Fairness in Kidney Exchange Programs: The Nash Social Welfare Program Perspective

Abstract

Patients suffering from end-stage kidney disease can see their life quality improved by receiving a transplant. Unfortunately, many patients do not have a compatible donor. Incompatible patient-donor pair can enter a Kidney exchange program (KEP) in the hope of exchanging donors in order to obtain a suitable kidney for the patient in the pair. These programs present number of challenges from the point of view of fairness. In this medical context, it must be decided which patient gets a compatible kidney under a limited amount of resources (i.e., compatible donors). First, how can we define fairness in this context? This question has already been explored in the literature. However, what is the cost of these fairness schemes with respect to utility? How can we systematize the selection of exchange plans that will balance both total utility and fairness? In this article, we provide an overview of various fairness schemes in KEPs and then apply the concept of Nash social welfare programs to KEPs, where we optimize over a utilitarian and a fairness components simultaneously.

Introduction

One in ten Canadian has kidney disease (Kidney Foundation 2020). This incurable illness leaves patients who have reached end-stage kidney disease with two availabe treatments: dialysis and transplantion. Since the former can prove very costly, transplantion is often the preferred method of treatment (Kidney Foundation 2020). Kidney exchange programs (KEP) provide a A Kidney Exchange Program (KEP) consists of a graph $G = (V, A)$, where $V = P \cup N$ is the set of vertices and A is the set of arcs. The sets P and N represent incompatible patient-donor pairs and altruistic donors, respectively. Incompatible pairs enter the program in the hope of finding a compatible donor for their patient through a bartering mechanism. A feasible exchange in the graph G is *(i)* a cycle (or closed chain) such that all its vertices belong to P, or *(ii)* a chain in a simple path with vertices belonging to P, except for the first one that must be in N. The objective of a KEP is to maximize the benefit of patients in selected exchanges. This results in the following integer program:

$$\max \sum_{c \in \mathcal{C}} w_c x_c$$

$$\sum_{c \in \mathcal{C} | i \in c} x_c \leq 1 \qquad \forall i \in V \tag{1}$$

$$x_c \in \{0, 1\} \qquad \forall c \in \mathcal{C}, \tag{2}$$

where $\mathcal{C}$ is the set of cycles and chains for the KEP graph and the w_c's are the weights of each cycle/chain. Typically, one sets $w_c = |c|$ for cycles and $w_c = |c| - 1$ for chains, and thus maximize the number of transplants. This formulation is called the *cycle formulation* (Abraham, Blum, and Sandholm 2007; Roth, Sönmez, and Ünver 2007). Others like the *position-indexed edge formulation* (Dickerson et al. 2016), associate binary variables to the arcs.

The aforementioned description of KEPs does not take into account notions of fairness. It simply maximizes the total utility. Thus, this procedure might prove undesirable for many vertices that cannot be in a solution that maximizes the total number of transplants. Would it be fair to leave those vertices out entirely? Especially, if including them does not sacrifice too much total utility, it might be desirable to give these vertices a chance of being selected in an exchange plan. In order to better understand how to balance fairness and utility, it is crucial to explore various notions of fairness in KEPs. This motivates the next section.

Fairness Schemes in KEPs

In order to compare various notions of fairness, we need to introduce some key definitions inspired by Bertsimas, Farias, and Trichakis (2011).

Definition 1. *A utility set U given by a vector-valued function f is defined as*

$$U = \left\{ u \in \mathbb{R}_+^{|P|} \middle| \exists x \in X : u_j = f_j(x), j \in P \right\},$$

where $X \subseteq \mathbb{R}_+^m$ is the ***resource set*** *(i.e. the constraints) and f_j is the utility function of patient j.*

In other words, a utility set is the set of achievable utilities over the patients. Using this concept, we can now define a fairness scheme.

Definition 2. *A fairness scheme $\mathcal{L} : 2^{\mathbb{R}_+^{|P|}} \to \mathbb{R}_+^{|P|}$ is defined by*

$$\mathcal{L}(U) \mapsto u$$

for some $u \in U$ for each utility set U.

A fairness scheme represents a method for selecting the distribution of utilities among the participants. If we think of resource allocation problems, this will determine the utility received by each individual.

Since each resource set leads to a utility set, through a fairness scheme one can select the allocation of utilities for each patient (rather, each pair) in the KEP. It is now interesting to measure the cost to social welfare of using such a scheme. This is called **price of fairness** (POF).

Definition 3. *The price of fairness function POF for a utlity set U and fairness scheme $\mathcal{L}$ is defined as*

$$POF(U, \mathcal{L}) = \frac{\sup_{u \in U} e^T u - e^T \mathcal{L}(U)}{\sup_{u \in U} e^T u},$$

where $e \in \mathbb{R}^{|P|}$ is a vector of ones.

We will use this definition to compare the various notions of fairness discussed in this article. Let $\mathcal{M}$ be the set of feasible matchings in the KEP graph.

Utilitarian principle

The utilitarian approach represents the most basic approach to fairness in KEPs: it optimizes a fairness criterion that aggregates the utilities of each participating pairs. Supposing that we have a utility score $u_i : \mathcal{M} \to \mathbb{R}$ for each pair $i \in V$, the optimization takes the form:

$$\max_{M \in \mathcal{M}} \sum_{i \in V} u_i(\mathcal{M}).$$

For example, in Dickerson, Procaccia, and Sandholm (2014); Zafar et al. (2017), the authors seek to maximize the number of hard-to-match patients in a selected exchange (often referred to as group fairness). Prior to the optimization, pairs are labelled as hard-to-match if the PRA score of its patient exceeds 80%. One can also seek to simply optimize the number of transplants in a selected exchange plan as the fairness criterion, thus leaving us with a KEP in its simplest form.

The main disadvantage of the utilitarian approach is that maximizing the social welfare can lead to the systematic exclusion of certain pairs. For instance, image a large cycle in the graph, where one vertex is in another much smaller edge-disjoint cycle. Then, we can choose either the large cycle or the small one. By the utilitarian principle, the large one will always be selected. This leaves out all the remaining vertices in the smaller cycle. They will never be chosen as part of a selected exchange.

Aristotle's equity principle

Aristotle's equity principle is based on individuals having pre-existing rights to resources. They should receive what is rightfully theirs based on these claims. A simple example where this principle arises is in the context of stock dividends: the dividends are paid out to the stockholders proportionally to the number of stocks owned by each individual (when all stocks are of the same type). Here, the number of shares is a direct measure of a person's claims on the distribution of the resource. However, in the context of KEPs, there are some immediate concerns. How should we prioritize patients over others? Should we even prioritize some patients and if so, when? What constitutes a stronger claim for transplantation over another one? One attempt to answer these questions regards the waiting time of patients in a KEP pool. A scoring mechanism for ranking patient priority should reflect a higher score for patients that have waited longer. The intuition behind this is that the health of patients worsens over time. Mathematically, the scoring mechanism ρ has the property

$$\forall i \in P, t_2 \geq t_1 \implies \rho(y_i, t_2) \geq \rho(y_i, t_1),$$

where y_i is the feature vector for patient i. So all else being equal, a patient will receive a higher score after waiting longer for a kidney. A simple objective will seek to maximize the sum of the $\rho(y_i, t)$ over all patients $i \in P$. An immediate way of having a scoring mechanism ρ would be to attribute weights to its arguments. However, attributing these weights in an ad hoc manner seems rather undesirable. We can always try to learn these weights but then, we are faced with another task in itself entirely. If the priority score of a vertex is very high, it might be chosen and thus lead to a high POF (see vertex i in Figure 1).

Nash standard of comparison

John Nash introduced his concept of fairness in Nash (1950). A transfer of resources is justified if the relative gain to certain participants is bigger than the relative loss to others. This fairness scheme satisfies the four axioms of *Pareto optimality*, *symmetry*, *affine invariance* and *independence of irrelevant alternatives*. Below, we introduce the fairness notion formally.

Definition 4. *Under **proportional fairness**, a fairness scheme $\mathcal{L}(U)$ is fair if*

$$\sum_{j \in P} \frac{u_j - \mathcal{L}(U)_j}{\mathcal{L}(U)_j} \leq 0,$$

for any $u \in U$, where P is the set of participants.

Thus, if a person has accumulated several resources (i.e. they have a high utility), redistributing some of those resources to others with much smaller utilities is justified. When selecting an exchange plan, if we can find a marginal increase in the utilities of the patients by modifying the exchange plan, then this change is warranted.

We can consider the case where the utility of a patient is 1 if they receive a kidney, otherwise $\epsilon > 0$ arbitrarily small. Under this utility function, proportional fairness will correspond to an allocation that also maximizes total utility. The POF will thus be 0 and we are left with the same issues that are plaguing the utilitarian approach. This counter-example may seem unrealistic because we attribute the same utility of

1 for all patients when receiving a kidney. It can be possible to break this symmetry by accounting for the organ quality and patient compatibility.

Rawlsian justice

In Rawls (1973), John Rawls introduces the concept of the veil of ignorance. Individuals "forget" about everything related to their own situation (i.e. personal biases). Only knowledge about the world remains. By using their rational minds, Rawls argues that individuals can negotiate the terms of a social contract that seems inherently fair. A common way of approaching such as decision is to seek to maximize the utility of the least well-off person. Since there would be no way for a person to know whether or not they will be that person under the veil of ignorance, it would seem a rational solution to adopt.

It is interesting to ask what is the price to optimality for this approach. The price of fairness can be very high for some examples, e.g., see Figure 1 where POF is high when choosing to include vertex i is the solution.

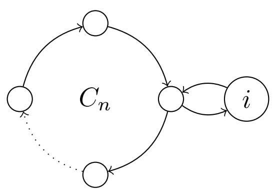

Figure 1: High POF when choosing vertex i

Shapley values

Shapley values come from the context of cooperative game theory (Owen 2013). Before discussing their use in KEPs, we will introduce them.

Definition 5. *For an n-player game and a set function $\nu : 2^N \to \mathbb{R}$, where $N = \{1, \ldots, n\}$, the value $\nu(S)$ of a colation $S \subseteq N$ represents the total utility achieved by the coalition S.*

Shapley's function is the unique (n)-vector valued function ϕ that satisfies the following three properties: efficiency, symmetry and linearity. Efficiency simply means that the sum of the Shapley values for the participants is the total value of the coalition. Symmetry ensures that if two participants contribute the same amount to any coalition when they are added, then their Shapley values will be the same. Finally, linearity is with respect to the functions ν, i.e. $\phi(\nu_1 + \nu_2) = \phi(\nu_1) + \phi(\nu_2)$.

Definition 6. *Shapley's function ϕ is defined as*

$$\phi_i(\nu) = \sum_{T \subseteq N | i \in T} \frac{(|T| - 1)!(n - |T|)!}{n!} (\nu(T) - \nu(T \setminus \{i\}))$$

Shapley values measure the marginal contribution of a player in a coalition under a random arrival order. The two main issues are related to real KEP instances are the following: 1) the arrival of pairs is not random but pre-determined, 2) Shapley values can be prohibitively expensive to compute. Nevertheless, they have been previously used in the context of multi-country KEPs where the number of countries (i.e. players) is relatively small (Biro et al. 2020).

Takeaway

In the previous sections, we have seen many fairness notions that can be applied to KEPs. The main issue with part of them is that they can lead to cases where optimizing fairness implies a significant utility reduction. What if it were possible to balance these two objectives? In other words, we can look for a solution that does well with respect to the utility measure and also a fairness component. In the next section, we propose such solution.

Methodology

In order to balance utility and fairness, we will use an idea from the field of multi-objective optimization: the *Nash social welfare program* (NSWP). Before we dive into its formulation, we will introduce a few key concepts in multi-objective optimization.

Definition 7. *Consider a vector-valued objective function $f : X \to \mathbb{R}^k$. A solution vector x (Pareto) dominates another solution y (for a maximization problem) if*

1. $f_i(x) \geq f_i(y) \quad \forall i \in \{1, \ldots, k\}$
2. $\exists j \in \{1, \ldots, k\}$ *s.t.* $f_j(x) > f_j(y)$.

We denote this dominance as $x \succ y$.

Definition 8. *A Pareto optimal solution is a solution x that is not dominated by any other solution y. In other words $\nexists y$ such that $y \succ x$.*

In multi-objective optimization, one seeks to find a solution from the set of Pareto optimal solutions.

Definition 9. *The set of Pareto optimal solutions is called the Pareto frontier (or Pareto front, Pareto set). It can also be written as*

$$P^* = \{y : \nexists z \text{ s.t. } z \succ y\}.$$

From the Pareto frontier, we have two special vectors that are given the names of ideal and nadir vectors.

Definition 10. *The ideal vector z^{ideal} is defined as*

$$z^{ideal} = \sup_{x \in P^*} \{f_i(x)\} \quad \forall i \in \{1, \ldots, k\}.$$

Definition 11. *The nadir vector z^{nadir} is defined as*

$$z^{nadir} = \inf_{x \in P^*} \{f_i(x)\} \quad \forall i \in \{1, \ldots, k\}.$$

We will also use the concep of **price of utility**. Its definition is similar to POF but instead of having a set U of feasible utility scores for each person, we have a set of feasible fairness scores F for each person.

Definition 12. *The price of utility function POU for a fairness set F and a utility-maximizing scheme $\mathcal{U}$ is defined as*

$$POU(F) = \frac{\sup_{\phi \in F} e^T \phi - e^T \mathcal{U}(F)}{\sup_{\phi \in F} e^T \phi}$$

where $\mathcal{U}$ is a utility-maximizing scheme given a set of feasible fairness scores F.

$\|P\|$	16	32	64	128	256
POF	0.34 ± 0.14	0.33 ± 0.09	0.36 ± 0.06	0.34 ± 0.07	0.42 ± 0.02
POU	0.43 ± 0.15	0.51 ± 0.07	0.56 ± 0.06	0.58 ± 0.04	0.63 ± 0.01

Table 1: Efficiency of the NSWP compared to the utilitarian approach

This definition essentially corresponds to the cost to fairness that one pays by maximizing utility. We can now proceed to discuss the NSWP and its application to fairness in KEPs.

The Nash social welfare program perspective

The Nash social welfare program (Charkhgarda et al. 2020) is mathematically represented as

$$\begin{aligned} \max \prod_{i=1}^{k} (f_i(x) - d_i)^{w_i} & \\ \text{s.t.} \quad x \in X & \\ f_i(x) \geq d_i \qquad & \forall i \in \{1, \dots, k\}, \end{aligned}$$

where d is a reference vector in $\mathbb{R}^k$ that is chosen beforehand. Likewise, the w_i are pre-determined weights that are applied to each objective. The NSWP has two key properties: it is both invariant to the scaling of the w_i's and the objectives f_i's.

If we restrict ourselves to two objectives and unitary weights (i.e. $w_1 = w_2 = 1$), we can apply the NSWP to KEPs. Instead of dealing with deterministic exchange plans, we will focus on distributions over feasible exchanges given a particular graph. We can think of the first objective (i.e. f_1) as the utilitarian component. We seek to maximize the total number of transplants in an exchange plan. The second objective can be a fairness metric such as L_1 or L_2, where the goal is to determine a selection strategy among distributions over feasible exchanges such that it minimizes the Lp-norm mean deviation of each participants (Farnadi et al. 2021). In the other word, this fairness notion (also referred to as individual fairness) balances the chance of each participant to be part of a selected exchange plan.

In what follows, we will describe the procedure with L_1. The NSWP takes the form

$$\begin{aligned} \max \left(\sum_{s \in \mathcal{S}} \delta_s \sum_{c \in \mathcal{C}} w_c s_c \right) \times & \qquad \text{(P)} \\ \left(- \sum_{i \in P} \left| \sum_{s \in \mathcal{S} | i \in s} \delta_s - \frac{1}{|P|} \sum_{i \in P} \sum_{s \in \mathcal{S} | i \in s} \delta_s \right| - d_2 \right) & \\ \text{s.t.} \quad \sum_{s \in \mathcal{S}} \delta_s = 1 & \\ \delta_s \geq 0 \quad \forall s \in \mathcal{S}, & \end{aligned}$$

where $\mathcal{S}$ is the set of feasible exchange plans. The variables δ_s correspond to the probability of having solution s under the distribution δ. A solution to (P) will be expected to balance both objectives *efficiently*.

Evaluation

In order to evaluate the NSWP method on KEPs, we compared it against the typical utilitarian method. We performed tests on graphs sizes 16, 32, 64, 128, and 256 (no altruistic donors). The instances were generated using the simulator found in (Dickerson et al. 2016).

From Table 1, it is possible to observe that the price of fairness for the NSWP approach is always lower than the price of utility for the utilitarian approach. In other words, by using the NSWP, we are able to perform less transplants in a selected exchange plan, but the fairness score is significantly higher than that of the utilitarian approach. The POU here is computed with respect to the NSWP, i.e. we compute the value of the fairness objective (f_2) of the utilitarian approach and compare it the same value for the NSWP. It is also worth to point out that even though the NSWP pays a price in terms of the number of transplantations, the POF is significantly lower than 50%. The NSWP effectively manages to find a middle-ground between the maximum fairness score and maximum number of transplants.

Conclusion

We have shown that the NSWP can prove to be an efficient tool in balancing fairness and utility. By not paying too high a price to utility, we can dramatically improve the fairness score of our selected exchange plans.

Further analysis needs to be done where we investigate the effect of choosing different reference points. Particularly, careful considerations must be given to the use of a second objective involving fairness. We need to ensure that the fairness component really captures the trade-off between sacrificing maximal utility to include more pairs in an exchange plan. In the current version of the NSWP, this might not always be the case. We will explore this idea in future research in order to place the NSWP on a stronger foothold when being applied to KEPs. The same thought process can be applied to using different weights w_i. These values could also be learned simultaneously, effectively combining learning and optimization.

Acknowledgements

Funding support for project activities has been provided by the Institut de valorisation des données and Fonds de Recherche du Québec through the FRQ–IVADO Research Chair in Data Science for Combinatorial Game Theory, Canada CIFAR AI Chair, Google scholar award, and NSERC Discovery Grants program.

References

Abraham, D. J.; Blum, A.; and Sandholm, T. 2007. Clearing Algorithms for Barter Exchange Markets: Enabling Nationwide Kidney Exchanges. In *Proceedings of the 8th ACM Conference on Electronic Commerce*, EC '07, 295–304. New York, NY, USA: ACM. ISBN 978-1-59593-653-0.

Bertsimas, D.; Farias, V.; and Trichakis, N. 2011. The Price of Fairness. *Operations Research*, 59: 17–31.

Biro, P.; Gyetvai, M.; Klimentova, X.; Pedroso, J. P.; Pettersson, W.; and Viana, A. 2020. Compensation Scheme With Shapley Value For Multi-Country Kidney Exchange Programmes. In *Proceedings of the 34th International ECMS Conference on Modelling and SimulationECMS 2020*, 129–136.

Charkhgarda, H.; Keshanianb, K.; Esmaeilbeigic, R.; and Charkhgardc, P. 2020. The Magic of Nash Social Welfare in Optimization: Do Not Sum, Just Multiply! *Optimization Online*.

Dickerson, J. P.; Manlove, D. F.; Plaut, B.; Sandholm, T.; and Trimble, J. 2016. Position-Indexed Formulations for Kidney Exchange. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, EC '16, 25–42. New York, NY, USA: Association for Computing Machinery. ISBN 9781450339360.

Dickerson, J. P.; Procaccia, A. D.; and Sandholm, T. 2014. Price of fairness in kidney exchange. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, 1013–1020. International Foundation for Autonomous Agents and Multiagent Systems.

Farnadi, G.; St-Arnaud, W.; Babaki, B.; and Carvalho, M. 2021. Individual Fairness in Kidney Exchange Programs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, 11496–11505.

Kidney Foundation. 2020. Facing the facts. https://kidney.ca/KFOC/media/images/PDFs/Facing-the-Facts-2020.pdf. Accessed: 2021-11-19.

Nash, J. F. 1950. The Bargaining Problem. *Econometrica*, 18(2): 155–162.

Owen, G. 2013. *Game Theory*, volume Fourth edition of *0*. Emerald Group Publishing Limited. ISBN 9781781905081.

Rawls, J. 1973. *A theory of justice*. Oxford University Press.

Roth, A. E.; Sönmez, T.; and Ünver, M. U. 2007. Efficient Kidney Exchange: Coincidence of Wants in Markets with Compatibility-Based Preferences. *American Economic Review*, 97(3): 828–851.

Zafar, M. B.; Valera, I.; Gomez-Rodriguez, M.; and Gummadi, K. 2017. Fairness Constraints: Mechanisms for Fair Classification. In *AISTATS*.