The Prize Collecting Connected Subgraph Problem - A New NP-Hard Problem arising in Snow Removal Routing

P O Lindberg[1] and Gholamreza Razmara[1]

Linköping University, Linköping, SE-58183, Sweden

Abstract. We consider a new NP-hard optimization problem, the Prize Collecting Connected Subgraph Problem, which appears as a subproblem in routing of snow plows during snow fall. In this problem we have a set of edges in an undirected network, with edge costs and edge times. Moreover there is a time budget. The problem is to find a connected subset (corresponding to a snowplow tour) of minimal cost subject to the budget constraint. This problem can be modeled using flow constraints or introducing valid inequalities. We exemplify computations for the classical Sioux Falls Network

1 Introduction

In an *undirected* network $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ with *edge costs* c_e, nonnegative (i.e. expenditures) or positive (i.e. incomes), and *edge times* t_e, the *Prize Collecting Connected Subgraph Problem (PCCSP)* is to find to find the connected subgraph that maximizes income minus expenditures subject to a time budget. This problem arises as a column generation subproblem in periodic snow removal routing problems, [4].

In this paper we introduce and study the PCCSP. In particular we show that it is NP-hard, by reducing the *Prize Collecting Steiner Tree Problem* to it. We give various formulations of the PCCSP, based on flow constraints or valid inequalities. We further suggest a computational procedure for its solution, and demonstrate it on the Sioux Falls network.

The terminology in the Prize Collecting (PC) area is not very standardized. As to PC Steiner Trees (e.g [3] and the references therein) it is not customary to introduce a budget constraint. For PC TSPs though it seems customary to have some buget/minimal income constraint. Feillet et al [1] term the budget-free PC TSP a *profitable tour* problem, and suggest a terminology for this class of problems. In this terminology the PCCSP would be something like the *Profitable Connected Subgraph Problem with One Additional Constraint.*

2 Background: Snow Removal Routing under Snowfall

Suppose we are given a *directed* network $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ with *arcs* $a \in \mathcal{A}$ corresponding to road segments to be plowed by snow clearing vehicles. For the

case during (a long) snowfall, we want to design a set of cyclic snow clearing routes that can be run perpetually during the snowfall, and that cover all arcs in $\mathcal{A}_p$, the set of arcs to be plowed. To avoid plowing too often, some arcs in some routes may be used for transport. Further, the arcs in $\mathcal{A}_p$ have *minimal time spans* $\underline{t}_a$ between successive plowings, smaller for more important arcs (i.e. road segments).
To be more specific, associate with all arcs plowing costs c_a^P and times t_a^P as well as transportation costs c_a^T and times t_a^T. To streamline the model we will duplicate each arc into one plowing arc and one transportation arc, with corresponding costs and times c_a and t_a.
Let a *route* r be a directed cyclic path in the network, and let the *duration* of r be $t_r =_{df} \sum_{a\in\mathcal{A}} t_a$ and its cost $c_r =_{df} \sum_{a\in\mathcal{A}} c_a$. To be consistent with the minimal time spans, each route will be assigned a *period* $p_r \in \underline{T} = \{\underline{t}_a\}_{a\in\mathcal{A}}$, which is how often it is repeated. To be *feasible* the route must fulfill on the one hand $t_r \leq p_r$, and on the other hand $\underline{t}_a \geq p_r, \forall a \in r$.
Let $\mathcal{R}$ be the set of feasible routes and let $A_{ar} = 1$ if a is plowed in r, and 0 else. Then the problem to find the minimal cost set of routes covering all arcs can be stated as

$$(P)\{\min_x \sum_{r\in\mathcal{R}} c_r x_r \text{ s.t. } \{\sum_{r\in\mathcal{R}} A_{ar}x_r \geq 1, a \in \mathcal{A}_p; x_r \in \{0,1\}, r \in \mathcal{R}$$

Remark. We have used a set covering rather than a set partitioning formulation, since if an arc is covered twice, say, one of the plowings can be changed into a transport, giving a cheaper route.
Since $\mathcal{R}$ typically is too large to be enumerated efficiently, we will have to use column generation. To this end, let the rows in (P) get multipliers λ. (For notational simplicity we let transport arcs have $\lambda_a = 0$.) Thus the reduced cost for route r is $\bar{c}_r = c_r - \sum_a A_{ar}\lambda_a = \sum_{a\in r} c_a - \sum_{a\in r} \lambda_a = \sum_{a\in r} \bar{c}_a$, where $\bar{c}_a = c_a - \lambda_a$ is the reduced cost of a. Thus we have the following

Column Generation Subproblem:
(P_{CG}) Find feasible cyclic route with minimal reduced cost $\bar{c}_r =_{df} \sum_{a\in r} \bar{c}_a$.

3 Subproblems as PCCSPs

In practical operations, all road segments are cleared by the same vehicle in both directions. Thus, we will assume the same. To further simplify the subproblem, we will assume that the same is true also about transports.
Assumption 1. In any route, all road segments plowed or used as transport, are plowed or transported in both directions.
Remark. These assumptions lead to an essentially undirected network. In practice there are small one-way segments. These have to be modeled by undirected edges. The ensuing errors can be made small, though, [4]. These simplifications, however, allow more efficient solution of the simplified subproblems below.
The simplifications just made allow simple representation of cyclic routes.