A Note on “Multistage Methods for Freight Train Classification”

Dirk Briskorn\textsuperscript{1}, Florian Jaehn\textsuperscript{2,*}

November, 2012

\textsuperscript{1} Universität Siegen, Lehrstuhl für Quantitative Planung, Hölderlinstraße 3, D-57076 Siegen, Germany, dirk.briskorn@uni-siegen.de

\textsuperscript{2} Universität Augsburg, Lehrstuhl für Sustainable Operations and Logistics, Universitätstraße 16, D-86159 Augsburg, Germany, florian.jaehn@wiwi.uni-augsburg.de

\textsuperscript{*}Corresponding author, phone +49 821 598-4044, fax +49 821 598-4215.

Keywords: Shunting Yard; Multistage Sorting; Complexity

Abstract

The paper “Multistage Methods for Freight Train Classification” by Jacob et al. ([2]) provides a great insight to the theory and practice of sorting procedures at shunting yards. In [2] many relevant shunting situations (e.g. single or multiple inbound trains, single or multiple outbound trains, (un)restricted number of tracks, (un)restricted track capacity) are formally specified as optimization problems. Then, for almost all of them either an exact polynomial time algorithm or an NP-hardness proof is provided. However, the case of multiple inbound trains, which is of high practical relevance, is left open. We close this gap by providing a proof of NP-hardness.

1 Introduction

We face the situation that $l \in \mathbb{N}$ trains arrive at a shunting yard (see [2] for details on layout and operations) and must be sorted in such a way that one outbound train with a predetermined order of the railcars is formed. W.l.o.g. we can number the railcars according to their position in the outbound train and the total number of railcars is denoted...
by \( n \in \mathbb{N} \). Thus, each inbound train consists of a subset of the railcars from the set \( \{1, \ldots, n\} \) in a given order. The objective we aim at is to minimize the number of so called track-pulls.

In [2] the authors show that a shunting strategy for multiple inbound trains can be seen as a two step procedure where first the inbound trains are concatenated in a certain order and then this single train is treated as a single inbound train. The same authors provide an efficient method for shunting a single inbound train thus the crucial step here is to find the sequence of inbound trains in which they are concatenated. Consequently, we we restrict ourselves to this issue in the following.

The concatenation of the inbound trains is a permutation of \( \{1, \ldots, n\} \). A break marks the occurrence of two consecutive railcars in the wrong partial order in this permutation.

**Definition 1.1**
Given a permutation \( \tau : \{1, \ldots, n\} \to \{1, \ldots, n\} \). A break is pair of consecutive integers (= railcars) \( i \) and \( i + 1 \), \( i \in \{1, \ldots, n - 1\} \) with \( \tau_i > \tau_{i+1} \).

Jacob et al. (Theorem 5 in [2]) show that the number of required track-pulls can directly be derived from the number of breaks in this permutation and the number of required track-pulls is non-decreasing in the number of breaks. Thus, the optimization problem considered in this paper is to find a concatenation of inbound trains such that the number of breaks is minimized. We formally define this problem using a slightly more simplified notation than the one used in [2], being sufficient considering this problem solely.

**Definition 1.2**
Given integers \( n, l \in \mathbb{N} \) and \( l \) sequences \( (T_1, \ldots, T_l) \) such that

- each integer \( i \in \{1, \ldots, n\} \) is contained in exactly one of these sequences and
- each sequence only contains elements from \( \{1, \ldots, n\} \),

\( \text{OPT-PERM} \) is the problem to find a permutation of the sequences such that the number of breaks in the concatenation is minimized.

## 2 Complexity of OPT-PERM

We will now show that \( \text{OPT-PERM} \) is NP-hard by reduction from the minimum feedback arc set problem (MIN-FAS):

**Definition 2.1**
Given a directed graph \( G = (V, E) \) with \( |V| = n \) vertices. For a given bijective function \( \sigma : V \to \{1, \ldots, n\} \), an edge \( (u, v) \in E \) is called a feedback arc, if \( \sigma(v) < \sigma(u) \). The problem to determine \( \sigma \) for a given graph \( G \) such that the number of feedback arcs is minimized is called feedback arc set problem (MIN-FAS).
Min-FAS is known to be NP-hard (see [1]). The minimum number of feedback arcs of a graph $G$ is denoted by $FA(G)$. An important correspondence between special MIN-FAS problems and OPT-PERM was shown by [2], which we adopt in order to keep this paper self-contained.

**Lemma 2.2 (see Lemma 3 in [2])**

There is a one-to-one correspondence between OPT-PERM and MIN-FAS for directed multigraphs the edges of which form an Eulerian path.

Thus, we only have to show that MIN-FAS remains NP-hard, even if the input is a multi-graph containing an Eulerian path, i.e. there is a path in the graph visiting each edge exactly once. Before doing so, we show that MIN-FAS is equivalent to the special case of MIN-FAS in which the input graph contains an Eulerian cycle, i.e. there is an Eulerian path in the graph which is a cycle.

**Lemma 2.3**

MIN-FAS remains NP-hard even for graphs containing an Eulerian cycle.

**Proof:** We reduce MIN-FAS to MIN-FAS on graphs containing an Eulerian cycle. Given an instance $I$ of MIN-FAS specified by an arbitrary graph $G = (V,E)$ we construct an instance $I'$ of MIN-FAS on a graph $G' = (V',E')$ containing an Eulerian cycle. Let the in-degree and the out-degree of each node $v \in V$ be denoted by $deg^-(v)$ and $deg^+(v)$, respectively. We construct a graph $G' = (V',E')$ as follows. Let $V' = V \cup \{v^*\}$ and $E' = E \cup \bar{E}$ such that $\bar{E}$ contains $\max\{0,deg^-(v) - deg^+(v)\}$ arcs from $v$ to $v^*$ and $\max\{0,deg^+(v) - deg^-(v)\}$ arcs from $v^*$ to $v$ for each $v \in V$. In $G'$ the out-degree of each node equals its in-degree and, therefore, $G'$ contains an Eulerian cycle.

Consider an optimum solution to the instance $I'$ given as a sequence of nodes in $V'$. Assume, for simplicity, that the nodes are numbered accordingly, that is the optimum sequence is $(1,\ldots,|V'|)$. Note that then the sequence $(2,\ldots,|V'|,1)$ is also optimal as in-degree and out-degree are the same for vertex 1 (and all other vertices). More generally speaking, each sequence $(k,\ldots,|V'|,1,\ldots,k-1)$ with $k \in V'$ is an optimum solution as well. Thus, there is an optimum sequence $\pi'$ to $I'$ where $v^*$ is the last node of the sequence. Now, consider the subsequence $\pi$ of the first $|V|$ nodes in $\pi'$ which, obviously, constitutes a feasible solution to $I$.

We now show that $\pi$ is an optimum solution to $I$. Consider an optimum solution $\pi^*$ to $I$. Concatenating $v^*$ to $\pi^*$ we obtain a feasible solution to $I'$ and, thus, $FA(G') \leq FA(G) + deg^+(v^*)$. Thus, $FA(G') - deg^+(v^*)$ is a lower bound on the optimum solution value for $I$ which is reached by $\pi$. □

**Theorem 2.4**

OPT-PERM is NP-hard.

**Proof:** MIN-FAS remains NP-hard even for graphs containing an Eulerian cycle and is generalized by MIN-FAS on graphs containing an Eulerian path. The theorem follows from the fact that OPT-PERM is equivalent to MIN-FAS with input graphs containing an Eulerian path (Lemma 3 in [2]). □
Theorem 2.4 closes the open question for the complexity of the problem to find a shunting strategy for multiple inbound trains and single or multiple outbound trains that minimizes the number of track-pulls.

References
