Crossings and Permutations^{*}

Therese Biedl¹, Franz J. Brandenburg², and Xiaotie Deng³

¹ School of Computer Science, University of Waterloo, ON N2L3G1, Canada
biedl@uwaterloo.ca

² Lehrstuhl für Informatik, Universität Passau, 94030 Passau, Germany
brandenb@informatik.uni-passau.de

³ Department of Computer Science, City University of Hong Kong,
83 Tat Chee Avenue, Kowloon Tong, Hong Kong, SAR, China
csdeng@cityu.edu.hk

Abstract. We investigate crossing minimization problems for a set of permutations, where a crossing expresses a disarrangement between elements. The goal is a common permutation π^* which minimizes the number of crossings. This is known as the Kemeny optimal aggregation problem minimizing the Kendall- τ distance. Recent interest into this problem comes from application to meta-search and spam reduction on the Web.

This rank aggregation problem can be phrased as a one-sided two-layer crossing minimization problem for an edge coloured bipartite graph, where crossings are counted only for monochromatic edges.

Here we introduce the max version of the crossing minimization problem, which attempts to minimize the discrimination against any permutation. We show the NP-hardness of the common and the max version for $k \geq 4$ permutations (and k even), and establish a $2-2/k$ and a 2-approximation, respectively. For two permutations crossing minimization is solved by inspecting the drawings, whereas it remains open for three permutations.

1 Introduction

One-sided crossing minimization is a major component in the Sugiyama algorithm. The one-sided crossing minimization problem has gained much interest and is one of the most intensively studied problems in graph drawing [8, 15]. For general graphs the crossing minimization problem is known to be NP-hard [13]. The NP-hardness also holds for bipartite graphs where the upper layer is fixed, and the graphs are dense with about $n_1 n_2 / 3$ crossings [10], or alternatively, the graphs are sparse with degree at least four on the free layer [17]. The special case with degree 2 vertices on the free layer is solvable in linear time, whereas the degree 3 case is open.

The rank aggregation problem finds a consensus ranking on a set of alternatives, based on preferences of individual voters. The roots for a mathematical

^{*} The work of the first author was supported by NSERC, and done while the author was visiting Universität Passau. The work of the second and third authors was partially supported by a grant from the German Academic Exchange Service (Project D/0506978) and from the Research Grant Council of the Hong Kong Joint Research Scheme (Project No. G_HK008/04).

investigation of the problem lie in voting theory and go back to Borda (1781) and Condorcet (1785). Rank aggregations occur in many contexts, including sport, voting, business, and most recently, the Internet. "*Who is the winner?*" In gymnastics, figure skating or dancing this is decided by averaging or ranking the points of the judges. In Formula 1 racing and similarly at the annual European Song Contest the winner is who has the most points. Is this scheme fair? Why not deciding the winner by the majority of first places?

Also, the organizers of GD2005 are confronted with our crossing minimization problem. They have to make many decisions. For example, which beer (wine, food) shall be served at the GD conference dinner? What is the best choice for the individual taste of the participants? Or, more specific: *which beer is the best?*

In their seminal paper from the WWW10 conference, Dwork et al. [9] have used rank aggregation methods for web searching and spam reduction. A search engine is called *good* if it behaves close to the aggregate ranking of several search engines. Besides experimental results they have investigated the theoretical foundations of the rank aggregation problem. One of the main results is the NP-hardness of computing a so-called Kemeny optimal permutation of just four permutations, here called PCM-4. However, the given proof has some flaws, and is repaired here. In addition, we show a relationship to the feedback arc set problem and establish a $2\frac{2}{k}$ approximation, which is achieved by the best input permutation.

The common rank aggregation methods take the sum of all disagreements over all permutations. Here we introduce the *maximum version*, $\text{PCM}_{\max-k}$, which expresses a fair aggregation and attempts to avoid a too severe discrimination of any participant or permutation. With the optimal solution, nobody should be totally unhappy. We show the NP-hardness of $\text{PCM}_{\max-k}$ for all $k \geq 4$ and establish a 2-approximation, which is achieved by any input permutation. This parallels similar results for the Kemeny aggregation problem [1, 9] and for the Coherence aggregation problem [5]. The case $\text{PCM}_{\max-2}$ with two permutations is efficiently solvable, whereas the case $k = 3$ remains open.

Besides the specific results, this work aims to bridge the gap between the combinatorics of rank aggregations and crossing minimizations in graph drawing, with a mutual exchange of notions, insights, and results.

In Section 2 we introduce the basic notions from graph drawing and rank aggregations, and show how to draw rank aggregations. In Section 3 we state the NP-hardness of the crossing minimization problems for just four permutations, and prove the approximation results, and in Section 4 we investigate the special cases with two and three permutations.

2 Preliminaries

Given a set of alternatives U , a *ranking* π with respect to U is an ordering of a subset S of U such that $\pi = (x_1, x_2, \dots, x_r)$ with $x_i > x_{i+1}$, if x_i is ranked higher than x_{i+1} for some total order $>$ on U .

For convenience, we assign unique integers to the items of U and let $U = \{1, \dots, n\}$. We call π a *(full) permutation*, if $S = U$, and a *partial permutation*, if $S \subseteq U$. A permutation is represented by an ordered list of items, where the rank of an item is given by its position in the ordered list, with the highest, most significant, or best item in first place.

The rank aggregation or the *crossings of permutations problem* is to combine several rankings $\pi_1, \dots, \pi_k$ on U , in order to obtain a common ranking π^* , which can be regarded as the compromise between the rankings. The goal is the best possible common ranking, where the notion of ‘better’ depends on the objective. It is formally expressed as a cost measure or a penalty between the π_i and π^* ; the *common* version takes the sum of the penalties, the *max* version is introduced here. Several of these criteria have a correspondence in graph drawing.

A prominent and frequently studied criterion is the Kendall- τ distance [3, 5, 9, 16]. The *Kendall- τ distance* of two permutations over $U = \{1, \dots, n\}$ measures the number of pairwise disagreements or inversions, $K(\pi, \tau) = |\{(u, v) \mid \pi(u) < \pi(v) \text{ and } \tau(u) > \tau(v)\}|$. This value is invariant under renaming, or the application of a permutation σ on both π and τ , and such that τ becomes the identity. For a set of permutations $P = \{\pi_1, \dots, \pi_k\}$ this generalizes by collecting all disagreements, $K(P, \pi^*) = \sum_{i=1}^k K(\pi_i, \pi^*)$.

The value $K(P, \pi^*)$ can be expressed in various ways. For every pair of distinct items (u, v) , the *agreement* $A_P(u, v)$ is the number of permutations from P which rank u higher than v , and the *disagreement* is $D_P(u, v) = k - A_P(u, v)$. Clearly, the agreement on (u, v) equals the disagreement on the reverse ordering (v, u) . For every (unordered) pair of items, let $\Delta(u, v) = |k - 2A_P(u, v)|$ express the difference between the agreement and the disagreement of u and v .

There is an established lower bound for the number of unavoidable crossings for the permutations of P , which is the sum over the least of the agreements and disagreements,

$$LB(P) = \sum_{u < v} \min\{A_P(u, v), D_P(u, v)\}.$$

Then the disagreement against a common permutation π^* is

$$K(P, \pi^*) = LB(P) + \sum_{\pi^*(u) < \pi^*(v) \text{ and } D_P(u, v) > A_P(u, v)} \Delta(u, v).$$

Thus $\Delta(u, v)$ is added as a penalty if π^* disagrees with the majority of the permutations. If there is a tie for the ranking of u and v in P , then just the term from the lower bound is taken into account.

Recall that for the crossing minimization problem of two layered graphs the agreement and disagreement of two free vertices u and v is the crossing number of the edges incident with u and v and placing u left of v , or vice versa. The so obtained lower bound is often ‘good’ and close to the optimum value [14].

Another popular measure for the distance between permutations is the *Spearman footrule distance*, which accumulates the linear arrangement or the length between two permutations over $\{1, \dots, n\}$ by $f(\pi, \tau) = \sum_i |\pi(i) - \tau(i)|$. Again this extends to a set P of permutations by summation $f(P, \pi^*) = \sum_{j=1}^k f(\pi_j, \pi^*)$.

These measures can be scaled by individual weights, and they can be extended to partial permutations $\pi_1, \dots, \pi_k$, where each permutation operates on its subset of the universe, see [9].

Given a set of (full or partial) permutations $P = \{\pi_1, \dots, \pi_k\}$ on a universe $U = \{1, \dots, n\}$, the *crossing number* of P is the number of crossings against the best permutation π^* with respect to the Kendall- τ -distance, i.e., $CR(P) = \min_{\pi^*} K(P, \pi^*)$. The crossing minimization problem is finding such a permutation π^* . We will refer to the crossing minimization problem of k permutations as the PCM- k problem.

A new cost measure is the *max crossing number*, which attempts to minimize the number of crossings for any permutation. For a set of k permutations P and a target permutation π^* let $K_{max}(P, \pi^*) = \max\{K(\pi_i, \pi^*) \mid \pi_i \in P\}$ and define the max crossing number of P by $CR_{max}(P) = \min_{\pi^*} K_{max}(P, \pi^*)$. The permutation π^* giving the value $CR_{max}(P)$ is a solution to the max crossing minimization problem. This problem is referred to as the PCM_{max}- k problem. One could similarly consider a maximum version for the Spearman footrule distance; we have not investigated the latter further.

The following fact is readily seen.

Lemma 1. *For a set of k permutations $P = \{\pi_1, \dots, \pi_k\}$,*

$$CR_{max}(P) \leq CR(P) \leq k \cdot CR_{max}(P).$$

The crossing number represents an aggregation, which is the best compromise for the given lists of preferences and minimizes the number of disagreements. The minimal number of crossings does not necessarily distribute them uniformly among the given permutations; one can construct examples where $CR_{max}(P) \geq \lceil CR(P)/2 \rceil$ and not $CR_{max}(P) = \lceil CR(P)/k \rceil$ as one would hope. The latter equation holds for $k = 2$. The objective behind the max crossing number is an aggregation, which is fair and treats every permutation equally well and minimizes the discrimination of each participant. Clearly, both objectives can be combined to the best possible permutation π^* which minimizes the sum of crossings and then balances their distribution.

2.1 Drawing Permutations

We now translate rank aggregations to graph drawing. Two permutations π and τ on a universe $U = \{1, \dots, n\}$ are drawn as a two-layer bipartite graph with the vertices $1, \dots, n$ on each layer in the order given by π and τ and a straight-line edge between the two occurrences of each item v on the two layers.

A set of k permutations $\pi_1, \dots, \pi_k$ and a common permutation π^* are represented by a sequence of pairs of permutations, where the lower layer is fixed in all drawings. For convenience, we let the lower layer be the identity with $\pi^*(i) = i$. We can merge the permutations into the *coloured permutation graph* G , which is a bipartite graph with k edge colours, such that there are vertices $1, \dots, n$ on each layer. There is an edge in the i -th colour between u on the upper layer and j on the lower layer if and only if $\pi_i(u) = j$. See also Fig. 1.

Fig. 1. Coloured permutation graph for $\pi_1 = (6, 3, 1, 4, 2, 5)$ (green and solid), $\pi_2 = (3, 5, 2, 6, 1, 4)$ (blue and dashed), and $\pi_3 = (4, 1, 5, 3, 6, 2)$ (red and dotted)

Obviously, for two full or partial permutations π and τ , the Kendall- τ distance $K(\pi, \pi^*)$ is the number of edge crossings in a straight-line drawing of their bipartite graph. It ranges between 0 and $n(n-1)/2$ and can be efficiently computed either by accumulating for every i the number of items, which are greater than i and occur to the left of i in π , provided π^* is the identity, or by techniques from counting crossings in two-layer graphs in [21].

Lemma 2. *The Kendall- τ distance $K(\pi, \pi^*)$ of two permutations over $U = \{1, \dots, n\}$ can be computed in $O(n \log n)$ time.*

2.2 Penalty Graphs

There is a direct relationship between the crossing minimization problem and the *feedback arc set problem*, which has been established at several places. Recall that the feedback arc set problem is finding the least number of arcs F in a directed graph $G = (V, E)$, such that every directed cycle contains at least one arc from F , i.e., the graph $G' = (V, E - F)$ is acyclic. In the more general weighted case, the objective is a set of arcs with least weight. In the two-layer crossing minimization problem, the *penalty graph* has arcs with weights corresponding to the difference between the number of crossings among the edges incident with two vertices u and v , if u is placed left of v , or vice versa.

In their seminal paper, Sugiyama et al. [20] have introduced the penalty digraph for the two-layer crossing minimization problem, and in [2] it is used for voting tournaments. Demetrescu and Finocchi [6] have used this approach for the two-sided crossing minimization problem and have tested several heuristics. Recently, Ailon et al. [1] have established improved randomized approximations for aggregation and feedback arc set problems.

For the crossing minimization problem for permutations, the penalty graph can be applied in the same spirit, but we use the difference in the majority counts $\Delta(u, v)$ as edge weights. Thus, for a set of permutations P over $\{1, \dots, n\}$ the *penalty digraph* of P is a weighted directed graph $H = (V, A, w)$ with a vertex for each item u and an arc (u, v) with weight $\Delta(u, v)$ if and only if a strict majority of permutations rank u higher than v , i.e., if $(u - v) \cdot (D_P(u, v) - A_P(u, v)) < 0$. Let $w(FAS(P))$ denote the *weight* of the optimum feedback arc set in the penalty digraph.

First, we establish the connection between the crossing number and the feedback arc set of the penalty graph. For the two-layer crossing minimization problem it was first observed by Sugiyama [20], and used in various places, [1, 6, 10, 17]. As a consequence, the crossing minimization problem can be reduced to a feedback arc set problem.

Theorem 1. *Let $P = \{\pi_1, \dots, \pi_k\}$ be a set of permutations. Then the crossing number of P equals the lower bound plus the weight of the feedback arc set*

$$CR(P) = LB(P) + w(FAS(P)).$$

Proof. For any permutation π there are $LB(P)$ unavoidable inversions or crossings and $K(P, \pi) = LB(P) + \sum_{\pi(u) < \pi(v) \text{ and } D_P(u, v) > A_P(u, v)} \Delta(u, v)$. Now, the deletion of all arcs (u, v) with $u < v$ and $\pi(u) > \pi(v)$ from the penalty digraph of P leaves an acyclic digraph, since there are no cycles in a single permutation π . If π is such that $K(P, \pi)$ is minimal, then the set of arcs removed from the penalty graph is a feedback arc set.

Conversely, consider the penalty graph of P and remove any set of arcs F to make the remainder acyclic. Consider any permutation π which is in conformity with a topological ordering. Then $K(P, \pi) \leq LB(P) + \sum_{f \in F} \Delta(f)$, and if F is such that its weight is $w(FAS(P))$, then π is such that $K(P, \pi)$ is minimal.

3 Complexity of Optimal Permutations

In this section we study the complexity of finding an optimal permutation for the common and the max crossing numbers. There are strong similarities to the one-sided crossing minimization problem, which go through to the number of permutations and the degrees of the free vertices.

Crossing minimization in graphs is NP-hard. This holds true for general graphs [13], and even for two-layer graphs with the upper layer fixed. These graphs may be dense [10] or sparse with degree $k = 4$ for the vertices on the free layer [17]. The case of degree 3-graphs for the free layer is still open.

Correspondingly, there are NP-hardness results for permutations. For many partial permutations with just two elements the crossing minimization problem is in one-to-one correspondence with the feedback arc set problem, where every two element permutation represents an arc, and thus is NP-hard [11, 12]. By a different reduction from the feedback arc set problem, Bartholdi et al. [3] have proved the NP-hardness of Kemeny optimal permutations for many permutations. In [2] the first NP-hardness proof is credited to Orlin (1981, unpublished manuscript).

A major strengthening has been claimed by Dwork [9] with a reduction from the feedback arc set problem to just four permutations. However, the construction in [9] has some flaws and needs some minor corrections.

Theorem 2. *The (common) crossing minimization problem PCM- k is NP-hard for k full permutations, where $k \geq 4$ and k even.*

Proof. (Sketch). We follow the construction in [9] and reduce from the feedback arc set problem. We only explain the case $k = 4$ here; for $k \geq 6$ use the technique of [9]. Let $G = (V, E)$ be a directed graph with $|V| = \{v_1, \dots, v_n\}$ and $|E| = m$ in which we want to find the smallest feedback arc set. For every vertex v let $out(v)$ be the sequence of outgoing edges in any order, and let $in(v)$ denote the sequence of incoming edges. Finally, for a sequence x let x^r denote its reversal, reading the elements right-to-left. Now, construct two pairs of permutations from the vertices and edges of G .

$$\begin{aligned}\pi_1 &= v_1, out(v_1), v_2, out(v_2), \dots, v_n, out(v_n), \\ \pi_2 &= v_n, out(v_n)^r, \dots, v_2, out(v_2)^r, v_1, out(v_1)^r, \\ \pi_3 &= in(v_1), v_1, in(v_2), v_2, \dots, in(v_n), v_n, \text{ and} \\ \pi_4 &= in(v_n)^r, v_n, \dots, in(v_2)^r, v_2, in(v_1)^r, v_1.\end{aligned}$$

In [9] the incoming edges are listed to the right of their vertices in π_3 and π_4 , but then the construction does not work.

Let $K' = 2\binom{n}{2} + 2\binom{m}{2} + 2m(n-1)$. The claim is now that G has a feedback set of size at most f iff $CR(P) \leq K = K' + 2f$. Dwork et al. [9] use a different value for K . We omit the (straightforward) proof of this claim for space reasons.

For the common crossing minimization problem we sum the number of crossings of monochrome edges. In the max problem we wish to minimize the maximal number of such crossings, i.e., we wish to treat every arrangement as fair as possible.

Theorem 3. *The max crossing minimization problem $PCM_{\max-k}$ is NP-hard for any $k \geq 4$ (full or partial) permutations.*

Proof. (Sketch) Consider the permutations $\pi_1, \dots, \pi_4$ from Theorem 2, and construct four new permutations over four copies of pairwise disjoint elements, namely

$$\sigma_1 = \pi_1 \cdot \pi_2 \cdot \pi_3 \cdot \pi_4, \quad \sigma_2 = \pi_2 \cdot \pi_3 \cdot \pi_4 \cdot \pi_1, \quad \sigma_3 = \pi_3 \cdot \pi_4 \cdot \pi_1 \cdot \pi_2, \quad \sigma_4 = \pi_4 \cdot \pi_1 \cdot \pi_2 \cdot \pi_3.$$

One can show that the permutation that minimizes the maximal number of crossings to $\sigma_1, \dots, \sigma_4$ solves again the feedback arc problem.

3.1 Approximation Algorithms

Since the crossing minimization problems are NP-hard for any (even) $k \geq 4$, we cannot hope to find the best solution in polynomial time, and hence study other ways to attack the problem. One easy way is to use integer programming; the problem can be formulated, in a relative straightforward way (we omit details) as a 0/1 program with $O(n^4 + k)$ variables and constraints. Another way is to consider approximation algorithms, which we study next.

There is a close connection between the number of crossings, i.e., the Kendall- τ distance and the Spearman-footrule distance, as established in [7]. For a pair of

permutations, every move induces a disarrangement and each crossings implies that at most two elements must move each by one position. Hence, $K(\pi, \tau) \leq f(\pi, \tau) \leq 2K(\pi, \tau)$ for full permutations π and τ . The optimal permutation for the Spearman-footrule distance can be computed by solving a weighted perfect bipartite matching problem, as explained in [9].

An alternative 2-approximation is obtained by choosing the best among the given permutations, see [1], and there is a simple 2-approximation for the coherence complexity [5]. We now show that the technique of choosing the best among the given permutations in fact gives an even better approximation, in particular for small values of k .

Theorem 4. *There is a $(2 - \frac{2}{k})$ -approximation for the (common) crossing minimization problem PCM- k .*

Proof. Let $P = \pi_1, \dots, \pi_k$ be the input permutations. For $a > d$ and $a + d = k$, let $E_{a,d}$ be those arcs $u \rightarrow v$ for which $A_P(u, v) = a$ and $D_P(u, v) = d$, i.e., u comes before v in a permutations, and after v in d permutations. Denote $m_{a,d} = |E_{a,d}|$.

Consider the k vertex orderings defined by the k permutations, and count the number of arcs that are reversed in them. For $a > d$, each arc in $E_{a,d}$ must be reversed in exactly d of the permutations, hence the total number of reversed arcs is

$$L = m_{k-1,1} + 2m_{k-2,2} + \dots + jm_{k-j,j} + \dots^1 = \sum_{a>d, a+d=k} dm_{a,d}. \quad (1)$$

By the pigeon hole principle, therefore in at least one of the permutations (say in π_1), the number of reversed arcs is at most $1/k$ th of Equation 1. Denote by $r_{a,d}$ the number of arcs in $E_{a,d}$ that are reversed in π_1 , then we therefore have

$$r_{k-1,1} + r_{k-2,2} + \dots + r_{k-j,j} + \dots \leq \frac{1}{k} (m_{k-1,1} + 2m_{k-2,2} + \dots + jm_{k-j,j} + \dots)$$

Each arc in $E_{a,d}$ has weight $a - d$ in the feedback arc set problem, so the weight of the feedback arc set solution defined by π_1 is

$$\begin{aligned} w(FAS) &= (k-2)r_{k-1,1} + (k-4)r_{k-2,2} + \dots + (k-2j)r_{k-j,j} + \dots \\ &\leq (k-2)r_{k-1,1} + (k-2)r_{k-2,2} + \dots + (k-2)r_{k-j,j} + \dots \\ &\leq (k-2)\frac{1}{k} (m_{k-1,1} + 2m_{k-2,2} + \dots + jm_{k-j,j} + \dots) = \frac{k-2}{k}L \end{aligned}$$

Now note that L of Equation 1 also exactly equals the lower bound $LB(P)$, since we only consider edges in $E_{a,d}$ with $a > d$. Therefore, the number of crossings obtained with π_1 is

$$LB(P) + w(FAS) \leq L + \frac{k-2}{k}L = (2 - \frac{2}{k})L \leq (2 - \frac{2}{k})OPT,$$

where OPT is the number of crossings in the optimal solution.

¹ The series ends for $j = \lceil (k-1)/2 \rceil$, but in order not to clutter the equations, we will not write this explicitly here.

We note here that if the target permutation is taken from the given set of permutations, the $(2 - \frac{2}{k})$ -approximation is best possible for PCM- k . Namely, let $\sigma_1, \dots, \sigma_k$ be k permutations (over distinct elements) of length $N = n/k$, and consider the following k permutations:

$$\begin{aligned}\pi_1 &= \sigma_1^r \cdot \sigma_2 \cdot \sigma_3 \cdot \dots \cdot \sigma_k \\ \pi_2 &= \sigma_1 \cdot \sigma_2^r \cdot \sigma_3 \cdot \dots \cdot \sigma_k \\ \pi_3 &= \sigma_1 \cdot \sigma_2 \cdot \sigma_3^r \cdot \dots \cdot \sigma_k \\ &\vdots \\ \pi_k &= \sigma_1 \cdot \sigma_2 \cdot \sigma_3 \cdot \dots \cdot \sigma_k^r\end{aligned}$$

Then $\pi^* = \sigma_1 \cdot \sigma_2 \cdot \dots \cdot \sigma_k$ achieves $k \binom{N}{2}$ crossings. However, any π_i disagrees with any π_j on the directions of both σ_i and σ_j , and hence creates $2(k-1) \binom{N}{2}$ crossings, which is $\frac{2k-2}{k} = 2 - \frac{2}{k}$ times the optimum.

Now we turn to approximation algorithms for the max version of the problem. Here, choosing any of the input permutations yields a 2-approximation, and again, this cannot be improved.

Theorem 5. *There is a 2-approximation for the max crossing minimization problem PCM_{max}- k .*

Proof. Let $\pi_1, \dots, \pi_k$ be a given set of permutations. We claim that any of these permutations is a 2-approximation, and prove this for π_1 .

Let π^* be the optimal permutation for the PCM_{max}- k problem, and let j^* be the index of the permutation where the maximum is achieved in the optimal solution, i.e.,

$$K(\pi_{j^*}, \pi^*) \geq K(\pi_i, \pi^*) \text{ for all } i.$$

Note that the optimal value OPT equals therefore $K(\pi_{j^*}, \pi^*)$. Now for any permutation π_i , we have

$$K(\pi_i, \pi_1) \leq K(\pi_i, \pi^*) + K(\pi^*, \pi_1) \leq K(\pi_{j^*}, \pi^*) + K(\pi^*, \pi_{j^*}) = 2\text{OPT},$$

so $\max_i K(\pi_i, \pi_1) \leq 2\text{OPT}$, and therefore π_1 is a 2-approximation for the max crossing number problem.

Clearly, if the target permutation is taken from the given set of permutations, the 2-approximation is best possible for PCM_{max}- k . To see this use any permutation π and its reversal π^r . Then $CR(\pi, \pi^r) = n(n-1)/2$ and $CR_{\max}(\pi, \pi^r) = \lceil n(n-1)/4 \rceil$.

It remains open whether the approximation bound could be improved by choosing some other permutations. Note that for the one-sided two-layer crossing minimization, the best approximation bound long stood at 2 as well [22], but was recently improved to 1.4664 [19]. Some randomized approximations have been established in [1].

4 The Small Cases

We now consider $\text{PCM-}k$ and $\text{PCM}_{\max}\text{-}k$ for small values of k . Clearly, for $k = 1$, a single user will take his preferences for the optimal arrangement, and then there are no crossings.

Consider the case $k = 2$. For bipartite graphs with vertices of degree 2 on the lower layer the one-sided crossing minimization problem is solvable in linear time by the barycenter heuristic, and due to the nesting structure of the neighbours on the upper layer determines the left-right positions in an optimal layout, see [17]. The main ingredient here is that the penalty digraph is acyclic.

Similarly, the permutation crossing number can be found easily for two permutations π_1 and π_2 ; π_1 itself is optimal with value $c = K(\pi_1, \pi_2)$. Many optimal permutations can be found from a straight-line drawing of π_1 and π_2 , see also Figure 2. Consider an arbitrary poly-line from left to right that crosses each straight line (v, v) for $v = 1, \dots, n$ exactly once (we call such a line a *pseudo-line*.) This yields a permutation π^* by listing the elements in the order in which they were crossed. Any permutation obtained in such a way is optimal for $\text{PCM-}2$.

For example, for $\pi_1 = (6, 3, 1, 4, 2, 5)$ and $\pi_2 = (3, 5, 2, 6, 1, 4)$, π_1 and π_2 themselves and also $(6, 3, 5, 2, 1, 4)$ are optimal, see Fig. 2.

Fig. 2. Crossings for 2 permutations

Using these “intermediate” permutations, the max crossing problem can be solved in polynomial time by a sweep-line technique. Since the sum of the number of crossings c is determined, the max crossing minimization problem is solved by distributing these crossings uniformly to either side such that $CR_{\max}(\pi_1, \pi_2) = \lceil c/2 \rceil$. An optimal permutation—which is best possible both for the sum and for the maximum—can be computed in $O(n + r) \log n$ time by a standard sweep-line technique, where r is the number of crossings, by searching among all pseudo-lines.

Now we address the case $k = 3$. Here, the complexity is open, both for permutations and for one-sided two-layered graphs with degree k on the free layer [17].

There is a 3-D drawing of the crossing minimization problem, where the permutations are represented on three piles in parallel to the Z -axis, and for every item i there is a triangle between the three occurrences of i . Whether such a drawing can be used to find the optimal solution (or even a good approximation), similar as for $k = 2$, remains open.

For the crossings of permutations problem the case with odd numbers is special. For every pair of items u and v there is a clear winner. There are no ties

and the penalty graph is a complete tournament, i.e., there is exactly one directed arc (u, v) or (v, u) between each pair of vertices. Then every cycle c has a subcycle of length three [18]. There are simple permutations including a cycle, e.g. $(1, 2, 3)$, $(2, 3, 1)$ and $(3, 2, 1)$. The feedback arc set problem in tournaments has been discussed at several places, see e.g. [1, 4]. It is NP-hard in the weighted version, and still open in the unweighted case.

5 Conclusion

In this paper, we investigated the problem of rank aggregation, which corresponds to find a permutation that minimizes the number of crossings with a given set of permutations. We introduced a variant that instead considers the maximum number of crossings among those permutations. We investigated complexity results and approximation algorithms.

This problem is a one-sided two-layer crossing minimization problem in an edge-coloured bipartite graph, where only crossings between equally coloured edges are counted. As such, it is not surprising that the complexity results for our problem mirror the ones for one-sided two-layer crossing minimization. We end by mentioning some of the numerous open problems that remain in this field:

1. How do the common techniques from one-sided two-layer crossing minimization, such as barycenter and median heuristics, sifting, or ILP approaches perform for the crossing minimization of permutations?
2. How can the Spearman footrule distance be used for the one-sided two-layer crossing minimization problem? How does it relate to sorting the barycenters?
3. Investigate the *max* versions, e.g., max Spearman footrule distance and the maximum number of crossings for any edge in the one-sided two-layer crossing minimization problem.
4. Improve the approximations and establish bounds for partial permutations.
5. The case $k = 3$ remains wide open. Is it NP-hard or polynomial?

Acknowledgments

The authors would like to thank Wolfgang Brunner, Christof König and Marcus Raitner for inspiring discussions.

References

1. N. Ailon, M. Charikar, and A. Newman. Aggregating inconsistent information: ranking and clustering. STOC (2005), 684–693.
2. J.P. Barthélemy, A. Guenoche, and O. Hudry. Median linear orders: heuristics and a branch and bound algorithm. Europ. J. Oper. Res. 42, (1989), 313–325.
3. J. Bartholdi III, C.A. Tovey, and M.A. Trick. Voting schemes for which it can be difficult to tell who won the election. Soc. Choice Welfare 6, (1989), 157–165.

4. I. Charon, A. Guenoche, O. Hudry, and F. Woigard. New results on the computation of median orders. *Discrete Math.* 165/166 (1997), 139–153.
5. F.Y.L. Chin, X. Deng, Q. Feng, and S. Zhu. Approximate and dynamic rank aggregation. *Theoret. Comput. Sci.* 325, (2004), 409–424.
6. C. Demetrescu and I. Finocchi. Breaking cycles for minimizing crossings. *Electronic J. Algorithm Engineering* 6, No. 2, (2001).
7. P. Diaconis and R. Graham. Spearman’s footrule as a measure for disarray. *Journal of the Royal Statistical Society, Series B*, 39, (1977), 262–268.
8. G. Di Battista, P. Eades, R. Tamassia, and I.G. Tollis. *Graph Drawing: Algorithms for the Visualization of Graphs*. Prentice Hall, (1999).
9. C. Dwork, R. Kumar, M. Noar, and D. Sivakumar. Rank aggregation methods for the Web. *Proc. WWW10* (2001), 613–622.
10. P. Eades and N.C. Wormald. Edge crossings in drawings of bipartite graphs. *Algorithmica* 11, (1994), 379–403.
11. G. Even, J. Naor, B. Schieber, and M. Sudan. Approximating minimum feedback sets and multicuts in directed graphs. *Algorithmica* 20, (1998), 151–174.
12. M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H. Freeman, San Francisco, (1979).
13. M.R. Garey and D.S. Johnson. Crossing number is NP-complete. *SIAM J. Alg. Disc. Meth.* 4, (1983), 312–316.
14. M. Jünger and P. Mutzel. 2-layer straightline crossing minimization: performance of exact and heuristic algorithms. *J. Graph Alg. Appl.* 1, (1997), 1–25.
15. M. Kaufmann and D. Wagner (Eds.). *Drawing Graphs: Methods and Models*, LNCS 2025, (2001).
16. J. G. Kemeny. *Mathematics without numbers*. *Daedalus* 88, (1959), 577–591.
17. X. Munos, W. Unger, and I. Vrto. One sided crossing minimization is NP-hard for sparse graphs. *Proc. GD 2001*, LNCS 2265, (2002), 115–123.
18. J.W. Moon. *Topics on Tournaments*. Holt, New York (1968).
19. H. Nagamochi. An Improved approximation to the One-Sided Bilayer Drawing. *Discr. Comp. Geometry* 33(4), (2005), 569–591.
20. K. Sugiyama, S. Tagawa, and M. Toda. Methods for visual understanding of hierarchical systems structures. *IEEE Trans. SMC* 11, (1981), 109–125.
21. V. Waddle and A. Malhotra. An $E \log E$ line crossing algorithm for leveled graphs. *Proc. GD 99*, LNCS 1731 (2000), 59–70.
22. A. Yamaguchi and A. Sugimoto. An approximation algorithm for the two-layered graph drawing problem. *Discrete Comput. Geom.* 33, (2005), 565–591.