Transformation of Non-feasible Inverse Maximum Flow Problem into a Feasible one by Flow Modification

ADRIAN DEACONU  
Transilvania Univ. of Brașov  
Theoretical Computer Science Dept.  
Brașov, Iuliu Maniu 50, cod 500091  
ROMANIA  
a.deaconu@unitbv.ro

ELEONOR CIUREA  
Transilvania Univ. of Brașov  
Theoretical Computer Science Dept.  
Brașov, Iuliu Maniu 50, cod 500091  
ROMANIA  
e.ciurea@unitbv.ro

CORNELIU MARINESCU  
Transilvania University of Brașov  
Electrical Eng. and Computer Sci.  
Brașov, Eroilor 29, 500036  
ROMANIA  
corneliu.marinescu@unitbv.ro

Abstract: A linear time method to decide if any inverse maximum flow (denoted IMFG) problem has solution is presented. For the case of IMFG not being feasible, a new inverse combinatorial optimization problem is introduced and solved. The problem is to modify as little as possible the flow so that the problem becomes feasible for the modified flow.

Key–Words: inverse combinatorial optimization, maximum flow.

1 Introduction

An inverse combinatorial optimization problem consists in modifying some parameters of a network such as capacities or costs so that a given feasible solution of the direct optimization problem becomes an optimal solution and the distance between the initial vector and the modified vector of parameters is minimum. Different norms such as $l_1$, $l_\infty$ and even $l_2$ are considered to measure this distance. In the last years many papers were published in the field of inverse combinatorial optimization [7]. Almost every inverse problem was studied considering $l_1$ and $l_\infty$ norms, resulting in different problems with completely different solution methods. Strongly polynomial time algorithms to solve the inverse maximum flow problem when $l_1$ norm is considered (denoted IMF) were presented by Yang, Zhang and Ma [12]. IMF is reduced to a minimum cut problem in an auxiliary network with finite and infinite arc capacities. The algorithm for IMF has an $O(n \cdot m \cdot \log(n^2/m))$ time complexity, where $m$ is the number of arcs and $n$ is the number of nodes.

The most general case (denoted GIMF) under $l_1$ norm is studied in [5], where the lower and upper bounds for the flow are changed. Strongly and weakly polynomial algorithms to solve GIMF are proposed. The strongly polynomial algorithms for GIMF have the same time complexity as the algorithms for IMF, but the minimum cut is searched in a network with fewer arcs. The weakly polynomial algorithms for GIMF have an $O(\min\{n^{2/3}, m^{1/2}\} \cdot m \cdot \log(n^2/m) \cdot \log(\max\{n, R\}))$ time complexity, where $R = \max\{c(x, y) - f(x, y) + f(y, x) - l(y, x)|x, y \in N\}$.

The inverse maximum flow under $l_\infty$ norm (denoted IMF$_\infty$) is considered in [4]. A very fast $O(m \cdot \log(n))$ time algorithm to solve this problem is presented.

The least number of modifications to the lower or/and upper bounds is considered in [6]. An $O(\min\{n^{2/3}, m^{1/2}\} \cdot m)$ time algorithm for solving this problem is presented.

Four inverse maximum flow problems are also studied by Liu and Zhang [8] under the sum-type and bottleneck-type weighted Hamming distance. Strongly polynomial algorithms to solve these problems are proposed.

In this paper, a theorem on the feasibility of the inverse maximum flow problems is presented. This theorem leads to an $O(m)$ time algorithm for deciding if an inverse maximum flow problem has solution or not. If a problem is not feasible what do we do in this situation? From the practical point of view it is not acceptable to give up. The problem must be solved somehow even if we have to make a compromise. The compromise consists in modifying (as little as possible) some parameters of the problem. Of course, this leads to new inverse combinatorial optimization problems.

In this paper a new inverse combinatorial optimization problem is introduced. This problem consists in modifying as little as possible the flow in order to transform the inverse maximum problem into a feasible problem.
2 The Inverse Maximum Flow Problems

Let $G = (N, A, l, c, s, t)$ be an $s$-$t$ network, where $N$ is the set of nodes, $A$ is the set of directed arcs, $l$ and $c$ are the lower and, respectively, the upper bound vectors for the flow, $s$ is the source and $t$ is the sink node.

If a network has more than a source or more than a sink node, it can be transformed into an $s$-$t$ network (introducing a super-source and a super-sink node) [1].

Let $f$ be a given feasible flow in the network $G$. It means that $f$ has to satisfy the flow balance condition and the capacity restrictions. The balance condition for the flow $f$ is:

$$\sum_{y \in N, (x, y) \in A} f(x, y) - \sum_{y \in N, (y, x) \in A} f(y, x) = v(f), \quad x = s$$

$$= \begin{cases} 
    v(f), & x = s \\
    -v(f), & x = t \\
    0, & x \in N - \{s, t\}
\end{cases},$$

where $v(f)$ is the value of the flow $f$ from $s$ to $t$.

The capacity restrictions are:

$$l(x, y) \leq f(x, y) \leq c(x, y), \quad \forall (x, y) \in A,$$

where $c(x, y) \geq l(x, y) \geq 0$, for every arc $(x, y) \in A$.

The maximum flow problem is:

$$\max \{ v(f) \} \quad f \text{ is a feasible flow in } G.$$

We shall introduce now the definition of the minimum cut $s$-$t$ in the network $G$. The set of arcs $[S, \bar{S}] = (S, \bar{S}) \cup (\bar{S}, S)$ is called an $s$-$t$ cut in $G$ if $S \cap \bar{S} = \emptyset, S \cup \bar{S} = N, s \in S$ and $t \in \bar{S}$, where $(S, \bar{S}) = \{(x, y) \in A | x \in S \text{ and } y \in \bar{S}\}$ is the set of direct arcs of the cut and $(\bar{S}, S) = \{(x, y) \in A | x \in \bar{S} \text{ and } y \in S\}$ is the set of the inverse arcs. The capacity of the $s$-$t$ cut $[S, \bar{S}]$ in $G$ is $c[S, \bar{S}] = c(S, \bar{S}) - l(\bar{S}, S) = \sum_{(x, y) \in (S, \bar{S})} c(x, y) - \sum_{(x, y) \in (\bar{S}, S)} l(x, y)$. An $s$-$t$ cut is a minimum cut in $G$ if its capacity is minimal in the set of $s$-$t$ cuts of the network $G$.

The residual network attached to the network $G$ for the flow $f$ is $G_f = (N, A_f, r, s, t)$, where for each pair of nodes $(x, y)$ the value of $r(x, y)$ is defined as follows:

$$r(x, y) = \begin{cases} 
    c(x, y) - f(x, y) + f(y, x) - l(y, x), & \text{if } (x, y) \in A \text{ and } (y, x) \in A \\
    c(x, y) - f(x, y), & \text{if } (x, y) \in A \text{ and } (y, x) \notin A \\
    f(y, x) - l(y, x), & \text{if } (x, y) \notin A \text{ and } (y, x) \in A \\
    0, & \text{otherwise}
\end{cases}$$

The set $A_l$ contains as arcs of the residual network only the pairs of nodes $(x, y) \in N \times N$ for which the residual capacity is positive, i.e., $r(x, y) > 0$.

An inverse maximum flow problem is to change as little as possible the lower and/or upper bound vectors $l$ and respectively $c$ so that the given feasible flow $f$ becomes a maximum flow in $G$.

An inverse maximum flow problem (denoted IMFG) can be formulated using the following mathematical model:

$$\min \ dist((l, c), (\bar{l}, \bar{c}))$$

where $f$ is a maximum flow in $G = \{N, A, \bar{l}, \bar{c}, s, t\}$

$$l(x, y) - \gamma(x, y) \leq \bar{l}(x, y) \leq \min\{\bar{c}(x, y), l(x, y) + \beta(x, y)\},$$

$$\forall (x, y) \in A$$

$$c(x, y) - \delta(x, y) \leq \bar{c}(x, y) \leq c(x, y) + \alpha(x, y), \quad \forall (x, y) \in A$$

In the model (5) different formulas to measure the distance between $(l, c)$ and $(\bar{l}, \bar{c})$ are considered, such as (weighted) $l_1$ norm, (weighted) $l_\infty$ norm, the (weighted) Hamming distance etc. So, (5) is a general model for any inverse maximum flow problem, where the lower and the upper bounds for the flow can be modified.

The values $\alpha(x, y), \beta(x, y), \gamma(x, y)$ and $\delta(x, y)$ are given non-negative integer numbers, where $\gamma(x, y) \leq l(x, y)$ and $\delta(x, y) \leq c(x, y)$, for each arc $(x, y) \in A$. These values show how much the bounds for the flow of the arcs can vary.

In the formula above by $(l, c)$ is denoted the vector obtained by adding the components of the vector $c$ at the end of $l$. Similarly, $(\bar{l}, \bar{c})$ is the vector obtained by putting together the components of the vectors $\bar{l}$ and $\bar{c}$.

In order to make the flow $f$ a maximum flow in the network $G$, the upper bounds of some arcs from $A$ must be decreased and/or the lower bounds of some arcs from $A$ must be increased. So, the conditions $\bar{c}(x, y) \leq c(x, y) + \alpha(x, y)$ and $\bar{l}(x, y) \geq l(x, y) - \delta(x, y)$...
\( \gamma(x, y) \), for each arc \((x, y) \in A\) have no effect and, instead of (5), the following mathematical model is considered:

\[
\begin{align*}
(5') \quad & \min \ dist((l, c), (\bar{l}, \bar{c})) \\
& f \text{ is a maximum flow in } \bar{G} = \{N, A, \bar{l}, \bar{c}, s, t\} \\
& \bar{l}(x, y) \leq \min \{\bar{c}(x, y), l(x, y) + \beta(x, y)\}, \\
& \forall (x, y) \in A \\
& c(x, y) - \delta(x, y) \leq \bar{c}(x, y), \forall (x, y) \in A
\end{align*}
\]

The inverse maximum flow problems where \( l(x, y) = 0, \forall (x, y) \in A \) and only the upper bounds for the flow can be modified are particular cases of IMFG. Indeed, if the lower bounds for the flow can not be modified in order transform the given flow \( f \) into a maximum flow, then in \((5')\) we can consider \( \beta(x, y) = 0, \forall (x, y) \in A \).

### 3 The Feasibility of IMFG

When solving IMFG, if the upper bound is changed on an arc \((x, y)\), then it will be decreased with the amount of \( c(x, y) - f(x, y) \). If not so, then there still is an augmenting path from \( s \) to \( t \) that contains the direct arc \((x, y)\) and the modification of the upper bound is useless. This means that if \( c(x, y) > f(x, y) + \delta(x, y) \) on an arc \((x, y)\), then, when solving IMFG, there is no need to change the upper bound on \((x, y)\).

Similarly, when solving IMFG, if the lower bound is changed on an arc \((x, y)\), then it will be increased with the amount of \( f(x, y) - l(x, y) \). If not so, then there still is an augmenting path from \( s \) to \( t \) that contains the arc \((x, y)\) in inverse direction and the modification of the lower bound is useless. This means that if \( f(x, y) > l(x, y) + \beta(x, y) \) on an arc \((x, y)\), then, when solving IMFG, there is no need to change the lower bound on \((x, y)\).

Let's determine the arcs in the network \( G \) on which the capacity will not be changed.

First, as it has been seen, changing the upper bound have no effect on an arc \((x, y)\) with \( c(x, y) > f(x, y) + \delta(x, y) \). So, there is no need to try changing the upper bounds of the arcs from the following set:

\[
(6) \quad \bar{A}_1 = \{(x, y) \in A | f(x, y) + \delta(x, y) < c(x, y)\}.
\]

Similarly, as it has been seen, changing the lower bound have no effect on an arc \((x, y)\) with \( f(x, y) > l(x, y) + \beta(x, y) \). So, there is no need to try changing the lower bounds of the arcs from the following set:

\[
(7) \quad \bar{A}_2 = \{(x, y) \in A | l(x, y) + \beta(x, y) < f(x, y)\}.
\]

It is easy to see that if there is a path from \( s \) to \( t \) in the network \( G \) that contains only direct arcs \((x, y)\) so that \( c(x, y) > f(x, y) + \delta(x, y) \) and/or inverse arcs \((y, x)\) with \( f(y, x) > l(y, x) + \beta(y, x) \), then IMFG has no solution.

A graph denoted \( \bar{G} = (N, \bar{A}) \) can be constructed to verify the feasibility of IMFG, where:

\[
(8) \quad \bar{A} = \bar{A}_1 \cup \{(x, y) \in N \times N | (y, x) \in A \text{ and } f(y, x) > l(y, x) + \beta(y, x)\}.
\]

We have the following theorem:

**Theorem 1** In the network \( G \), IMFG has optimal solution for the given flow \( f \), if and only if there is no directed path in the graph \( \bar{G} \) from the node \( s \) to the node \( t \).

**Proof:** Let \( G' = (N, A, l', c') \) be a network for which the last conditions from \((5')\) hold: \( l(x, y) \leq l'(x, y) \leq \min \{c'(x, y), l(x, y) + \beta(x, y)\} \) and \( c(x, y) - \delta(x, y) \leq c'(x, y), \forall (x, y) \in A \). Let \( G'_f = (N, A'_f, r'_f) \) be the residual network attached to the network \( G' \) for the flow \( f \). It is easy to see that \( r'_f(x, y) > 0, \forall (x, y) \in A \) due to the restriction on the upper bound vector for the arc \((x, y)\) of \( G \) \( c(x, y) > f(x, y) + \delta(x, y) \Rightarrow c'(x, y) > f(x, y) \) or because \( (y, x) \in A \) and \( f(y, x) > l(y, x) + \beta(y, x) \Rightarrow f(y, x) > l'(y, x) \). This means that \( \bar{A} \subseteq \bar{A}_f \).

If IMFG is a feasible problem, then it means that there is a vector \((\bar{l}, \bar{c})\) with \( l(x, y) \leq \min \{\bar{c}(x, y), l(x, y) + \beta(x, y)\} \) and \( c(x, y) - \delta(x, y) \leq \bar{c}(x, y), \forall (x, y) \in A \) and for which the flow \( f \) is a maximum flow in the network \( \bar{G} = (N, A, \bar{l}, \bar{c}) \). Since \( \bar{A} \subseteq \bar{A}_f \), if it exists a directed path in \( \bar{G} \) from \( s \) to \( t \), it corresponds to a directed path in \( G_f \), which leads to an augmentation to the flow \( f \) in \( G \) (contradiction).

Now, for the inverse implication we construct the following upper and lower bound vectors for the arcs of the network \( G \):

\[
c''(x, y) = \begin{cases} 
c(x, y), & c(x, y) > f(x, y) + \delta(x, y) \\
f(x, y), & \text{otherwise}
\end{cases}
\]
\[ l''(x, y) = \begin{cases} l(x, y), f(x, y) > l(x, y) + \beta(x, y) \\ f(x, y), \text{ otherwise} \end{cases} \]

It is easy to see that \( l''(x, y) \leq l(x, y) + \beta(x, y) \) and \( c(x, y) - \delta(x, y) \leq c''(x, y), \forall (x, y) \in A \). In the residual network \( G_{f'}'' = (N, A_{f''}', r'') \) attached to \( G'' = (N, A, l'', c'') \) and to the flow \( f \) we have \( r''(x, y) = 0 \), for all \( (x, y) \in (N \times N) - \tilde{A} \). Since \( \tilde{A} \subseteq A_{f''}' \), it means that \( \bar{A} = A_{f''}' \) (see (4)). Therefore, because there is no path from \( s \) to \( t \) in the graph \( \tilde{G} \), it results that there is no directed path from \( s \) to \( t \) in \( G_{f''}' \). This implies that the flow \( f \) is a maximum flow in the network \( G'' = (N, A, l'', c'') \). It means that \((l'', c'')\) is a feasible solution for \( \text{IMFG} \).

In \( \text{IMFG} \), the feasible region for the vector \((\bar{l}, \bar{c})\) can be reduced to \( l \leq l \leq \bar{l} + \beta \) and \( c - \delta \leq \bar{c} \leq c \) (from (5') and because when solving \( \text{IMFG} \) there is no need to increase the upper bounds for the flow and there is no need to decrease the lower bounds for the flow), which is a compact region. So, because \( \text{IMFG} \) has a feasible solution, it results that \( \text{IMFG} \) has optimal solution.

The verification of \( \text{IMFG} \) being feasible can be done in \( O(\bar{m}) \) time complexity, using a graph search algorithm in \( \tilde{G} \), where \( \bar{m} \) is the number of arcs in the set \( \bar{A} \) with \( \bar{m} \leq 2m \). So, this test of feasibility can be applied to any inverse maximum flow problem.

4 The Modification of Flow

In this section for a non-feasible \( \text{IMFG} \) problem in the network \( G \) for the given flow \( f \) we shall modify the flow \( f \) so that the inverse maximum flow problem in \( G \) for the modified flow becomes feasible and the distance between the value of the initial flow and the value of the modified flow is minimum.

The following inverse optimization problem (denoted \( \text{TFIMF} \)) is obtained:

\[
\begin{cases}
\min |v(f') - v(f)| \\
\text{\( f' \) is a feasible flow in \( G \)} \\
\text{The inverse maximum flow problem is feasible in \( G \) for \( f' \)}
\end{cases}
\]

(10)

So, the problem is to find a feasible flow \( f' \) in the network \( G \) so that the inverse maximum flow problem in \( G \) for the flow \( f' \) is feasible and \( |v(f') - v(f)| \) is minimum. For a feasible flow \( f' \) the test of feasibility of \( \text{IMFG} \) can be done in \( O(m) \) time (as we have seen in section three, theorem 1). The first thing we have to do is to construct all the feasible flows in \( G \) with the value equal to \( v(f) \). If there is such a flow then any of these flows is solution of problem (10).

The problem of maximum flow and, consequently, the inverse minimum flow problem is with integer values (for the lower bounds, upper bounds, flow, value of the flow and the restrictions to variation of the bounds). That is why we can think to a strategy of solving the problem (10) as follows:

If there is no feasible feasible flow \( f_0 \) in \( G \) so that IMFG is feasible for \( f_0 \) in \( G \) and \( v(f_0) = v(f) \) then we look for any feasible flow \( f_1 \) in \( G \) so that IMFG is feasible for \( f_1 \) in \( G \) and \( |v(f_1) - v(f)| = 1 \) and so on. Finally, after \( k \) iterations, \( k \leq |V - v(f)| + 1 = V - v(f) + 1 = V \) (is the value of any maximum flow in \( G \)) we shall find a feasible flow \( f_k \) in \( G \) so that IMFG is feasible for \( f_k \) in \( G \) and \( |v(f_k) - v(f)| = k - 1 \). This flow is the solution of problem (10).

In order to solve the problem (10) in the manner described above we have to find a method to generate all the feasible flows in the network \( G \) for a given integer value \( v \).

In paper [9] a method for finding all the maximum flows in a given network \( G \) is presented.

We are interested in finding all the feasible flows in the network \( G \) for a given integer value \( v \). In order to do that we transform the network \( G \) as follows:

We introduce in \( G \) a new node denoted \( s' \) and the arc \((s', s)\) with the upper bound equal to \( v \), the lower bound equal to 0 and restrictions to the variation of bounds equal to 0, i.e., \( c(s', s) = v \) and \( l(s', s) = \alpha(s', s) = \delta(s', s) = 0 \). The new node \( s' \) becomes the only source node in the modified network denoted \( G_v \). It is easy to see that any feasible flow in \( G \) with the value equal to \( v \) is a feasible flow in \( G_v \) and any feasible flow in \( G_v \) with the value equal to \( v \) is a feasible flow in \( G \). Moreover, a feasible flow with the value equal to \( v \) in \( G_v \) is a maximum flow in \( G_v \) because the total flow that exits the source node \( s' \) can not exceed \( v = c(s', s) \). Consequently, for a given value \( v \) we can find all the maximum flows in \( G_v \) using the algorithm from (9) and these flows are all the feasible flows with value \( v \) in \( G \) if we ignore the arc \((s', s)\).

We are able now to present the algorithm for solving the problem (10):

PROGRAM SolvingTFIMF;
BEGIN
\[ v := v(f); \]
\[ v' := v(f); \]
Find the maximum flow \( F \) in \( G \);
\begin{align*}
V & := v(F);
\text{WHILE } v \leq V \text{ DO} \\
& \text{BEGIN} \\
& \quad \text{Construct the network } G_v; \\
& \quad \text{Find all max. flows } f_i \ (i = 1..p) \text{ in } G_v; \\
& \quad \text{FOR } i := 1 \text{ TO } p \text{ DO} \\
& \qquad \text{IF IMFG is feasible in } G \text{ for } f_i \text{ THEN} \\
& \qquad \qquad f' := f_i; \\
& \qquad \text{STOP;} \\
& \quad \text{END IF;} \\
& \quad \text{END FOR;} \\
& \quad \text{IF } v' \geq 0 \text{ THEN} \\
& \quad \quad \text{Find all max. flows } f_i \ (i = 1..p) \text{ in } G_{v'}; \\
& \quad \quad \text{FOR } i := 1 \text{ TO } p \text{ DO} \\
& \qquad \text{IF IMFG is feasible in } G \text{ for } f_i \text{ THEN} \\
& \qquad \qquad f' := f_i; \\
& \qquad \text{STOP;} \\
& \quad \text{END IF;} \\
& \quad \text{END FOR;} \\
& \quad v := v + 1; \\
& \quad v' := v' - 1; \\
& \text{END WHILE;} \\
& \text{END.}
\end{align*}

**Theorem 2** The program "SolvingTFIMF" finds an optimum solution \( f' \) of the problem (10).

**Proof:** It is easy to see that any maximum flow \( F \) in \( G \) is a feasible solution for (10). That is because for \( F \) the lower bound vector \( l \) and the upper bound vector \( c \) forms the optimal solution for IMFG in the network \( G \) (we have to make no modification to \( l \) or/and \( c \) so that \( F \) becomes a maximum flow in \( G \)).

The algorithm constructs sequentially all the flows in \( G \) with the value equal to \( v(f) \), then all the flows in \( G \) equal to \( v(f) + 1 \), then all the flows in \( G \) equal to \( v(f) - 1 \), then all the flows in \( G \) equal to \( v(f) + 2 \), all the flows in \( G \) equal to \( v(f) - 2 \) and so on. The algorithm stops (after at most \(|V - v(f)| + 1 = V - v(f) + 1 \) iterations of the "WHILE ... DO" loop) when a flow \( f' \) is found for which the problem IMFG is feasible in the network \( G \). Of course, \( f' \) is a feasible solution of problem (10).

We suppose that \( f' \) (found by the algorithm) is not an optimum solution for (10). This means that there exists a flow \( f'' \) for which IMFG is feasible in \( G \) and \(|v(f'') - v(f)| < |v(f') - v(f)| \). We denote by \( k' \) the iteration of the algorithm when \( f' \) is found (and the algorithm stops). It is easy to see that \( k' = |v(f') - v(f)| + 1 \).

We denote \( k'' = |v(f'') - v(f)| + 1 \). Since \( k'' < k' \), it results that the flow \( f'' \) was one of the feasible flows in \( G \) constructed by the algorithm in the iteration \( k'' \), previous to iteration \( k' \) and, since \( f'' \) is a feasible solution for the problem (10), the algorithm had to stop in iteration \( k'' \) not in iteration \( k' \) (contradiction).

So, \( f' \) (found by the algorithm) is an optimum solution of the problem (10).

\[ \blacksquare \]

**Acknowledgements.** The research was supported by Grant PNII no. 22134/2008.

**References:**


