2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
July 6-11, 2014, Beijing, China

Nearest Neighbour-Guided Induced OWA and its Application to Journal Ranking

Pan Su, Tianhua Chen, Changjing Shang and Qiang Shen
Department of Computer Science
Institute of Mathematics, Physics and Computer Science
Aberystwyth University, UK
E-mail: {pas23, tic4, cns, qqs}@aber.ac.uk

***Abstract*—Aggregation operators are useful tools which summarise multiple inputs to a single output. In practice, inputs to such operators are variables which represent different criteria, measurements, or opinions from experts. In this paper, a nearest neighbour-guided induced OWA operator, abbreviated as kNN-IOWA, is proposed as a special case of the generic induced OWA where the input arguments are ordered by the average distances to their k nearest neighbours. The weighting vectors in kNN-IOWA are defined, which are used to interpret the overall behaviour of the operator's reliability. kNN-IOWA is applied for building aggregated fuzzy relations between academic journals, based on their indicator scores. It combines the similarities between academic journals to assess their performance with respect to different journal impact indicators. The work is compared against different types of aggregation operator and tested on six bibliometric datasets. The results of experimental evaluation demonstrate that kNN-IOWA outperforms other aggregation operators in terms of standard accuracy and within-1 accuracy. The proposed method also exhibits the advantages of being more intuitive and interpretable.**

I. INTRODUCTION

Aggregation of several input values into a single output value is an indispensable tool in a wide range of applications such as human resource management [1], group decision making [2] and journal ranking [3]. Different types of aggregation operator have been proposed in the literature. A popular aggregation method is the Ordered Weighted Averaging (OWA) operator originally introduced in [4]. It provides a parameterised family of aggregation operators, including as special cases the maximum, the minimum and the average calculus [5].

Academic journal ranking is a specific application problem addressed here in which OWA may also play a significant role. The most recent methods for the ranking of academic journals require automated assessment of journal quality. Many on-line academic publication databases do offer various information about journals' impact indicators such as the *Thomson Reuters Impact Factor* (IF) [6], the Eigenfactor [7] and the 5-year IF [8]. However, each indicator has its own strengths and limitations, and their results can be quite diverse [9]. An intuitive way to improve the reliability of such indicators is the integration of multiple metrics. For instance, the Choquet integral classifier [10] has been employed to combine different indicator scores that are reported in the *Thomson Reuters Journal Citation Report* (JCR) [11] in order to predict the journal ranks (as published in the ERA 2010 [12]). Instead of direct aggregation of the individual scores, another direction is to fuse the distances measured over journals that are placed in a multi-dimensional space with each dimension representing a certain impact indicator [13]. OWA can also be employed to aggregate fuzzy similarities between journals in terms of their impact indicators, thereby generating clusters of journals that reflect their indicator scores [3]. Whilst promising, there is much to be done in making these techniques more robust and generic in order to support activities such as research excellence assessment [14].

To enhance the reliability in performing aggregation of publication impact indicators for the task of academic journal ranking, this paper proposes a nearest neighbour-guided induced OWA operator, denoted as kNN-IOWA hereafter, for developing aggregated fuzzy relations between journals, based on their impact indicator scores. The proposed operator is a special case of the Induced OWA (IOWA) [15], [16], [17], with two characters that distinguish it from other IOWA operators: 1) the elements of the order inducing vector that is associated with the arguments represent their relative reliabilities, and 2) the value of the reliability measure depends on the distribution of the arguments. That is, for each individual argument, its average distance to the other k nearest arguments [18] is calculated and transformed into its corresponding element in the order inducing vector.

The resulting kNN-IOWA is employed to support the classic c-means clustering algorithm where clusters of journals are constructed according to their kNN-IOWA aggregated relations over the indicator scores. Such derived clusters are more interpretable and intuitive than the original indicator scores. The proposed method is tested on six datasets of journals from different research areas. The results demonstrate that kNN-IOWA helps perform journal quality clustering, revealing the relative impact of academic journals effectively.

The remainder of this paper is organised as follows. Section II introduces the basics of the OWA and IOWA aggregation operators. Sections III and IV define kNN-IOWA and describe its application to c-means clustering for journal ranking, respectively. Section V presents the experimental evaluation of the proposed approach and discusses the results. Finally, Section VI concludes the paper and points out directions for further development.

II. Preliminaries

A. OWA Aggregation

Definition 1: [4] A mapping $A_{\text{owa}} : \Re^n \rightarrow \Re$ is called an OWA operator if

$$A_{\text{owa}}(a_1, \cdots, a_n) = \sum_{i=1}^{n} w_i a_{\pi(i)}$$

where $a_{\pi(i)}$ is a permutation of $a_i \in \Re, i = 1, \cdots, n$, which satisfies that $a_{\pi(i)}$ is the i-th largest amongst all a_i, and $w_i \in [0, 1], i = 1, \cdots, n$ is a collection of weights that satisfies $\sum_{i=1}^{n} w_i = 1$.

For simplicity, the arguments and weights of an OWA operator are hereafter denoted as the argument vector $A = (a_1, \cdots, a_n)$ and the weighting vector $W = (w_1, \cdots, w_n)$, respectively.

Different choices of the weighting vector W can lead to different aggregation results. The ordering of individual inputs in the argument vector A presents OWA with an inherent nature of nonlinearity. Three special cases of the OWA operator are the classical *mean*, *max* and *min* operators. The *mean* operator results by setting $W_{mean} : w_i = 1/n$, the *max* by $W_{max} : w_1 = 1$ and $w_i = 0$ for $i \neq 1$, and the *min* by $W_{min} : w_n = 1$ and $w_i = 0$ for $i \neq n$. Obviously, an important feature of OWA is that it is an operator which satisfies

$$\min\{a_1, \cdots, a_n\} \leq A_{\text{owa}}(A) \leq \max\{a_1, \cdots, a_n\}.$$

Such an operator provides aggregation between the maximum and minimum of the values that the arguments may take. It is also idempotent; that is, if all $a_i = a$ then $A_{\text{owa}}(a_1, \cdots, a_n) = a$.

A property that is commonly used to interpret the overall behaviour of an OWA aggregation operator is *orness/andness* [4]. It gives an indication of whether an OWA aggregation behaves similarly to conjunction (influenced by smaller inputs) or disjunction (influenced by larger inputs). In particular, the orness measure of an OWA operator with the weighting vector W is defined by

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

It can be calculated that $\text{orness}(W_{mean}) = 0.5$, $\text{orness}(W_{max}) = 1$ and $\text{orness}(W_{min}) = 0$. A useful method for generating the OWA weights is by the use of a so-called stress function [19], enabling formal characterisation of the resulting OWA aggregation operator. This can be accomplished using a function $h : [0, 1] \rightarrow \Re^+$ to stress the places where to obtain significant values for the weighting vector. Formally, a weighting vector of OWA is defined by a stress function h as follows.

Definition 2: [19] Let $h : [0, 1] \rightarrow \Re^+$ be a non-negative function on the unit interval. The OWA weights $W = (w_1, \cdots, w_i, \cdots, w_n)$ can be defined as:

$$w_i = \frac{h(\frac{i}{n})}{\sum_{i=1}^{n} h(\frac{i}{n})}. \quad (2)$$

This method of obtaining the OWA weighting vector has a number of useful features. For instance, the $h(x)$ values associated with the lower portion of the left side of $[0, 1]$ reflect those weights associated with the larger argument values, while the values associated with the right side of the unit interval reflect the weights associated with the smaller values in the aggregation. Other properties are omitted here but can be found in [19], [20], [21].

B. Induced OWA Aggregation

A key step of OWA aggregation is the ordering of the arguments which transforms the original argument vector $(a_1, \cdots, a_i, \cdots, a_n)$ into an ordered argument vector $(a_{\pi(1)}, \cdots, a_{\pi(i)}, \cdots, a_{\pi(n)})$. The ordering used in OWA depends upon the actual value of the arguments as $a_{\pi(i)}$ is the i-th largest of the arguments. A more general strategy towards the ordering of the arguments has been proposed in [15]. This has led to the development of a generalised approach to OWA aggregation, termed the Induced OWA (IOWA). In IOWA, each of the input values is represented as a two-tuple $\langle u_i, a_i \rangle$ that is referred to as an OWA pair. The input arguments $(a_1, \cdots, a_i, \cdots, a_n)$ are ordered on the basis of the values u_i. In particular, the procedure for calculating the IOWA aggregation over these OWA pairs is defined by

$$A_{\text{iowa}}(\langle u_1, a_1 \rangle, \cdots, \langle u_n, a_n \rangle) = \sum_{i=1}^{n} w_i a_{\pi'(i)}$$

where $a_{\pi'(i)}$ is from the permutation of $\langle u_i, a_i \rangle$ which satisfies that $\langle u_{\pi'(i)}, a_{\pi'(i)} \rangle$ has the i-th largest amongst all u_i, and $w_i \in [0, 1], i = 1, \cdots, n$ is a collection of weights which satisfies that $\sum_{i=1}^{n} w_i = 1$. $U = (u_1, \cdots, u_n)$ is called the order inducing vector. The bounding property exhibited by IOWA aggregation is similar to that by OWA: $\min\{a_1, \cdots, a_n\} \leq A_{\text{iowa}}(\langle u_1, a_1 \rangle, \cdots, \langle u_n, a_n \rangle) \leq \max\{a_1, \cdots, a_n\}$. Idempotency also holds in IOWA: If all $a_i = a$ then $A_{\text{iowa}}(\langle u_1, a_1 \rangle, \cdots, \langle u_n, a_n \rangle) = a$, no matter which order inducing vector U and weighting vector W are used. Note that if two or more OWA pairs have identical values of u_i, their argument values are averaged firstly before aggregation.

The introduction of inducing vector helps improve the flexility of the ordering process in OWA aggregation. OWA operators can be rephrased as special cases of IOWA operators where $u_i = a_i$ for all $i = 1, \cdots, n$. In IOWA, different order inducing vectors can lead to different results of aggregation. Hence, the interpretation of orness of the weighing vectors is also dependent on the choice of a certain order inducing vector.

In the following, a novel approach to developing IOWA is proposed by exploiting the neighbourhood information of a given argument. In particular, the k-Nearest-Neighbour Induced OWA (kNN-IOWA) is introduced, where the k nearest neighbours of an argument a_i are employed to generate its order inducing value u_i, and the orness of W is interpreted as its reliability [22].

III. k-Nearest-Neighbour Induced OWA

The weighting vectors in OWA and IOWA are normally argument-independent as they are not necessarily related to the inputs they are applied to. However, with the argument-dependent approach, weights are indeed determined on the basis of the input arguments. In particular, the Depend OWA

(DOWA) operators [23] utilise weighting vectors that are derived in accordance with the values of arguments. In kNN-DOWA [22] for example, the *reliability* of an argument is defined as the appropriateness of using that argument as the aggregated outcome, aiming to decrease the effect of potential outliers in input arguments.

The concept of reliability has a strong initiative appeal. This is because it helps differentiate amongst a collection of arguments such that an argument whose value is similar to its k neighbours [18] is deemed reliable and can be assigned with a higher weight. In contrast, an argument that is largely different from its neighbours is discriminated as an unreliable member. Formally, the reliability measure of an argument a_i, $i = 1, \cdots, n$ in kNN-DOWA is defined as:

$$R_i^k = 1 - \frac{\sum_{t=1}^{k} |a_i - n_t^{a_i}|}{\max_{i \in \{1,\cdots,n\}} \{\sum_{t=1}^{k} |a_i - n_t^{a_i}|\}} \tag{3}$$

where $n_t^{a_i}$, $t = 1, \cdots, k$ is the t-th nearest neighbour of the argument a_i, and the distance measure d used to perform neighbour-searching is $d(a_i, a_j) = |a_i - a_j|$, where $i, j = 1, \cdots, n$. Note that other distance metrics may be used for this. However, for computational simplicity, the absolute distance metric is used here.

Having obtained the reliability values of all arguments concerned, they are normalised to form the weighing vectors in kNN-DOWA. Given the reliability value of each argument R_i^k, the corresponding kNN-DOWA operator $A_{\text{dowa}}^k : \Re^n \rightarrow \Re$ can be specified by

$$A_{\text{dowa}}^k(a_1, \cdots, a_n) = \sum_{i=1}^{n} w_i^k a_i \tag{4}$$

where $w_i^k = R_i^k / \sum_{j=1}^{n} R_j^k$. kNN-DOWA is order independent (termed *neat* in the literature) [24], as it generates the same outcome regardless of the order of argument values.

kNN-DOWA has been applied to the task of alien detection, where different similarity measures of textual entities are combined. One crucial assumption in kNN-DOWA is that arguments which have high reliability values should be highly weighted. However, empirical results have shown that in certain situations, dependent weights do not always perform as expected. Besides, retaining more diversity of base members in the aggregated output is sometimes preferable [25], [26]. Inspired by these observations, and in order to generalise the dependent determination of the weighting vectors in kNN-DOWA, the k-Nearest-Neighbour-Induced OWA is herein proposed.

kNN-IOWA is designed to be a special case of IOWA, where each input two-tuple is $\langle R_i^k, a_i \rangle$ with R_i^k representing the reliability measure of a_i as with kNN-DOWA, where k is a predefined number of nearest neighbours to be considered. Particularly, the input arguments $(a_1, \cdots, a_i, \cdots, a_n)$ are ordered with respect to their induced values $(R_1^k, \cdots, R_i^k, \cdots, R_n^k)$. Formally, kNN-IOWA is a mapping $A_{\text{iowa}}^k : \Re^n \rightarrow \Re$ and the kNN-IOWA aggregation over the given arguments is calculated as follows:

$$\begin{aligned} A_{\text{iowa}}^k(a_1, \cdots, a_n) &= A_{\text{iowa}}(\langle R_1^k, a_1 \rangle, \cdots, \langle R_n^k, a_n \rangle) \\ &= \sum_{i=1}^{n} w_i a_{\pi^k(i)} \end{aligned} \tag{5}$$

where $a_{\pi^k(i)}$ is from the permutation of OWA pairs $\langle R_{\pi^k(i)}, a_{\pi^k(i)} \rangle$ which satisfies that $R_{\pi^k(i)}$ has the i-th largest amongst all R_i^k, and $w_i \in [0, 1], i = 1, \cdots, n$ is a collection of weights that satisfies $\sum_{i=1}^{n} w_i = 1$.

As a special case of IOWA, the bounding property of kNN-IOWA is similar to that of the IOWA operators: $\min\{a_1, \cdots, a_n\} \leq A_{\text{iowa}}^k(a_1, \cdots, a_n) \leq \max\{a_1, \cdots, a_n\}$. The idempotency also holds: If all $a_i = a$ then $A_{\text{iowa}}^k(a_1, \cdots, a_n) = a$. Note that if two or more arguments have an identical value of the reliability measure, their argument values are averaged before being aggregated.

Interestingly, the weights in kNN-IOWA are independent of the argument values. Any weights that satisfy $\sum_{i=1}^{n} w_i = 1$ can be employed in the process of aggregating the sorted arguments. This flexibility in weight determination offers a degree of freedom to control the behaviour of the resulting kNN-IOWA aggregation operator. The stress function which is designed for obtaining weights in OWA can be employed in kNN-IOWA in a similar way as with the existing work, in implementing the control of the *reliablility* of kNN-IOWA. This work has an intuitive appeal in that high weights are associated with large reliability values. The reverse holds also; if high weights are associated with small reliability values the aggregated outcome will then be not reliable or trustworthy. In the situation where users have no a-priori knowledge of weight settings, both the weighting vectors of high reliability and those of low reliability need to be tested first, and the one with better performance is selected. After that, a limited local search may be applied to decide on the appropriate element values of the weighting vector.

IV. kNN-IOWA Aggregation of Journal Similarities

With the aid of on-line academic publication databases such as Web of Knowledge, IEEE Xplore, and DBLP [27], the calculation of individual journal impact indicators can be carried out efficiently. A number of indicators are widely accepted and applied by scholars, which typically aim to evaluate a single journal or usually focus on one particular aspect of journal quality. Most of the journal impact indicators assign a score (a numerical value) to the journal under evaluation. Although many journal impact indicators have been proposed and applied to the evaluation of journal quality, none of these are sufficient to characterise all aspects of journal impact by itself in the real-world. Novel evaluation tools will help.

Note that when human experts assess the quality of academic journals, comparison of the scores is commonly and sensibly used to support their judgement. Based on this observation, the similarities between journals with respect to the indicators, rather than the raw scores, are here considered in ranking journals in terms of their quality. To compensate for the potential bias of using single indicators, thereby enriching the reliability of fuzzy similarity relations amongst journals,

kNN-IOWA is herein employed to integrate fuzzy similarity measures. This also offers a useful testbed to examine the utility of the above-proposed kNN-IOWA aggregation.

Given a set of academic journals $J = \{j_1, \cdots, j_n\}$ and a journal impact indicator $I : J \rightarrow \Re$, fuzzy similarity measures can be employed to perform pairwise comparison of journal indicator scores into a similarity relation $S_I : J \times J \rightarrow [0, 1]$. For example:

$$S_I(j_x, j_y) = 1 - \frac{|I(j_x) - I(j_y)|}{\max\limits_{j_i \in J}\{I(j_i)\} - \min\limits_{j_i \in J}\{I(j_i)\}} \quad (6)$$

$$S_I(j_x, j_y) = \exp\left(- \frac{(I(j_x) - I(j_y))^2}{2\delta^2} \right) \quad (7)$$

$$S_I(j_x, j_y) = \max\left(\min\left(\frac{I(j_y) - (I(j_x) - \delta)}{I(j_x) - (I(j_x) - \delta)}, \frac{(I(j_x) + \delta) - I(j_y)}{(I(j_x) + \delta) - I(j_x)} \right), 0 \right) \quad (8)$$

where $I(j_x)$ and $I(j_y)$ denote the scores of journal $j_x, j_y \in J$ assigned by the indicator I, and δ^2 is the variance of the scores $\{I(j_i)|j_i \in J\}$. Other definitions for implementing the transformation can be found in the literature [28].

More generally, given J and a set of journal impact indicators $I = \{I_1, \cdots, I_n\}$, the fuzzy similarity between two journals $j_x, j_y \in J$ with respect to the indicator $I_i \in I$ is represented by $S_{I_i}(j_x, j_y)$, and the kNN-IOWA aggregation of these similarities between j_x and j_y can be computed by

$$S^k(j_x, j_y) = A^k_{\text{iowa}}(S_{I_1}(j_x, j_y), \cdots, S_{I_n}(j_x, j_y)) \quad (9)$$

where the weighting vector may be defined offline (say, by the user) or learned from historical data, and $S_{I_i}(j_x, j_y)$ are ordered with respect to their reliability values which are subsequently based on their k nearest neighbours. The transformation from individual indicator scores $I_i(j_x), I_i(j_y)$ to the similarity relations $S_{I_i}(j_x, j_y)$ can be achieved in a straightforward manner, using either of Eqns. (6)–(8).

To illustrate the computation process of $S^k(j_x, j_y)$, suppose that three journals are individually evaluated using four separate indicators: IF (I_1), 5-year IF (I_2), Eigenfactor (I_3) and Immediacy Index (I_4), as listed in Table I. Also, without losing generality, suppose that the fuzzy similarity relation with respect to each indicator is evaluated by the use of Eqn. (6). This leads to the following similarities between journal j_1 and j_2, which are each assigned with respect to one of the four individual indicators: $S_{I_1}(j_1, j_2) = 0.46$, $S_{I_2}(j_1, j_2) = 0.50$, $S_{I_3}(j_1, j_2) = 0$ and $S_{I_4}(j_1, j_2) = 0.43$. Suppose that 2NN-IOWA (i.e., $k = 2$) is adopted to perform aggregation. Two nearest neighbours are therefore considered in calculating the reliability of arguments. This results in the ordered argument vector of $(0.46, 0.43, 0.50, 0)$, with the corresponding order inducing vector $(0.92, 0.87, 0, 0.88)$. Given that the weighting vector in A^2_{iowa} is $W = (0.40, 0.30, 0.20, 0.10)$, the aggregation result of the four fuzzy similarities between j_1 and j_2 is $S^2(j_1, j_2) = 0.410$. Comparatively, with the same W, the aggregated similarity between j_1 and j_2 using the original OWA operator is 0.422, which is closer to the argument of largest value (0.50, given by the 5-year IF) rather than the argument which has the largest reliability (0.46, given by the IF). Intuitively, in the tasks such as journal ranking, a reliable aggregated output is preferable to the aggregated output that is simply close to a single extreme argument.

TABLE I. EXAMPLES OF JOURNALS

	I_1	I_2	I_3	I_4
j_1	7.806	10.716	0.00571	0.867
j_2	5.027	7.228	0.05002	0.591
j_3	2.683	3.752	0.00895	0.387

Generally speaking, the pairwise relations obtained by the application of kNN-IOWA can be utilised in a variety of similarity/distance-based learning algorithms. In this paper, it is employed to aid in performing c-means clustering of academic journals with respect to their scores against different impact indicators. The task of clustering is to assign objects to groups (namely clusters) such that objects in the same group are similar to each other, and dissimilar to those in the other clusters [29]. Journal clustering is no exceptional, seeking to partition a collection of academic journals using the fuzzy relations between journals that are aggregated by kNN-IOWA. Note that a number of generic clustering methods may be employed to implement this. Amongst them, the c-means algorithm is popular, due to its simplicity and success in solving real-world problems [30], [31], [32]. Thus, c-means is herein integrated with kNN-IOWA to perform academic journal clustering.

It is worth indicating that the computational results from applying the proposed aggregation operator are easier to interpret (than the concept of orness), owing to the use of the reliability measure for order inducing. This is of particular significance to performing journal ranking and assessment, as it mirrors the way that human experts make such decisions, where multiple indicators are necessary whilst only one overall impact value (that takes into consideration of the multiple indicators) is ultimately employed when judging a journal's standing.

V. EXPERIMENT AND EVALUATION

A. Experimental Set-up

In order to evaluate the performance of different aggregation operators for journal ranking, their clustering results are compared with human expert decisions as reflected in the Ranked Journal List (RJL) that is provided by the ERA 2010 [10]. RJL has involved a large group of scholars to rank a large number of academic journals. Although many debates surrounded the end result of RJL, it has been employed by scholars as a benchmark to compare journal ranking outcomes. Following this, in the present experiments, the result of RJL (2010) is assumed to be the ground truth in comparing the "accuracy" of different methods. Each journal in RJL has a rank in the domain $Labels$ = {A*, A, B, C}, where the label A* indicates top journals in a certain research area, and the significance of journals decreases from it down to the label C. Each journal studied in the experiments below is therefore assigned a label also taken from this domain.

In collecting sets of data to carry out the experiments, journals from six areas in the JCR Science Edition 2010 are selected:

- Agriculture (Agricultural Economics & Policy, Agricultural Engineering, Dairy & Animal Science, Multidisciplinary);
- Chemistry (Analytical, Applied, Inorganic & Nuclear, Medicinal, Multidisciplinary, Organic, Physical);
- Computer Science (Artificial Intelligence, Cybernetics, Hardware & Architecture, Information Systems, Interdisciplinary Applications, Software Engineering, Theories & Methods);
- Materials Science (Biomaterials, Ceramics, Characterization & Testing, Coatings & Films, Composites, Multidisciplinary, Paper & Wood, Textiles);
- Medicine (General & Internal, Legal, Research & Experimental, Medical Ethics, Medical Informatics, Medical Laboratory Technology);
- Physics (Applied, Atomic, Molecular & Chemical, Condensed Matter, Fluids & Plasmas, Mathematical, Multidisciplinary, Nuclear, Particles & Fields).

Amongst them, only those journals that are ranked both in RJL and indexed by the JCR are considered as valid data objects (in order to have the ground truth to entail comparison). If a journal is missed from the JCR, then it is removed from the experimental data. A summary of the resulting datasets is shown in Table II.

Scores for seven indicators as reported in the JCR Science Edition 2010 are selected to generate fuzzy similarities amongst journals. These indicators are: Total Cites (number of times a journal being cited in 2010); IF; 5-year IF; Immediacy Index (ratio of cites to the current articles over the number of those articles); Cited Half-Life (median age of the articles cited); Eigenfactor; and Article Influence (ratio of the Eigenfactor score to the total number of articles considered). All these indicators are normalised to $[0, 1]$ before they are employed to generate similarity relations between journals.

Two criteria, "accuracy" and "within-1 accuracy" [10] are adopted in order to analyse the consistency between the proposed approach and RJL. Here, accuracy is defined as the ratio of correctly clustered objects to the total number of objects in the dataset. The label of majority journals in each cluster is deemed as the rank of the journals within it. The correctly clustered objects are in turn deemed to be the journals whose assigned ranks are consistent with their ranks in the expert-devised RJL. Within-1 accuracy is a relaxed version of the standard accuracy measure; it is often adopted when cluster labels can be ordered. Following this criterion, an A*-rated journal is regarded to be a correctly classified object if it is classified as A* or A. Similarly, an A-rated journal is deemed correctly classified if the result is A*, A, or B, and so on.

B. Results and Discussion

To examine the relationship between journal clustering accuracy and the reliability of the weighting vectors for kNN-IOWA (that is equivalent to the orness for OWA), twenty weighting vectors are generated using linear stress functions with the orness values approximately uniformly distributed from zero to one. Figure 1 shows the change of accuracy (Y-axis) with respect to the orness(W) of the weighing vectors (X-axis) in both kNN-IOWA and OWA. Each point in Fig. 1 is an averaged value of 50 random centroid initialisation, and Eqn. (6) is employed to generate the similarity between journals regarding each indicator. To facilitate comparison, DOWA [23] and kNN-DOWA are also implemented, with their results shown in Fig. 1 as straight dot-lines. Further, the other two similarity measures Eqn. (7) and Eqn. (8) are also employed to carry out clustering to enrich the comparison. The best averaged results are reported in Table III, where the results on each dataset are given in the format of "aggregation operator-accuracy(%)-reliability/orness".

TABLE II. A SUMMARY OF THE DATASETS USED

Number of Instances	A*	A	B	C	Total
Agriculture	3	35	39	31	108
Chemistry	37	70	95	143	345
Computer Science	44	101	108	67	320
Material Science	26	61	80	61	228
Medicine	20	39	73	107	239
Physics	30	50	73	56	209

For five of the six datasets, the accuracies achieved by the use of non-dependent aggregation operators (kNN-IOWA and OWA) generally increase along with the increase of the reliability/orness of the weighting vectors. The performance of kNN-IOWA in relation to the weighting vectors of extreme reliability values is more stable than that of OWA. This indicates that the use of nearest neighbours as guidance for ordering arguments entails more reliable output in aggregation operators, which in turn allows the generation of better results in journal ranking. Figure 1 also shows that the kNN-IOWA is not very sensitive to the selection of k on the tested datasets. Expect on the agriculture dataset, the results of $k = 1, 3, 5$ have similar trends when orness(W) is changed and their accuracies start to show differences only when orness$(W) \approx 0.75$.

Note that the outcomes of using dependent weighting vectors in DOWA are not so good as those of using dependent weighting vectors in kNN-DOWA. This may be due to the fact that the kNN-based operators, including both kNN-DOWA and kNN-IOWA, are able to assign high weights to arguments which are close to the other relevant arguments, while DOWA only emphasises on the arguments close to their means. Thus, if individual journal indicators focus on rather different aspects, say the calculation of IF and five-year IF includes self-citations while that of Eigenfactor and Article Influence excludes self-citations, then kNN-based methods can achieve better results than DOWA. However, the accuracy reachable by using dependent weighting vectors is not so high as that achievable by the use of carefully selected weighting vectors. This shows that although dependent methods can help aggregation operators to learn weights from arguments, human intervention for carefully choosing the weights is still necessary in situations where higher accuracies are required.

Generally, the weighting vectors which have orness$(W) > 0.5$ achieved the best results in terms of both the standard accuracy and within-1 accuracy. This indicates that the weight-

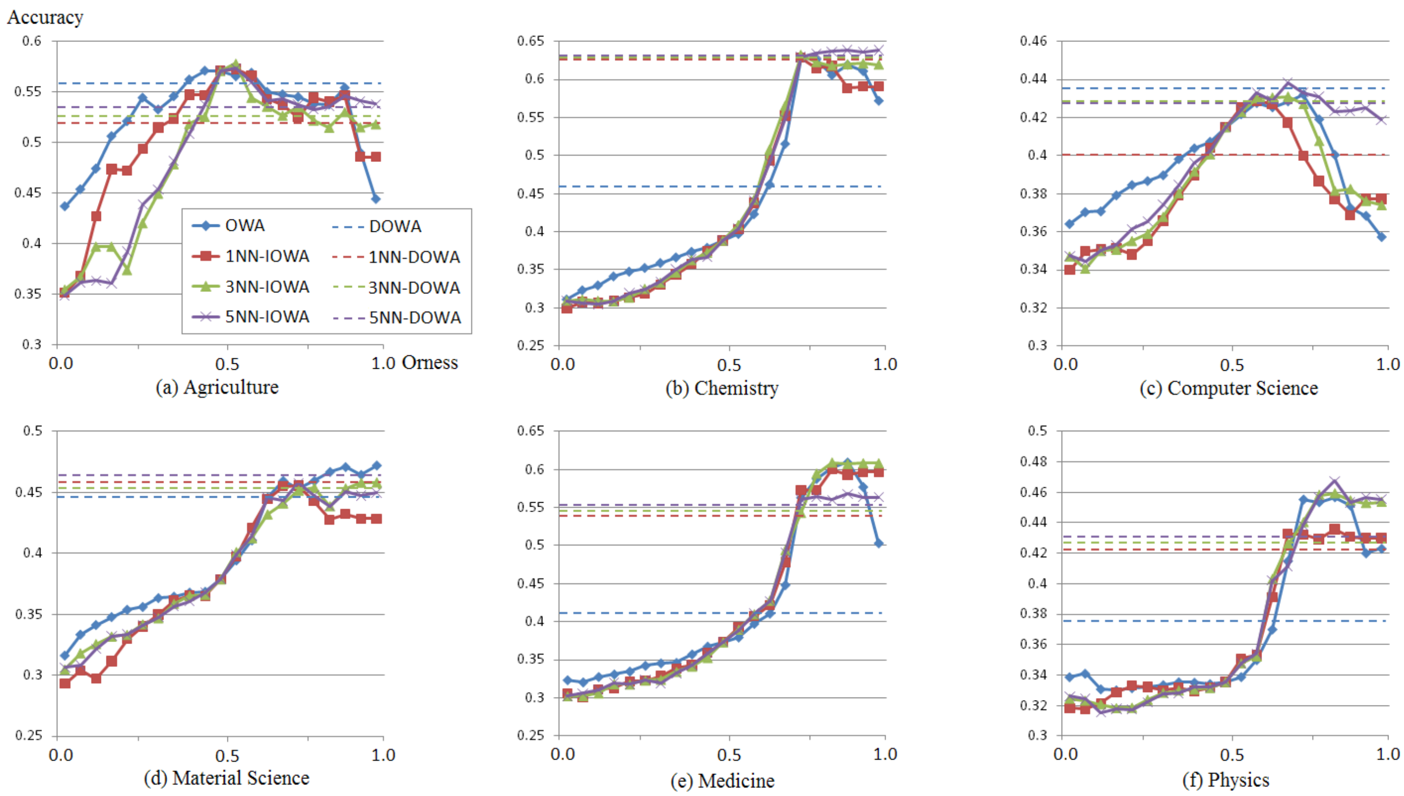

Fig. 1. Trend of Accuracy Change against Reliability

TABLE III. COMPARISON OF BEST ACHIEVED RESULTS

	Eqn. (6)		Eqn. (7)		Eqn. (8)	
	Accuracy	Within-1	Accuracy	Within-1	Accuracy	Within-1
Agriculture	3NN-57.93-0.527	5NN-96.22-0.527	3NN-58.12-0.516	3NN-96.88-0.833	5NN-57.60-1.000	5NN-96.78-1.000
Chemistry	5NN-63.85-0.889	5NN-98.46-0.944	5NN-64.38-0.944	5NN-98.13-0.944	1NN-64.13-0.944	5NN-97.15-0.622
Computer Science	5NN-43.84-0.667	OWA-86.89-0.667	5NN-44.47-0.423	OWA-89.74-0.833	3NN-44.41-1.000	OWA-86.84-0.333
Material Science	OWA-47.19-1.000	5NN-91.41-0.622	1NN-47.20-0.722	1NN-90.93-0.667	OWA-47.89-0.722	5NN-91.18-0.503
Medicine	OWA-60.95-0.889	OWA-92.01-0.889	OWA-61.72-0.889	OWA-91.49-0.944	1NN-58.12-0.889	OWA-88.75-0.944
Physics	5NN-46.75-0.833	3NN-91.00-0.722	5NN-45.35-0.889	1NN-94.01-0.833	1NN-47.85-0.577	1NN-93.88-0.722

ing vectors that exhibit a high reliability are preferable when kNN-IOWA/OWA operators are used for journal clustering. The results also show that if journals have high similarities for *more than one* indicator, the aggregated impact scores of journals may also be similar. This may be expected as there are only seven individual indicators considered.

It is interesting to note that the aggregated fuzzy relation has shown a higher accuracy and within-1 accuracy when compared with the use of Manhattan distance [33] (which is commonly adopted in classical clustering algorithms). In fact, if the fuzzy similarity for each indicator is generated using Eqn. (6) with orness$(W) = 0.5$, the accuracies of all non-dependent aggregation operators are identical to those obtained using Manhattan distance-based c-means, which is clearly reflected by the intersections on orness$(W) = 0.5$ in Fig. 1.

VI. CONCLUSION

This paper has presented a nearest neighbour-guided induced OWA operator, kNN-IOWA. The proposed aggregation operator has the strength of controlling the reliability of the aggregated output. It has been applied for building aggregated fuzzy relations between academic journals, on the basis of the individual indicator scores of the journals concerned. Three fuzzy similarity measures are used to construct different similarity measures between journals, and twenty weighting vectors which cover the range of orness(W) from 0 to 1 are compared. Experimental results on six data sets indicate that the proposed approach can outperform OWA, DOWA and kNN-DOWA, if appropriate weighting vectors are selected.

Note that a group of journals of a certain rank may often be heavily overlapped with journals of other ranks. Therefore, the low accuracy of journal ranks using clustering is not unexpected. After all, most of the journals are not obviously better or worse than others, although their ranks are more likely to be affected by the preference of the human assessors. Besides, the assumed ground truth is itself not necessarily accurate. In light of this, it may be interesting to develop an objective means for determining the relative ranking positions of academic journals using only aggregation operators on those indicators. Also, there are many other similarity measures and clustering methods available in the literature than what have been utilised in the present work. These may be employed as alternative. Such work remains active research.

REFERENCES

[1] L. Canós and V. Liern, "Soft computing-based aggregation methods for human resource management," *European Journal of Operational Research*, vol. 189, no. 3, pp. 669–681, 2008.

[2] Z. Xu, "Induced uncertain linguistic owa operators applied to group decision making," *Information Fusion*, vol. 7, no. 2, pp. 231–238, 2006.

[3] P. Su, C. Shang, and Q. Shen, "Owa aggregation of fuzzy similarity relations for journal ranking," in *Fuzzy Systems (FUZZ), 2013 IEEE International Conference on*, 2013, pp. 1–7.

[4] R. Yager, "On ordered weighted averaging aggregation operators in multicriteria decisionmaking," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 18, no. 1, pp. 183–190, 1988.

[5] T. Calvo, G. Mayor, and R. Mesiar, *Aggregation operators: new trends and applications*. Springer, 2002, vol. 97.

[6] J. Stegmann, "How to evaluate journal impact factors," *Nature*, vol. 390, no. 6660, p. 550, 1997.

[7] C. Bergstrom, J. West, and M. Wiseman, "The eigenfactor metrics," *The Journal of Neuroscience*, vol. 28, no. 45, pp. 11 433–11 434, 2008.

[8] E. Garfield, "The history and meaning of the journal impact factor," *JAMA: The Journal of the American Medical Association*, vol. 295, no. 1, pp. 90–93, 2006.

[9] R. Rousseau, "Journal evaluation: technical and practical issues," *Library Trends*, vol. 50, no. 3, pp. 418–439, 2002.

[10] G. Beliakov and S. James, "Citation-based journal ranks: the use of fuzzy measures," *Fuzzy Sets and Systems*, vol. 167, no. 1, pp. 101–119, 2011.

[11] J. Stegmann and G. Grohmann, "Citation rates, knowledge export and international visibility of dermatology journals listed and not listed in thejournal citation reports," *Scientometrics*, vol. 50, no. 3, pp. 483–502, 2001.

[12] S. Cooper and A. Poletti, "The new era of journal ranking," *Australian Universities Review*, vol. 53, no. 1, pp. 57–65, 2011.

[13] P. Su, C. Shang, and Q. Shen, "Link-based approach for bibliometric journal ranking," *Soft Computing*, vol. 17, no. 12, pp. 2399–2410, 2013.

[14] Research Excellence Framework 2014. [Online]. Available: http://www.ref.ac.uk/

[15] R. Yager and D. Filev, "Induced ordered weighted averaging operators," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 29, no. 2, pp. 141–150, 1999.

[16] J. M. Merigó and A. M. Gil-Lafuente, "The induced generalized owa operator," *Information Sciences*, vol. 179, no. 6, pp. 729–741, 2009.

[17] G. Beliakov and S. James, "Induced ordered weighted averaging operators," in *Recent Developments in the Ordered Weighted Averaging Operators: Theory and Practice*. Springer, 2011, pp. 29–47.

[18] ——, "Using choquet integrals for knn approximation and classification," in *Fuzzy Systems, 2008. FUZZ-IEEE 2008. (IEEE World Congress on Computational Intelligence). IEEE International Conference on*. IEEE, 2008, pp. 1311–1317.

[19] R. Yager, "Using stress functions to obtain owa operators," *Fuzzy Systems, IEEE Transactions on*, vol. 15, no. 6, pp. 1122–1129, 2007.

[20] G. Beliakov, S. James, and G. Li, "Learning choquet-integral-based metrics for semisupervised clustering," *Fuzzy Systems, IEEE Transactions on*, vol. 19, no. 3, pp. 562–574, 2011.

[21] R. Yager, "Norms induced from owa operators," *Fuzzy Systems, IEEE Transactions on*, vol. 18, no. 1, pp. 57–66, 2010.

[22] T. Boongoen and Q. Shen, "Nearest-neighbor guided evaluation of data reliability and its applications," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 40, no. 6, pp. 1622–1633, 2010.

[23] Z. Xu, "Dependent owa operators," in *Modeling Decisions for Artificial Intelligence*. Springer, 2006, pp. 172–178.

[24] R. Yager, "Families of owa operators," *Fuzzy Sets and Systems*, vol. 59, no. 2, pp. 125–148, 1993.

[25] L. Kuncheva and C. Whitaker, "Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy," *Machine Learning*, vol. 51, no. 2, pp. 181–207, 2003.

[26] R. Diao, F. Chao, T. Peng, N. Snooke, and Q. Shen, "Feature selection inspired classifier ensemble reduction," *Cybernetics, IEEE Transactions on*, vol. PP, no. 99, pp. 1–1, 2013.

[27] M. Ley, "The dblp computer science bibliography: Evolution, research issues, perspectives," in *String Processing and Information Retrieval*. Springer, 2002, pp. 481–486.

[28] L. Dengfeng and C. Chuntian, "New similarity measures of intuitionistic fuzzy sets and application to pattern recognitions," *Pattern Recognition Letters*, vol. 23, no. 1, pp. 221–225, 2002.

[29] A. Jain, M. Murty, and P. Flynn, "Data clustering: a review," *ACM Computing Surveys (CSUR)*, vol. 31, no. 3, pp. 264–323, 1999.

[30] T. Boongoen, C. Shang, N. Iam-On, and Q. Shen, "Extending data reliability measure to a filter approach for soft subspace clustering," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 41, no. 6, pp. 1705–1714, 2011.

[31] J. Górriz, J. Ramírez, E. Lang, and C. Puntonet, "Hard c-means clustering for voice activity detection," *Speech Communication*, vol. 48, no. 12, pp. 1638–1649, 2006.

[32] A. Graaff and A. Engelbrecht, "Clustering data in stationary environments with a local network neighborhood artificial immune system," *International Journal of Machine Learning and Cybernetics*, vol. 3, no. 1, pp. 1–26, 2012.

[33] P. Su, Y. Li, Y. Li, and S.-K. Shiu, "An auto-adaptive convex map generating path-finding algorithm: Genetic convex a*," *International Journal of Machine Learning and Cybernetics*, vol. 4, no. 5, pp. 551–563, 2013.