

On a Labeled Vehicle Routing Problem

Hatem Chatti, Laurent Gourvès, and Jérôme Monnot

LAMSADE, CNRS FRE 3234 & Université de Paris-Dauphine
{hatem.chatti, laurent.gourves, monnot}@lamsade.dauphine.fr

Abstract. In this paper, we study the complexity and (in)approximability of the *minimum label vehicle routing* problem. Given a simple complete graph $G = (V, E)$ containing a special vertex 0 called the depot and where the edges are colored (labeled), the minimum label k -vehicle routing problem consists in finding a k -vehicle routing E' , i.e. a collection of cycles of size at most $k + 1$ which all contain the depot 0, and such that every customer $v \in V \setminus \{0\}$ is visited once, minimizing the number of colors used.

1 Introduction

In many graph connectivity problems each edge is associated with a numerical attribute, which may represent length, weight or cost, depending on the related real-life context, and the task is to identify a minimum cost subgraph satisfying given connectivity requirements. In contrast with this standard framework, *labeled optimization* supposes that the set of available edges is partitioned into classes, each of which can be purchased in its entirety or not at all. A convenient representation of such a model couples each edge with a *label*, or color, that specifies its class, and a subset of labels forms a feasible solution when the edges whose labels belong to this subset induce a subgraph satisfying the given connectivity requirements. The objective is to find a solution that optimizes the number of picked labels.

The main fundamental labeled connectivity problems, namely spanning tree, s - t -path, matching, traveling salesman problems have been studied in the literature from a complexity and approximation theories point of view, see for instance [2–4, 9]. In all these labeled problems, if for example every color represents a technology consulted by a different vendor, then we wish to use as few colors as possible, so as to diminish incompatibilities among different technologies.

We are interested in studying the complexity and approximability of a *vehicle routing* problem. Vehicle routing problems that involve the periodic collection and delivery of goods and services such as mail delivery or trash collection are of great practical importance. Usually there is a constraint on the number of customers visited by a vehicle. This constraint reflects the assumption that the vehicle has a finite capacity and that it *collects* from the customers (or *distributes* among them) a commodity.

Simple variants of these real problems can be modeled naturally with graphs. Unfortunately even simple variants of vehicle routing problems are **NP**-hard [1].

For the well known metric k -vehicle routing problem (METRIC k VRP in short), we are given a complete graph K_{n+1} of $(n + 1)$ vertices $\{0, \dots, n\}$ containing a special vertex 0 (the depot), an integer $k \geq 1$ and a distance d between pair of vertices satisfying the triangular inequality. The objective is to find a collection of cycles C_1, \dots, C_p of K_{n+1} minimizing $\sum_{i=1}^k \sum_{e \in C_i} d(e)$ where vertices of $\{1, \dots, n\}$ are visited once by the collection of cycles, each cycle C_i is of size at most $k + 1$ and contains the depot. It is easy to see that 2VRP is polynomial time solvable. For $k \geq 3$, METRIC k VRP was proved **NP**-hard in [6]. In [7], the authors gave a $(\frac{5}{2} - \frac{3}{2k})$ -approximation for METRIC k VRP. In [1], an improvement to $\frac{197}{99}$ is proposed for METRIC 3VRP and some other approximation bounds are presented using the differential measure. To the best of our knowledge, these performance ratios are the best known for any fixed $k \geq 3$.

This paper deals with a labeled version of k VRP. In the *minimum label vehicle routing* problem, n customers have to be served by *vehicles* of limited capacity from a common *depot*. A solution consists of a set of routes, where each starts at the depot and returns there after visiting a subset of customers, such that each customer is visited exactly once. In the model studied in this paper, each route has a label and we seek solutions which use a minimum number of distinct labels.

The problem arises in *multimodal transportation networks* [11]. In such problems, it is desirable to provide a complete service using the minimum number of companies. The multimodal transportation network is represented by a graph where each edge is assigned a label, denoting a different company managing that edge. Another example is the following. Suppose that the customers are distributed over an area (e.g. a map). This area is partitioned into zones which are owned by some entities (e.g. countries). An entity can own several zones. To enter a zone, one needs to get an authorization from its owner. It is assumed that an authorization concerns all zones owned by an entity and not only a subset. One can model the situation as a labeled vehicle routing problem where there are as many labels as entities. A trip between two points (two customers or the depot and a customer) has label ℓ_e if one needs to enter a zone owned by entity e . If each authorization induces a cost or a delay, the goal is to minimize their number. In other words, a routing which minimizes the number of labels is sought. In this paper, we consider that all authorizations have the same cost and every trip has a unique label but natural extensions can be investigated as a future work.

Contribution and organization of the paper. In Section 2 we formally define the labeled vehicle routing problem and give some properties that are often used in the proofs. Some simple polynomially solvable cases are identified in Section 3. Section 4 and 5 are devoted to hardness and inapproximability results which draw a rather complete picture of the complexity of the labeled vehicle routing problem. Before we conclude and list future directions in Section

7, some approximation results are given Section 6. Due to space limitation, some proofs are put in the appendix.

2 Definitions, notations and some properties

Given a simple graph $G = (V, E)$ where $V = \{v_1, \dots, v_n\}$, a *path* P of G is a sequence $P = (v_{r_1}, \dots, v_{r_{k+1}})$ where $[v_{r_i}, v_{r_{i+1}}] \in E$ for $i = 1, \dots, k$ and $v_{r_i} \neq v_{r_j}$ for $i \neq j$. A *cycle* C of G is a sequence $P = (v_{r_1}, \dots, v_{r_{k+1}}, v_{r_1})$ where $[v_{r_i}, v_{r_{i+1}}] \in E$ for $i = 1, \dots, k$, $[v_{r_{k+1}}, v_{r_1}] \in E$ and $v_{r_i} \neq v_{r_j}$ for $1 \leq i < j \leq k+1$. For a path P or cycle C , $V(P)$ and $V(C)$ denote the set of the vertices of P and C respectively. The *length* of a path P or cycle C (denoted by $|P|$ and $|C|$ respectively) is the number of its edges. So, with the previous notations, we get $V(P) = V(C) = \{v_{r_1}, \dots, v_{r_{k+1}}\}$, $|P| = k = |V(P)| - 1$ and $|C| = k + 1 = |V(C)|$.

Given a complete graph K_{n+1} of $(n+1)$ vertices $\{0, \dots, n\}$ containing a special vertex 0 (the depot) and an integer $k \geq 1$, a k -vehicle routing E' of K_{n+1} is a collection of cycles C_1, \dots, C_p of K_{n+1} such that

- (a) $\forall i = 1, \dots, p, 0 \in V(C_i)$,
- (b) $\cup_{i=1}^p V(C_i) = \{0, \dots, n\}$,
- (c) $V(C_i) \cap V(C_j) = \{0\}$ for $1 \leq i < j \leq p$,
- (d) $\forall i = 1, \dots, p, |V(C_i)| \leq k + 1$.

The vertices of $V(K_{n+1}) \setminus \{0\}$ will be called *the customers*. In other words, each vehicle starts at the depot 0, visits at most k customers and returns to the depot. Each customer is visited exactly once. The specific solution $E_0 = \{C_1, \dots, C_n\}$ where $C_i = (0, i, 0)$ will be called the *star* of K_{n+1} and it will be extensively studied in this paper (see Figure 1). When $C_i = (0, i, 0)$ is used, the edge $[0, i]$ and its color will be counted once.

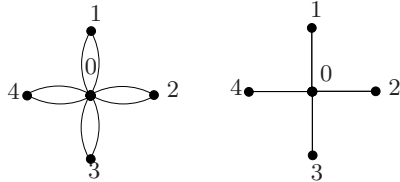


Fig. 1. Two equivalent representations of the star E_0 .

In the *minimum label k -vehicle routing* problem for $k \geq 1$ (LVRP(k) in short), we are given a complete graph K_{n+1} of $(n+1)$ vertices $\{0, \dots, n\}$ containing a special vertex 0 (the depot) and an edge-labeling function $\mathcal{L} : E(K_{n+1}) \rightarrow L = \{\ell_1, \dots, \ell_q\}$. The objective is to find a k -vehicle routing E' of K_{n+1} minimizing

the number of colors used by E' , i.e., $|\mathcal{L}(E')|$, where $\mathcal{L}(E') = \{\mathcal{L}(e) : e \in E'\}$. The *minimum label vehicle routing* problem (LVRP in short) is the restriction of the minimum label k -vehicle routing problem when $k \geq n$.

Given $I = (K_{n+1}, \mathcal{L})$ instance of LVRP(k), the *frequency* of color ℓ_i , denoted by $f(\ell_i)$, is the number of times that color ℓ_i appears in I . For $L' \subseteq L = \{\ell_1, \dots, \ell_q\}$, the *frequency* of L' , denoted by $f(L')$ is the maximum frequency of colors in L' , i.e., $f(L') = \max\{f(\ell) : \ell \in L'\}$. The *frequency* of I is the maximum number of times that a color appears in I , i.e., $f(I)$. Using abusive notations, the frequency of a set E' of edges, denoted by $f(E')$, is the maximum frequency of colors used by E' , i.e., $f(E') = f(\mathcal{L}(E'))$. The restriction of LVRP(k) to instances where the frequency is upper bounded by f is denoted by LVRP(k, f).

Given an instance $I = (K_{n+1}, \mathcal{L})$ of LVRP(k) and a feasible (resp., optimal) k -vehicle routing $E' = \{C_i : i = 1, \dots, p\}$ (resp., $E^* = \{C_i^* : i = 1, \dots, p^*\}$) of I , s'_i (resp., s_i^*) for $i = 2, \dots, k+1$ denotes the number of cycles of E' (resp., E^*) of size exactly i . We have the following properties:

Property 1. We can always assume that:

- (i) $s'_3 = 0$.
- (ii) For each cycle $C_i = (0, x_1, x_2, \dots, x_{p-1}, x_p, 0)$ of size $p+1 \geq 4$ of E' , the color(s) $\mathcal{L}([x_1, x_2])$ and $\mathcal{L}([x_{p-1}, x_p])$ appears at least twice in E' .
- (iii) $s'_2 + \sum_{i=4}^{k+1} (i-1)s'_i = n$.

Obviously Property 1 also holds if we consider the optimal solution E^* .

3 Polynomial cases

We present some simple polynomial cases of LVRP(k, f). In each case, the star E_0 will be an optimal solution.

Theorem 1. *The following cases are polynomial:*

- LVRP(k, f) if $f = 1$ or $k = 1, 2$.
- LVRP(3, 2) if $f(E_0) = 1$.

4 Hardness results

All the hardness results presented here will be done from the k -path partition problem (denoted by k -PPP). In k -PPP, we are given a simple graph $G = (V, E)$ with $|V| = kq$ and we want to decide whether a collection of q vertex-disjoint paths, each of length exactly $k-1$, exists.

This problem is **NP**-complete for any $k \geq 3$, and polynomial otherwise, [5, 8]. More recently, k -PPP has been proved **NP**-complete in bipartite graphs of maximum degree 3 for any $k \geq 3$, [10].

4.1 When $f(E_0) = 1$

From Theorem 1, we know that LVRP(3, 2) when $f(E_0) = 1$ is polynomial (we recall $f(E_0) = 1$ means that for each edge $[0, x]$, the color $\mathcal{L}([0, x])$ appears exactly once in the instance). Now, we prove that it is not the case for LVRP(3, 3).

Theorem 2. LVRP(3, 3) with $f(E_0) = 1$ is **NP-hard**.

When we deal with LVRP(k) with $k \geq 4$, we can obtain stronger hardness results since similar results hold with a frequency equal to 2.

Theorem 3. For any $k \geq 4$, LVRP(k , 2) with $f(E_0) = 1$ is **NP-hard**.

Proof. The proof is very similar to Theorem 2, except that we start from k -PPP and we duplicate the graph twice instead of 3 times. Formally, fix $k \geq 4$ and let $G = (V, E)$ with $|V| = kq$ be an instance of k -PPP. We build an instance $I = (K_{2kq+1}, \mathcal{L})$ of LVRP(k , 2) as follows:

- Color each edge of G with a unique color.
- Make two copies of G , denoted by $G_i = (V_i, E_i)$ for $i = 1, 2$.
- Add a depot 0 and complete the graph into K_{2kq+1} by adding a unique color per missing edge.

Obviously, I is an instance of LVRP(k , 2) with $f(E_0) = 1$.

We claim that G admits a collection $\mathcal{P} = \{P_1, \dots, P_q\}$ of q vertex-disjoint paths with $|P_i| = k - 1$ iff there is a k -vehicle routing E' of I with $|\mathcal{L}(E')| \leq q(k + 3)$.

Let $\mathcal{P} = \{P_1, \dots, P_q\}$ be a collection of paths of G , vertex-disjoint and such that $|P_i| = k - 1$. Consider $E' = \{C_i : i = 1, \dots, 2q\}$ where $C_{i+(j-1)q}$ is the cycle in the j -th copy G_j for $j = 1, 2$, isomorphic to $(0, P_i, 0)$. E' is a feasible solution of I and since $f(E_0) = 1$, $|\mathcal{L}(E')| = |\mathcal{L}(E' \cap E_0)| + |\mathcal{L}(E' \setminus E_0)| = 4q + (k - 1)q = q(k + 3)$.

Conversely, let $E' = \{C_1, \dots, C_r\}$ be a k -vehicle routing of I such that $|\mathcal{L}(E')| \leq q(k + 3)$. Since the frequency of I is 2 and the problem studied is LVRP(k), we deduce that $|\mathcal{L}(E' \setminus E_0)| \geq (\sum_{i=4}^{k+1} (i - 2)s'_i)/2$. On the other hand, by hypothesis, we have $f(E_0) = 1$; so, $|\mathcal{L}(E' \cap E_0)| = s'_2 + 2 \sum_{i=4}^{k+1} s'_i$. Finally, using arguments similar to those given in Theorem 2 (in this case, equality (iii) of Property 1 gives $s'_2 = 2qk - \sum_{i=4}^{k+1} (i - 1)s'_i$), we obtain:

$$|\mathcal{L}(E')| = |\mathcal{L}(E' \setminus E_0)| + |\mathcal{L}(E' \cap E_0)| \geq 2qk - \frac{\sum_{i=4}^{k+1} (i - 4)s'_i}{2} \quad (1)$$

Using inequality (1) and $|\mathcal{L}(E')| \leq q(k + 3)$, we obtain:

$$\sum_{i=4}^{k+1} (i - 4)s'_i \geq 2q(k - 3) \quad (2)$$

Since $(i - 4) \leq k - 3$ for $i \in \{4, \dots, k + 1\}$, we deduce from inequality (2), the bound

$$\sum_{i=4}^{k+1} s'_i \geq 2q \quad (3)$$

On the other hand, since $\sum_{i=4}^{k+1} (i - 4)s'_i = \sum_{i=4}^{k+1} (i - 1)s'_i - 3 \sum_{i=4}^{k+1} s'_i$, by using equality (iii) of Property 1, i.e. $\sum_{i=4}^{k+1} (i - 1)s'_i = 2qk - s'_2$, we obtain:

$$\sum_{i=4}^{k+1} (i - 4)s'_i = 2qk - s'_2 - 3 \sum_{i=4}^{k+1} s'_i \quad (4)$$

Using inequalities (2) and (4), we deduce:

$$6q \geq s'_2 + 3 \sum_{i=4}^{k+1} s'_i \quad (5)$$

From inequalities (3) and (5), we obtain $s'_2 \leq 0$, that is $s'_2 = 0$. Then, we also deduce that $\sum_{i=4}^{k+1} s'_i = 2q$. Using this last equality and $2qk = \sum_{i=4}^{k+1} (i - 1)s'_i$ (given by (iii) of Property 1) we obtain $\sum_{i=4}^{k+1} (i - 1)s'_i = 2qk = k \sum_{i=4}^{k+1} s'_i$ or equivalently $\sum_{i=4}^{k+1} (k + 1 - i)s'_i$. This implies that $s'_i = 0$ for $i = 4, \dots, k$ and then in conclusion, we get:

$$s'_{k+1} = 2q \quad (6)$$

Hence, from equality (6), we deduce that $|\mathcal{L}(E')| = q(k + 3)$. This means that the edges of $E' \setminus E_0$ corresponds to edges of G and in particular $E' \cap E_1$ (the edges of E' in the first copy G_1) is a collection of q vertex-disjoint paths, each of length exactly $k - 1$ in G_1 and then in G . \square

4.2 LVRP(3) with frequency 2

We conclude this section by studying the complexity of LVRP(3) with frequency 2. Hence, we will get a complete description of the complexity of LVRP(k) following the parameters k , $f(E_0)$ and the frequency.

Theorem 4. LVRP(3, 2) is **NP-hard**.

Proof. We polynomially reduce 3-PPP to LVRP(3, 2). Let $G = (V, E)$ with $|V| = 3q$ and $E = \{e_1, \dots, e_m\}$ be an instance of 3-PPP. We build an instance $I = (K_{3q+m+1}, \mathcal{L})$ of LVRP(3, 2) as follows:

- Color each edge e_i of G with a unique color $\mathcal{L}(e_i)$.
- Add m new vertices $3q + i$ for $i = 1, \dots, m$ and set $\mathcal{L}([0, 3q + i]) = \mathcal{L}(e_i)$
- Add a depot 0 and complete the graph into K_{3q+m+1} by adding a unique color per missing edge.

Obviously, I is an instance of $\text{LVRP}(3, 2)$.

We claim that G admits a collection $\mathcal{P} = \{P_1, \dots, P_q\}$ of q vertex-disjoint paths with $|P_i| = 2$ iff there is a 3-vehicle routing E' of I with $|\mathcal{L}(E')| \leq 2q + m$.

Let $\mathcal{P} = \{P_1, \dots, P_q\}$ be a collection of vertex-disjoint paths of G such that $|P_i| = 2$. Consider $E' = \{C_i : i = 1, \dots, q + m\}$ where for $i \leq q$, $C_i = (0, P_i, 0)$ while for $i = q + 1, \dots, q + m$, $C_i = (0, 3q + i, 0)$. E' is a 3-vehicle routing of I and $|\mathcal{L}(E')| = 2q + m$.

Conversely, let $E' = \{C_1, \dots, C_r\}$ be a 3-vehicle routing of I such that $|\mathcal{L}(E')| \leq 2q + m$. Let us prove that the following property holds:

Property 2. We can always assume that $\{(0, 3q + i, 0) : i = q + 1, \dots, q + m\} \subset E'$.

Proof. By contradiction, assume that some vertex $3q + i$ is contained in a cycle C_i of E' with $|C_i| > 2$. Using (i) of Property 1 and the fact that E' is a 3-vehicle routing, we get $|C_i| = 4$. Since the color of every edge incident to $3q + i$, except $\mathcal{L}([0, 3q + i])$ appears once, we obtain a contradiction with (ii) of Property 1. \square

Now, since $|\mathcal{L}(E' \cap E_0)| \leq |\mathcal{L}(E')| \leq 2q + m$ and $|\mathcal{L}((E' \cap E_0) \setminus E)| = m$ (from Property 2 where we recall that E is the edge set of G), we deduce that $|\mathcal{L}(E' \cap E_0 \cap E)| \leq 2q$ because $f(E_0 \cap E) = 1$. On the other hand, $|\mathcal{L}(E' \cap E_0 \cap E)| \geq 2|V|/3 = 2q$ since E' is a 3-vehicle routing. Thus, $|\mathcal{L}(E' \cap E_0 \cap E)| = 2q$, and then every cycle of E' which is in the complete subgraph induced by $V \cup \{0\}$, has a length 4. We also deduce that $|\mathcal{L}(E' \setminus E_0)| = |\mathcal{L}(E' \cap E)| = m$. Thus $E' \setminus E_0$ are edges of G . In conclusion, G admits a collection of q vertex-disjoint paths, each of length exactly 2. \square

5 Inapproximation results

We now present for some value of k , some inapproximation results of $\text{LVRP}(k)$, that is, we produce some lower bounds that the performance ratio of any approximation algorithms can not reach unless $\mathbf{P}=\mathbf{NP}$. For this, we will apply a gap-reduction from $\text{MAX}\mathbf{P}_k\text{PACKING}$. This problem consists, given a simple graph $G = (V, E)$, of finding a maximum number of vertex-disjoint paths of length $k - 1$. In [10], it is proved that $\text{MAX}\mathbf{P}_k\text{PACKING}$, for $k \geq 3$, admits a constant $\varepsilon_k > 0$, such that for every bipartite graph $G = (V, E)$ of maximum degree 3, it is \mathbf{NP} -hard to decide between $\text{opt}(G) = \frac{|V|}{k}$ and $\text{opt}(G) \leq (1 - \varepsilon_k) \frac{|V|}{k}$. Here $\text{opt}(G)$ is the value of a maximum \mathbf{P}_k -Packing on G . All these results hold if $|V|$ is assumed to be even.

Theorem 5. *There is a constant $\varepsilon_3 > 0$, such that for all $\varepsilon > 0$, $\text{LVRP}(3, 2)$ is not $(\frac{13+2\varepsilon_3}{13} - \varepsilon)$ -approximable unless $\mathbf{P}=\mathbf{NP}$.*

The same kind of result holds if we consider $\text{LVRP}(4)$.

Theorem 6. *There is a constant $\varepsilon_4 > 0$, such that for all $\varepsilon > 0$, $\text{LVRP}(4, 2)$ is not $(\frac{7+\varepsilon_4}{7} - \varepsilon)$ -approximable unless $\mathbf{P}=\mathbf{NP}$.*

If we study $\text{LVRP}(k)$ where k depends on the number of customers, we can obtain stronger results. For instance, for the labeled vehicle routing problem LVRP , i.e. without any constraint on the length of each cycle, we prove that LVRP is not $n^{1-\varepsilon}$ -approximable, for all $\varepsilon \in (0; 1)$. On the other hand, any Hamiltonian cycle of K_{n+1} , which is a feasible solution of $I = (K_{n+1}, \mathcal{L})$, guarantees the performance ratio $n + 1$. Indeed the Hamiltonian cycle uses at most $n + 1$ colors while $\text{opt}(I) \geq 1$.

Theorem 7. *For all $\varepsilon \in (0; 1)$, for any $k \geq n^\varepsilon$, $\text{LVRP}(k)$ is not $n^{1-\varepsilon}$ -approximable unless $\mathbf{P}=\mathbf{NP}$, where n is the number of customers.*

Proof. Let $\varepsilon > 0$ and let $G = (V, E)$ be an instance of the Hamiltonian s - t -path problem on a graph with two specified vertices $s, t \in V$ having degree 1 in G . The Hamiltonian s - t -path problem is defined as follows: given a graph $G = (V, E)$ with two specified vertices $s, t \in V$, decide whether G has a Hamiltonian path from s to t (see [5]). The restriction of the Hamiltonian s - t -path problem on graphs where vertices s, t are of degree 1 remains \mathbf{NP} -complete.

Let $p = \lceil \frac{1}{\varepsilon} \rceil - 1$. We construct the following instance I of $\text{LVRP}(k)$ where $k \geq q^\varepsilon$, q is the number of customers of the resulting instance: take a graph consisting of n^p copies of G and add a depot 0, where the i -th copy is denoted by $G_i = (V_i, E_i)$ and s_i, t_i are the corresponding copies of vertices s and t . Set $\mathcal{L}(e) = c_0$ for every $e \in \cup_{i=1}^{n^p} E_i$, $\mathcal{L}([s_i, 0]) = c_0$ and $\mathcal{L}([t_i, 0]) = c_0$ for all $i = 1, \dots, n^p$. Complete this graph by taking a new color per remaining edge. This construction can obviously be done in polynomial time, and the resulting graph has $n^{p+1} + 1$ vertices. Moreover since $q = n^{p+1}$, then $k \geq n$. Let $I = (K_{n^{p+1}+1}, \mathcal{L})$ be the resulting instance of $\text{LVRP}(k)$ and let E^* be an optimal k -vehicle routing of I with value $\text{opt}(I) = |\mathcal{L}(E^*)|$.

- If G has an Hamiltonian s - t -path P , then let P_i be the Hamiltonian s - t -path in G_i . By setting $E^* = \{(0, P_i, 0) : i = 1, \dots, n^p\}$, we get $\text{opt}(I) = 1$.
- Otherwise, G (and then, each copy G_i) has no Hamiltonian path for any pair of vertices since vertices $s, t \in V$ have a degree 1 in G . Hence $\text{OPT}(I) \geq n^p$, because on the one hand, for each copy G_i , there is at least one vertex v_i which is incident to an edge $e_i \in E^*$ with $\mathcal{L}(e_i) \neq c_0$, and on the other hand there are n^p copies.

We deduce that it is \mathbf{NP} -complete to distinguish between $\text{OPT}(I) = 1$ and $\text{OPT}(I) \geq |V(K_{n^{p+1}+1}) \setminus \{0\}|^{1-\frac{1}{p+1}} \geq |V(K_{n^{p+1}+1}) \setminus \{0\}|^{1-\varepsilon}$. \square

6 A simple approximation

We first analyze how far the star E_0 , i.e. the solution where every customer is covered by a cycle of length 2, is from an optimal k -vehicle routing. We provide tight bounds on the approximation ratio. Next we show that local improvements made on the star E_0 lead to better worst case performances.

Theorem 8. *For all $f \geq 2$ and $k \geq 3$, The star E_0 is a $\frac{f(k-2)+1}{k-1}$ -approximation for LVRP(k, f) when $f \leq \frac{k+1}{2}$ and it is a $\frac{f(k-2)+k+1}{k+1}$ -approximation for LVRP(k, f) when $f \geq \frac{k+1}{2}$.*

The following proposition shows the upper bounds on the approximation ratio of the star are tight.

Proposition 1. *Given $f \geq 2$ and $k \geq 4$, the star E_0 is at most a $\frac{f(k-2)+1}{k-1}$ -approximation of the optimum when $f \leq \frac{k+1}{2}$ and at most a $\frac{f(k-2)+k+1}{k+1}$ -approximation of the optimum when $f \geq \frac{k+1}{2}$.*

We recall that the surplus of any k -vehicle routing E'' is defined as $R(E'') = |E'' \cap E_0| - |\mathcal{L}(E'' \cap E_0)|$ where $E_0 = \{(0, v, 0) : v \in V(K_{n+1}) \setminus \{0\}\}$ is the star of I (see Theorem 8). For instance, if $f = 2$, then for any $A \subseteq E(K_{n+1})$, $R(A)$ counts the number of colors ℓ which appears twice in A and such that the two edges of color ℓ are incident to the depot. One can see that the following property holds:

Property 3. For any couple of sets A, B of E , we have:

$$A \subseteq B \Rightarrow R(A) = R(A \cap E_0) \leq R(B \cap E_0) = R(B) \quad (7)$$

Proof. Actually, if $e \notin E_0$ (or $e \in E_0$ and $\mathcal{L}(e) \notin \mathcal{L}(A \cap E_0)$) then $R(A) = R(A \cup \{e\})$; otherwise, $R(A \cup \{e\}) = R(A) + 1$. \square

Now, we focus on a restriction of LVRP($3, f$) for $f \geq 2$ where no two edges incident to the depot have the same color. In other words, there are n colors incident to the depot, or equivalently we assume $R(E_0) = 0$. In the light of Theorems 3 and 4, this restriction remains **NP**-hard for any $f \geq 2$ and for instance, the star E_0 is exactly a $\frac{3}{2}$ -approximation for LVRP($3, 2$) when $R(E_0) = 0$ (see the tightness in Proposition 1). More generally, Theorem 8 states that the star E_0 is, in this case, a $\left(1 + \frac{f}{4}\right)$ -approximation of the optimum. Below we analyze an algorithm which, starting from E_0 , resorts to local improvements to try to reduce the number of labels. A simple (informal) description of the algorithm (called here LOC IMPROV) is the following: Start with $L' := \mathcal{L}(E_0)$. While there exists $\ell \in L'$ such that the set of edges having a color in $L' \setminus \{\ell\}$ contains a 3-vehicle routing, do $L' := L' \setminus \{\ell\}$. Return a 3-vehicle routing E' such that $\mathcal{L}(E') = L'$. Formally, the algorithm maintains a feasible 3-vehicle routing E' which is initialized to E_0 . While it is possible, a label is removed from $\mathcal{L}(E')$ by replacing three cycles of length two by one cycle of length four: Take three nodes u, v and w which are all covered by a cycle of length 2 in E' , then do $E' \leftarrow E' \cup \{[u, v], [v, w]\} \setminus \{[0, v]\}$ if $\mathcal{L}(E') \supset \mathcal{L}(E' \cup \{[u, v], [v, w]\} \setminus \{[0, v]\})$. At the end, the algorithm returns a 3-vehicle routing E' , a local minimum, which uses $|\mathcal{L}(E')|$ labels. LOC IMPROV is clearly polynomial and it provides a $\left(1 + \frac{f}{4} \times \frac{f+3}{f+4}\right)$ -approximation of the optimum for LVRP($3, f$) when $R(E_0) = 0$.

Theorem 9. For every $f \geq 2$, LOC IMPROV is a $\left(1 + \frac{f(f+3)}{4(f+4)}\right)$ -approximation for LVRP(3, f) when $R(E_0) = 0$.

Proof. Let $f \geq 2$ and let $I = (K_{n+1}, \mathcal{L})$ be an instance of LVRP(3, f) such that $R(E_0) = 0$. Let $E' = \mathcal{C}'_2 \cup \mathcal{C}'_4$ and $E^* = \mathcal{C}^*_2 \cup \mathcal{C}^*_4$ be the approximate and an optimal solutions respectively where \mathcal{C}'_i and \mathcal{C}^*_i for $i = 2, 4$ are the cycles of size i of E' and E^* respectively (using (i) of Property 1, we know that $\mathcal{C}^*_3 = \mathcal{C}'_3 = \emptyset$). Moreover, using the previous notations, we have $s'_i = |\mathcal{C}'_i|$ and $s^*_i = |\mathcal{C}^*_i|$ for $i = 2, 4$. By construction, the algorithm may only delete some colors which appear once in the star E_0 ; hence, we get $|\mathcal{L}(E')| = |\mathcal{L}(E' \cap E_0)|$. Since, by hypothesis, $R(E_0) = 0$, inequality (7) of Property 3 leads to the conclusion that $R(E'') = 0$ for every $E'' \subseteq E(K_{n+1})$. In particular, $R(E') = R(E^*) = 0$. Thus, $|\mathcal{L}(E' \cap E_0)| = |E' \cap E_0| = n - s'_4$ and we obtain:

$$\text{apx}(I) = |\mathcal{L}(E' \cap E_0)| = n - s'_4 \quad (8)$$

Concerning the optimal solution E^* , let $\mathcal{C}^*_{4,0} = \{C \in \mathcal{C}^*_4 : \mathcal{L}(C) \subseteq \mathcal{L}(E_0)\}$ and $\mathcal{C}^*_{4,1} = \mathcal{C}^*_4 \setminus \mathcal{C}^*_{4,0}$. This means that for each cycle $C = (0, x, y, z, 0) \in \mathcal{C}^*_{4,1}$, we have $\mathcal{L}([x, y]) \not\subseteq \mathcal{L}(E_0)$ or $\mathcal{L}([y, z]) \not\subseteq \mathcal{L}(E_0)$ or both. Let $s^*_{4,i} = |\mathcal{C}^*_{4,i}|$ for $i = 0, 1$; by construction, $s^*_4 = s^*_{4,0} + s^*_{4,1}$. We have the following inequality:

$$|\mathcal{L}(E^*) \setminus \mathcal{L}(E_0)| \geq \frac{s^*_{4,1}}{f} \quad (9)$$

Actually, $|\mathcal{L}(E^*) \setminus \mathcal{L}(E_0)|$ is the number of colors of the cycles of \mathcal{C}^*_4 which do not belong to $\mathcal{L}(E_0)$. By construction, these cycles are in $\mathcal{C}^*_{4,1}$ and each one has at least one edge with a color in $\mathcal{L}(E^*) \setminus \mathcal{L}(E_0)$. Finally, since each color appears at most f times, the result follows.

On the other hand, we also have $|\mathcal{L}(E^*) \cap \mathcal{L}(E_0)| \geq |\mathcal{L}(E^* \cap E_0)| = |E^* \cap E_0| = n - s^*_4$. Hence, using this latter inequality, inequality (9) and $\text{opt}(I) = |\mathcal{L}(E^*) \setminus \mathcal{L}(E_0)| + |\mathcal{L}(E^*) \cap \mathcal{L}(E_0)|$, we obtain:

$$\text{opt}(I) \geq n - s^*_4 + \frac{1}{f}s^*_{4,1} \quad (10)$$

Since E^* contains at least $4s^*_4$ edges and since each color appears at most f times, we deduce that:

$$\text{opt}(I) \geq \frac{4}{f}s^*_4 = \frac{4}{f}(s^*_{4,0} + s^*_{4,1}) \quad (11)$$

Now, we prove the main inequality:

Property 4. The following inequality holds:

$$s^*_{4,0} \leq (f+4)s'_4 \quad (12)$$

Using equality (8) and inequalities (10), (11), (12), we obtain:

$$\begin{aligned}
 apx(I) &= n - s'_4 \\
 &\leq n - \frac{s_{4,0}^*}{f+4} \\
 &\leq \left(n - s_{4,0}^* + \frac{1}{f} s_{4,1}^* \right) + \frac{(f+3)}{f+4} s_{4,0}^* + \frac{(f-1)}{f} s_{4,1}^* \\
 &\leq opt(I) + \frac{f+3}{f+4} (s_{4,0}^* + s_{4,1}^*) \\
 &\leq opt(I) + \frac{f}{4} \times \frac{f+3}{f+4} opt(I) \\
 &\leq \left(1 + \frac{f(f+3)}{4(f+4)} \right) opt(I)
 \end{aligned}$$

The result follows. □

Below, an instance proving that E' is exactly a $\left(1 + \frac{f(f+3)}{4(f+4)} \right)$ -approximation for LVRP(3, f) when $R(E_0) = 0$ for $f = 2$. In Figure 3, only the edges of $E' \cup E^*$ are indicated. We complete the graph by adding a new color for each missing edge.

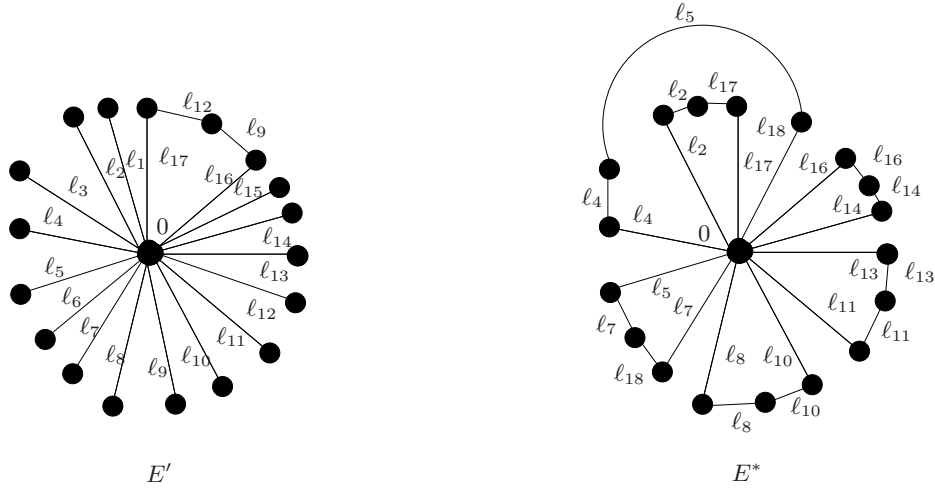


Fig. 2. An instance of LVRP(3, 2) with $R(E_0) = 0$ where E' is a $\frac{17}{12}$ -approximation of the optimum. The value of approximate solution E' is $apx = |\mathcal{L}(E')| = 17$ where $\mathcal{L}(E') = \{\ell_1, \dots, \ell_{17}\}$ while the value of optimal solution E^* is $opt = |\mathcal{L}(E^*)| = 12$ where $\mathcal{L}(E^*) = \{\ell_2, \ell_4, \ell_5, \ell_7, \ell_8, \ell_{10}, \ell_{11}, \ell_{13}, \ell_{14}, \ell_{16}, \ell_{17}, \ell_{18}\}$.

7 Conclusion

The results presented in this article give a good picture of the computational complexity of the problem. Indeed LVRP(k, f) is polynomial when $k = 1, 2$ or $f = 1$. LVRP(3, 2) is polynomial if $f(E_0) = 1$ and **NP**-hard otherwise. In addition

LVRP($k, 2$) for $k > 3$ and LVRP($3, f$) for $f > 2$ are **NP**-hard, even if $f(E_0) = 1$. Without any bound on f , the problem is - from an approximability point of view - closed since the approximation guarantee of any Hamiltonian cycle is almost the best we can expect.

In this paper we provide a non trivial analysis of simple approximation solutions like the star (or minimal solutions with respect to the colors used in some particular cases) but it would be interesting to investigate more elaborate approximation algorithms.

As a future work, it would be interesting to study the case when every label has a weight. The goal is to minimize the total weight of the labels used by a feasible solution, not the cardinality.

In another related problem, there is a known positive cost $c_{i,j}$, called *reload cost*, associated with every change from label i to label j . The goal is not the minimization of the number of labels but to minimize the sum of reload costs induced by the k -vehicle routing. Reload costs captures situations where the travel time crucially depends on the number of times the transportation mode is changed.

References

1. C. Bazgan, R. Hassin, and J. Monnot. Approximation algorithms for some routing problems. *Discrete Applied Mathematics*, 146:3–26, 2005.
2. H. Broersma and X. Li. Spanning trees with many or few colors in edge-colored graphs. *Discussiones Mathematicae Graph Theory*, 17(2):259–269, 1997.
3. R.-S. Chang and S.-J. Leu. The minimum labeling spanning trees. *Information Processing Letters*, 63(5):277–282, 1997.
4. B. Couetoux, L. Gourvès, J. Monnot, and O. Telelis. On labeled traveling salesman problems. In *Proceedings of ISAAC. 08, LNCS 5369*, pages 776–787, 2008.
5. M. R. Garey and D. S. Johnson. *Computers and intractability. A guide to the theory of NP-completeness*. CA, Freeman, 1979.
6. M. Haimovich and A. H. G. Rinnooy Kan. Bounds and heuristics for capacitated routing problems. *Mathematics of Operations Research*, 10:527–542, 1985.
7. M. Haimovich, A. H. G. Rinnooy Kan, and L. Stougie. *Analysis of Heuristics for Vehicle Routing Problems*, chapter in *Vehicle Routing Methods and Studies* (Golden, Assad editors), pages 47–61. Elsevier, 1988.
8. D. G. Kirkpatrick and P. Hell. On the completeness of a generalized matching problem. In *Proceedings of STOC'78*, pages 240–245, 1978.
9. J. Monnot. The labeled perfect matching in bipartite graphs. *Information Processing Letters*, 96:81–88, 2005.
10. J. Monnot and S. Toulouse. The path partition problem and related problems in bipartite graphs. *Operations Research Letters*, 35:677–684, 2007.
11. R. Van-Nes. *Design of multimodal transport networks: A hierarchical approach*. PhD thesis, Delft University Press, 2002.

Appendix

Property 1. *We can always assume that:*

- (i) $s'_3 = 0$.
- (ii) For each cycle $C_i = (0, x_1, x_2, \dots, x_{p-1}, x_p, 0)$ of size $p + 1 \geq 4$ of E' , the color(s) $\mathcal{L}([x_1, x_2])$ and $\mathcal{L}([x_{p-1}, x_p])$ appears at least twice in E' .
- (iii) $s'_2 + \sum_{i=4}^{k+1} (i-1)s'_i = n$.

Proof. For (i). By replacing each cycle $C_i = (0, x, y, 0)$ (of size 3) of E' by the two cycles $C'_x = (0, x, 0)$ and $C'_y = (0, y, 0)$, we obtain a new k -vehicle routing E'' with $|\mathcal{L}(E'')| \leq |\mathcal{L}(E')|$. In E'' , we get $s''_3 = 0$.

For (ii). Let $C_i = (0, x_1, x_2, \dots, x_{p-1}, x_p, 0)$ of size $p + 1 \geq 4$ of E' such that $\mathcal{L}([x_1, x_2])$ or $\mathcal{L}([x_{p-1}, x_p])$ appears exactly once in E' ; wlog., assume that it is $\mathcal{L}([x_1, x_2])$. By replacing C_i by the two cycles $C'_x = (0, x_1, 0)$ and $C_i = (0, x_2, \dots, x_{p-1}, x_p, 0)$, we obtain a new k -vehicle routing E'' with $|\mathcal{L}(E'')| \leq |\mathcal{L}(E')|$.

For (iii). By using properties (a) to (d) of a k -vehicle routing, we deduce that the quantity $\sum_{i=2}^{k+1} (i-1)s'_i$ counts the number of customers of I (i.e., $|V(K_{n+1}) \setminus \{0\}| = n$). Property 1 (i) completes the proof. \square

Theorem 1. *The following cases are polynomial:*

- LVRP(k, f) if $f = 1$ or $k = 1, 2$.
- LVRP(3, 2) if $f(E_0) = 1$.

Proof. The results are straightforward when we deal with LVRP($k, 1$) or LVRP(1, f) since the star E_0 is the unique optimal solution. Using Property 1 (i), we deduce that LVRP(2, f) is polynomially solvable. Consider now an instance $I = (K_{n+1}, \mathcal{L})$ of LVRP(3, 2) with $f(E_0) = 1$. We recall that $f(E_0) = 1$ means that the color of edge $[0, i]$ appears exactly once in I , for every $i \in \{1, \dots, n\}$. In particular, we get $|\mathcal{L}(E_0)| = n$. Let E^* be an optimal k -vehicle routing of I . By hypothesis, we have $|\mathcal{L}(E^*)| = |\mathcal{L}(E^* \cap E_0)| + |\mathcal{L}(E^* \setminus E_0)| = s_2^* + 2s_4^* + |\mathcal{L}(E^* \setminus E_0)|$. Using Property 1 (ii) and $f = 2$, we deduce that $|\mathcal{L}(E^* \setminus E_0)| = s_4^*$. Thus, $|\mathcal{L}(E^*)| = s_2^* + 3s_4^* = n$ by Property 1 (iii). In conclusion, the star E_0 is an optimal solution. \square

Theorem 2. LVRP(3, 3) with $f(E_0) = 1$ is **NP-hard**.

Proof. We polynomially reduce 3-PPP to LVRP(3, 3) with $f(E_0) = 1$. Let $G = (V, E)$ with $|V| = 3q$ be an instance of 3-PPP. We build an instance $I = (K_{9q+1}, \mathcal{L})$ of LVRP(3, 3) as follows:

- Color each edge of G with a unique color.
- Make 3 copies of G denoted by $G_i = (V_i, E_i)$ for $i = 1, 2, 3$.

- Add a depot 0 and complete the graph into K_{9q+1} by adding a unique color per missing edge.

Note that the color of an edge in a copy of G appears 3 times (one time in each copy). The color of any other edge appears once. In particular, each edge incident to the depot appears once, i.e. $f(E_0) = 1$. Thus, I is an instance of LVRP(3, 3) with $f(E_0) = 1$.

We claim that G admits a collection $\mathcal{P} = \{P_1, \dots, P_q\}$ of q vertex-disjoint paths with $|P_i| = 2$ iff there is a 3-vehicle routing E' of I with $|\mathcal{L}(E')| \leq 8q$.

Let $\mathcal{P} = \{P_1, \dots, P_q\}$ be a collection of paths of G , vertex-disjoint and such that $|P_i| = 2$. Consider $E' = \{C_i : i = 1, \dots, 3q\}$ where $C_{i+(j-1)q}$ is the cycle in the j -th copy G_j for $j = 1, 2, 3$, isomorphic to $(0, P_i, 0)$. E' is a feasible solution of I and by construction of I , $|\mathcal{L}(E')| = |\mathcal{L}(E' \cap E_0)| + |\mathcal{L}(E' \setminus E_0)| = 6q + 2q = 8q$.

Conversely, let $E' = \{C_1, \dots, C_r\}$ be a 3-vehicle routing of I such that $|\mathcal{L}(E')| \leq 8q$. Since the frequency of I is 3 and the problem studied is LVRP(3), we deduce that $|\mathcal{L}(E' \setminus E_0)| \geq 2s'_4/3$. On the other hand, by hypothesis, we have $f(E_0) = 1$; so, $|\mathcal{L}(E' \cap E_0)| = s'_2 + 2s'_4$ and $\mathcal{L}(E' \setminus E_0) \cap \mathcal{L}(E' \cap E_0) = \emptyset$. In conclusion, from $|\mathcal{L}(E')| = |\mathcal{L}(E' \cap E_0)| + |\mathcal{L}(E' \setminus E_0)|$, the two previous inequalities and equality (iii) of Property 1 where $n = 9q$, we get:

$$|\mathcal{L}(E')| \geq 9q - \frac{s'_4}{3} \quad (13)$$

Using inequality (13) and $|\mathcal{L}(E')| \leq 8q$, we obtain $s'_4 \geq 3q$. On the other hand, since the instance I has $9q$ customers, we also have $3s'_4 \leq 9q$. Thus,

$$s'_4 = 3q \quad (14)$$

From equality (14), inequality (13) and $|\mathcal{L}(E')| \leq 8q$, we deduce that $s'_2 = 0$, $|\mathcal{L}(E')| = 8q$ and then $|\mathcal{L}(E' \setminus E_0)| = 2q$. In particular, each cycle C'_i (for $i = 1, \dots, r = s'_4$) of E' is of length 4 and each edge $[x, y] \in E' \setminus E_0$ has a frequency 3 in E' . This means that the edges of $E' \setminus E_0$ correspond to edges of G and in particular $E' \cap E_1$ (the edges of E' in the first copy G_1) is a collection of q vertex-disjoint paths, each of length exactly 2 in G_1 and then in G . \square

Theorem 5. *There is a constant $\varepsilon_3 > 0$, such that for all $\varepsilon > 0$, LVRP(3, 2) is not $(\frac{13+2\varepsilon_3}{13} - \varepsilon)$ -approximable unless $\mathbf{P}=\mathbf{NP}$.*

Proof. We prove that a slightly modification of the reduction given in Theorem 4 is indeed a gap-reduction. Let $G = (V, E)$ be a bipartite graph of maximum degree 3 on n vertices (with n even) and m edges, instance of MAX \mathbf{P}_3 PACKING. Notice that by hypothesis, we get $2m \leq 3n$ since the maximum degree of G is 3. We apply the same transformation as the one given in Theorem 4. Then, we add $3n/2 - m$ new vertices and we complete this graph by adding a new color per missing edge. Let $I = (K_{5n/2+1}, \mathcal{L})$ be the resulting instance of LVRP(3, 2). We mainly prove that if $opt(G)$ is the value of a maximum \mathbf{P}_3 -Packing on G

and if $opt(I)$ is the value of an optimal 3-vehicle routing of I , then we have the following equality:

$$opt(I) = \frac{5n}{2} - opt(G) \quad (15)$$

Let $\mathcal{P}^* = \{P_1^*, \dots, P_r^*\}$ be an optimal \mathbf{P}_3 -Packing on G (with $|P_i^*| = 2$); thus $opt(G) = r$. We build a 3-vehicle routing E' of I containing the cycles $(0, P_i^*, 0)$ for $i = 1, \dots, r$ and the cycles $(0, x, 0)$ for the vertices not belonging to $V(\cup_{i=1}^r P_i^*)$. We deduce that $opt(I) \leq |\mathcal{L}(E')| = |\mathcal{L}(E' \cap E_0)| = 2r + (n - 3r) + 3n/2 = 5n/2 - opt(G)$.

Conversely, let $E^* = \{C_1^*, \dots, C_q^*\}$ be an optimal 3-vehicle routing of I with value $opt(I) = |\mathcal{L}(E^*)|$. Obviously, E^* satisfies (ii) of Property 1 and Property 2. This means that if $\{C_1^*, \dots, C_r^*\}$ are the cycles of size 4 of E^* with $C_i^* = (0, P_i, 0)$ then the colors of edges of $\cup_{i=1}^r P_i$ appears twice in E^* ; thus, we deduce that on the one hand $\mathcal{L}(\cup_{i=1}^r P_i) \subseteq \mathcal{L}(E^* \cap E_0)$ and on the other hand $\cup_{i=1}^r P_i \subseteq E$, the edge set of G . Hence, $\mathcal{P} = \{P_1, \dots, P_r\}$ is a \mathbf{P}_3 -Packing on G and we get $opt(G) \geq r$. We also have $opt(I) = 2r + (n - 3r) + 3n/2 = 5n/2 - r \geq 5n/2 - opt(G)$.

Using the two previous inequalities, equality (15) holds. Now, we know that there exists a constant $\varepsilon_3 > 0$, such that it is **NP**-hard to decide between $opt(G) = \frac{n}{3}$ and $opt(G) \leq (1 - \varepsilon_3)\frac{n}{3}$. Thus:

- If $opt(G) = \frac{n}{3}$, then using equality (15), we get $opt(I) = \frac{13n}{6}$.
- If $opt(G) \leq (1 - \varepsilon_3)\frac{n}{3}$, then from equality (15), we obtain $opt(I) \geq \frac{(13+2\varepsilon_3)n}{6}$.

The result follows. \square

Theorem 6. *There is a constant $\varepsilon_4 > 0$, such that for all $\varepsilon > 0$, LVRP(4, 2) is not $(\frac{7+\varepsilon_4}{7} - \varepsilon)$ -approximable unless $\mathbf{P}=\mathbf{NP}$.*

Proof. We prove that the reduction given in Theorem 3 for $k = 4$ is indeed a gap-reduction. Let $G = (V, E)$ be a bipartite graph of maximum degree 3 on n vertices, instance of $\text{MAX}\mathbf{P}_4\text{PACKING}$. We apply the same transformation as the one given in Theorem 4. Let $I = (K_{2n+1}, \mathcal{L})$ be the resulting instance of LVRP(4, 2). We mainly prove that if $opt(G)$ is the value of a maximum \mathbf{P}_4 -Packing on G and if $opt(I)$ is the value of an optimal 4-vehicle routing of I , then we have the following equality:

$$opt(I) = 2n - opt(G) \quad (16)$$

Let $\mathcal{P}^* = \{P_1^*, \dots, P_r^*\}$ be an optimal \mathbf{P}_4 -Packing on G (with $|P_i^*| = 3$) with value $opt(G) = r$. For every $i = 1, \dots, r$, let $P_i^{*,j}$ be the copy of path P_i^* in the j -th copy G_j where $j = 1, 2$. We build a 4-vehicle routing E' of I containing the cycles $(0, P_i^{*,j}, 0)$ for $i = 1, \dots, r$ and $j = 1, 2$. E' also contains the cycles $(0, x, 0)$ for the vertices not belonging to $V(\cup_{i=1}^r (P_i^{*,1} \cup P_i^{*,2}))$. By construction, we get

$|\mathcal{L}(E' \cap E_0)| = 4r + 2(n - 4r)$ and $|\mathcal{L}(E' \setminus E_0)| = 3r$. Hence, since $f(E_0) = 1$, we deduce $\text{opt}(I) \leq |\mathcal{L}(E')| = |\mathcal{L}(E' \cap E_0)| + |\mathcal{L}(E' \setminus E_0)| = 2n - r = 2n - \text{opt}(G)$.

Conversely, let $E^* = \{C_1^*, \dots, C_q^*\}$ be an optimal 4-vehicle routing of I with value $\text{opt}(I) = |\mathcal{L}(E^*)|$. We say that a cycle C_i^* of E^* is *good* if $|C_i^*| = 5$ and every color of $\mathcal{L}(C_i^* \setminus E_0)$ appears twice in E^* . Let $s_{5,1}^*$ be the number of good cycles. Using (ii) of Property 1 (with $p = 4$), we know that $s_5^* - s_{5,1}^*$ is the number of cycles C_i^* of size 5 where exactly 3 edges of C_i^* have their colors which appear once in E^* (actually, it is the two edges of $C_i^* \cap E_0$ and the middle edge of the path $C_i^* \setminus E_0$). Thus, on the one hand using (i) of Property 1 we get $|\mathcal{L}(E^* \cap E_0)| = 2s_5^* + 2s_4^* + s_2^*$ since $f(E_0) = 1$ and on the other hand, $|\mathcal{L}(E^* \setminus E_0)| = (s_5^* - s_{5,1}^*) + (2(s_5^* - s_{5,1}^*) + 3s_{5,1}^* + 2s_4^*)/2 = 2s_5^* + s_4^* - s_{5,1}^*/2$. Finally, using (iii) of Property 1, that is $4s_5^* + 3s_4^* + s_2^* = 2n$, and since $|\mathcal{L}(E^*)| = |\mathcal{L}(E^* \cap E_0)| + |\mathcal{L}(E^* \setminus E_0)|$ because $f(E_0) = 1$, we get that:

$$\text{opt}(I) = 2n - \frac{s_{5,1}^*}{2} \quad (17)$$

The good cycles $C_i^* = (0, P_i, 0)$ are by construction such that $P_i \subseteq E_1 \cup E_2$ where we recall that E_j is the edge set of G_j for $j = 1, 2$. Let r_j for $j = 1, 2$ be the number of good cycles in the copy G_j and assume that $r_1 \geq r_2$. This means that in G we can find r_1 vertex-disjoint paths of length 4. Hence, $\text{opt}(G) \geq r_1 \geq (r_1 + r_2)/2 = s_{5,1}^*/2$. Plugging this last inequality in equality (17), we obtain $\text{opt}(I) \geq 2n - \text{opt}(G)$.

Using the two main previous inequalities, equality (16) holds. Now, we know that there exists a constant $\varepsilon_4 > 0$, such that it is **NP**-hard to decide between $\text{opt}(G) = \frac{n}{4}$ and $\text{opt}(G) \leq (1 - \varepsilon_4)\frac{n}{4}$. Thus:

- If $\text{opt}(G) = \frac{n}{4}$, then using equality (16), we get $\text{opt}(I) = \frac{7n}{4}$.
- If $\text{opt}(G) \leq (1 - \varepsilon_4)\frac{n}{4}$, then from equality (15), we obtain $\text{opt}(I) \geq \frac{(7 + \varepsilon_4)n}{4}$.

The result follows. \square

Theorem 8. *For all $f \geq 2$ and $k \geq 3$, The star E_0 is a $\frac{f(k-2)+1}{k-1}$ -approximation for LVRP(k, f) when $f \leq \frac{k+1}{2}$ and it is a $\frac{f(k-2)+k+1}{k+1}$ -approximation for LVRP(k, f) when $f \geq \frac{k+1}{2}$.*

Proof. Let $I = (K_{n+1}, \mathcal{L})$ be an instance of LVRP(k, f). We produce two different bounds for the performance ratio of the star: $\frac{f(k-2)+1}{k-1}$ and $\frac{f(k-2)+k+1}{k+1}$. The first one is better than the second one iff $f \leq \frac{k+1}{2}$.

Let E^* be an optimal k -vehicle routing of $I = (K_{n+1}, \mathcal{L})$ with value $\text{opt}(I) = |\mathcal{L}(E^*)|$. The approximate solution under consideration is a star, i.e. each node is covered by a cycle of length two.

CASE $f \leq \frac{k+1}{2}$: The number of edges of the optimal solution is $|E^*| = s_2^* + \sum_{i=4}^{k+1} i s_i^*$. Thus we deduce the following bound on the number of colors used by E^* , since each label appears at most f times:

$$f \text{opt}(I) \geq s_2^* + \sum_{i=4}^{k+1} i s_i^* \quad (18)$$

For any solution, E'' , let $R(E'') = |E'' \cap E_0| - |\mathcal{L}(E'' \cap E_0)|$. This quantity indicates the *surplus* of labels of $\mathcal{L}(E'' \cap E_0)$. If label $\ell \in \mathcal{L}(E'' \cap E_0)$ appears $p \geq 1$ times in $E'' \cap E_0$, then its surplus is $p - 1$. Since $\text{opt}(I) \geq |\mathcal{L}(E^* \cap E_0)|$ and $|E^* \cap E_0| = s_2^* + 2 \sum_{i=4}^{k+1} s_i^*$, it follows that

$$\text{opt}(I) \geq s_2^* + 2 \sum_{i=4}^{k+1} s_i^* - R(E^*) \quad (19)$$

Since $|\mathcal{L}(E_0)| \leq |\mathcal{L}(E^* \cap E_0)| + |\mathcal{L}(E_0 \setminus E^*)|$, $|\mathcal{L}(E_0 \setminus E^*)| \leq |E_0 \setminus E^*|$ and $|E_0| = n = s_2^* + \sum_{i=4}^{k+1} (i-1) s_i^*$ (from (i) of Property 1), we get

$$|\mathcal{L}(E_0)| \leq s_2^* + \sum_{i=4}^{k+1} (i-1) s_i^* - R(E^*) \quad (20)$$

Let $\alpha = f \frac{k-2}{k-1}$ and $\beta = \frac{1}{k-1}$. It is not difficult to see that $0 < \alpha$ and $0 < \beta$ since $k \geq 3$. Then, adding inequality (18) multiplied by α to inequality (19) multiplied by β we obtain:

$$(\alpha + \beta) \text{opt}(I) \geq \left(\frac{\alpha}{f} + \beta\right) s_2^* + \sum_{i=4}^{k+1} \left(\frac{\alpha}{f} + 2\beta\right) s_i^* - \beta R(E^*) \quad (21)$$

Since $\frac{\alpha}{f} + \beta \geq 1$, $\frac{\alpha}{f} + 2\beta \geq i - 1$ for all $i \in \{4, \dots, k+1\}$ and $\beta \leq 1$, we get

$$\left(\frac{\alpha}{f} + \beta\right) s_2^* + \sum_{i=4}^{k+1} \left(\frac{\alpha}{f} + 2\beta\right) s_i^* - \beta R(E^*) \geq s_2^* + \sum_{i=4}^{k+1} (i-1) s_i^* - R(E^*)$$

Using this inequality with inequalities (21) and (20), we obtain

$$\frac{f(k-2)+1}{k-1} \text{opt}(I) = (\alpha + \beta) \text{opt}(I) \geq |\mathcal{L}(E_0)|.$$

CASE $f \geq \frac{k+1}{2}$: It follows from inequality (18) that

$$f \text{opt}(I) \geq \sum_{i=4}^{k+1} i s_i^* \quad (22)$$

because $s_2^* \geq 0$. Let $\eta = |\mathcal{L}(E^* \cap E_0)|$. We deduce that

$$\text{opt}(I) \geq \eta \quad (23)$$

and

$$|\mathcal{L}(E_0)| \leq \eta + \sum_{i=4}^{k+1} (i-3) s_i^* \quad (24)$$

where $\sum_{i=4}^{k+1} (i-3)s_i^* = |E_0 \setminus E^*|$. We have $\sum_{i=4}^{k+1} (i-3)s_i^* = \sum_{i=4}^{k+1} \frac{i(i-3)}{i} s_i^* \leq \frac{k-2}{k+1} \sum_{i=4}^{k+1} i s_i^*$. Using (22), (23) and (24), we get

$$|\mathcal{L}(E_0)| \leq \eta + \frac{k-2}{k+1} \sum_{i=4}^{k+1} i s_i^* \leq \text{opt}(I) + f \frac{k-2}{k+1} \text{opt}(I) = \frac{f(k-2) + k+1}{k+1} \text{opt}(I)$$

and the result follows. \square

Proposition 1. *Given $f \geq 2$ and $k \geq 4$, the star E_0 is at most a $\frac{f(k-2)+1}{k-1}$ -approximation of the optimum when $f \leq \frac{k+1}{2}$ and at most a $\frac{f(k-2)+k+1}{k+1}$ -approximation of the optimum when $f \geq \frac{k+1}{2}$.*

Proof. For each case we describe an instance of LVRP(k, f).

CASE $f \leq \frac{k+1}{2}$: The optimum is composed of $f-1$ cycles of length $k+1$ and $k+1-2f$ pendant edges. We have $k+1-2f \geq 0$ because $f \leq \frac{k+1}{2}$. The cycles are denoted by C_1, \dots, C_{f-1} and the pedant edges are denoted by e_1, \dots, e_{k+1-2f} . The two edges of a cycle C_i which are incident to the depot have the labels ℓ_i and ℓ'_i respectively. We suppose that ℓ_i and ℓ'_i both appear f times in C_i . Each label ℓ''_j , for $j = 1, \dots, k+1-2f$, appears once in each cycle. The label of e_j is ℓ''_j . Finally, suppose that any other edge has a label which appears once in the instance. Remark that each color appear at most f times.

The optimum uses $k+1-2f+2(f-1) = k-1$ labels. The star uses $k-1+(f-1)(k-2) = f(k-2)+1$ labels.

CASE $f \geq \frac{k+1}{2}$: The optimum is composed of f cycles of length $k+1$. The cycles are denoted by C_1, \dots, C_f . Each cycle has one edge with label ℓ_1 to ℓ_{k+1} . Actually all cycles are identical except that the depot is not at the same place. Place the depot arbitrarily for C_1 and suppose w.l.o.g. that the two edges of C_1 which are incident to the depot have colors ℓ_1 and ℓ_2 . For C_2 , the depot is shifted twice so that the two edges incident to the depot have colors ℓ_3 and ℓ_4 . Formally the two edges of a cycle C_i which are incident to the depot have the labels $\ell_{(2i-2 \bmod k+1)+1}$ and $\ell_{(2i-1 \bmod k+1)+1}$ respectively. Finally, suppose that any other edge has a label which appears once in the instance. The optimum uses $k+1$ labels and the star uses $k+1+f(k-2)$ labels since $f \geq \frac{k+1}{2}$. \square

Property 4. *The following inequality holds:*

$$s_{4,0}^* \leq (f+4)s'_4 \quad (25)$$

Proof. By contradiction, suppose that $s_{4,0}^* \geq (f+4)s'_4 + 1$. Now, let us introduce some useful notations in the following: G' is the complete subgraph induced by the vertices incident to C'_2 , V_4 denotes the customers of C'_4 and V'_4 are the customers of V_4 which are not incident to the depot in the subgraph induced by C'_4 (for instance, if $C = (0, x, y, z, 0) \in C'_4$, then $\{x, y, z\} \subseteq V_4$, $y \in V'_4$ and

$x, z \in V_4 \setminus V'_4$). Finally, let $E'_4 = \{[0, v] : v \in V'_4\}$ and $E''_4 = E' \setminus E_0$; thus $|V'_4| = |E'_4| = s'_4$, $|V_4| = 3s'_4$ and $|E''_4| = 2s'_4$. We also partition the cycles of $\mathcal{C}_{4,0}^*$ into $\mathcal{C}_{4,0}^{*,1}$ and $\mathcal{C}_{4,0}^{*,2}$ where each cycle $C \in \mathcal{C}_{4,0}^{*,1}$ contains at least a vertex of V_4 and $\mathcal{C}_{4,0}^{*,2} = \mathcal{C}_{4,0}^* \setminus \mathcal{C}_{4,0}^{*,1}$. In other words, the cycles of $\mathcal{C}_{4,0}^{*,2}$ are completely contained in G' . See Figure 3 for an illustration where the solution proposed is E' .

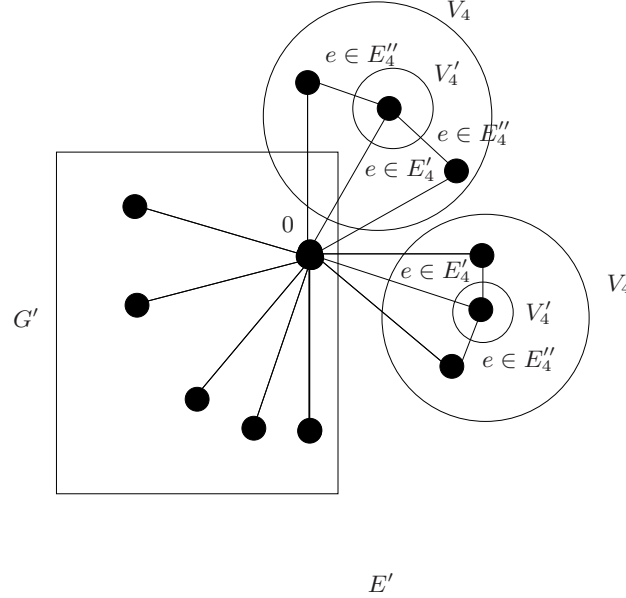


Fig. 3. The solution E' , the sets V_4 , V'_4 , E'_4 , E''_4 and a partial subgraph of the complete subgraph G' induced by the vertices incident to \mathcal{C}'_2 .

With these notations in mind, we have $|\mathcal{C}_{4,0}^{*,1}| \leq |V_4| = 3s'_4$; then using the assumption, we get $|\mathcal{C}_{4,0}^{*,2}| = |\mathcal{C}_{4,0}^*| - |\mathcal{C}_{4,0}^{*,1}| \geq (f+4)s'_4 + 1 - 3s'_4 = (f+1)s'_4 + 1$. How many cycles of $\mathcal{C}_{4,0}^{*,2}$ may contain some colors in $\mathcal{L}(E'_4)$? We have $|\mathcal{L}(E'_4)| = |E'_4| = s'_4$ (since $R(E'_4) = 0$) colors and any color appears at most f times. However, each such a color appears once outside in G' (in E'_4). Thus, at most $(f-1)|\mathcal{L}(E'_4)|$ edges with colors in $\mathcal{L}(E'_4)$ may appear in G' . Since, we are interested in cycles of $\mathcal{C}_{4,0}^{*,2}$ which contain some colors in $\mathcal{L}(E'_4)$, there is at most $(f-1)s'_4$ such cycles. In $\mathcal{C}_{4,0}^{*,2}$, it may also exist some cycles $C = (0, x, y, z) \in \mathcal{C}_{4,0}^{*,2}$ such that $\mathcal{L}([0, y]) \in \mathcal{L}(E''_4)$ where we recall that $E''_4 = E' \setminus E_0$. In the worst case, there are $|E''_4| = 2s'_4$ such edges (and then, such cycles) since by hypothesis $R(E_0) = 0$ (ie., any pair of edges incident to the depot have distinct colors). In conclusion, there is at least one cycle $C = (0, x, y, z, 0)$ of $\mathcal{C}_{4,0}^{*,2}$ included in the subgraph G' such that on the one hand $\mathcal{L}(C) \cap \mathcal{L}(E'_4) = \emptyset$ (in other words $\mathcal{L}(C) \subseteq \mathcal{L}(E')$) and on the hand $\mathcal{L}([0, y]) \notin \mathcal{L}(E''_4)$. Then, the color $\mathcal{L}([0, y])$ appears one time

in $\mathcal{L}(E')$ (since $R(E_0) = 0$). Hence, the algorithm LOC IMPROV should delete color $\mathcal{L}([0, y])$ (and take C), contradiction. Inequality (25) is proved.