RAPPORT DE RECHERCHE

L R I

IS THE POLYTOPE ASSOCIATED WITH A TWO STAGE STOCHASTIC PROBLEM TDI ?

LETOURNEL M / LISSER A / SCHULZ R

Unité Mixte de Recherche 8623
CNRS-Université Paris Sud – LRI

07/2010

Rapport de Recherche N° 1535

CNRS – Université de Paris Sud
Centre d'Orsay
LABORATOIRE DE RECHERCHE EN INFORMATIQUE
Bâtiment 490
91405 ORSAY Cedex (France)

Is the polytope associated with a two stage stochastic problem TDI?

Marc Letournel Abdel Lisser Rüdiger Schulz

June 25, 2010

Abstract

The maximum weight forest problem (MWFP) in a graph is solved by the famous greedy algorithm due to Edmonds (1971) where every edge has a known weight. In particular, the system of constraints on the set of edges is TDI (totally dual integral), since the set of independent edges, i.e., of acyclic subsets of edges, is a matroïd. We extend this approach to the case of two-stage maximum weight forest problems. The set of edges is composed of first stage edges with known weights and second stage ones where the weights are known a priorily in terms of discrete random variables. As the probability distribution is discrete, we transform the stochastic problem into a deterministic equivalent problem. In this article, we prove TDIness for the two stage maximum weight forest problem in the following cases: Two scenarios with reduced number of first stage variables and we propose an efficient greedy algorithm for solving this problem. We provide a counter example to prove that the problem is not anymore TDI for more than two scenarios.

1 Introduction, notation and generalities

1.1 Introduction

We consider a graph $=(V, E)$ where E is a set of edges of cardinality N, and a cost function c defined on edges. The edges are indexed by $i \in [1, N]$ and for any subset F of edges, we call $c(F)$ the sum $\sum_{j \in F} c_j x_j$ where $x_j = 1$ if $j \in F$ and $x_j = 0$ otherwise. A subset F is said independant if there is no cycle in F. The maximization problem of $c(F)$ for F independant is well known to be connected to matroïds and is efficiently solved in the case of a fixed cost value for every edge (Nemhauser and Wolsey, 1999). When every edge has a fixed value, we say that the problem is deterministic and since independant sets are a matroïd, the greedy algorithm is efficient to provide the maximisation problem of $c(F)$ polynomialy. The matroïd structure of independant sets involves the fact that the rank function $r(F)$ (the maximum cardinality of any independant subset included in F) is submodular. In this case, the polytope associated to independance constraint is Totally Dual Integral. This is a very nice property which allows to apply algebraic methods to solve maximization. The problem we introduce in this paper is to understand what happens when cost function is not deterministic but follows a discrete stochastic distribution π. In our problem, the edges are splitted into two subsets $E = X \cup Y$. In stochastic programming, the first subset X has a deterministic cost function, this set is called first stage, we set $card(X) = n$, whereas in second subset called second stage, $card(Y) = q$.
Cost function depends on $K \geq 2$ scenarios and cost values are given by a probability distribution $\pi = (\pi_1, \ldots, \pi_K)$. This problem is turned into a deterministic formulation by splitting any second stage edge $x_j \in Y$ into K equivalent new edges x_j^k with $k \in \{1, \ldots, K\}$,

connecting the same vertices (multigraph) with a fixed cost $\pi_k c_j(k)$. The whole set of edges becomes $V = X \cup Y_1 \cup \ldots \cup Y_K$. We consider that a first stage edge belongs to every scenario $\mathscr{S}_i$ for $i \in \{1, \ldots, K\}$ while a second stage belongs to only one scenario. The formulation of this problem is :

$$z_{IP} = max\{\sum_{j \in X} c_j x_j + \sum_{k=1}^{k=K} \sum_{j \in Y} \pi_k c_j(k) x_j^k : x \in \rho(r)\} \tag{1}$$

$$\rho(r) = \{x \in \{0,1\}^{n+Kq} : \sum_{j \in S \cap \mathscr{S}_i} x_j \leq r(S_i) \; for \; i \in \{1, \ldots, K\} \quad \forall S \subseteq E\} \tag{2}$$

In our formulation, there exist $n + Kq$ edges where constraints express choice of independant edges. Since these new edges of second stage don't exist simultaneously, the set of contraints involves only edges of same scenario, and the whole set of contraint can be expressed as a block matrix system.
The most powerfull case is when the deterministic problem is associated with a matrix constraint which is Totally Unimodular (TU). Introducing multi stages and dubbing sub matrices of second stage edges is studied by (Kong et al., 2007). In our case, initial properties are weaker and are formulated in terms of TDI system. Operations that preserve TDI properties are presented in (Cook, 1983) and we will outline connection with our formulation and matroïds intersection, the reader will refer to Franck (1981) for further details. In section 2, we reproduce the basic results of the determinist case. The reader will refer to (Nemhauser and Wolsey, 1999). In the greedy algorithm, we outline the importance of the closure of a subset and the dynamical point of view of the building scheme of the dual solution. In section 3, we deal with the case of two scenarios with a small number of first stage edges. We sketch an approach to deal with any situation with only two scenarios. In section 4, we show by an example to the contrary that the TDI properties are not preserved in case of strictly more than two scenarios.

1.2 Preliminaries

We begin by reminding classical notations and properties:

1.2.1 Matroids

Definition 1.1. *Let E a finite set of cardinality N, and $\mathscr{F}$ a set of subsets of E,we say that $\mathscr{F}$ is an independence system if:*

$$\forall (F_1, F_2) \subset E^2, F_1 \in \mathscr{F} \quad and \quad F_2 \subset F_1 \Rightarrow F_2 \in \mathscr{F}$$

In this article, E is the set of edges of a graph G, and $\mathscr{F}$ are the acyclic subset of edges in G.
The elements of $\mathscr{F}$ are called *independent sets.*

Definition 1.2. *Given an independence system $\mathscr{F}$ in E, we say that $F \in \mathscr{F}$ is a maximal independent set if*

$$\forall j \notin F, F \cup \{j\} \notin \mathscr{F}$$

For a subset $T \subset E$, we consider the independent sets F included in T, and especialy the maximal independent sets in T. We define the rank function r as

$$r(T) = max\{|S| : S \in \mathscr{F} \text{ for } S \subset T\}$$

It is not always the case that every maximal independent set F in T has exactly the same cardinality. This is the situation of a *matroïd*:

Definition 1.3. *A matroïd is an independent set $\mathscr{F}$ for which every $T \subset E$, every maximal independent set $\mathscr{F}$ in T has exactly the same cardinality - which is $r(T)$.*

When $\mathscr{F}$ is defined as the acyclic set of edges of E, it is a matroïd. It is interesting to notice that the rank function is given by the following property:

Proposition 1.4. *$r(T)$ is the cardinal of the set of covered vertices by edges in T minus the number of connected components of the subgraph.*

Theorem 1.5. *(Nemhauser and Wolsey, 1999)*
Let consider a finite set E and an independence system $\mathscr{F}$, and the (rank) function r defined above, $\mathscr{F}$ is a matroïd $\Leftrightarrow r$ is submodular.

We outline that there exist some submodular functions different from r, for instance there exist modular functions which are not *nondecreasing* (while r is always nondecreasing).

1.3 Polytopes and problems associated

We introduce the notations:

$$\pi(r) = \{x \in \{0,1\}^n : \sum_{j \in S} x_j \leq r(S) \quad \forall S \subseteq E\} \tag{3}$$

$$z_{IP} = max\{c^t x : x \in \pi(r)\} \tag{4}$$

The polytope associated to the matroïd:

$$P(r) = \{x \in (R^+)^n : \sum_{j \in S} x_j \leq r(S) \quad \forall S \subseteq E\} \tag{5}$$

and the linear program associated

$$z_{LP} = max\{c^t x : x \in P(r)\} \tag{6}$$

The dual problem associated with z_{IP} is :

$$z_{LD} = min \sum_{S \subseteq E} r(S) y_S \tag{7}$$

$$st \begin{cases} \sum_{S:j \in S} y_S \geq c_j \quad \forall j \in E \\ y_S \geq 0 \quad \forall S \subset E \end{cases} \tag{8}$$

2 Deterministic case

In this section we remind the results given in Nemhauser and Wolsey (1999). We outline a dynamical point of view for the construction of the greedy solution by watching carefully how the closure of a subset of chosen edges is increasing.
The Problem of finding a maximum-weight independant set in $\mathscr{F}$ is formulated by (4);
We show that $z_{IP} = z_{LD}$, by exhibiting the optimal solution for z_{LD} which is indeed the greedy solution for z_{IP}. The conclusion is that $\pi(r)$ is TDI and the duality gap is equal to 0.

2.1 The greedy algorithm

This algorithm is due to Edmonds (Nemhauser and Wolsey (1999)).
Rank the elements of N so that $c_1 \geq c_2 \geq \dots c_n$
Let begin by $J_0 = \emptyset$, $t = 1$
Iteration t : If $c_t \leq 0$ then stop and $S^G = J_{t-1}$.
If $c_t > 0$ and $J_{t-1} \cup \{t\} \in \mathscr{F}$, then set $J_t = J_{t-1} \cup \{t\}$.
If $c_t > 0$ and $S_{t-1} \cup \{t\} \notin \mathscr{F}$ then set $J_t = J_{t-1}$.
if $t = N$ stop and $S^G = J_t$
set t to $t+1$
The greedy solution is $\{x_i / i \in S^G\}$
We now present the main result connecting the greedy solution to the Dual Problem, and we emphasize on the growing closure mechanism.

Definition 2.1. *The closure or span of a set A is $sp(A) = \{i \in N : r(A \cup \{i\}) = r(A)\}$.*

Theorem 2.2. *The dual problem z_{LD} and the greedy solution of z_{IP} have the same objective value. When $\mathscr{F}$ is a matroïd, $P(r)$ is a integral polytope. And the system defining $P(r)$ is TDI.*

Proof. First, we still assume that indexation of edges is equal to their rank according to a decreasing weight. We dynamicaly construct both greedy solution and dual solution equals at every step. The first step is the choice of the most weighted edge and we set $J_1 = 1$. We call $j_1 = 1$ the first chosen edge and we notice that $K_1 = sp(J_1)$. We now choose the second edge j_2 with the highest rank among all remaining edges such that $j_2 \notin K_1$. Up to this point of construction, obviously $j_2 = 2$ since there is no cycle with only two edges. We build now J_{i+1}:

Let $J_i = (j_1, \dots, j_i)$ be the set of chosen edges at step $i \geq 2$, and $K_i = sp(J_i)$.
Choose j_{i+1} with the highest rank among all remaining edges such as $j_{i+1} \notin K_i$ and $c_{j_{i+1}} \geq 0$. We call p the number of chosen edges and $S^G = J_p$.

The optimal solution to z_{LD} is :

$y_{K_t} = c_{j_t} - c_{j_{t+1}} \quad for\, t = 1, \dots, p-1$
$y_{K_p} = c_{j_p}$
$y_S = 0 \quad otherwise$

First, we need to check that for every $j \in N$: $\sum\limits_{S/j \in S} y_S \geq c_j$.
it is obviously the case for every $j_i \in S^G$ since $j_i \in J_t$ for $i \leq t$.
For any j which has not been choosen during the greedy algorithm, j belongs to some $sp(J_t) - sp(J_{t-1})$
So $\sum\limits_{S/j \in S} y_S = c_{j_t} \geq c_j$.

Secondly, we need to compute the sum : $\sum\limits_{S \subseteq E} r(S) y_S = \sum\limits_{t=1}^{p-1} t(c_{j_t} - c_{j_{t+1}}) + p c_{j_p}$

this leads by splitting in two sums and reindexing to $\sum\limits_{S \subseteq E} r(S) y_S = \sum\limits_{t=1}^{p} c_{j_t}$.
This shows that z_{IP} and z_{LD} have the same objective value. More precisely, the greedy solution to z_{IP} and z_{LD} have the same objective value. We conclude that the set of inequalities defining $P(r)$ is TDI, and $P(r)$ is an integral polytope. $\square$

In the next section we introduce two stage stochastic problem, where the cost of several edges can have stochastic values. We consider the case of discrete probabilities, where costs belong to a finite set of values. The values of each edges are sampled simultaneously so that there exists a finite number of scenarios. We construct a greater graph where there exist as many exemplary of edges as possible scenarios, and we aim to formulate the same constraints as in the deterministic case : avoiding cycle in any scenario. From a stochastic point of view, it is clearly not possible to formulate constraints including edges of different scenarios together. In the sense of our formulation, this turns into the fact that some constraints that would limit the summits of the polyedra don't exist in the inequality system. That means that there exist less constraints than in an equivalent deterministic graph: some edges belong to different scenarios and it is possible to create cycles built with edges of different scenarios without violating contraints.

3 Two stage Stochastic Problem: two scenarios case

We consider a two stage stochastic program and we try to see if we keep some properties of TDI systems. The edges of E are classified according to the scenarios $\mathscr{S}_1$ and $\mathscr{S}_2$:
The edges of first stage identified are common to both scenarios $\mathscr{S}_1$ and $\mathscr{S}_2$;
The edges of second stage belong to specific scenario $\mathscr{S}_1$ xor $\mathscr{S}_2$.
Any subset S of edges is $S = S_1 \cup S_2$ where S_i are all the edges in S of the scenario i, we use the notation $S_i = S \cap \mathscr{S}_i$ and we outline that $S_1 \cap S_2 \neq \emptyset$ but is the subset of edges of S belonging to first stage.
The interpretation is that we consider a graph where some edges have a fixed value (first stage) while edges from the second stage have two possible values according to a stochastic distribution. We define a weight function c on E of cardinality $N = n + 2q$.

3.1 Notations

There is a possible confusion in the notation. Usualy, it would seem simplest to note the edges with an increasing indexation begining with first stage from x_1 to x_n and indexation for second stage from y_{n+1} up to y_{n+2q} (with a different name and a different set of indexation to notify the scenario number). But most of time, we need to run a greedy algorithm and proofs are easier to formulate with an indexation corresponding to a decreasing weight $c_1 \geq ... \geq c_{n+2q}$ regardless the scenario nor the stage they belong to. Moreover, the variable y is used to formulate the dual problem, so this is what we did :

We note x_j^0 for $j \in [1 \ldots n]$ with a null upper index for an edge of the first stage.
And x_j^1 or x_j^2 for $j \in [1 \ldots q]$ for edges of the second stage.
In the same manner, we note c_j^k a weight associated to x_j^k where $k \in \{0, 1, 2\}$.
We write $x \in \{0, 1\}^{n+2q}$ the vector associated with the edges, with no specification on upper index for any componant in a sum.

The equivalent problem to the deterministic case is to consider in E the subsets F such that there is no cycle in F when only considering edges separately in scenario $\mathscr{S}_1$ and scenario $\mathscr{S}_2$. These subsets are said to be acyclic.
Such subsets belong to the family $\mathscr{F}$ we call the *independent sets.*

$$\mathscr{F} = \{F \subset E : \ F \cap \mathscr{S}_1 \text{ and } F \cap \mathscr{S}_2 \text{ are acyclic}\}$$

We consider the problem of maximizing $c(F)$ for $F \in \mathscr{F}$.
We now introduce the corresponding notation to section 2.

3.2 Polytopes and problems associated with multistage case

$$\rho(r) = \{x \in \{0,1\}^{n+2q} : \sum_{j \in S \cap \mathscr{S}_i} x_j \leq r(S_i) \ for \ i \in \{1,2\} \quad \forall S \subseteq E\} \tag{9}$$

$$z_{IP} = max\{c^t x : x \in \rho(r)\} \tag{10}$$

The polytope associated with relaxing variables:

$$R(r) = \{x \in (R^+)^{n+2q} : \sum_{j \in S \cap \mathscr{S}_i} x_j \leq r(S_i) \ for \ i \in \{1,2\} \quad \forall S \subseteq E\} \tag{11}$$

and the linear program associated

$$z_{LP} = max\{c^t x : x \in R(r)\} \tag{12}$$

The dual problem associated with z_{IP} is :

$$z_{LD} = min \sum_{S \subseteq V} r(S_1) y_{S_1} + r(S_2) y_{S_2} \tag{13}$$

$$st \begin{cases} \sum_{S/j \in S} y_S \geq c_j & \forall j \in E \\ y_S \geq 0 & \forall S \end{cases} \tag{14}$$

3.3 The case of only one edge in the first level

We consider the case where only one edge x_1^0 belongs to first stage, we are going to see that the dual problem has an optimal value equal to a particular solution of a double greedy algorithm lead separately into scenarios $\mathscr{S}_1$ and $\mathscr{S}_2$.
First of all, we rank separately in both scenarios the elements of N so that $c_1^i \geq c_2^i \geq \dots c_q^i$ for $i \in \{1,2\}$. Concerning the first stage edge, we introduce a variable value $\delta \in [-c^0/2, c^0/2]$ set initialy to 0 and we split the total weight of x^0 into $c^{01} = c^0/2 + \delta$ and $c^{02} = c^0/2 - \delta$ and we introduce the first edge x_0^1 with these two values respectively into scenario $\mathscr{S}_1$ and $\mathscr{S}_2$. Now we run separately the greedy algorithm on both scenarios.

Lemma 3.1. *When running separately greedy algorithm on both scenarios - considering x_0^1 with a splitted cost - , we get the same status for x_0^1 (i.e chosen in both scenarios or left in both scenarios), the dual problem z_{LD} has the same value than the merge of both greedy solutions.*

Proof. Suppose that in scenario $\mathscr{S}_1$, x^0 belongs to the closure K_k^1 of a chosen edge x_k^1 and in the same time in scenario 2, x^0 belongs to the closure K_m^2 of a chosen edge x_m^2. That means that it is no use collecting x^0 in the greedy solution. We just gather greedy solutions into scenarios $\mathscr{S}_1$ and $\mathscr{S}_2$ and in the corresponding dual formulation we set dual variables $y_{S_i} = y_{K_t^i}$ exactly in the same manner as with deterministic case. We clearly check that:
- for any edge of second stage in $\mathscr{S}_1$ for example, we have $\sum_{S/j \in S_1} y_{S_1} \geq c_j^1$.
- for the first stage edge, since it belongs to K_k^1 we just get $\sum_{S/x^0 \in S_1} y_{S_1} \geq c_k^1 \geq c^{01}$, but there is another sum with scenario $\mathscr{S}_2$: $\sum_{S/x^0 \in S_2} y_{S_2} \geq c_m^2 \geq c^{02}$ so that finaly $\sum_{S/x^0 \in S} y_S \geq c^0$.
Now we just check $\sum_{S \subseteq V} r(S_1) y_{S_1} + r(S_2) y_{S_2} = \sum_{t=1}^{p_1} c_{j_t}^1 + \sum_{t=1}^{p_2} c_{j_t}^2$.

This equality proves that the dual problem z_{LP} has the same value than the greedy solution.

Suppose that in both scenarios $\mathscr{S}_1$ and $\mathscr{S}_2$, the first stage edge x^0 is chosen during the greedy algorithms. The situation is exactly the same, except the fact that x^0_1 belongs to the greedy solution. We have in $\mathscr{S}_1$: $\sum_{S/x^0 \in S_1} y_{S_1} = c^{01}$, and in $\mathscr{S}_2$: $\sum_{S/x^0 \in S_2} y_{S_2} = c^{02}$ so that finaly $\sum_{S/x^0 \in S} y_S = c^0$.

□

Theorem 3.1. *With only one edge in first stage, the primal problem z_{IP} and dual problem z_{LP} have the same integer value. This entails that the system is TDI.*

Proof. The last issue is that x^0_1 is not chosen during greedy algorithm in scenario $\mathscr{S}_1$ but is chosen in scenario $\mathscr{S}_2$ or opposite case. There exists k such that x^0 belongs to the closure K^1_k of a chosen edge x^1_k. We remark that if we change value δ, the greedy algorithm is disturbed but only at a certain point, we have two situations:
$c^{01} = c^0/2 + \delta \leq c^1_k$ implies $x^0 \in K^1_k$ and changes should occur only for edges with $c^1_j \leq c^{01}$;
$c^{01} = c^0/2 + \delta > c^1_k$ implies that x^0 becomes a chosen edge in the greedy algorithm run on first scenario $\mathscr{S}_1$.
Just observe that these changes don't affect the greedy algorithm on remaining edges in the sense that there will be certainly changes among chosen edges in $\mathscr{S}_1$, for $j \geq k$, K^1_j becomes K'^1_j and the number of chosen edges should change to p'_1, but even in this case the dual problem has a value still equal to the new greedy solution:

$$\sum_{K^1 \in \{K^1_1, \ldots, K'^1_k, \ldots, K'^1_{p'_1}\}} r(K^1) y_{K^1} = \sum_{t=1}^{p'_1} c^1_{j_t}$$

We summarize these considerations:
For $c^{01} \in [0, c^1_k]$ and $\delta \in [-c^0/2, c^0/2]$, x^0 is not chosen during greedy algorithm on $\mathscr{S}_1$ and belongs to the closure K^1_k.
For $c^{01} > c^1_k$ and $\delta \in [-c^0/2, c^0/2]$, x^0 is chosen during greedy algorithm on $\mathscr{S}_1$ and becomes part of new closure K'^1_k.
Meanwhile, in the second scenario, x^0 has been chosen during greedy algorithm and we call m the rank of choice so that the closure set is K^2_m. Modifying δ implies similary changes on ranking costs. But increasing c^{02} won't change the status of x^0 on scenario $\mathscr{S}_2$, it will be chosen during the greedy algorithm since it doesn't belong to any closure of preceding chosen edges. On the opposite, decreasing c^{02} should provide change as soon as it decreases under any value of a non covered edge $x^2_{j_0} \notin K^2_m$ (it is still possible that x^0 should be chosen at a further step). In this case $x^2_{j_0}$ would become next chosen edge and K^2_m is modified into K'^2_m as a consequence.
The question is to find a value for δ such as x^0 has the same status in scenario $\mathscr{S}_1$ and $\mathscr{S}_2$. The answer is to increase δ untill one of these situations occurs:

- c^{01} reach and overstep value c^1_k and c^{02} is still greater than $c^2_{j_0}$. In this case x^0 is chosen in both scenarios.
- c^{01} increases and don't overstep value c^1_k and c^{02} has decreased enough for x^0 to be covered by a new closure K'^2_j with $j \geq m$. In this case x^0 is not taken in greedy solutions.

- c^{01} increases up to the point that $\delta = c^0/2$ without overstepping c_k^1 and the whole cost of x^0 is covered by first scenario $\mathscr{S}_1$, $c^{02} = 0$ and it is no use taking x^0 in the greedy solutions.

□

3.4 The case of two edges in the first stage

We consider the case where two edges x_1^0 and x_2^0 belong to first stage with costs $c_1^0 \geq c_2^0$, and we introduce two associated variables δ_i as above; $\delta_1 \in [-c_1^0/2, c_1^0/2]$ set initialy to 0 and we apply the the method described in 3.3. So we first set δ_1 in order to place x_1^0 in both scenarios with the same status (taken in both scenarios or not taken in any scenario). Meanwhile, we choose an initial value for δ_2 such as x_2^0 stays initialy beyond x_1^0 in both scenarios: $c_1^{01} = c_1^0/2 + \delta_1$ and $c_1^{02} = c_1^0/2 - \delta_1$; $c_2^{01} = c_2^0/2 + \delta_2$ and $c_2^{02} = c_2^0/2 - \delta_2$; $c_1^{01} \geq c_2^{01}$ and $c_1^{02} \geq c_2^{02}$. There are several cases and first lemma deals withs simplest cases:

Lemma 3.2. *When there exist two values δ_1 and δ_2 such that respectively x_1^0 and x_2^0 get the same status in both scenarios, then the dual problem z_{LD} and the merge of both greedy solutions have the same value*

Proof. The point to be discussed is to check if it is possible to give the same status for each first stage edge in both scenarios by adapting the fractinonal part of their weights. The relative ranking between first stage edges is indeed not significant.

- According to the method described in 3.3, it is possible to adapt δ_2 in a way that x_2^0 get the same status in scenario $\mathscr{S}_1$ and $\mathscr{S}_2$ with respect to $c_1^{01} \geq c_2^{01}$ and $c_1^{02} \geq c_2^{02}$. Then any status for x_1^0 is changed and both first stage edges are correctly inserted into the greedy solutions. The same computation as for a single first stage edge holds.

- According to the method described in 3.3, it is possible to adapt δ_2 in a way that x_2^0 get the same status in scenario $\mathscr{S}_1$ and $\mathscr{S}_2$ with respect to $c_1^{01} \leq c_2^{01}$ and $c_1^{02} \geq c_2^{02}$ or opposite configuration. But these changes are made without modifying the status of x_1^0. Then both first stage edges are correctly inserted into the greedy solutions.

□

Theorem 3.2. *With only two edges in first stage, the primal problem z_{IP} and dual problem z_{LP} have the same integer value. This entails that the system is TDI.*

Proof. We need to examine the case where first having introduced the first edge of first stage, the relative balance of the second edge can't be realized without modifying the status of the first edge:
Suppose that x_1^0 is chosen in both greedy solutions and x_2^0 is chosen only in scenario $\mathscr{S}_2$, and that there is no way to change these status except for c_2^{01} to overstep c_1^{01}. We summarize this case by:
In $\mathscr{S}_1$, x_1^0 is chosen in greedy algorithm but x_2^0 belongs to some closure K_k^1.
In $\mathscr{S}_2$, x_1^0 is chosen in greedy algorithm and x_2^0 is equaly chosen.
There is no way to change this situation without changing the status of the first edge x_1^0 : as long as $c_1^{01} \geq c_2^{01}$, x_1^0 is chosen during greedy algorithm and x_2^0 is not chosen during greedy algorithm on $\mathscr{S}_1$; and x_1^0 and x_2^0 are both chosen during greedy algorithm on $\mathscr{S}_2$. As soon as $c_1^{01} < c_2^{01}$, x_1^0 is not chosen any longer during greedy algorithm on $\mathscr{S}_1$. We see that as long as $c_2^{01} > c_1^{01}$ and in consequence that x_1^0 is not chosen while x_2^0 is chosen in $\mathscr{S}_1$, then $c_2^{02} < c_1^{02}$ with both x_1^0 and x_2^0 chosen in $\mathscr{S}_2$. So we decrease c_1^{02} down to change the

status of x_1^0 in$\mathscr{S}_2$: not chosen or down to zero. But it is not possible to decrease to zero unless that $c_2^{02} = 0$ so that $c_1^{01} \geq c_2^{01}$. We can conclude that there are two possible issues: $c_2^{01} > c_1^{01}$ and x_2^0 chosen while x_1^0 not chosen in $\mathscr{S}_1$, while $c_1^{02} > c_2^{02}$ but x_1^0 not chosen and x_2^0 chosen in $\mathscr{S}_2$, or $c_1^{01} > c_2^{01}$ and $c_2^{02} = 0$ and x_1^0 chosen in both scenarios while x_2^0 not chosen in any scenario.
This dynamical balance shows that a first stage edge can disappear in greedy formulation. The last case is when x_1^0 is not chosen before introducing second edge x_2^0 and is treated exactly in the same way. □

3.5 going towards more than two edges

Some considerations on duality will be developped further to show that with any number of first stage edges, in the case of two scenarios, the system is TDI.

4 Multiple scenarios-The case of more than two scenarios

In this section, we study the case where there exist more than two scenarios in the second stage. We will change our point of view by exhibiting an example to the contrary where a fractional solution for x still compatible with all requirements leads to a higher value than for all integer vectors x. We present a graph where there exist 3 first stage edges not directly connected, and 6 second stage edges.
For all first stage edges, the cost values are $c_1^0 = c_2^0 = c_3^0 = 5$.
In scenario $\mathscr{S}_1$, the cost function for second stage edges is $c_1^1 = c_2^1 = 6$ and $c_3^1 = c_4^1 = c_5^1 = c_6^1 = 0$.
In scenario $\mathscr{S}_2$, the cost function for second stage edges is $c_3^2 = c_4^2 = 6$ and $c_1^2 = c_2^2 = c_5^2 = c_6^2 = 0$.
In scenario $\mathscr{S}_3$, the cost function for second stage edges is $c_5^3 = c_6^3 = 6$ and $c_1^3 = c_2^3 = c_3^3 = c_4^3 = 0$.
There is no integer solution x where it is possible to take all second stage edges with positive strictly cost and strictly more than only one first stage edge, otherwise there would be a cycle (see figure 1). We can afford that best integer value is less or equal than $6*6+5 = 41$.
Now we propose to take $x_1^0 = x_2^0 = x_3^0 = \dfrac{1}{2}$ and in second stage only edges with strictly positive cost: $x_1^1 = x_2^1 = 1$; $x_3^2 = x_4^2 = 1$; $x_5^3 = x_6^3 = 1$. This fractionnal solution gives a positive value of $6 * 6 + 3 * 5 * \dfrac{1}{2} = 43,5$. This clearly shows that system is not TDI.

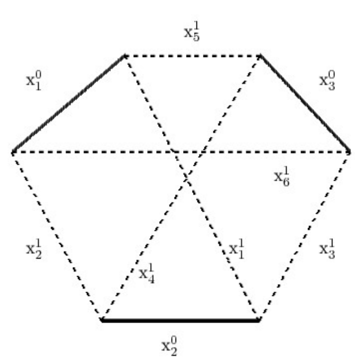

Figure 1: Complete graph for one scenario

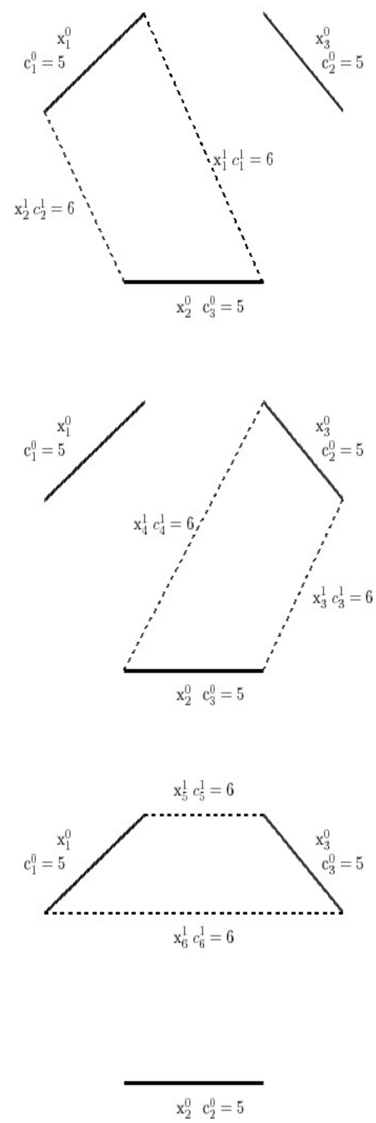

Figure 2: three scenarios with cost function

References

Cook, W. (1983). operations that preserve total dual integrality. *Elsevier Science Publishers B.V*, 2:31–35.

Franck, A. (1981). A weighted matroïd intersection. *Journal of Algorithms*, 2:328–336.

Kong, N., Schaeffer, A., and Shabbir, A. (2007). Totally unimodular programs. *?*, 1.

Nemhauser, G. and Wolsey, L. (1999). *Integer And Combinatorial Optimization*. Wiley-Interscience series in Discrete Mathematics and Optimization.