Automedians sets of permutation: extended abstract ⋆

Charles Desharnais and Sylvie Hamel

DIRO - Université de Montréal,
C. P. 6128 Succursale Centre-Ville, Montréal, Québec,
Canada, H3C 3J7,
{charles.desharnais, sylvie.hamel}@umontreal.ca

Abstract. Given a set $\mathcal{A} \subseteq \mathbb{S}_n$ of m permutations of $\{1, 2, \ldots, n\}$ and a distance function d, the **median** problem consists of finding the set $\mathcal{M}(\mathcal{A})$ of all the permutations that are the "closest" of the m given permutations. Here, we study the automedian case of the problem, i.e. when $\mathcal{A} = \mathcal{M}(\mathcal{A})$, under the Kendall-$\tau$ distance which counts the number of order disagreements between pairs of elements of permutations.

1 Introduction

The problem of finding medians of a set of m permutations of $\{1, 2, \ldots, n\}$ under the Kendall-τ distance [8, 12] is often cited in the literature as the Kemeny Score Problem [7]. Giving an ordering of n political candidate by m voters, the problem consists of finding a *Kemeny consensus*: an order of the candidates that agrees the most with the order of the m voters. This problem is polynomial-time solvable for $m = 2$, has been proved to be NP-complete when $m \geq 4$, m even (first proved in [4], then corrected in [3]), but its complexity remains unknown for $m \geq 3$, m odd. Some approximate and fixed parameters algorithms have been derived [1, 2, 6, 9–11, 13]. Here we investigate automedian sets of permutations (when the sets of permutations $\mathcal{A}$ is equal to the set of its medians $M(\mathcal{A})$) by building on a previously published extended abstract [5] and constructing new variants of automedian sets of permutations using some combinatorial properties of those sets. Note that all proofs are omitted here but will be included in the full version of this article.

2 Median of permutation

2.1 Definitions and notations

A **permutation** π is a bijection of $[n] = \{1, 2 \ldots, n\}$ onto itself. The set of all permutations of $[n]$ is denoted $\mathbb{S}_n$. As usual we denote a permutation π of $[n]$ as $\pi = \pi_1\pi_2 \ldots \pi_n$. The **identity permutation** corresponds to the identity bijection of $[n]$ and is denoted $\imath = 12 \ldots n$. Given two permutations π and $\sigma \in \mathbb{S}_n$, $\pi\sigma$ (or $\pi \circ \sigma$) denote the usual **product of permutations** *i.e.* their composition as functions. Let $\mathcal{A} \subseteq \mathbb{S}_n$ be a set of permutations of $[n]$; we will denote its **cardinality** by $\#\mathcal{A}$.
The **Kendall-τ distance**, denoted d_{KT}, counts the number of order disagreements between pairs of elements of two permutations and can be defined formally as follows: for permutations π and σ of $[n]$, we have that $d_{KT}(\pi, \sigma) = \#\{(i, j) | i < j \textit{ and } [(\pi_i^{-1} < \pi_j^{-1} \textit{ and } \sigma_i^{-1} > \sigma_j^{-1})$ or $(\pi_i{-}1 > \pi_j{-}1 \textit{ and } \sigma_i^{-1} < \sigma_j^{-1})]\}$. Given any set of permutations $\mathcal{A} \subseteq \mathbb{S}_n$ and a permutation π, we have $d_{KT}(\pi, \mathcal{A}) = \sum_{\sigma \in \mathcal{A}} d_{KT}(\pi, \sigma)$.

The **problem of finding a median of a set of permutations $\mathcal{A}$ under the Kendall-τ distance** is the following. Given $\mathcal{A} \subseteq \mathbb{S}_n$, we want to find a permutation π^*

⋆ supported by NSERC through an Individual Discovery Grant (Hamel)

of $[n]$ such that $d_{KT}(\pi^*, \mathcal{A}) \leq d_{KT}(\pi, \mathcal{A}),\ \forall\ \pi \in \mathbb{S}_n$. Finally, let us define $\mathcal{M}(\mathcal{A})$ as **the set of all medians of** $\mathcal{A}$ i.e. $\mathcal{M}(\mathcal{A}) = \{\pi \in \mathbb{S}_n | \forall \sigma \in \mathbb{S}_n, d_{KT}(\pi, \mathcal{A}) \leq d_{KT}(\sigma, \mathcal{A})\}$. In this article, we are interested to investigate automedian sets of permutations i.e sets of permutations, $\mathcal{A}$, that are equal to their set of medians, $\mathcal{M}(\mathcal{A})$.

2.2 Previous work on automedian sets

In a previous extended abstract [5], we gave two special simple cases of automedian sets. The first case (see Theorem 1) is when $\mathcal{A}$ is composed of a permutation and its "circular rotations". The second one (see Theorem 2) is when $\mathcal{A}$ consist of permutations composed of a common S_k kernel with some (or none) shared fixed points.

Theorem 1 ([5]). *Let* $\pi = \pi_1 \dots \pi_n$ *be a permutation of* $[n]$ *and let* $\uparrow_i \pi = \pi_{i+1} \dots \pi_n \pi_1 \dots \pi_i$. *If* $\mathcal{A} = \{\uparrow_i \pi \mid 0 \leq i \leq n-1\}$ *then* $\mathcal{A} = \mathcal{M}(\mathcal{A})$ *and,* $\forall i$, $0 \leq i \leq n-1$, $d_{KT}(\uparrow_i \pi, \mathcal{A}) = \frac{(n+1)n(n-1)}{6}$.

Theorem 2 ([5]). *If* $\mathcal{A} = \{\pi = 1\,2 \dots \ell\ \sigma\ \ell{+}k{+}1 \dots n \mid \sigma \text{ is any permutation of } \{\ell+1, \dots, \ell+k\}\}$, *for any* $0 \leq \ell \leq n-1$, $1 \leq k \leq n$, *then* $\mathcal{A} = \mathcal{M}(\mathcal{A})$ *and for* $\pi \in \mathcal{A}$, $d_{KT}(\pi, \mathcal{A}) = \frac{k(k-1)k!}{4}$. *Note that* $\mathcal{A} = \mathbb{S}_n$, *is the special case when* $\ell = 0$ *and* $k = n$.

3 Group action and morphism

Let $\mathcal{P}(\mathbb{S}_n)$ be the set of all subsets of $\mathbb{S}_n$. Composition of permutations turns $\mathbb{S}_n$ into a group and we get a **left group action** on $\mathcal{P}(\mathbb{S}_n)$:

$$\begin{aligned} \circ : \mathbb{S}_n \times \mathcal{P}(\mathbb{S}_n) &\longrightarrow \mathcal{P}(\mathbb{S}_n) \\ (\pi\ ,\ \mathcal{A}) &\mapsto \pi \circ \mathcal{A}, \end{aligned}$$

by setting $\pi \circ \mathcal{A} = \{\pi \circ \sigma \,|\, \sigma \in \mathcal{A}\}$. We will first show two small lemmas about the invariability of the Kendall-τ distance under this action. Note that we denote $\pi \circ \sigma$ by $\pi\sigma$ in the rest of this article.

Lemma 1. *Let* π, σ, ψ *be permutations of* $\mathbb{S}_n$. *Then* $d_{KT}(\pi\sigma, \pi\psi) = d_{KT}(\sigma, \psi)$.

Lemma 2. *Let* $\mathcal{A} \subseteq \mathbb{S}_n$ *be a set of permutations and let* π *be any permutation of* $\mathbb{S}_n$. *Then,* $d_{KT}(\pi\sigma, \pi A) = d_{KT}(\sigma, A)$.

We now have everything in place to show the following:

Theorem 3. *Let* $\mathcal{M} : \mathcal{P}(\mathbb{S}_n) \longrightarrow \mathcal{P}(\mathbb{S}_n)$ *be the function that maps a set of permutations* $\mathcal{A} \subseteq \mathbb{S}_n$ *to the set of its medians* $\mathcal{M}(\mathcal{A})$. *The function* $\mathcal{M}$ *is a morphism of action i.e.* $\pi\mathcal{M}(\mathcal{A}) = \mathcal{M}(\pi\mathcal{A})$.

Recall that the **orbit** of $\mathcal{A}$ under the action $\circ$ is the set of all $\sigma\mathcal{A}$, $\sigma \in \mathbb{S}_n$. We have the following first result on automedian sets:

Corollary 1. *Let* $\mathcal{A}$ *and* $\mathcal{B}$ *be two sets of permutations of* $\mathbb{S}_n$ *in the same orbit of the action* $\circ : \mathbb{S}_n \times \mathcal{P}(\mathbb{S}_n) \to \mathcal{P}(\mathbb{S}_n)$. *Then* $\mathcal{B} = \mathcal{M}(\mathcal{B}) \iff \mathcal{A} = \mathcal{M}(\mathcal{A})$.

4 Decomposable permutations and direct sum

Now that we have derived some basic properties of the sets of medians, let us explore ways to built new automedian sets of decomposable permutations, using the direct sum of already known sets. First, let us recall some definitions and set some notations.

4.1 Definitions and notations

Definition 1. *A permutation $\pi \in \mathbb{S}_n$ will be called* **k-decomposable***, $1 \le k \le n-1$, if $i > k \Leftrightarrow \pi_i > k, \forall i \in \{1, 2, \ldots, n\}$.*

Definition 2. *A permutation $\pi \in \mathbb{S}_n$ will be called* **indecomposable***, if it is not* **k-decomposable** *for any $k \in \{1, 2, \ldots, n-1\}$.*

Definition 3. *A set of permutations $\mathcal{A} \subseteq \mathbb{S}_n$ will be called* **k-decomposable***, $1 \le k \le n-1$, if all of its permutations are k-decomposable.*

Definition 4. *Let $\mathcal{A} \subseteq \mathbb{S}_n$ be a k-decomposable set, $1 \le k \le n-1$, and let $\pi \in \mathbb{S}_n$ be any permutation. The set $\pi\mathcal{A}$ is called* **k-separable***. If $\mathcal{A}$ is not* **k-separable** *for any k, then it we be called* **inseparable***.*

Note that to lighten notation, if there are many strictly ascending values k_1, k_2... k_m such that a permutation or a set of permutations is k_i-decomposable (resp. k_i-separable), $1 \le i \le m$, we will say that it is $k_1; k_2; ...k_m$-decomposable (resp. $k_1; k_2; ...k_m$-separable).

Example 1. Let $\pi = 32145$, then π is 3;4-decomposable. Let $\sigma = 54321$, then σ is indecomposable.

Definition 5. *Let $\pi \in \mathbb{S}_k$ and $\sigma \in \mathbb{S}_\ell$ be two permutations of length k and ℓ, respectively. The* **direct sum** *of π and σ, denoted $\pi \oplus \sigma$, is the permutation of length $k + \ell$ defined as $\pi \oplus \sigma = \pi_1\pi_2...\pi_k(\sigma_1 + k)(\sigma_2 + k)...(\sigma_\ell + k)$.*

Example 2. Let $\pi = 132$ and $\sigma = 2143$, then $\pi \oplus \sigma = 1325476$.

Note: A permutation $\pi \in \mathbb{S}_n$ that is k_1, k_2... k_m-decomposable can be written as the direct sum of $m+1$ indecomposable permutations of length $k_i - k_{i-1}$, for $1 \le i \le m+1$, with $k_0 = 0$ and $k_{m+1} = n$.

Example 3. The permutation $\pi = 32145 \in \mathbb{S}_5$ is $3; 4$-decomposable. It can be written as $\pi = 321 \oplus 1 \oplus 1$, where $321 \in \mathbb{S}_3$ and $1 \in \mathbb{S}_1$ are indecomposable permutations.

We can extend the definition of direct sum to sets of permutations in the following way:

Definition 6. *Let $\mathcal{A} \subseteq \mathbb{S}_k$ and $\mathcal{B} \subseteq \mathbb{S}_\ell$ be two sets of permutations of length k and ℓ, respectively. The* **direct sum** *of $\mathcal{A}$ and $\mathcal{B}$, denoted $\mathcal{A} \oplus \mathcal{B}$, is defined as $\mathcal{A} \oplus \mathcal{B} = \{\pi \oplus \sigma \mid \pi \in \mathcal{A}$ and $\sigma \in \mathcal{B}\}$.*

Note that by definition, $\mathcal{A} \oplus \mathcal{B}$ is k-decomposable.

Example 4. Let $\mathcal{A} = \{132, 312\}$ and $\mathcal{B} = \{2143, 2314, 2431\}$. Then $\mathcal{A} \oplus \mathcal{B}$ is 3-decomposable and $\mathcal{A} \oplus \mathcal{B} = \{1325476, 1325647, 1325764, 3125476, 3125647, 3125764\}$. Let $\pi = 3271653$, then $\pi(\mathcal{A} \oplus \mathcal{B}) = \{3726145, 3726451, 3726514, 7326145, 7326451, 7326514\}$ is 3-separable.

It is easy to show that the direct sum $\oplus$ is an associative operation and so we have, for any permutations π, σ and ϕ, that $(\pi \oplus \sigma) \oplus \phi = \pi \oplus (\sigma \oplus \phi)$. This associativity easily extends to the direct sums of sets of permutations.

4.2 Theorems and properties

Let us first show this easy result about the Kendall-τ distance of permutations that are direct sums of smaller permutations.

Lemma 3. *Giving ρ and ψ, two permutations of $\mathbb{S}_k$, ϕ and θ, two permutations of $\mathbb{S}_\ell$, let π and σ be the two permutations of $\mathbb{S}_n$, $n = k + \ell$, such that $\pi = \rho \oplus \phi$ and $\sigma = \psi \oplus \theta$. Then, $d_{KT}(\pi, \sigma) = d_{KT}(\rho, \psi) + d_{KT}(\phi, \theta)$.*

Now, let us explore some properties connecting the direct sums of sets of permutations and the median of those sets.

Theorem 4. *Let $\mathcal{A} \subseteq \mathbb{S}_k$ and $\mathcal{B} \subseteq \mathbb{S}_\ell$, $k + \ell = n$, be two sets of permutations. Then, $\mathcal{M}(\mathcal{A} \oplus \mathcal{B}) = \mathcal{M}(\mathcal{A}) \oplus \mathcal{M}(\mathcal{B})$.*

We can use Theorem 4 and the next Lemma to obtain new automedian sets by taking the direct sums of already known ones. This implies that automedian sets of permutations are closed under the operation direct sum.

Lemma 4. *Let $\mathcal{A}, \mathcal{C} \subseteq \mathbb{S}_k$ and $\mathcal{B}, \mathcal{D} \subseteq \mathbb{S}_\ell$ be four sets of permutations. If $\mathcal{A} \oplus \mathcal{B} = \mathcal{C} \oplus \mathcal{D}$, then $\mathcal{A} = \mathcal{C}$ and $\mathcal{B} = \mathcal{D}$.*

Corollary 2. *Let $\mathcal{A} \subseteq \mathbb{S}_k$ and $\mathcal{B} \subseteq \mathbb{S}_\ell$ be two sets of permutations. Then $\mathcal{A} = \mathcal{M}(\mathcal{A})$ and $\mathcal{B} = \mathcal{M}(\mathcal{B}) \iff \mathcal{A} \oplus \mathcal{B} = \mathcal{M}(\mathcal{A} \oplus \mathcal{B})$*

Now, we have everything in place to show that any k-decomposable automedian set of permutations comes from the direct sums of automedian sets of smaller permutations.

Theorem 5. *Let $\mathcal{C} \subset \mathbb{S}_n$ be a k-decomposable automedian set. Then there exist automedian sets of permutations $\mathcal{A} \subseteq \mathbb{S}_k$ and $\mathcal{B} \subseteq \mathbb{S}_\ell$, $k + \ell = n$, such that $\mathcal{A} \oplus \mathcal{B} = \mathcal{C}$.*

Corollary 3. *Let $\mathcal{C} \subset \mathbb{S}_n$ be a k-separable automedian set. Then there exist $\pi \in \mathbb{S}_n$ and automedian sets of permutations $\mathcal{A} \subseteq \mathbb{S}_k$ and $\mathcal{B} \subseteq \mathbb{S}_\ell$, $k + \ell = n$, such that $\pi(\mathcal{A} \oplus \mathcal{B}) = \mathcal{C}$.*

Lemma 5. *Let $\mathcal{C} \subset \mathbb{S}_n$ be a $k_1; k_2; \ldots; k_m$-decomposable automedian set. Then there exist automedian sets of permutations $\mathcal{A}_1 \subseteq \mathbb{S}_{k_1}$, $\mathcal{A}_2 \subseteq \mathbb{S}_{k_2 - k_1}$, $\ldots$, $\mathcal{A}_m \subseteq \mathbb{S}_{k_m - k_{m-1}}$, $\mathcal{A}_{m+1} \subseteq \mathbb{S}_{n-k_m}$, such that $\mathcal{A}_1 \oplus \mathcal{A}_2 \oplus \ldots \oplus \mathcal{A}_m \oplus \mathcal{A}_{m+1} = \mathcal{C}$.*

Corollary 4. *Let $\mathcal{C} \subset \mathbb{S}_n$ be a $k_1; k_2; \ldots; k_m$-separable automedian set. Then there exist $\pi \in \mathbb{S}_n$ and automedian sets of permutations $\mathcal{A}_1 \subseteq \mathbb{S}_{k_1}$, $\mathcal{A}_2 \subseteq \mathbb{S}_{k_2 - k_1}$, $\ldots$, $\mathcal{A}_m \subseteq \mathbb{S}_{k_m - k_{m-1}}$, $\mathcal{A}_{m+1} \subseteq \mathbb{S}_{n-k_m}$, such that $\pi(\mathcal{A}_1 \oplus \mathcal{A}_2 \oplus \ldots \oplus \mathcal{A}_m \oplus \mathcal{A}_{m+1}) = \mathcal{C}$.*

4.3 Counting automedian sets

Corollary 4 shows us that any separable automedian set of permutations $\subseteq \mathbb{S}_n$ can be written as a composition of a permutation $\pi \in \mathbb{S}_n$ with the direct sum of inseparable automedian sets. It follows that to solve the problem of finding all automedian sets of permutations of $\mathbb{S}_n$, we only need to find inseparable ones because we can get all the others by combining smaller inseparable sets. To count all the automedian sets of $\mathbb{S}_n$, let $\mathcal{AM}_n$ be the set of all automedian sets of $\mathbb{S}_n$, i.e. $\mathcal{AM}_n = \{\mathcal{A} = \mathcal{M}(\mathcal{A}) \mid \mathcal{A} \subseteq \mathbb{S}_n\}$ and let $\mathcal{I}_n$ be the set of all inseparable ones, i.e., $\mathcal{I}_n = \{\mathcal{A} = \mathcal{M}(\mathcal{A}) \mid \mathcal{A} \subseteq \mathbb{S}_n \text{ and } \mathcal{A} \text{ inseparable}\}$.

Now, an automedian set of permutations $\mathcal{A} \subseteq \mathbb{S}_n$ is either inseparable, and there are $\#\mathcal{I}_n$ of those sets or it is separable. In that case, let $i, 1 \leq i \leq n-1$ be the smallest value for which $\mathcal{A}$ is i-separable. Everything we have built so far gives us that $\#\mathcal{AM}_n = \#\mathcal{I}_n + \sum_{i=1}^{n-1} \binom{n}{i} \times \#\mathcal{I}_i \times \#\mathcal{AM}_{n-i}$, where the $\binom{n}{i}$ in the formula counts the number of ways to pick the i elements among n that are in the inseparable part of $\mathcal{A}$. We still do not know how to generate all inseparable automedian sets but we did count them for small n and got the following values of $\#\mathcal{AM}_n$ and $\#\mathcal{I}_n$:

n	$\#\mathcal{I}_n$	$\#\mathcal{AM}_n$
1	1	1
2	1	3
3	3	15
4	27	117
5	≥ 429	≥ 1389
6	≥ 8889	≥ 23667

5 Conclusion and future works

The median problem has been essentially studied for the permutation group $\mathcal{S}_n$ and under the Kendall-τ distance. From a combinatorial point of view, it would be very interesting to generalize the median problem to the signed permutations group $\mathcal{B}_n$, as well as to other Coxeter groups which are known to behave in similar manner to these two groups for a large number of combinatorial questions.

References

1. N. Ailon, M. Charikar, and N. Newman. Aggregating inconsistent information: ranking and clustering. *Journal of the ACM*, 55(5):1–27, 2008.
2. N. Betzler, J. Guo, C. Komusiewick, and R. Niedermeier. Average parameterization and partial kernelization for computing medians. *Journal of Computer and System Sciences*, 77(4):774–789, 2011.
3. T. Bield, F.J. Brandenburg, and X. Deng. Crossings and Permutations. In P. Healy and N.S. Nokolov, editors, *Graph Drawing*, volume 3843 of *Lecture Notes in Computer Science*, pages 1–12. Springer Berlin Heidelberg, 2005.
4. C. Dwork, R. Kumar, M. Naor, and D. Sivakumar. Rank Aggregation Methods for the Web. In *Proceedings of the 10th international conference on World Wide Web*, WWW '01, pages 613–622. ACM, 2001.
5. S. Hamel and R. Milosz. Medians of permutations: when $\mathcal{A} = \mathcal{M}(\mathcal{A})$. In *Proceedings of the 12th Permutation Pattern Conference*, PP'2014, pages 68–71, 2014.
6. M. Karpinski and W. Schudy. Faster Algorithms for Feedback Arc Set Tournament, Kemeny Rank Aggregation and Betweenness Tournament. In K.-Y. Chwa O. Cheong and K. Park, editors, *Algorithms and Computation*, volume 6506 of *Lecture Notes in Computer Science*, pages 3–14. Springer Berlin Heidelberg, 2010.
7. J.G. Kemeny. Mathematics without numbers. *Daedalus*, 88(4):577–591, 1959.
8. M.G. Kendall. A New Measure of Rank Correlation. *Biometrika*, 30(1-2):81–93, 1938.
9. C. Kenyon-Mathieu and W. Schudyr. How to rank with few errors. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, STOC '07, pages 95–103. ACM, 2007.
10. N. Nishimura and N. Simjour. Parameterized enumeration of (locally-) optimal aggregations. In R. Solis-Oba F. Dehne and J.R. Sack, editors, *Algorithms and Data Structures*, volume 8037 of *Lecture Notes in Computer Science*, pages 512–523. Springer Berlin Heidelberg, 2013.
11. N. Simjour. Improved parameterized algorithms for the kemeny aggregation problem. In J. Chen and F.V. Fomin, editors, *Parameterized and Exact Computation*, volume 5917 of *Lecture Notes in Computer Science*, pages 312–323. Springer Berlin Heidelberg, 2009.
12. M. Truchon. An Extension of the Condorcet Criterion and Kemeny Orders. Internal Report, Université Laval, 16 pages, 1998.
13. A. vanZuylen and D.P. Williamson. Deterministic pivoting algorithms for constrained ranking and clustering problems. *Mathematics of Operations Research*, 34(3):594–620, 2009.