Maximum weight cycle packing in directed graphs, with application to kidney exchange programs

Péter Biró, David F. Manlove, and Romeo Rizzi

1 Department of Computing Science, University of Glasgow, Glasgow G12 8QQ, UK.
Email: {pbiro,davidm}@dcs.gla.ac.uk.
2 Dipartimento di Matematica ed Informatica (DIMI), Università di Udine, Italy.
Email: romeo.rizzi@dimi.uniud.it.

Abstract

Centralised matching programs have been established in several countries to organise kidney exchanges between incompatible patient-donor pairs. At the heart of these programs are algorithms to solve kidney exchange problems, which can be modelled as cycle packing problems in a directed graph, involving cycles of length 2, 3, or even longer. Usually the goal is to maximise the number of transplants, but sometimes the total benefit is maximised by considering the differences between suitable kidneys. These problems correspond to computing cycle packings of maximum size or maximum weight in directed graphs. Here we prove the APX-completeness of the problem of finding a maximum size exchange involving only 2-cycles and 3-cycles. We also present an approximation algorithm and an exact algorithm for the problem of finding a maximum weight exchange involving cycles of bounded length. The exact algorithm has been used to provide optimal solutions to real kidney exchange problems arising from the National Matching Scheme for Paired Donation run by NHS Blood and Transplant, and we describe practical experience based on this collaboration.

Keywords: Living kidney donation; Optimal exchange; APX-completeness; Approximation algorithm; Exact algorithm; NHS Blood and Transplant.

1 Introduction

Given a simple directed graph $D = (V, A)$, an exchange is a permutation $\pi$ of $V$ such that, for each $v \in V$, $\pi(v) \neq v$ implies $(v, \pi(v)) \in A$. If a vertex $v$ is involved in a cycle of length at least 2 in $\pi$, then $v$ is said to be covered. The size of an exchange $\pi$ is the number of vertices covered by $\pi$. In a $\leq k$-way exchange there is no cycle of length more than $k$. Note that a $\leq k$-way exchange is equivalent to a vertex-disjoint packing of directed cycles of length at most $k$ in $D$.

Let MAX SIZE $\leq k$-WAY EXCHANGE denote the problem of finding an $\leq k$-way exchange of maximum size. Furthermore, if $w : A(D) \rightarrow \mathbb{R}_+$ is a weight function on the arcs, then let MAX ARC WEIGHT $\leq k$-WAY EXCHANGE denote the problem of finding a maximum weight $\leq k$-way exchange, where the weight of the exchange is equal to the sum of weights

*Supported by EPSRC grant EP/E011993/1.
†Supported by OTKA grant K69027.
of the arcs in the exchange. Finally, if \( C \) is the set of directed cycles in \( D \), then let \( w_C : C \rightarrow \mathbb{R}_+ \) be a weight function on the cycles of \( D \), and define **max cycle weight** \( \leq k \)-way exchange to be the problem of finding a maximum weight \( \leq k \)-way exchange, where the weight of an exchange is equal to the sum of the weights of the cycles in the exchange.

In the case where we seek an *unbounded exchange* (i.e., there is no upper bound on the lengths of the cycles), we simply omit \( \leq k \)-way from the names of the **max size** \( \leq k \)-way exchange, **max arc weight** \( \leq k \)-way exchange and **max cycle weight** \( \leq k \)-way exchange problems when referring the corresponding problems in this more general context.

The above graph problems are useful in order to model the *kidney exchange problem*, where the vertices in \( V(D) \) correspond to the incompatible patient-donor pairs and we have an arc \((u, v) \in A(D)\) if the kidney of \( u \)'s donor is suitable for \( v \)'s patient. Let us describe the nature of this application in more detail.

Transplantation is the most effective treatment that is currently known for kidney failure. A patient who requires a transplant may have a willing donor but often he/she is unable to donate a kidney to the intended recipient because of immunological incompatibilities. However these incompatible patient-donor pairs may be able to exchange kidneys with other pairs in a similar position. Kidney exchange programs have already been established in several countries, for example the USA [33, 2, 35], the Netherlands [24, 25], South Korea [38, 26, 37], Romania [29, 28] and the UK [34, 22].

In the case of many of the current programs [41, 40, 42, 43, 44], the organisers wish to maximise the number of patients who receive a kidney in the exchange by considering only the suitability of the grafts. Other schemes consider also distinctions between suitable kidneys [34]. In this context, some models [39, 10, 11, 6, 8] require, as their primary optimality property, the stability of the solution under various criteria. In a third approach, an exchange is said to be optimal if the sum of the benefits is maximal. This model was described in [45] for **pairwise exchanges** (i.e., exchanges involving only 2-cycles).

The lengths of the cycles in the exchanges that arise in kidney exchange programs are bounded in practice, since all operations along a cycle have to be performed simultaneously. Most programs allow only pairwise exchanges, however sometimes 3-cycles are also possible, such as in the New England Program [33] and also in the National Matching Scheme for Paired Donation (NMSPD) administered by the Organ Donation and Transplantation Directorate, NHS Blood and Transplant (NHSBT) [34]. Furthermore, the Proposal for a National Paired Donation Program in the US [46] declares as a goal the possibility to construct 3-cycles. Moreover, longer cycles may be considered as well: an exchange along a 6-cycle has already been performed in the US [4]. In this paper, we study the problem of maximising the total score of an exchange in situations where 2-cycles, 3-cycles and longer cycles are also allowed. This particular problem variant was also studied independently by Abraham *et al.* [1].

From the above description it is obvious that if the goal of the kidney exchange program is to maximise the number of patients who receive a kidney in the exchange by allowing cycles of length at most \( k \) only then we get the problem **max size** \( \leq k \)-way exchange. In some schemes scores are used for measuring the expected utility of the particular donations and the total score of a cycle is calculated simply as the sum of the scores of the donations involved in that cycle. If, in this case, the organisers want to find a maximum score solution then we obtain **max arc weight** \( \leq k \)-way exchange. However, the total score of a cycle may be calculated differently. For example, whereas the total score of a 2-cycle may be obtained as the sum of the scores of the two respective donations, it may be appropriate to calculate the total score of a 3-cycle to be less than the sum of the scores
of the three donations involved, since conducting exchanges along a 3-cycle is more risky than along a 2-cycle (see Section 6 for further details). This motivates the problem \textsc{Max Cycle Weight} ≤ k-way exchange.

Our contribution in this paper is the following. In Section 2 we give some preliminary results for the cases that only 2-cycles are allowed and when we have no restriction on the length of the cycles, and we also survey some other related results. Next, in Section 3, we prove that \textsc{Max Size} ≤ k-way exchange is APX-complete for each fixed k ≥ 3. We note that the NP-hardness of these problems was proved by Abraham \textit{et al.} [1]. Then in Section 4, we give a (2 + ε)-approximation algorithm for \textsc{Max Cycle Weight} ≤ k-way exchange. This performance ratio is the best known for this particular problem. The algorithms of Arkin and Hassin [3] and of Berman [5] also achieve this ratio; we claim that our alternative method and analysis is simpler. In Section 5, we present an exact algorithm for \textsc{Max Cycle Weight} ≤ 3-way exchange, leading to a parameterised complexity result for this problem. This algorithm has been implemented and successfully used for solving kidney exchange problem instances arising from the NMSPD; this practical experience is described in Section 6. Finally we describe some future work directions in Section 7.

2 Preliminaries

Here we describe some straightforward results about the cases when either only pairwise exchanges are allowed or we have no restriction on the length of the exchange. We also describe existing results in the literature that are related to the problems that we consider in this paper.

2.1 Pairwise exchanges

It turns out that \textsc{Max Size} ≤ 2-way exchange, \textsc{Max Arc Weight} ≤ 2-way exchange and \textsc{Max Cycle Weight} ≤ 2-way exchange are all solvable in polynomial time. For, if only 2-cycles are allowed in a given exchange, the problem of finding a so-called optimal pairwise exchange becomes a matching problem in an undirected graph $G$ with the same vertex set as $D$. Here, an edge links two vertices if there is a 2-cycle involving the corresponding two pairs in $D$. So $\{u, v\} \in E(G)$ if both $(u, v)$ and $(v, u) \in A(D)$. A maximum size pairwise exchange in $D$ corresponds to a maximum cardinality matching in $G$. The fastest current implementation of Edmonds’ algorithm for maximum matching [15] is due to Micali and Vazirani [32], and has $O(\sqrt{nm})$ complexity, where $n = |V|$ and $m = |E|$. The use of maximum matching in an undirected graph was described as a possible solution strategy for kidney exchange programs in [41]. Similarly, a maximum weight pairwise exchange in $D$ is equivalent to a maximum weight matching in $G$ with weights $w(\{u, v\}) = w(u, v) + w(v, u)$ or $w(\{u, v\}) = w_C(u, v)$, according to whether we are considering arc weights or cycle weights, respectively. The fastest current implementation of Edmonds’ algorithm for maximum weight matching [16] is due to Gabow [18], and has complexity $O(n(m + n \log n))$. A detailed description of this problem model for kidney exchange can be found in [45].

2.2 Unbounded exchanges

The \textsc{Max Size Exchange} and \textsc{Max Arc Weight Exchange} problems are solvable in polynomial time, using a reduction to the maximum weight perfect matching problem in a bipartite graph. This reduction is decribed in full in our technical report [7]; it was also observed by Abraham \textit{et al.} [1], and resembles a well-known reduction due to Tutte (c.f. [27, p.385]). We remark that in some recent papers, general integer programming
(IP) methods were used to solve the above versions of the kidney exchange problem (see [40, 44]).

In the case of cycle weights, however, the problem of finding a maximum weight unbounded exchange becomes NP-hard. To see this, consider MAX SIZE ≤ 3-WAY EXCHANGE, which is NP-hard as mentioned in Section 1. An instance $D$ of this problem may be transformed to an instance of MAX CYCLE WEIGHT EXCHANGE by creating a weight function $w_C$ on the set of directed cycles $C$ in $D$ as follows. Let $C'$ be a cycle in $D$. If $C'$ is a 2-cycle, set $w_C(C') = 2$; if $C'$ is a 3-cycle, set $w_C(C') = 3$; otherwise set $w_C(C') = 0$. Clearly $\pi$ is a maximum size ≤ 3-way exchange in $D$ if and only if $\pi$ is a maximum weight exchange in $D$ according to the cycle weight function $w_C$.

2.3 Other related work

We emphasise that our focus is on digraphs, and thus in this context we are interested only in directed cycles. However in undirected graphs, variants of MAX WEIGHT ≤ 3-WAY EXCHANGE, especially where $k = 3$, have been the subject of much prior work, which we survey in this subsection.

Let $G = (V, E)$ be a simple and undirected graph. A triangle of $G$ is any induced subgraph of $G$ having precisely 3 edges and 3 vertices. A family of triangles $T_1, \ldots, T_l$ of $G$ is called a vertex-packing of triangles if $T_1, \ldots, T_l$ are vertex-disjoint. The size of this packing is $l$. The problem of finding a maximum size vertex-packing of triangles in a given graph $G$, called VERTEX-DISJOINT TRIANGLE PACKING (vDTP), is NP-hard [19, pp. 68-69]. Moreover, vDTP is APX-complete [23], even if $G$ has maximum degree 4. A tighter lower bound for the general case appears in [14].

As far as upper bounds for vDTP are concerned, Hurkens and Schrijver’s $(\frac{3}{2} + \varepsilon)$-approximation for the more general 3-SET-PACKING problem (see Section 4) is still the state of the art, however improvements are possible for bounded degree graphs or other special classes of graphs [9, 30]. A natural generalisation of vDTP occurs when each edge of $G$ has a weight, and we seek a maximum weight triangle packing in $G$. For this problem, recent work has enabled the performance ratio to drop below 2 [20, 13].

3 APX-completeness of MAX SIZE ≤ $k$-WAY EXCHANGE

In this section, we show that MAX SIZE ≤ $k$-WAY EXCHANGE is APX-complete for any constant $k \geq 3$. We begin in the following subsection by giving some definitions and results relating to the problem that we reduce from.

3.1 Preliminaries

We firstly remark that MAX SIZE ≤ $k$-WAY EXCHANGE belongs to APX for each $k \geq 3$, as we will demonstrate in Section 4. We will give an L-reduction [36] that establishes APX-completeness for MAX SIZE ≤ $k$-WAY EXCHANGE for each $k \geq 3$. This reduction starts from vDTP, which was defined in Section 2.3. Kann [23] showed that this problem is APX-complete in the case that $G$ is a 3-partite graph, that is, the vertex set $V$ of $G$ can be partitioned into three disjoint colour classes $V = A \cup B \cup C$, and no edge of $G$ has its two end-vertices in the same colour class. Notice that when $G$ has this special structure, every triangle of $G$ must have precisely one vertex in each one of the three colour classes.
3.2 Inapproximability of MAX SIZE \( \leq 3\)-WAY EXCHANGE

In this subsection we show that MAX SIZE \( \leq 3\)-WAY EXCHANGE is APX-complete and then in the next subsection we generalise the construction to establish the APX-completeness of MAX SIZE \( \leq k\)-WAY EXCHANGE for \( k \geq 3 \). Let \( G = (V, E) \) be a 3-partite graph given as an instance of vDTP. Let \( A, B \) and \( C \) be the 3 colour classes in which \( V \) is partitioned. We construct a digraph \( D = (V, A) \) by simply orienting the edges of \( G \) in a circular way as follows:

\[
A(D) := \{(u,v) \mid uv \in E(G), u \in A, v \in B\} \cup \{(u,v) \mid uv \in E(G), u \in B, v \in C\} \cup \{(u,v) \mid uv \in E(G), u \in C, v \in A\}.
\]

Clearly, the digraph \( D \) can be constructed in polynomial time starting from the graph \( G \). Moreover, the following lemma says that the above is an objective function preserving reduction (a primary case of an \( L \)-reduction), from which the claimed APX-completeness result follows.

**Lemma 1.** The graph \( G \) admits a packing of vertex-disjoint triangles covering \( t \) vertices if and only if the digraph \( D \) admits a \( \leq 3 \)-way exchange covering \( t \) vertices.

**Proof.** Three vertices \( u, v, z \in V \) induce a triangle in \( G \) if and only if they induce a directed cycle with length 3 in \( D \). As a consequence, a packing of triangles in \( G \) can be regarded as a 3-way exchange covering precisely the same set of vertices. In the other direction, notice that \( D \) contains no 2-cycles. Thus, any \( \leq 3 \)-way exchange contains only cycles of length 3 and can hence be regarded as a packing of triangles in \( G \) covering precisely the same set of vertices. \( \square \)

3.3 Inapproximability of MAX SIZE \( \leq k\)-WAY EXCHANGE for \( k \geq 3 \)

The trick to generalise the above reduction to a generic \( k \geq 3 \) is as follows. Once \( D \) has been obtained, we obtain a second digraph \( D' \) from \( D \) as follows. For each arc \( a = (u,v) \) of \( D \) with \( u \in C \) and \( v \in A \), add the vertices \( w_{a,1}, \ldots, w_{a,k-3} \) and replace the arc \( a = (u,v) \) with the arcs \((u, w_{a,1}), (w_{a,k-3}, v)\) and \((w_{a,i}, w_{a,i+1})\) for \( i = 1, \ldots, k - 4 \). To summarise, \( V(D') := V(D) \cup \{w_{a,i} \mid a = (u,v) \in A(D), u \in A, v \in C, i = 1, \ldots, k - 3\} \), and

\[
A(D') := \{(u,v) \mid uv \in E(G), u \in A, v \in B\} \cup \{(u,v) \mid uv \in E(G), u \in B, v \in C\} \cup \{(u, w_{a,1}), (w_{a,k-3}, v) \mid uv \in E(G), u \in C, v \in A\} \cup \{(w_{a,i}, w_{a,i+1}) \mid uv \in E(G), u \in C, v \in A, i = 1, \ldots, k - 3\}.
\]

Clearly, the digraph \( D' \) can be constructed in polynomial time starting from the graph \( G \). Moreover, the following lemma says that the above is an \( L \)-reduction, which implies that MAX SIZE \( \leq k\)-WAY EXCHANGE is APX-complete for any constant \( k \geq 3 \).

**Lemma 2.** The graph \( G \) admits a packing of vertex-disjoint triangles covering \( t \) vertices if and only if the digraph \( D' \) admits a \( \leq k \)-way exchange covering \( tk/3 \) vertices.

**Proof.** Notice that \( D' \) contains no directed cycle of length less then \( k \). Moreover, three vertices \( a \in A, b \in B \) and \( c \in C \) induce a triangle in \( G \) if and only if the arc \( f = (c,a) \) belongs to \( D \) and the vertices \( a, b, c, w_{f,1}, \ldots, w_{f,k-4}, w_{f,k-3} \) induce a directed cycle in \( D' \). As a consequence, a packing of triangles in \( G \) can be regarded as an \( \leq k \)-way exchange in \( D' \) covering precisely \( k/3 \) as many vertices. In the other direction, since \( D' \) contains
no directed cycle of length less then $k$, then any $\leq k$-way exchange contains only cycles of length $k$ and can hence be regarded as a packing of triangles in $G$ covering precisely $3/k$ as many vertices.

Lemmas 1 and 2 imply the following theorem.

**Theorem 3.** MAX SIZE $\leq k$-WAY EXCHANGE is APX-complete for any integer $k \geq 3$.

We remark, that MAX ARC WEIGHT $\leq k$-WAY EXCHANGE and MAX CYCLE WEIGHT $\leq k$-WAY EXCHANGE are also APX-complete, since MAX COVER $\leq k$-WAY EXCHANGE is their special case with unit weights.

## 4 Approximation algorithm for MAX CYCLE WEIGHT $\leq k$-WAY EXCHANGE

In this section we give a $(k - 1 + \varepsilon)$-approximation algorithm for the MAX CYCLE WEIGHT $\leq k$-WAY EXCHANGE problem (and hence the MAX ARC WEIGHT $\leq k$-WAY EXCHANGE problem) for any $\varepsilon > 0$ and for any $k \geq 3$. We begin by showing that MAX CYCLE WEIGHT $\leq k$-WAY EXCHANGE can be reduced to the maximum weight matching problem in a hypergraph. Let $H = (V, E)$ be defined on the same vertex set as $D$. A hyperedge $e_X$ corresponds to a set of vertices $X \subseteq V(D)$ if $|X| \leq k$ and there exists a directed cycle on $X$ that covers precisely the vertices in $X$. Let the weight of $e_X$ in $H$, denoted by $w_H(e_X)$, be equal to the weight of a maximum weight cycle in $D$ that can be formed on $X$. Obviously, there is a one-to-one correspondence between the $\leq k$-way exchanges in $D$ and the matchings in $H$. Moreover, the weights of the maximum weight solutions are equal, so that MAX WEIGHT $\leq k$-SET PACKING, defined as follows, is a generalisation. Given a hypergraph $H = (V, E)$ where every hyperedge has size at most $k$ and has a non-negative weight $w_H(e) \in \mathbb{R}_+$ associated to each hyperedge $e$, define MAX WEIGHT $\leq k$-SET PACKING to be the problem of finding a maximum weight matching $M$ of $H$. \(^1\)

The MAX WEIGHT $\leq k$-SET PACKING problem can be further reduced to the MAX WEIGHT INDEPENDENT SET problem (the problem of finding a maximum weight independent set) in $(k + 1)$-claw free graphs\(^2\), using the intersection graph of the hypergraph. In this simple graph, denoted by $L(H)$, the vertices of $L(H)$ are the edges of $H$, and two vertices of $L(H)$ are adjacent if the corresponding edges of $H$ intersect. The fact that $H$ contains only edges with size at most $k$ implies that $L(H)$ is a $(k + 1)$-claw free graph. We define a weight function on the vertices of $L(H)$ in a natural way: given an edge $X$ of $H$, let $v_X$ denote the corresponding vertex in $L(H)$. The weight of $v_X$ in $L(H)$, denoted by $w_L(v_X)$, is equal to $w_H(X)$. Obviously, a maximum weight independent set in $L(H)$ corresponds to a maximum weight matching in $H$ (which in turn corresponds to a maximum weight exchange in $D$). Our goal is to approximate the MAX WEIGHT INDEPENDENT SET problem in $(k + 1)$-claw free graphs.

### 4.1 Local search method

Let $I$ be an independent set in $L(H)$. A natural idea to improve a sub-optimal solution is the $t$-local search technique (see [3]). In this method, we attempt to add an independent set $X$ to $I$ with cardinality at most $t$ and remove the subset of $I$ that is in $N(X)$, so that

\(^1\)The 3-SET PACKING problem referred to in Section 2.3 is the variant of MAX WEIGHT $\leq k$-SET PACKING in which every hyperedge in $H$ has size exactly 3 and weight 1.

\(^2\)A graph is $(k + 1)$-claw-free if it does not contain $K_{1,k+1}$ as an induced subgraph.
the total weight increases. If no such t-local improvement exists then the solution is a t-local optimum.

Note, that if we compare two disjoint independent sets, say $I$ and $I_{opt}$, then these sets can be viewed as the two sides of a bipartite subgraph of $L(H)$. Moreover, each vertex in this subgraph has degree at most $k$, by the $(k+1)$-claw freeness of $L(H)$. That is why the conditions of the following theorem can describe the relation of a t-local optimum $I$ and a global optimum $I_{opt}$.

**Theorem 4 ([3]).** For any given $k$ and $t$ and every instance $G = (A, B, E)$ satisfying the following three conditions:

- $|N(a)| \leq k$ for each $a \in A$;
- $|N(b)| \leq k$ for each $b \in B$;
- any subset $X \subseteq A$ of at most $t$ vertices satisfies $w(X) \leq w(N(X))$;

we have

$$\frac{w(A)}{w(B)} \leq k - 1 + \frac{1}{t}.$$

Applying the above theorem to the case where $A = I_{opt}$ and $B = I$ thus implies that a t-local optimum approximates the global optimum within $k - 1 + \frac{1}{t}$ for any $t$. It follows that an algorithm based on t-local search also approximates the max cycle weight $\leq k$-way exchange problem within a factor of $k - 1 + \varepsilon$ for any $\varepsilon > 0$.

### 4.2 Local search via augmenting paths

Here we show, that the same performance ratio can be reached by using only a particular t-local search. The approach is straightforward and the same performance ratio as stated in the previous subsection for general local search can be established using a relatively simple argument (the proof in [3] runs to several pages).

The t-augmenting path search is a specific type of t-local search, where the new set $X$ is chosen along an alternating path in $L(H)$. Formally, let $X = x_1, x_2, \ldots, x_s$, where $s \leq t$ and there is a subset $Y$ of the actual solution $I$, such that $|Y| = s - 1$ and $X = x_1, y_1, x_2, y_2, \ldots, x_{s-1}, y_{s-1}, x_s$ is a path in the intersection graph, $L(H)$.

**Theorem 5.** For any given $k$ and $t$ and every instance $G = (A, B, E)$ satisfying the following three conditions:

- $|N(a)| \leq k$ for each $a \in A$;
- $|N(b)| \leq k$ for each $b \in B$;
- any subset $X \subseteq A$ of at most $t$ vertices, where $X$ is a set of alternate vertices in an alternating path of $G$, satisfies $w(X) \leq w(N(X))$;

we have

$$\frac{w(A)}{w(B)} \leq k - 1 + \frac{2}{t}.$$

This theorem implies an alternative proof for the existence of an approximation algorithm – that use “only” augmenting path searches – with factor $k - 1 + \varepsilon$ for any $\varepsilon > 0$.

We use the following well-known lemma in the proof of Theorem 5:

**Lemma 6.** If $G = (A, B, E)$ is a $k$-regular bipartite graph, then the set of edges, $E(G)$ can be partitioned into $k$ disjoint perfect matchings.
Proof. (of Theorem 5) We complete $G$ into a $k$-regular graph $G'$ by adding dummy vertices with zero weight and some edges. By Lemma 6 we can partition the set of edges into $k$ disjoint perfect matchings. Two of these perfect matchings form a 2-factor, that is a perfect covering by a set of disjoint alternating cycles. Let us fix such a 2-factor $C$.

The main idea of the proof is that the cycles in $C$ of length exceeding $2t$ can be cut into alternating paths of length at most $2$ such that the total weight of the end-vertices of these paths in $B$ is at most $\frac{2}{t}w(B)$. We can achieve this in the following way. Consider a cycle $C_i = (X_i|Y_i) = (x_0^i, y_0^i, x_1^i, \ldots, x_{c_i}^i, y_{c_i}^i)$ from $C$, with $|X_i| = |Y_i| = c > t$. We show that we can always find a set of vertices $R_i \subseteq Y_i$ (the so-called cut vertices of $Y_i$) with cardinality $\lceil \frac{c}{t} \rceil$, that satisfies the following properties:

\[ w(R_i) \leq w(Y_i) \]

(Here, property (1) ensures that between two consecutive vertices in $R_i$ the distance is at most $2t$. Property (2) says that the average weight of the vertices in $R_i$ is less than or equal to the average weight of the vertices in $Y_i$.)

The proof of the existence is straightforward: first we choose a set of vertices from $Y_i$ such that its cardinality satisfies property (1), then we rotate this set along the cycle $c$ times (increasing every index one by one) and we select the set of vertices of minimum total weight to be $R_i$.

Using property (2) and $c > t$, we obtain

\[ w(R_i) \leq |R_i| \leq w(Y_i) = \frac{c}{t}w(Y_i) < \frac{2}{t}w(Y_i). \]

If we consider the subset of cycles in $C_i \in C$ of length more than $2t$, then for $R = \bigcup R_i$ we get

\[ w(R) \leq \sum_{i : |X_i| > t} w(R_i) < \sum_{i : |X_i| > t} \frac{2}{t}w(Y_i) \leq \frac{2}{t}w(B). \]

Now, we create a partition of $A$ the following way. First we remove $R$ together with the incident edges from $C$, and then we also remove the additional vertices and edges (i.e. $G' \setminus G$). The remaining graph, denoted by $\mathcal{P}$, consists of disjoint paths and cycles, where the cycles have length at most $2t$ and each path contains at most $t$ vertices from $A$. Now, we form partition $A = X_1 \cup X_2 \cup \ldots \cup X_p$ such that each set $X_i \subseteq A$ consists of vertices that are in the same component of $\mathcal{P}$. Therefore, each set $X_i$ is a set of consecutive vertices in some alternating path of $G$.

To show that the set of inequalities $w(X_i) \leq w(N(X_i))$, for $1 \leq i \leq p$, imply the statement of the theorem, first we note that any vertex $b \in B$ belongs to at most $k$ sets of the form $N(X_i)$, for $1 \leq i \leq p$. Moreover, if $b \in B \setminus R$ then $b$ belongs to at most $k - 1$ sets of the form $N(X_i)$, for $1 \leq i \leq p$, since either $b$ has at most one neighbour in $\mathcal{P}$ which implies $|N(b)| \leq k - 1$, or $b$ has two neighbours in $\mathcal{P}$ which means that these neighbours are in the same class $X_i$ of $A$. Therefore, by summing up these inequalities and using (3) (since only the vertices of $R$ can be counted $k$ times, from each of their $k$ neighbours), we obtain:

\[ w(A) = \sum_{X_i \in A} w(X) \leq \sum_{X_i \in A} w(N(X)) \]
\[(k - 1)w(B \setminus R) + kw(R) \leq (k - 1)w(B) + w(R)\]
\[< \left( k - 1 + \frac{2}{t} \right) w(B). \]

The proof is complete. \qed

### 4.3 Other methods

Chandra and Halldórsson [12] gave a \((2(k + 1)/3 + \varepsilon)\)-approximation algorithm for the Max Weight Independent Set problem in \((k + 1)\)-claw free graphs by combining a local search method with the greedy algorithm, which we now describe. They start with an independent set \(I\) obtained by the greedy algorithm (this algorithm repeatedly chooses a vertex of maximum weight, deleting both it and its neighbours from the graph). Then \(I\) is improved by a special type of local search, where the new additional set \(X\) is chosen from the neighbours of a vertex from \(I\), in such a way that the ratio of the weights of \(X\) and the removed set, \(N(X) \cap I\), is always maximal.

Finally, the best approximation known so far for the Max Weight Independent Set problem in \((k + 1)\)-claw free graphs was given by Berman [5]. He showed that an \(((k + 1)/2 + \varepsilon)\)-approximation is possible by a polynomial time algorithm that is based on local improvement by considering the squared total weights. This provides the same performance ratio for \(k = 3\) as our local search algorithm via augmenting path method, but has better performance for larger \(k\). Again, the advantage of our method for the case that \(k = 3\) is that the derivation of the performance ratio is much simpler as compared to [5].

### 4.4 Approximability of MAX SIZE \(\leq k\)-WAY EXCHANGE

For the approximability of Max Size \(\leq k\)-Way Exchange, a general result of [21] leads to a polynomial time \((k/2 + \varepsilon)\)-approximation algorithm (for any \(\varepsilon > 0\)) for any fixed \(k\). In the special case of \(k = 3\), this gives a \((3/2 + \varepsilon)\)-approximation algorithm. An approximation algorithm improving this ratio could be directly translated into an algorithm for VDTP that improves on the best known performance ratio for this problem.

### 5 Exact algorithm and parameterised complexity

In this section we give an exact algorithm for Max Cycle Weight \(\leq 3\)-Way Exchange, leading to a parameterised complexity result for this problem.

As described in [42, 44], the Max Size \(\leq 3\)-Way Exchange problem is currently solved in the New England Program for Kidney Exchange by IP-based methods. Recently, Abraham et al. [1] implemented a specialised IP-heuristic for the Max Cycle Weight \(\leq 3\)-Way Exchange problem that would be capable of handling the data of a future national kidney exchange program in the USA (for up to approximately 10,000 couples) according to their simulations. Despite these excellent empirical results, it is still an interesting question to construct an exact combinatorial algorithm for this problem, and we describe such a method in this section. One source of motivation for this is that the IP-technique does not give any guarantee for the running time in a theoretical sense. Another motivating factor is that our alternative technique may also be used as a heuristic in conjunction with other methods.

To describe our method, suppose that we are given an instance of Max Cycle Weight \(\leq 3\)-Way Exchange in a digraph \(D\), and denote by \(\pi^*\) an optimal solution. Let \(\{C_1, C_2, \ldots, C_l\}\) be the 3-cycles in \(\pi^*\) and let \(Y\) be a set of arcs \(\{a_1, a_2, \ldots a_l\}\) where \(a_i \in C_i\) (\(1 \leq i \leq l\)
so that $Y$ contains one arc from each 3-cycle. We show that with this knowledge of $M$, we can efficiently find a maximum weight 3-way exchange (i.e. either $\pi^*$ or another optimal solution).

We transform our instance of Max Cycle Weight $\leq 3$-Way Exchange to a maximum weight matching problem in an undirected graph $G_Y$ in the following way. We denote by $V(Y)$ the set of vertices in $D$ that are covered by $Y$. Let $y_{i,j} \in V(G_Y)$ if $(v_i, v_j) \in Y$, otherwise let $x_i \in V(G_Y)$ if $v_i \in V(D) \setminus V(Y)$. Let $\{x_i, x_j\} \in E(G_Y)$ if both $(v_i, v_j)$ and $(v_j, v_i) \in A(D)$, and let $\{x_k, y_{i,j}\} \in E(G_Y)$ if both $(v_i, v_j)$ and $(v_j, v_k) \in A(D)$. Considering the weights, $w'(\{x_i, x_j\}) := w(v_i, v_j) + w(v_j, v_i)$ and $w'(\{x_k, y_{i,j}\}) := w(v_k, v_i) + w(v_i, v_j) + w(v_j, v_k)$ for the Max Arc Weight $\leq 3$-Way Exchange problem, and $w'(\{x_i, x_j\}) := w_C(v_i, v_j)$ and $w'(\{x_k, y_{i,j}\}) := w_C(v_k, v_i, v_j)$ for the Max Cycle Weight $\leq 3$-Way Exchange problem.

Obviously, a matching $M$ in $G_Y$ corresponds to a $\leq 3$-way exchange $\pi$ in $D$, in such a way that $\{x_i, x_j\} \in M$ if and only if $v_i \in Y$ and $v_j \in Y$ or another optimal $\pi^*$, then we can always find a suitable subset $Y_S = \{a_1, a_2, \ldots, a_l\}$ from $S$, as required. Therefore, if $s = |S|$, then the number of guesses is at most $2^s$. (Moreover, since this subset $Y_S$ must contain independent arcs, we need to check only the matchings of $S$.)

Let us choose $S$ to be of minimum size such that $S$ covers at least one arc from each 3-cycle of $D$, and consider $s = |S|$ as a parameter. Using the $O(n(m + n \log n))$ time maximum weight matching algorithm (see [18]) as a subroutine, we obtain the following parameterised complexity of Max Cycle Weight $\leq 3$-Way Exchange (in the statement of the following theorem, we use $O^*(c^k)$ to refer to $O(\Theta^k f(n))$ in relation to the parameterised complexity of a given problem, where $c$ is a constant, $n$ is the input size and $f(n)$ is a polynomial in $n$ [17, p. 12].)

**Theorem 7.** Max Cycle Weight $\leq 3$-Way Exchange can be solved in $O^*(2^s)$ time, where $s$ is the minimum number of arcs that cover at least one arc from each 3-cycle of $D$.

Finding a set $S$ of minimum size satisfying the conditions described above may well be an NP-hard problem in itself. However, we can give an $O^*(3^s)$ parameterised algorithm for this, where $s = |S|$, as follows. First we take a 3-cycle randomly. At the branching stage we select one arc from the three, adding this arc to $S$ and removing it from the graph. Then we continue this brute-force search until no 3-cycle remains. It is easy to see that $s$ will contain at least one arc from each 3-cycle. Furthermore, the depth of the search tree is $s$, and the detection of a 3-cycle can be carried out in $O(n + m)$ time, therefore we obtain the $O^*(3^s)$ running time.

Finally, we note that $s \leq \frac{m}{2}$ always holds, since every directed graph becomes acyclic.
by removing at most half of its arcs. As a result, the method described prior to Theorem 7 gives an $O(2^{\frac{m}{2}})$ time exact algorithm for MAX CYCLE WEIGHT $\leq$ 3-WAY EXCHANGE.

6 Practical experience: NHS Blood and Transplant

As part of their administration of the NMSPD, NHSBT maintain a database of incompatible patient-donor pairs who would be willing to participate in a live-donor kidney exchange with one or more other patient-donor pairs. At regular intervals (every three months at the time of writing), a matching run is carried out in which an optimal exchange is constructed from the dataset. At present, exchanges involving only 2-cycles and 3-cycles are sought, though (again at the time of writing), an exchange along a cycle with more than two couples has yet to be carried out in the UK. The term optimal refers to the fact that the overriding constraint is to maximise the number of transplants that can be carried out, and subject to this, to maximise the overall score of the exchange. The score of an exchange is based on the points system that NHSBT employs for couples involved in the process – see their web page [34] for further details. This optimisation problem can be reduced to a maximum weight exchange problem by increasing the weight of each arc by an extra weight that is greater than the maximum total weight of any exchange.

We implemented our exact algorithm as described in Section 5 for computing an optimal $\leq$ 3-way exchange in C++, using a LEDA implementation of an algorithm for the maximum weight matching problem [31, p. 443].

We tested this implementation on random samples from a pool of 392 incompatible patient-donor pairs whose data was collected by NHSBT. In particular, we took 10 random samples of $n$ patient-donor pairs from this pool, for $n$ in the range 30 to 50 (in intervals of 5). For each sample, we found an optimal pairwise exchange, an optimal $\leq$ 3-way exchange, and an optimal unbounded exchange. Details of the results of these trials can be found in [7].

We also used our exact algorithm to find optimal exchanges for NHSBT for the quarterly matching runs of the NMSPD from April 2008 to October 2009 inclusive, and the results corresponding to these input datasets are contained in Table 1. The column corresponding to a given matching run shows the following data. Rows 3-6 indicate the number of vertices (patient-donor pairs) in $D$, the number of arcs in $D$, and the number of 2-cycles and 3-cycles in $D$. Rows 7-9 indicate the number of 2-cycles in an optimal pairwise exchange, together with the size and weight of such an exchange. Rows 10-13 show the number of 2-cycles and 3-cycles in an optimal $\leq$ 3-way exchange, together with the size and weight of such an exchange. The size of $S$ (as described in Section 5), together with the number of subsets $Y$ of $S$ involved in computing an optimal $\leq$ 3-way exchange are shown in rows 14-15, together with the time taken by the algorithm in row 16. Rows 17-19 show the size and weight of an optimal unbounded exchange, together with the length of the longest cycle that was computed in such an exchange. The final four rows indicate the exchange that was ultimately chosen by NHSBT – this will be described in more detail below. All entries in Table 1 are rounded to the nearest integer, except for the running times which are rounded to 1 decimal place.

As can be seen from the table, the total number of 2-cycles and 3-cycles in each of the digraphs corresponding to the April and July 2008 instances were small, and in each case NHSBT were able to find an optimal exchange, using our exact algorithm to verify the optimality of the solution (this was essentially trivial in the case of the July 2008 dataset).

\footnote{To find a set $S$, as described in Section 5, we used the following heuristic. We selected the arcs to be added to $S$ one-by-one, by choosing the arc involved in the largest number of 3-cycles and removing it from the graph, until no 3-cycle remained.}
<table>
<thead>
<tr>
<th>Matching run</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Apr</td>
<td>Jul</td>
</tr>
<tr>
<td># pairs</td>
<td>76</td>
<td>85</td>
</tr>
<tr>
<td># possible donations</td>
<td>287</td>
<td>235</td>
</tr>
<tr>
<td>Total #</td>
<td>76</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>2-cycles</td>
<td>5</td>
</tr>
<tr>
<td># 3-cycles</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Pairwise exchanges</td>
<td>#2-cycles</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>size</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>weight</td>
<td>91</td>
</tr>
<tr>
<td>≤3-way exchanges</td>
<td>#2-cycles</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>#3-cycles</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>size</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>weight</td>
<td>620</td>
</tr>
<tr>
<td>the exact algorithm</td>
<td>size of $S$</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td># $Y \subseteq S$</td>
<td>24</td>
</tr>
<tr>
<td>Running time (sec)</td>
<td>0.3</td>
<td>0.0</td>
</tr>
<tr>
<td>Unbounded exchanges</td>
<td>size</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>weight</td>
<td>857</td>
</tr>
<tr>
<td></td>
<td>longest c.</td>
<td>20</td>
</tr>
<tr>
<td>Chosen solution (NHSBT)</td>
<td>#2-cycles</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>#3-cycles</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>size</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>weight</td>
<td>620</td>
</tr>
</tbody>
</table>

Table 1: Results arising from matching runs from April 2008 to October 2009.

However the digraphs corresponding to the October 2008, January 2009, April 2009 and July 2009 datasets were much richer in terms of the numbers of 2-cycles and 3-cycles. A consequence of this was that the problem instances were much harder to solve by hand, and therefore the results were provided by the exact algorithm in each of these two cases. By contrast, and perhaps surprisingly, the digraph corresponding to the October 2009 dataset turned out to be very sparse.

In general the real instances were sparser than the random samples described above, because of the large number of highly sensitised patients in the actual pools drawn from the quarterly matching runs. As a consequence, the size of $S$ remained relatively small, so that fewer subsets $Y$ of $S$ needed to be considered, and therefore the optimisation problems in each of the cases of the October 2008, January 2009 and April 2009 data were comfortably tractable by our exact algorithm. On the other hand the July 2009 dataset did cause the algorithm to take somewhat longer (around 25 minutes).

We finally note that, in the cases of datasets between October 2008 and October 2009, the actual solutions chosen by NHSBT (shown in the last four rows in Table 1) were different from the optimal (according to the original criteria) ≤3-way exchanges. This stemmed from the fact that none of the four 3-cycles identified in April 2008 led to transplants, for a range of reasons (e.g., a positive crossmatch being discovered late in the process following more detailed tests, a patient and/or a donor becoming ill, a patient deciding to proceed with an antibody incompatible transplantation, etc. – for further details about this, see [22]). Due to the risk involved with 3-cycles, our exact algorithm was already computing optimal ≤ 3-way exchanges subject to the additional constraint.
that the exchange could involve at most $k$ 3-cycles, for $0 \leq k \leq k'$, where $k'$ was the number of 3-cycles in an optimal $\leq 3$-way exchange without this constraint. As such, we were able to compute an optimal $\leq 3$-way exchange that contained the minimum number of 3-cycles.

However from October 2008, NHSBT decided that this was not sufficient, and changed their policy as follows. What is now sought is an optimal $\leq 3$-way exchange $\pi$ with the additional constraint that $\pi$ contains $k$ 2-cycles, where $k$ is the number of 2-cycles in an optimal pairwise exchange. That is, 3-cycles are permitted so long as they do not lead to a reduction in the maximum possible number of 2-cycles. Finding $\pi$ is not, however, simply a matter of adding 3-cycles to an optimal pairwise exchange, since NHSBT permit the total weight of the 2-cycles involved in an optimal pairwise exchange to decrease in order to accommodate the additional 3-cycles. Bearing this in mind, we were able to compute the alternative solutions shown under “chosen solutions” in Table 1 corresponding to the October 2008, January 2009 and April 2009 datasets using our exact algorithm. In July 2009 and October 2009, our exact algorithm constructed solutions that were a starting point. However NHSBT then extended these solutions by trying to augment a pairwise exchange of the form $(P_1, P_2)$ to 3-way exchanges of the form $(P_1, P_2, P_3)$. The intuition behind this was that if $P_3$ subsequently dropped out then the original pairwise exchange could still be possible, whereas if either $P_1$ or $P_2$ dropped out then the original pairwise exchange would have been failed anyway. Thus a 3-way exchange of this type was seen as carrying little additional risk over and above the original pairwise exchange.

7 Future work

Finding improved approximation algorithms, for each variation of the problems we studied in this paper, remains an important theoretical challenge. It may be worth trying to tackle these problems directly rather than transforming them to the maximum weight independent set problem in $(k + 1)$-claw free graphs.

It would be interesting to see whether our graph-based exact algorithm could be speeded up by additional techniques. One possible approach might be to use a special branch and bound method as follows. Given a subset $Y$ of $S$ and a current optimum found so far during the search, let us find a maximum weight exchange such that each arc of $Y$ is involved in an unbounded exchange. If this relaxed optimum is less than or equal to the current optimum then we can cut the search, since we do not need to test any superset $Y' \supseteq Y$. Furthermore, we would like to compare the running time of our exact algorithm with that of other methods, e.g. the IP-heuristic implemented by Abraham et al. [1]. This was shown to be very powerful for generated instances but might be less successful for real data combined with specific optimisation criteria such finding a maximum weight maximum cardinality matching.

Regarding the practical application, one of the current challenges is to incorporate the possibility of altruistic chains. Such a chain starts with an altruistic donor who donates her kidney to a patient $p$, whilst $p$’s incompatible donor continues the chain, with the final kidney being donated to the deceased donor list. Basically, we can model this extension using our existing techniques by introducing a dummy patient $p'$ for each altruistic donor, such that $p'$ is compatible with every donor in the pool. However, when considering the optimisation criteria, we may need to distinguish between an altruistic chain and a normal cycle of the same length, given that the risk and benefit in each case may differ (for further description about altruistic chains, see e.g., [44]). Another possible change in the application could be the introduction of 4-cycles, which is a long-term plan of NHSBT, giving rise to further algorithmic challenges.
Acknowledgements

We would like to thank Rachel Johnson and Joanne Allen of the Organ Donation and Transplantation Directorate, NHS Blood and Transplant, for helpful discussions and for the ongoing collaboration. We would also like to thank Kirstin MacDonald for useful observations concerning this paper. Further, we would like to thank Ildikó Schlotter for valuable remarks concerning the parameterised complexity of the exact algorithm in Section 5. Finally, we would like to thank an anonymous referee for helpful comments on an earlier version of this paper.

References


