Combining Multiple Ranking Systems on the Generalized Permutation Rank Space

X. Zhong*, L. Hurley*, S. Sirimulla†, C. Schweikert‡, D. F. Hsu*
* Laboratory of Informatics and Data Mining, Department of Computer and Information Sciences, Fordham University, New York, NY † Department of Pharmaceutical Sciences, School of Pharmacy, University of Texas at El Paso, El Paso, TX ‡ Division of Computer Science, Mathematics and Science, St John's University, Queens, NY

***Abstract*—Given t ranking systems of n objects, consensus ranking (CR) aims to derive a ranking which best represents the consensus goal of these ranking systems. Since a ranking system of n objects is equivalent to a permutation of the n natural numbers $[1, n]$ from 1 to n, the CR problem has been studied on the bubble-sort graph rank space B_n which consists of the set of all permutations of order n. However, it remains a challenging issue when a combination of ranking systems (consensus ranking) has ties. But B_n does not include permutation (or rank systems) with ties. In this paper, we propose a multi-layer combinatorial fusion algorithm for combining multiple ranking systems on the generalized permutation rank space where ties are allowed. Using two simulated data sets and an empirical data set in the molecular docking domain, we demonstrate the robustness of our approach. This study also provides a viable approach to data analytics, machine learning, and combinatorial fusion in the non-parametric rank space.**

***Keywords*—Bubble-sort graph, Combinatorial Fusion Algorithm, ER-algorithm, generalized permutation rank space, rank aggregation, scoring system combination**

I. Introduction

The bubble-sort graph B_n is a Cayley Graph with the symmetric group S_n as the node set and adjacency is generated by a subset of adjacent swaps. The n-dimensional bubble-sort graph B_n has many good combinatorial properties including $(n-1)$-regular, bipartite, $(n-1)$-connected, Hamiltonian Laceable, fault tolerant, and recursively constructible [1], [2], [9], [13], [23], [30], [46], [47], [49]. Combinatorial Fusion, proposed in Hsu, Shapiro, and Taksa [21], has been used to combine multiple scoring or ranking systems using the rank-score characteristic (RSC) function and cognitive diversity [18]–[20]. It improves data analytic and machine learning results in a variety of domains including biomedical informatics [7], [32], [33], [41], [42], [44], virtual screening [8], [12], [50], cognitive neuroscience [5], [43], [45], predictive analytics [15], [29], [34]–[36], [38], [39], text categorization [31], and portfolio management [48].

The Cayley graph B_n as a metric space has the distance between two nodes P, Q, $d(P, Q)$, the Kendall's τ distance τ_a [27] which is equal to the number of pairwise interchanges when applying bubble-sort sorting algorithm to P in order to obtain Q [9]. Kendall [26] defined another distance measurement τ_b which can handle the weak orders, the ranking order with ties. However, as pointed out by Emond and Mason [10], Kendall's τ_b distance does not satisfy the triangle inequality. Hence it is limited as a practical metric space. Kemeny and Snell [25] proposed a distance metric, we call the KS-metric, using absolute values, which satisfies several good axioms. However, since it uses summation of absolute values, it is difficult to come up with an accurate calculation for large sample spaces.

Emond and Mason [10] proposed a rank correlation τ_x which is useful in the rank space (H_n, τ_x), where H_n is a generalized bubble-sort graph rank space and equal to $B_n \cup \{\text{rank orders with ties}\}$. In this paper, we use the generalized Bubble-sort graph rank space with H_n as our rank space and d_H as our distance metric obtained from the rank correlation τ_x.

We propose a combinatorial fusion algorithm, called ER-algorithm for expansion and reduction, on the generalized rank space H_n with distance metric d_H. Let A be a set of t weak rank orders or permutations with or without ties, $A_1, A_2, \ldots, A_t$, where normalized d_H distances to the identity permutation I_n are within the interval [0.4,0.5). We apply the ER-algorithm to the set of initial t rank orders. The resulting t rank orders are used again in a multi-layer process which aims to move these t nodes closer and closer to the identity permutation I_n. This process terminates when the results at a layer are not improved from the previous layer. We illustrate the robustness of the ER-algorithm and the multi-layer process with two simulated data sets from H_{10} and H_{300}. We also apply the multi-layer ER-algorithm to a set of 12 scoring functions which are the best 12 neural networks obtained by a deep learning process in molecular docking function development [16]. Our results demonstrate good improvement over the given 12 scoring systems.

The current study will also provide a computational framework for a variety of data analytic and machine learning tasks such as prediction, rank aggregation, joint decision making, deep learning, preference learning [4], [10], [11], [17], [19], [27], [37], [51] and target search, in the non-parametric rank space.

II. Multi-layer Combinatorial Fusion

In this section, we describe the rank space H_n with distance metric τ_x to measure the distance between two nodes. We also describe combinatorial fusion algorithms which combine multiple scoring or ranking systems with weights. We

978-1-7281-4117-6/19/$31.00 ©2019 IEEE
DOI 10.1109/DataCom.2019.00027

then define the Expansion-Reduction (ER-) algorithm and describe the multi-layer combinatorial fusion process.

A. The Rank Space H_n with Distance Metric d_H

Let $[a, b]$, where $0 < a \leq b$, be the set of positive integers between a and b inclusively. Suppose we are given a group of t rank orders, ranking systems, or permutations on a set of n data items (objects, subjects, or labels) $D = \{d_1, d_2, \ldots, d_n\}$. Let S_n be the symmetric group of order n which consists of all permutations of $[1, n]$ or all the rank orders of the data items $d_1, d_2, \ldots, d_n$. Cayley graph $\text{Cay}(S_n, T_n)$ consists of the node set S_n and the subset T_n of adjacent transpositions. Two nodes A_1 and A_2 are adjacent to each other if $A_1 = A_2 \circ t_n$ or $A_2 = A_1 \circ t'_n$ where t_n and t'_n are two permutations which are adjacent transposition in T_n, and $t_n \circ t'_n = t'_n \circ t_n = I_n$, the identity permutation [9]. These graphs are also known as the Bubble-sort graph B_n of order n because the adjacency so defined leads to the distance metric counting the number of interchanges between two nodes, a permutation, using the Bubble-sort sorting algorithm. B_n has many excellent graph properties including regularity, biparticity, connectedness, Hamiltonian properties, and recursively constructable [1], [9], [13], [23], [28], [30], [46], [47]. We illustrate B_3 and B_4 as follows:

Example 1: B_3 consists of 6 permutations $(123), (132), (213), (231), (312), (321)$. B_4 has $4! = 24$ nodes, degree = 3, diameter = 6, and 36 edges. Two B_4 graphs are shown in Fig. 1. The graph B_4 in Fig. 1(b) shows that B_4 can be recursively constructed from 4 copies of B_3 and it has maximal connectivity of 3. B_4 in Fig. 1(c) is constructed as a metric space with S_4 as the node set and Kendall τ_a rank distance as the metric. In both B_4 in Fig. 1(b) and 1(c), there exist 3 disjoint paths between any pair of two nodes. For example, between 1432 and 4123, the three disjoint paths are: $P_1 = 1432, 4132, 4123$; $P_2 = 1432, 1423, 4123$; and $P_3 = 1432, 1342, 3142, 3412, 4312, 4321, 4231, 4213, 4123$.

Since the Bubble-sort distance is equivalent to the Kendall's rank distance resulting from τ_a correlation, the Bubble-sort graph B_n is turned into a metric space $B_n = \text{G}(S_n, d_K)$ where d_K is the Kendall rank distance. But Kendall's τ_a correlation cannot handle ranking with ties [26]. Kemeny and Snell [25] proposed a distance metric d_{KS} obtained from a score matrix of a rank order as was done in τ_b which can handle ties. But since d_{KS} is represented as sums of absolute values, it does not lead to practical applications. Emond and Mason [10] reexamined the concept of the half-flip proposed in [25] and defined a new rank correlation τ_x. They then proved that the half-flips metric T_h, the Kemeny-Snell metric, and the τ_x rank correlation coefficients are equivalent in the sense that they all satisfy the Kemeny-Snell axioms. The τ_x rank correlation coefficient is defined as follows where ties and partial orders are allowed in a rank order or permutation.

Definition 1 [10]: Suppose that a weak order A, a rank order with ties allowed, of n data items $D = \{d_1, d_2, \ldots, d_n\}$, is represented as an $n \times n$ score matrix a_{ij} as follows [10]:

$$a_{ij} = \begin{cases} 1 & \text{if data item } d_i \text{ is ranked ahead of} \\ & \text{or tied with data items } d_j; \\ -1 & \text{if } d_i \text{ is ranked behind } d_j; \\ 0 & \text{if } i = j. \end{cases}$$

Rank correlation τ_x between two weak orders A and B is given by the inner product of their score matrix [10]:

$$\tau_x(A, B) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \cdot b_{ij}}{n \cdot (n-1)}$$

The relationship between the rank correlation and rank distance d_H metric is:

$$\tau_x(A, B) = 1 - \frac{d_H(A, B)}{\frac{n(n-1)}{2}}$$

So the rank distance matrix d_H metric is:

$$d_H(A, B) = n(n-1) \cdot (\frac{1 - \tau_x}{2}) \quad (1)$$

This rank distance metric d_H defined in the node set $H_n = S_n \cup \{\text{rank orders with ties allowed}\}$ defines a rank space (H_n, d_H). H_3 is illustrated as follows in Fig. 2 [24], [25]:

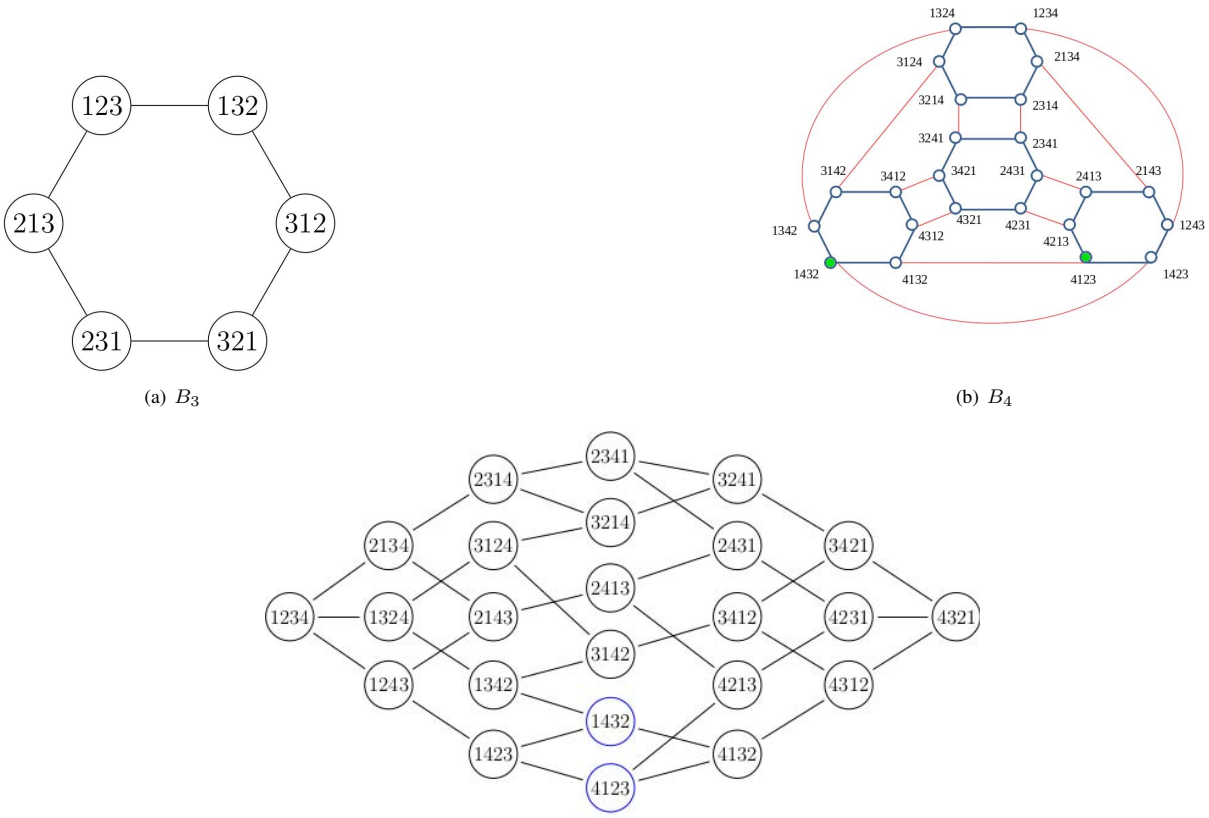
Example 2: The rank space H_3 has 13 nodes where 6 nodes are rank orders without ties, 6 nodes with 2 ties, and 1 node with three ties. In Figure 2, the node $\binom{a}{b}{c}$ means the data items a, b, c, are ranked in order 1,2,3. The node $\binom{a}{b-c}$ is equivalent to the ranking of data items as $\{1, 2.5, 2.5\}$.

We note that the number of nodes f_n in H_n is greater than or equal to the number of nodes in B_n, which is $n!$. Although an asymptotic function is given by Gross [14], the exact number is not known. We will elaborate more on this issue in concluding remark section.

B. Combinatorial Fusion Algorithm

Let $D = \{d_1, d_2, \ldots, d_n\}$ be a set of data items. Let A be a scoring system consisting of a score function s_A from D to R, the set of real numbers, and a rank function r_A from D to N, $N = [1, n]$, obtained by sorting the score values into descending order and assigning a rank $i \in N$ to each of the elements (data items) in D with respect to that order.

Let s_A and s_B be score functions and r_A and r_B be rank functions of two scoring systems A and B, respectively. Average score combination C and average rank combination D are calculated as follows: $s_C(d_i) = \frac{(s_A(d_i) + s_B(d_i))}{2}$, $s_D(d_i) = \frac{(r_A(d_i) + r_B(d_i))}{2}$. Rank functions r_C and r_D are obtained by sorting the score values in descending order based on the score functions s_C and s_D accordingly. We note that the two final rank orders r_C and r_D may not be the same. It was shown in [22] that under certain conditions, related to diversity or dissimilarity, r_D has more predictive power than r_C (also see [19]). Let $\{A_1, A_2, \ldots, A_t\}$ be t scoring systems on the set of data items D. Each scoring system A has a score function $s_A : D \to \text{R}$ and a rank function $r_A : D \to N$. For each score function, after normalization to $[0, 1]$, the score value vector $(s_A(d_1), s_A(d_2), \ldots, s_A(d_n))$ is in $[0, 1]^n$. The rank value vector $(r_A(d_1), r_A(d_2), \ldots, r_A(d_n))$ is in N^n. In this



(a) B_3 (b) B_4

(c) S_5

Fig. 1. The Bubble-sort graph B_3 and B_4. B_4 has connectivity = 3.

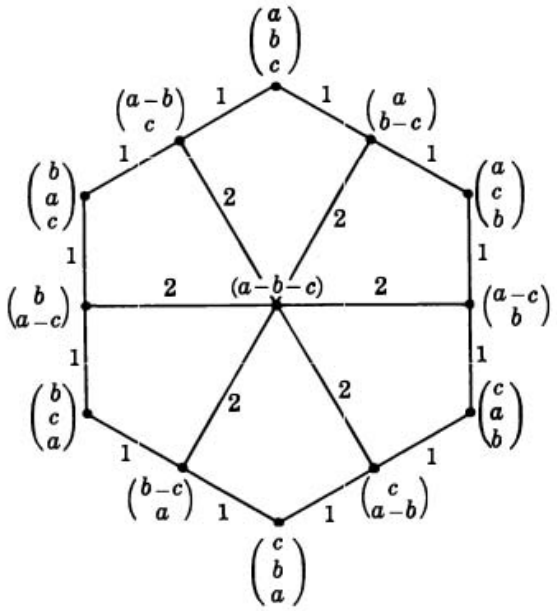


Fig. 2. The Rank Space $H_3 = \mathrm{G}(H_3, d_H)$ [24], [25].

paper, we view the rank order of A on the data set items in D as the rank order $(r_A(d_1), r_A(d_2), \ldots, r_A(d_n))$, where $r_A(d_i) \in [1, n]$ which is equivalent to the permutation of the n numbers in $[1, n]$. For a set of t rank orders (permutation) on the data items $D = \{d_1, d_2, \ldots, d_n\}$, there are totally $2^t - 1 - t$ combinations with k-combination where $k = 2, 3, \ldots, t$. For each of the k-combinations A_k, we have

$$s_{A_k}(d_i) = \sum_{j=1}^{k} r_{A_{m_j}}(d_i), \text{where } \{m_1, m_2, \ldots, m_j\} \subseteq [1, t] \tag{2}$$

C. The Multi-layer ER-Algorithm

The expansion-reduction algorithm consists of the expansion step using three combinatorial fusion algorithms: the weighted combination by performance, combination by geometric mean [6], and the mixed group rank (MGR) combination algorithm [40]. At the first step, the set of t rank orders is expanded to $2 \times (2^t - 1 - t) + 1 = 2^{t+1} - 2t - 1$ rank orders using the three methods. The second step takes these q rank orders (ranking system) which are better than the initial t rank orders. In the third step, these q systems are ordered in the decreasing order of two attributes (a) performance and (b) diversity strength calculated by average of diversity between

each ranking system (rank order) and other $q-1$ rank systems. We use a sliding rule on these two rankings to pick the top t rank orders which appear in both lists. Note that if the resulting number of ranking systems q is less than or equal to t, we just use these q systems in next layer. The ER-algorithm is summarized as follows:

ER-Algorithm: Let $A_1, A_2, \ldots, A_t$ be a group of t rank orders with ties allowed. Let $p(A_i), i = 1, 2, \ldots, t$, be the performance of A_i as measured as A_i's d_H distance to the identity permutations (including ties) I_n. The following steps will generate a new group of t or less than t rank orders which hopefully will be closer to I_n than the previous group of t rank orders.

Step 1:

- Step 1.a: Generate all the weighted $2^t - t - 1$ combinations (2-comb, 3-comb up to t-comb) using each of the combinations by performance and by geometric mean. For the k-combination r_{Ak}, we have the weighted combination W_k using performance P_{Am_j} of the rank system A_{m_j}:

$$W_k(d_i) = \sum_{j=1}^{k} \frac{r_{Am_j}(d_i)}{P_{Am_j}(d_i)}$$

  where $k = 2, 3, \ldots, t$
- Step 1.b: Generate the mixed group rank combinations [40].

Step 2: Pick top q rank orders, which are better than all the t rank orders, from the $2(2^t - 1 - t) + 1$ rank orders obtained from Steps 1.a and 1.b.

- Step 2.a: Calculate diversity (using d_H) between each pair of the $\frac{q \cdot (q-1)}{2}$ pairs of rank orders.
- Step 2.b: Calculate the performance and diversity strength of each of these q systems.
- Step 2.c: if $q < t$, we stop.

Step 3: Use the sliding rule to pick the top t rank orders which has the highest diversity strength and highest performance.

Now we apply the ER-algorithm to the initially given t rank orders continuously until it converges where there is no improved combination, $q < t$, or it reaches the identity permutation I_n.

III. Simulated and Empirical Results

In this section, we run three experiments to illustrate our proposed multi-layer ER-algorithm on the rank space H_n with d_H as distance metric of the simulated case to identity, and empirical case to target. Two simulated cases are in H_{10} and H_{300} while the empirical case is in H_{300}. In all three examples, the group of initial rank orders has $t = 6$ [3].

A. Data Sets and the Computing Environment

For the simulated cases in H_{10} and H_{300}, we initially take 125 rank orders from the Bubble-sort graph rank space S_{10} and S_{300} in H_{10} and H_{300} with the rank distance metric d_H in the interval $[0.4, 0.5)$ to the identity rank order I_{10} and I_{300}, respectively. For example in S_{300}, we first generate all 1,485,000 rank orders where d_H rank distance to I_{300} are three numbers in $[0.4, 0.5)$. Then 125,000 rank orders (permutations) are derived with uniform distribution. In this set of rank orders in S_{300}, we randomly select 150 as our pool of initial rank orders.

For the empirical case, 12 top scoring functions are taken from the top 100 neural networks resulting from a deep learning study which predicts protein-ligand binding affinities [16]. The distances of each of the 12 rank orders to the target rank order is in $[0.126, 0.135]$. This range of distances is quite close to the target in the scale from 0 to 1 as they are the results from a deep learning computing experiment [16].

All simulations and empirical results were conducted on an Ubuntu 18.04.2 server on an Intel Xeon® CPU E5-2670, 2.30Ghz, with 24 cores and 32GB of RAM.

B. Simulated Results in H_{10} and in H_{300}

In the rank space H_{10}, we divide the 150 rank orders into 28 groups of 6 rank orders with few duplicates. We randomly choose 8 groups from these 28 groups in Table I. The results demonstrate that distances of these initial 6 rank orders are in $[0.444, 0.489]$, distance of the output in $[0.0889, 0.4222]$, number of layers $[1, 5]$, and time spent (sec) in $[1.47, 7.68]$. Table I lists eight groups of the initial six rank orders from the 125 rank orders and their outputs with the number of layers and time spent.

In the rank space H_{300}, we use the 150 rank orders to construct 50 groups of 6 each with some duplicates. We choose 5 groups from these 50 groups in Table II. The results demonstrate that distances of the initial 6 rank orders in $[0.49949, 0.49996]$, distance of the outputs in $[0.46253, 0.48495]$, number of layers in $\{2, 3\}$, and time spent (min) in $[20.29, 31.64]$, respectively. Table II lists five groups of the initial six rank orders from the 150 rank orders and their outputs with the number of layers and time spent.

C. Empirical Results in H_{300}

The top 12 rank orders obtained from the results of the best 100 neural networks [16] have rank distances d_H to the target ranking in $[0.126466, 0.134448]$. We choose 5 groups of 6 rank orders from the 12: best 6, worst 6, random six (case A), random six (case B), and random six (case C). The distances of the output rank orders to the target ranking $[0.117659, 0.120223]$, number of layers $\{2, 3\}$, and time spent (in mins) $[39.64, 47.74]$ are listed in Table III.

D. Discussion

In the simulated case H_{10}, our ER-algorithm improves the d_H rank distance (to the identity permutation) from 0.489 to 0.0889 as demonstrated in Table I. This amounts to 36 steps out of the 90 steps which is the diameter of H_{10}. In the case of H_{300}, our multi-layer combinatorial fusion algorithm using the ER-algorithm improves the d_H rank distance from 0.49996

TABLE I
H_{10} SIMULATED CASE

Group	Index	Input	Input	Input	Input	Input	Input	Output	Layer	Time (sec.)
G_1	73,74,75,76,77,78	0.489	0.444	0.489	0.444	0.467	0.489	0.0889	5	5.93
G_2	7,32,57,82,107,132	0.489	0.489	0.467	0.467	0.444	0.444	0.0889	4	7.68
G_3	16,41,66,91,116,141	0.489	0.489	0.467	0.467	0.489	0.444	0.1333	4	6.90
G_4	49,50,51,52,53,54	0.444	0.467	0.467	0.467	0.444	0.489	0.2000	4	4.40
G_5	25,26,27,28,29,30	0.467	0.489	0.467	0.444	0.444	0.467	0.2000, 0.2222, 0.2444	3	5.50
G_6	15,40,65,90,115,140	0.444	0.467	0.489	0.467	0.467	0.467	0.2667, 0.2889, 0.3111	2	2.61
G_7	133,134,135,136,137,138	0.444	0.489	0.489	0.489	0.444	0.489	0.3111,0.3333	2	2.30
G_8	8,33,58,83,108,133	0.489	0.489	0.489	0.489	0.444	0.444	0.4000, 0.4222	1	1.47

TABLE II
H_{300} SIMULATED CASE

Group	Index	Input	Input	Input	Input	Input	Input	Output	Layer	Time (min.)
G_1	19,20,21,22,23,24	0.49996	0.49993	0.49971	0.49973	0.49953	0.49953	0.46253, 0.46446, 0.46462, 0.46524, 0.46537, 0.46540	3	22.90
G_2	13,14,15,16,17,18	0.49996	0.49996	0.49973	0.49971	0.49953	0.49953	0.47061, 0.47166, 0.47374, 0.47670, 0.47737, 0.47819	2	25.14
G_3	25,26,27,28,29,30	0.49993	0.49996	0.49971	0.49976	0.49953	0.49949	0.47458, 0.48486, 0.48495	2	20.29
G_4	1,2,3,4,5,6	0.49993	0.49993	0.49973	0.49973	0.49949	0.49951	0.47594, 0.47679, 0.47690	3	20.98
G_5	7,8,9,10,11,12	0.49996	0.49993	0.49971	0.49971	0.49951	0.49949	0.48370, 0.48375	3	31.64

TABLE III
H_{300} EMPIRICAL CASE

Group	Input	Input	Input	Input	Input	Input	Output	Layer	Time (min.)
$G_1(B6)$	0.126466	0.129164	0.130903	0.131037	0.131371	0.132308	0.117659, 0.118149, 0.118172, 0.118239, 0.118305	3	47.74
$G_2(W6)$	0.132486	0.132642	0.133289	0.133846	0.133846	0.134448	0.118707	2	39.64
$G_3(R6A)$	0.132308	0.131371	0.133846	0.132486	0.131037	0.130903	0.117793, 0.117837, 0.119576, 0.119755, 0.119777, 0.119866	3	44.87
$G_4(R6B)$	0.131037	0.132642	0.134448	0.132486	0.130903	0.133846	0.119086, 0.119420, 0.119955, 0.120045, 0.120201, 0.120223	3	46.34
$G_5(R6C)$	0.131037	0.133846	0.131371	0.129164	0.132642	0.134448	0.11922	2	45.53

to 0.46253 in Table II. This is equivalent to 3357 steps out of the diameter $300 \times 299 = 89700$ steps. Comparing the two rank spaces H_{10} and H_{300}, the complexity of more than 300! nodes in H_{300} increases tremendously compared to that of more than 10! nodes in H_{10}.

In the empirical case H_{300}, the multilayer CFA improves the d_H rank distance from 0.126466 to 0.117659 which is equivalent to 790 steps out of the diameter of 89700 steps.

It is interesting to note that although the improvement percentage 40.01% in H_{10} decreases to that of 3.743% in H_{300}, the time spent from 7.68 sec in H_{10} to 31.64 min=1898.4 sec in H_{300} is proportional to 300. The proportional increase of the time for the multi-layer CFA is reasonable with respect to the architecture of the generalized rank space. However, the performance improvement depends on the initial six ranking systems. In particular, the location of these initial ranking systems (or permutations) would dictate the diversity between them which is crucial in the expansion phase of the ER-algorithm. We will elaborate more on this issue in the next section.

IV. CONCLUDING REMARK AND FURTHER WORK

In this paper, we propose a multi-layer combinatorial fusion algorithm for t ranking systems of n objects, to derive a ranking system which best represents the consensus goal of these ranking systems. The proposed ER-algorithm works on the generalized rank space H_n with the node set $H_n = B_n \cup \{\text{ranking systems with ties allowed}\}$ and edge set defined by the rank distance d_H which is related to the rank correlation τ_x [10]. Using two simulated data sets in H_{10} and F_{300} and one empirical data set in H_{300} resulting from a deep learning study in the protein-ligand docking scoring function development domain, we demonstrated the robustness of the proposed process.

The study will also provide a new approach to perform big data analytics, deep learning, consensus ranking, decision making, and model fusion in the non-parametric rank space [10], [17], [19], [27]. In particular, the space H_n of ranking systems, with ties allowed, and the rank distance d_H provide a good learning and fusion space. The thrust of our approach is the combination of statistical techniques such as rank correlation and combinatorial techniques such as rank distance. In addition, we are able to extend the bubble-sort graph rank space B_n to the generalized rank space H_n which includes ranking systems with ties allowed. Since B_n (also S_n) is a subset of H_n, the many good combinatorial properties in B_n can be used to facilitate the ER-algorithm in the CFA process. Table IV gives a comparison between rank correlation and rank distance as used in statistics and in combinatorics, respectively:

In the following, we list some issues which are worthy of further investigation.

1) Which t initial ranking systems we should use? Ideally, these t initial ranking systems should be good enough and have diversity between them [18], [19], [50]. This translates to first choosing those t initial ranking systems from the left half of the rank space (see Fig. 1(c)) as we did to pick ranking systems with rank distance d_H in [0.4 to 0.5). Secondly, these t ranking systems have to be diverse. In our future work, we will pick those 150

TABLE IV
COMBINATORIAL DISTANCE VS STATISTICAL CORRELATION

Combinatorial properties	distance(A,B) adjacency(A, B)	d_K $B_n(\mathbb{S}_n, \tau_a)$	d_{KS} (H_n, T_h)	d_H (H_n, T_h) [10]
Statistical Properties	correlation(A,B) matrix fun A	τ_a [29] a_{ij} [25]	τ_b [26], d_{KS} [25] a_{ij} [26]	τ_x [10] a_{ij} [10]

TABLE V
THE NUMBER OF NODES IN H_n, f_n, AS COMPARED TO $n!$

n	2	3	4	5	6	7	8	9	10
f_n	3	13	75	541	4,683	47,293	545,835	7,087,261	102,247,563
$n!$	2	6	24	120	720	5,040	40,320	362,880	3,628,800

ranking systems in H_{10} and H_{300} (section III(A)) with high diversity to facilitate the formation of the t ranking systems.

2) What is the optimal number of initial ranking systems? In this paper we use $t = 6$. We could have used $t = 5$ or $t = 7$. Ideal case should be at least 3 or 4 and at most 10 or 11 [3], [19], [50] to have a proper balance between performance of each of the ranking systems and diversity between any pair of two systems. In the future, we will investigate the cases $t = 5, 7,$ and 8.
3) What are the combination methods we should use? In theory, we should use the combination method which expands the rank space to include ranking systems which are better than the existing ranking systems. In this paper, we use three kinds of combination methods in our multi-layer CFA process: performance weighted rank combination, rank combination using geometric mean, and the mixed group rank method [40]. We will explore other combination methods in the generalized rank space H_n.
4) What is the exact number of nodes f_n in H_n? Gross [14] gave a recurrence relation and an asymptotic function for f_n. The first few numbers produced are listed in Table V as compared to $n! = |B_n|$, where B_n is the bubble-sort Cayley Graph. Armed with the recurrence relation for f_n and the combinatorial structure of H_n, one might be hopeful to construct an exact formula for f_n.

REFERENCES

[1] S. Akers and B. Krishnamurthy, "A group-theoretic model for symmetric interconnection networks," *IEEE Transactions on Computers*, vol. 38, no. 4, pp. 555–566, Apr 1989.

[2] T. Araki and Y. Kikuchi, "Hamiltonian laceability of bubble-sort graphs with edge faults," *Information Sciences*, vol. 177, no. 13, pp. 2679–2691, Jul 2007.

[3] J. S. Armstrong, "Combining forecasts," in *Principles of forecasting*, ser. International Series in Operations Research & Management Science. Springer US, 2001, pp. 417–439.

[4] K. A. Bantel and S. E. Jackson, "Top management and innovations in banking: Does the composition of the top team make a difference?" *Strategic Management Journal*, vol. 10, no. S1, pp. 107–124, 1989.

[5] A. Batallones, K. Sanchez, B. Mott, C. Coffran, and D. F. Hsu, "On the combination of two visual cognition systems using combinatorial fusion," *Brain Informatics*, vol. 2, no. 1, pp. 21–32, Feb 2015.

[6] J. Bedö and C. S. Ong, "Multivariate spearman's ρ for aggregating ranks using copulas," *Journal of Machine Learninf Research*, vol. 17, pp. 7072–7101, 2016.

[7] S. Brown, D. F. Hsu, C. Schweikert, and Z. Tang, "ChIP-seq analytics: Methods and systems to improve ChIP-seq peak identification," in *Systems Biology*. World Scientific, Feb 2012, pp. 87–112.

[8] Y. F. Chen, K. C. Hsu, P. T. Lin, D. F. Hsu, B. S. Kristal, and J. M. Yang, "LigSeeSVM: Ligand-based virtual screening using support vector machines and data fusion," *International Journal of Computational Biology and Drug Design*, vol. 4, no. 3, p. 274, 2011.

[9] P. Diaconis, *Group representations in probability and statistics*, ser. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11. Hayward, CA: Institute of Mathematical Statistics, 1988.

[10] E. J. Emond and D. W. Mason, "A new rank correlation coefficient with application to the consensus ranking problem," *Journal of Multi-Criteria Decision Analysis*, vol. 11, no. 1, pp. 17–28, Jan 2002.

[11] J. Fürnkranz and E. Hüllermeier, Eds., *Preference Learning*. Springer Berlin Heidelberg, 2011.

[12] C. M. Ginn, P. Willett, and J. Bradshaw, "Combination of molecular similarity measures using data fusion," in *Virtual Screening: An Alternative or Complement to High Throughput Screening?* Kluwer Academic Publishers, pp. 1–16.

[13] M. D. Grammatikakis, D. F. Hsu, and M. Kraetzl, *Parallel System Interconnections and Communications*. CRC Press, Dec 2000.

[14] O. A. Gross, "Preferential arrangements," *The American Mathematical Monthly*, vol. 69, no. 1, p. 4, Jan 1962.

[15] Y. Hao, B. Kristal, and D. F. Hsu, "Predication of NCAA bracket using recurrent neural network and combinatorial fusion," in *4th IEEE Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech)*. IEEE, Aug 2018.

[16] M. Hassan, D. C. Mogollon, O. Fuentes, and S. Sirimulla, "DLSCORE: A deep learning model for predicting protein-ligand binding affinities," 2018.

[17] W. J. Heiser and A. D'Ambrosio, "Clustering and prediction of rankings within a kemeny distance framework," in *Algorithms from and for Nature and Life*. Springer International Publishing, 2013, pp. 19–31.

[18] D. F. Hsu, Y.-S. Chung, and B. S. Kristal, "Combinatorial fusion analysis: methods and practice of combining multiple scoring systems." in *Advanced Data Mining Technologies in Bioinformatics*. Idea Group Publisher Global, 2006.

[19] D. F. Hsu, B. S. Kristal, Y. Hao, and C. Schweikert, "Cognitive diversity: A measurement of dissimilarity between multiple scoring systems," *Journal of Interconnection Networks*, vol. 19, no. 01, p. 1940001, 2019.

[20] D. F. Hsu, B. S. Kristal, and C. Schweikert, "Rank-score characteristics (RSC) function and cognitive diversity," in *Brain Informatics*. Springer Berlin Heidelberg, 2010, pp. 42–54.

[21] D. F. Hsu, J. Shapiro, and I. Taksa, *Methods of data fusion in information retrieval: rank vs. score combination*, DIMACS; Rutgers University, 2002.

[22] D. F. Hsu and I. Taksa, "Comparing rank and score combination methods for data fusion in information retrieval," *Information Retrieval*, vol. 8, no. 3, pp. 449–480, Jan 2005.

[23] K. Kaneko and Y. Suzuki, "Node-to-node internally disjoint paths problem in bubble-sort graphs," in *10th IEEE Pacific Rim International Symposium on Dependable Computing, 2004 Proceedings.* IEEE, 2004.

[24] J. G. Kemeny, "Mathematics without numbers," *Quantity and Quality*, vol. 88, no. 4, pp. 577–591, Fall 1959.

[25] J. G. Kemeny and J. L. Snell, "Preference rankings: An axiomatic approach," in *Mathematical Models in the Social Sciences*. Ginn and Company; Blaisdell Publishing Company, 1962, pp. 9 – 23.

[26] M. G. Kendall, "A new measure of rank correlation," *Biometrika*, pp. 81–93, 1938.
[27] ——, "Rank correlation methods." 1948.
[28] Y. Kikuchi and T. Araki, "Edge-bipancyclicity and edge-fault-tolerant bipancyclicity of bubble-sort graphs," in *8th International Symposium on Parallel Architectures,Algorithms and Networks (ISPAN'05)*. IEEE, 2005.
[29] J. H. Lee, "Analyses of multiple evidence combination," *ACM SIGIR Forum*, vol. 31, no. SI, pp. 267–276, Dec 1997.
[30] T. Lengauer, *Combinatorial Algorithms for Integrated Circuit Layout*. Vieweg±Teubner Verlag, 1992.
[31] Y. Li, D. F. Hsu, and S. M. Chung, "Combining multiple feature selection methods for text categorization by using rank-score characteristics," in *2009 21st IEEE International Conference on Tools with Artificial Intelligence*. IEEE, Nov 2009.
[32] C. Y. Lin, K.-L. Lin, C.-D. Huang, H.-M. Chang, C. Y. Yang, C.-T. Lin, C. Y. Tang, and D. Hsu, "Feature selection and combination criteria for improving predictive accuracy in protein structure classification," in *Fifth IEEE Symposium on Bioinformatics and Bioengineering (BIBE'05)*. IEEE, 2005.
[33] K.-L. Lin, C.-Y. Lin, C.-D. Huang, H.-M. Chang, C.-Y. Yang, C.-T. Lin, C. Y. Tang, and D. F. Hsu, "Feature selection and combination criteria for improving accuracy in protein structure prediction," *IEEE Transactions on Nanobioscience*, vol. 6, no. 2, pp. 186–196, Jun 2007.
[34] C.-Y. Liu, C.-Y. Tang, and D. F. Hsu, "Comparing system selection methods for the combinatorial fusion of multiple retrieval systems," *Journal of Interconnection Networks*, vol. 14, no. 01, p. 1350003, Mar 2013.
[35] H. Liu, Z. Wu, and D. F. Hsu, "Combination of multiple retrieval systems using rank-score function and cognitive diversity," in *2012 IEEE 26th International Conference on Advanced Information Networking and Applications*. IEEE, Mar 2012.
[36] H. Liu, Z. Wu, X. Zhang, and D. F. Hsu, "An information-theoretic feature selection method based on estimation of markov blanket," in *2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI∗CC)*. IEEE, Jul 2015.
[37] T.-Y. Liu, *Learning to Rank for Information Retrieval*. Springer Berlin Heidelberg, 2011.
[38] D. M. Lyons and D. F. Hsu, "Combining multiple scoring systems for target tracking using rank–score characteristics," *Information Fusion*, vol. 10, no. 2, pp. 124–136, Apr 2009.
[39] D. M. Lyons, D. F. Hsu, Q. Ma, and L. Wang, "Selection of fusion operations using rank-score diversity for robot mapping and localization," in *Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2007*, B. V. Dasarathy, Ed. SPIE, Apr 2007.
[40] O. Melnik, Y. Vardi, and C.-H. Zhang, "Mixed group ranks: preference and confidence in classifier combination," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 8, pp. 973–981, Aug 2004.
[41] C.-H. Peng, J.-T. Hsu, Y.-S. Chung, Y.-J. Lin, W.-Y. Chow, D. F. Hsu, and C. Y. Tang, "Identification of degenerate motifs using position restricted selection and hybrid ranking combination," *Nucleic Acids Research*, vol. 34, no. 22, pp. 6379–6391, Nov 2006.
[42] C. Schweikert, S. Brown, Z. Tang, P. R. Smith, and D. F. Hsu, "Combining multiple ChIP-seq peak detection systems using combinatorial fusion," *BMC Genomics*, vol. 13, no. Suppl 8, p. S12, 2012.
[43] C. Schweikert, L. Gobin, S. Xie, S. Shimojo, and D. F. Hsu, "Preference prediction based on eye movement using multi-layer combinatorial fusion," in *Brain Informatics*. Springer International Publishing, 2018, pp. 282–293.
[44] C. Schweikert, D. Mulia, K. Sanchez, and D. F. Hsu, "The diversity rank-score function for combining human visual perception systems," *Brain Informatics*, vol. 3, no. 1, pp. 63–72, Feb 2016.
[45] C. Schweikert, S. Shimojo, and D. F. Hsu, "Detecting preferences based on eye movement using combinatorial fusion," in *2016 IEEE 15th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC)*. IEEE, Aug 2016.
[46] L.-M. Shih and J. J. Tan, "Fault-tolerant maximal local-connectivity on the bubble-sort graphs," in *2009 Sixth International Conference on Information Technology: New Generations*. IEEE, 2009.
[47] Y.-K. Shih, C.-K. Lin, D. F. Hsu, J. J. M. Tan, and L.-H. Hsu, "The construction of mutually independent hamiltonian cycles in bubble-sort graphs," *International Journal of Computer Mathematics*, vol. 87, no. 10, pp. 2212–2225, Aug 2010.
[48] H. D. Vinod, D. F. Hsu, and Y. Tian, "Combining multiple criterion systems for improving portfolio performance," *SSRN Electronic Journal*, 2008.
[49] S. Wang and Y. Yang, "Fault tolerance in bubble-sort graph networks," *Theoretical Computer Science*, vol. 421, pp. 62–69, Mar 2012.
[50] J.-M. Yang, Y.-F. Chen, T.-W. Shen, B. S. Kristal, and D. F. Hsu, "Consensus scoring criteria for improving enrichment in virtual screening," *Journal of Chemical Information and Modeling*, vol. 45, no. 4, pp. 1134–1146, Jul 2005.
[51] J. Zhang, "Binary choice, subset choice, random utility, and ranking: A unified perspective using the permutahedron," *Journal of Mathematical Psychology*, vol. 48, no. 2, pp. 107–134, 2004.