NEW INSIGHTS ON PRIZE COLLECTING PATH PROBLEMS

Edcarllos Santos, Alfredo Candia-Véjar, Luiz Satoru Ochi, Luidi Simonetti, Uéverton S. Souza[1]

[1]Instituto de Computação – UFF, Niterói, Brasil
Facultad de Ingeniería - Universidad de Talca, Talca, Chile
Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil

Abstract. *Given a graph G and a pair s,t in V(G), where each edge e has a weight t(e) and each vertex v has a value p(v) such that t(e) represent a transportation time and p(v) a prize collecting. Prize Collecting Path (PCP) consists of finding a (s,t)-path that minimizes the total transportation time minus the total prize of nodes in such path. PCP is at core of numerous relevant applications in several fields like telecommunications, transportation and logistics. In this paper, the complexity behavior of the problem is analyzed. For some cases we prove that PCP is NP-complete, these results lead to the generation of new sets of benchmark instances that are computationally hard according to natural characteristics of the problem. In addition, polynomial time algorithms are described for other cases and a mathematical formulation is introduced to solve general instances of PCP.*

Prize Collecting Path problem

Let $G = (V, A)$ be a directed and connected graph, where V is the set of nodes and A is the set of arcs. We assume that there exist n nodes and m arcs. Associated with the set of nodes there is a prize function $\mathbf{p} : V \to \mathbb{R}_{>0}$. Likewise, associated with the set of arcs there is a transportation time function $\mathbf{t} : A \to \mathbb{R}_{>0}$. Node $s \in V$ and $t \in V$ correspond, respectively, to source and target node. Let $\mathcal{P}_{st}$ be the set of all (s, t)-paths in G connecting s and t. The PCP consists on finding a (s, t)-path that minimizes the total transportation time cost minus the total prize of the nodes belonging to the (s, t)-path. PCP can be also defined as the following 0-1 integer programming problem:

$$\min \quad \sum_{(ij)\in A} t_{ij}x_{ij} - \sum_{j\in V} p_j z_j \tag{1}$$

$$\text{s.t} \quad \mathbf{x} \in \mathcal{P}_{st}, \qquad 0 \leq x_{ij} \leq 1 \tag{2}$$

Variables are represented by the binary vectors $\mathbf{x} \in \{0,1\}^{|A|}$ and $\mathbf{z} \in \{0,1\}^{|V|}$; such that $x_{ij} = 1$ if arc $(ij) \in A$ is used by a (s, t)-path and $x_{ij} = 0$ otherwise, and $z_i = 1$ if node $i \in V$ is visited by a (s, t)-path, and $z_i = 0$ otherwise. A feasible (s, t)-path is induced by a vector $\mathbf{x}$ that belongs to the following set:

$$\mathcal{P}_{st} = \left\{ \mathbf{x} \in \{0,1\}^{|A|} \;\middle|\; \begin{array}{c} \sum_{(sj)\in\delta^+(s)} x_{sj} = 1 \\ \sum_{(jh)\in\delta^+(t)} x_{jh} - \sum_{(kj)\in\delta^-(j)} x_{kj} = 0, \forall j \in V \setminus \{s,t\} \\ \sum_{(jt)\in\delta^-(t)} x_{jt} = 1 \end{array} \right\} \tag{3}$$

For a given vector $\mathbf{x} \in \mathcal{P}_{st}$, variables $\mathbf{z}$ are related as follows:

$$z_j \leq \sum_{(ij)\in\delta^-(j)} x_{ij}, \forall j \in V \setminus \{s,t\} \tag{4}$$

$$z_s \geq 1 \tag{5}$$

$$z_t \geq 1 \tag{6}$$

NP-completeness

Definition 1 *Given a graph G, a (s, t)-hamiltonian path is a simple path between two nodes (source and target) that visits each node exactly once. The (s, t)-Hamiltonian Path problem (for short, HP) is about determine whether a given graph contains a Hamiltonian path starting in s and finishing in t.*

Theorem 1.1 *HP $\propto$ PCP.*

Proof. Given a graph $G = (V, E)$ where V is the set of nodes and E is the set of edges. Let nodes $\{s, t\} \in V$, represents the source and target nodes, respectively. We construct an instance G' of PCP as follows: (i) set $G' = G$; (ii) for each edge e of G' we associate a transportation time t_e of value 1; and (iii) for each node v of G' we assign a prize p_v of value 2 ($p_s = p_t = 0$).

($\Rightarrow$) Since Hamiltonian path visits all nodes of a graph, all possible (s, t)-hamiltonian paths in G is composed by $|V|$ nodes and $|V| - 1$ edges, such way in G' the sum of transportation time for edges of these paths is equal $|V| - 1$ and the sum of prizes for nodes is $2(|V| - 2)$. Thus such (s, t)-hamiltonian paths have cost according to Equation 1 equal to $-|V| + 3$ in G'.

($\Leftarrow$) Let $p \in \mathcal{P}_{st}$ be a solution of PCP with $-|V| + 3$ cost in G' and $I_p = V(p) \setminus \{s, t\}$. Since all paths in $\mathcal{P}_{st}$ are simple path, *i.e.*, each node is visited just once. For each node $v \in I_p$, there is exactly one incoming edge $e_i \in E$ used in p. As any node $v \in I_p$ has prize $p_v = 2$, then each node v contribute with $(t_{e_i} - p_v) = -1$ in PCP objective function. Hence p has cost $-|I_p| + 1$, once t has no prize associated. Consequently $-|I_p| + 1 = -|V| + 3$ and $|I_p| = |V| - 2$ which implies that path p is hamiltonian in G. ∎

Corollary 1.2 *PCP is NP-hard.*

Proof. Follows from Theorem 1.1 and the fact that Hamiltonian Path problem is NP-complete [Karp(1972)].

Definition 2 *Given a acyclic digraph $D = (V, A)$ where V is the set of nodes and E is the set of arcs. Let nodes $\{s, t\} \in V$, represents the source and target nodes, respectively. For each arc $a \in A$ there is a arbitrary weight w_a (not necessarily greater than zero). The Shortest Path with Arbitrary Weights problem (SPAW) is about to determine the shortest simple path from s to t in G.*

Theorem 1.3 *PCP $\propto$ SPAW.*

Proof. Given a graph $G = (V, E)$ where $V(G)$ is the set of nodes and $E(G)$ is the set of edges. Let $\{s, t\} \in V(G)$ be nodes representing the source and target vertices, respectively. Each edge $e \in E(G)$ has an associated transportation time t_e, and every node $v \in V(G)$ has an assigned prize p_v, where $p_s = p_t = 0$.

From G we construct an instance G' of SPAW as follows: (i) set $G' = G$; (ii) each undirected edge $e = (u, v) \in E(G)$ is converted in two arcs $a_1 = (u, v), a_2 = (v, u) \in E(G')$ of opposite ways, where the weights w_{a_i} of a_i ($i \in \{1, 2\}$) is defined by $(t_e - p_{h(e)})$, where $p_{h(e)}$ is the prize assigned to the head node of a_i.

($\Rightarrow$) Let $p = s, v_1, v_2, \ldots, v_k, t \in \mathcal{P}_{st}$ be a solution of PCP in G. The cost for path p in G, according to Equation 1, is

$$\sum_{e \in E(p)} t_e - \sum_{v \in V(p)} p_v.$$

In G', the cost of p is defined by

$$\sum_{a \in A(p)} w_a$$

that is equivalent to

$$\sum_{e \in E(p)} (t_e - p_{h(e)})$$

which, by additivity, is equal to

$$\sum_{e \in E(p)} t_e - \sum_{e \in E(p)} p_{h(e)}.$$

As source and target nodes have null prizes then p has the same cost in both problems.

($\Leftarrow$) By construction, any (s, t)-path p of G' is also an (s, t)-path of G, and as shown previously, p has the same cost in both instances. ∎

By Theorem 1.1 and Theorem 1.3 HP $\propto$ PCP $\propto$ SPAW. This gives us a mapping of the complexity of the PCP problem, providing sufficient conditions for the problem becomes polynomial or NP-hard. Figure 1 illustrates the relation of complexity between the problems. More precisely, Theorem 1.1 and Theorem 1.3 implies the following corollaries.

Corollary 1.4 *For any graph class $\mathcal{C}$ such that* HP *on $\mathcal{C}$ is* NP*-hard, the* PCP *problem on $\mathcal{C}$ is also* NP-hard.

Corollary 1.5 *Any instance $\mathcal{I}$ of* PCP *can be solved in polynomial time whether the instance g($\mathcal{I}$) of* SPAW *can be solved in polynomial time, where g returns a digraph constructed as described in the proof of Theorem 1.3.*

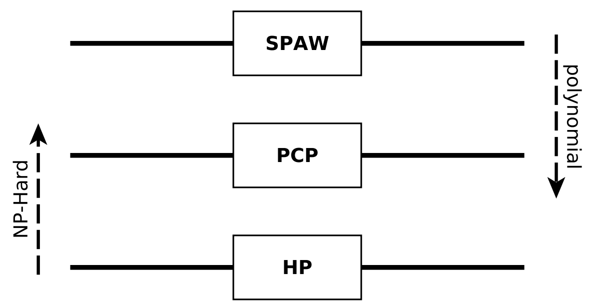

Figure 1. Some considerations for PCP and related problems

Corollary 1.6 PCP *can be solved in polynomial time when the input G does not contain cycles $C = v_1, v_2, \ldots, v_k$ such that $\sum(t_e - p_{v_j}) < 0.e = v_i, v_j$, $1 \leq i \leq k$, $j = ((i+1) \mod (k+1))$.*

Proof. Follows from Lemma 1.5 and the result for minimum path of graphs with no negative cycles.

Now, we study the complexity of the problem in a special graph class.

Grid graphs

Definition 3 *Let G^∞ be the infinite graph whose vertex set consists of all points of the plane with integer coordinates and in which two vertices are connected if and only if the Euclidean distance between them is equal 1. A Grid graph is a node-induced finite subgraph of the infinite grid. It is rectangular if its set of nodes is the product of two intervals.*

Lemma 1.7 [Itai et al.(1982)Itai, Papadimitriou, e Szwarcfiter] *Hamiltonian Path problem in grid subgraphs is NP-complete.*

Corollary 1.8 *(s, t) Longest Path problem in rectangular grid graphs is NP-complete.*

Theorem 1.9 *Prize Collecting Path problem in general grid graphs is NP-complete.*

References

Itai, A., Papadimitriou, C. H., e Szwarcfiter, J. L. (1982). Hamilton Paths in Grid Graphs. *SIAM Journal on Computing*, 11(4):676–686. ISSN 0097-5397.

Karp, R. M. (1972). Reducibility among Combinatorial Problems. In *Complexity of Computer Computations*, p. 85–103. Springer US, Boston, MA.