbf{A}_T$ with T_{SS}^6	**29.90** %	16.53 %

conorm T_M^* perform very badly compared to the baseline[12]. This is not surprising since, as we mentioned earlier, they represent extreme point of views and their behavior are too conjunctive or too disjunctive. They are not well-adapted for metasearch tasks. Furthermore, one could have tried to have a trade-off between the t-norm and the t-conorm such as the so-called convex-linear compensatory operator defined in [33] as $(1-\beta)T(S_1(i),\ldots,S_M(i)) + \beta T^*(S_1(i),\ldots,S_M(i))$. But we can see that this would not work out neither.

As a consequence, and this is confirmed from our results, we need more compensatory aggregation operators when dealing with metasearch problems. Indeed, the other aggregation techniques namely the power mean, the TOWA and our proposal can all perform better than the baseline given by $\mathrm{CombSUM}$.

Nevertheless, the results between all three aforementioned techniques are very close to each other and from our experiments we cannot conclude that one approach consistently outperforms the other ones. This is even true as each of them has different parameters to tune. Finally the three techniques offer three different models.

First the weighted quasi-arithmetic means allow one to assign different weights to experts that must reflect the importance of the latter. However, they do not permit one to model the relationships between them. Furthermore, it is not an easy task to set the function f or even the parameter p in the case of power means, since there is no particular semantic associated to such functions.

On the contrary, $\mathbf{TOWA}_T$ and $\mathbf{A}_T$ allow us to take into account different kinds of relationships between experts on the basis of consensus events and this is done by using t-norms among subsets of experts. For TOWA operators, these subsets of experts are different for each object and they are given by the non-increasing order of the scores. In the case of $\mathbf{A}_T$, these subsets are constituted of all pairs of experts and are independent of any object. Finally, despite quite similar results, both methods are different and further experiments need to be pursued to better understand the pros and cons of those two techniques.

7. Conclusion

In this paper, we have addressed data fusion problems. We have been particularly concerned with unsupervised rank aggregation tasks or metasearch tasks in information retrieval field. The general aim of this work is to investigate the use of some aggregation operators defined in information fusion and multicriteria decision making fields in the context of metasearch problems. We have recalled and tested classical aggregation operators such as weighted quasi-arithmetic means, t-norms, t-conorms and TOWA functions. We have moreover introduced a new aggregation method $\mathbf{A}$, its logical foundations and its combinatorial properties.

Our experiments using the Letor 2.0 package show that t-norms and t-conorms are not good aggregation operators for that kind of tasks since they are not compensatory. On the contrary, more compensative operators such as weighted quasi-arithmetic means, TOWA and $\mathbf{A}$ functions can improve the baseline results. The two latter approaches allow to consider the relationships between search engines through two different manners. Because of this general property we think that they are interesting techniques to further study.

As far as our future work is concerned, we intend to apply the studied aggregation operators to "learning to rank" tasks which correspond to supervised rank aggregation problems in information retrieval. Indeed, the different approaches that we have proposed to tackle the metasearch problems in information retrieval need all some paramaters to be tuned. In this paper, we have been in an unsupervised setting and we have set the parameters arbitrarily but to increase the performances of the different non-linear aggregation

[12] Other t-norms or t-conorms perform similarly as T_M and T_M^*

models, one could estimate their respective parameters by using a training set and machine learning techniques. In that perspective and regarding the different models we have proposed in this paper, one could better capture the efficiency of the search engines or evaluate the relationships between experts and use those pieces of information to better fit the models to the particular data and tasks a user would need to deal with. Thereby, in our future work, we intend to work on the concepts of search engines reliability on the one hand, and learning to model the relationships between search engines on the other hand. Furthermore, more experiments in that context would help us to better compare the different aggregation operators from a practical viewpoint.

References

[1] J. Ah-Pine. *Sur des aspects algébriques et combinatoires de l'Analyse Relationnelle*. PhD thesis, University of Pierre and Marie Curie (Paris 6), 2007.

[2] J. Ah-Pine. Data fusion in information retrieval using consensus aggregation operators. In *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence*, 2008.

[3] R. E. Bellman and L. A. Zadeh. Decision-making in a fuzzy environment. *Management Science*, 17:141–164, 1970.

[4] I. Bloch. Information Combination Operators for Data Fusion: A Comparative Review with Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 26(1):52–67, 1996.

[5] J. Borda. *Mémoire sur les élections au scrutin*. Histoire de l'Académie Royale des Sciences, 1781.

[6] B. Bouchon-Meunier and J. Kacprzyk, editors. *Aggregation and Fusion of Imperfect Information*. Physica-Verlag, 1998.

[7] T. Calvo, A. Kolesárová, M. Komorníková, and R. Mesiar. Aggregation operators: properties, classes and construction methods. In *Aggregation operators: new trends and applications*, pages 3–104. Physica-Verlag GmbH, Heidelberg, Germany, Germany, 2002.

[8] L. Comtet. *Analyse Combinatoire, tome 1 et tome 2*. Presses Universitaires de France, 1970.

[9] M. M. d. Condorcet. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Paris, 1785.

[10] A. De, E. D. Diaz, and V. V. Raghavan. On fuzzy result merging for metasearch. In *FUZZ-IEEE*, pages 1–6, 2007.

[11] D. Dubois and H. Prade. On the use of aggregation operations in information fusion processes. *Fuzzy Sets and Systems*, 142(1):143–161, 2004.

[12] C. Dwork, R. Kumar, M. Naor, and D. Sivakumar. Rank aggregation methods for the web. In *WWW*, pages 613–622, 2001.

[13] R. Fagin, R. Kumar, and D. Sivakumar. Comparing top k lists. *SIAM Journal on discrete mathematics*, 17:134–160, 2003.

[14] M. Farah and D. Vanderpooten. An outranking approach for rank aggregation in information retrieval. In *Proceedings of the 30th annual international ACM SIGIR Conference*, 2007.

[15] E. Fox and J. Shaw. Combination of multiple searches. In *Proceedings of the 3rd NIST TREC Conference*, pages 105–109, 1994.

[16] E. Herrera-Viedma, J. L. Gijón, S. Alonso, J. Vílchez, C. García, L. Villén, and A. G. López-Herrera. Applying aggregation operators for information access systems: An application in digital libraries. *Int. J. Intell. Syst.*, 23(12):1235–1250, 2008.

[17] C. Jordan. Problèmes de la probabilité des épreuves répétées dans le cas général. *Bulletin de la S.M.F.*, 67:223–242, 1939.

[18] E. Klement, R. Mesiar, and E. Pap. *Triangular norms*. Kluwer Academic Pub, 2000.

[19] A. Klementiev, D. Roth, and K. Small. An unsupervised learning algorithm for rank aggregation. In *Proceedings of the European Conference on Machine Learning (ECML)*, September 2007.

[20] J. H. Lee. Analyses of multiple evidence combination. *SIGIR Forum*, 31(SI):267–276, 1997.

[21] T.-Y. Liu, T. Qin, J. Xu, W. Xiong, and H. Li. Letor: Benchmark dataset for research on learning to rank for information retrieval. In *LR4IR 2007 in conjunction with SIGIR 2007*, 2007.

[22] J. Marcotorchino and P. Michaud. *Optimisation en analyse ordinale des données*. Masson, 1978.

[23] P. Michaud and J. Marcotorchino. Modèles d'optimisation en analyse des données relationnelles. *Mathématiques et Sciences Humaines*, 17(67):7–38, 1979.

[24] M. Montague and J. A. Aslam. Condorcet fusion for improved retrieval. In *CIKM '02: Proceedings of the eleventh international conference on Information and knowledge management*, pages 538–548, New York, NY, USA, 2002. ACM.

[25] M. Renda and U. Straccia. Web metasearch : rank vs score based rank aggregation methods. In *Proceedings of the 2003 ACM symposium on applied computing*, pages 841–846, 2003.

[26] W. Riker. *Liberalism against populism*. Waveland Press, 1982.

[27] B. Roy. The outranking approach and the foundations of ELECTRE methods. *Theory and decisions*, 31:49–73, 1991.

[28] B. Schweizer and A. Sklar. *Probabilistic metric spaces*. North-Holland, 1983.

[29] R. R. Yager. General multiple-objective decision functions and linguistically quantified statements. *Int. J. Man-Mach. Stud.*, 21(5):389–400, 1984.

[30] R. R. Yager. On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Trans. Syst. Man Cybern.*, 18(1):183–190, 1988.

[31] R. R. Yager. Extending multicriteria decision making by mixing t-norms and owa operators. *Int. J. Intell. Syst.*, 20(4):453–474, 2005.

[32] R. R. Yager and L. Troiano. On some properties of mixing owa operators with t-norms and t-conorms. In *EUSFLAT Conf.*, pages 1206–1212, 2005.

[33] H. Zimmermann and P. Zysno. Latent connectives in human decision making. *Fuzzy Sets and Systems*, 4:37–51, 1980.