An efficient label-setting Algorithm for the Bi-Objective Shortest Path problem
Antoine Giret, Emmanuel Neron, Gaël Sauvanet, Yannick Kergosien

To cite this version:
Antoine Giret, Emmanuel Neron, Gaël Sauvanet, Yannick Kergosien. An efficient label-setting Algorithm for the Bi-Objective Shortest Path problem. 5th International Conference on Operations Research and Enterprise Systems, 2016, Rome, France. hal-03018287

HAL Id: hal-03018287
https://hal.science/hal-03018287
Submitted on 1 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
An efficient label-setting Algorithm for the Bi-Objective Shortest Path problem

A. Giret1, Y. Kergosien1, G. Sauvanet2 and E. Neron1

1 Université François Rabelais de Tours, CNRS, LIA EA 6300, OC ERL CNRS 6305, 64 avenue Jean Portalis, 37200 Tours, France
{antoine.giret, yannick.kergosien, emmanuel.neron}@univ-tours.fr

2 La Compagnie des Mobilités, 30 rue André Theuriet, 37000 Tours, France
gael.sauvanet@geovelo.fr

Keywords: label-setting Algorithm, pareto Front, Bi-Objective Shortest Path problem, cycling, exact method

Abstract: In this paper, we consider a classical Bi-objective Shortest Path problem (BSP) that takes into account both distance and insecurity criteria. This work is in collaboration with an enterprise who provide a web platform called Géovélo that aims to propose routes for cycling. We propose a new exact method to solve a BSP, called Label Setting algorithm with Dynamic update of Pareto Front (LSDPF), which aims to find all non-dominated solutions of the problem. Different exploration strategies have been proposed and tested. Numerical experiments on real data sets and on instances of the literature were conducted. Comparison with recent benchmarks algorithms solving BSP - the bounded Label Setting algorithm by (Raith, 2010) and the pulse algorithm by (Duque et al., 2015) - shows that our method outperform these benchmarks algorithms.

1 INTRODUCTION

For economical and ecological reasons, alternative transportation such as cycling is expanding. Public actors have initiated several programs to adapt the road network for cycling (e.g., greenway, cycling lane), because users seek to be able to move safely. For a cyclist, travelling time and distance are not the only criteria to consider; they may also consider safety, difficulty and tourist attractions.

This work is in collaboration with La Compagnie des Mobilités who provide a web platform called Géovélo (http://www.geovelo.fr). Géovélo aims to propose routes for cycling. These routes consider two criteria which are safety (minimising lack of security), and distance. Both criteria are very conflicting, mainly because most road networks are designed for cars.

The studied problem is a Bi-Objective Shortest Path Problem: let $G = (V,A)$ be a directed graph where $V$ is the set of nodes ($|V| = n$) and $A$ the set of arcs ($|A| = m$). $(c_{ij}^1, c_{ij}^2)$ denotes the costs associated with each arc $(i,j)$, where $c_{ij}^1$ represents its distance and $c_{ij}^2$ its insecurity. The problem consist in finding a path $P$ from a source node $s$ to a target node $t$ minimizing both criteria $\{\sum_{(i,j)\in P} c_{ij}^1, \sum_{(i,j)\in P} c_{ij}^2\}$. The result of the problem is a set of strictly non-dominated solutions also called Pareto front defined in (Ulungu and Teghem, 1994).

The Bi-Objective Shortest Path Problem, noted BSP, is a specific case of the Multi-Objective Shortest Path Problem, that is one of the most studied problems among the Multi-Objective Combinatorial Optimisation (MOCO) problems. (Serafini, 1987) has proved that the BSP is $\mathcal{NP}$-Hard. (Tarapata, 2007) presents a state-of-the-art MOSP problem.

The first methods solving BSP are label-correcting and label-setting methods proposed by (Hansen, 1980) and (Martins, 1984). Labelling methods can be seen as a generalization of existing methods solving the mono-objective problem (Bertsekas, 1998) and they can be decomposed into two types: label-correcting and label-setting. The difference between these types is how the label queue is managed. Label-setting methods often use a lexicographic order to ensure that an explored label corresponds to a non-dominated path (see (C.T. Tung, 1992)). Label-correcting methods use a simple order to manage label queue exploration (e.g., FIFO) to simplify the data structure of the queue (see (Brumbaugh-Smith and Shier, 1989) and (Skriver and Andersen, 2000)). There are two different selection types for the label-correcting methods: a label is managed separately in label se-
lection, whereas all labels on a selected node are extended at the same time in node selection. (Guerrero and Musmanno, 2001) compares different node and label selection strategies for the label-correcting method. Node selection is generally used with label-correcting methods, whereas label selection must be used with label-setting methods.

Another type of method to solve BSP is Ranking method (Climaco and Martins, 1982) that are based on the $k$th-shortest path problem (Carlyle and Wood, 2005). These methods find the shortest path optimizing one objective, then the second shortest path, and so on. In addition to labelling and ranking methods, the two-phase approach is also commonly used to solve BSP problems. (Ulunngu and Teghem, 1995) introduced this concept. The two-phase method is basically used to compute separately supported (i.e., located on the boundary of the convex hull) and non-supported solutions. All solutions in the first phase can be computed by a single objective method with a weighted sum objective. In the second phase, a bi-objective method enumerates all non-supported solutions. During this phase, the search space is restricted by the solutions computed during the first phase. (Raith, 2009) has shown that this approach is efficient and proposed a comparison of the solution strategies for finding all efficient paths. (Raith, 2010) focuses on label-correcting and label-setting algorithm (bLSET) and propose an acceleration technique, improving the efficiency of these algorithms. He showed that it is not always necessary to continue the search to the target node to confirm that it is dominated. According to (Demeyer and al., 2013), bLSET algorithm presented better computational times among labelling approaches proposed during this period.

Recently, (Duque et al., 2015) proposed a new exact method, called Pulse algorithm, for the BSP and large-scale road networks. Pulse algorithm is based on recursive method using pruning strategies that accelerate the graph exploration. The results shown that the proposed algorithm outperform the bLSET algorithm on very-large scale instances from the DIMACS dataset.

This work is an extension of (Sauvanet and Neron, 2010) and aims to propose a new exact method to solve BSP. The proposed method, called Label Setting algorithm with Dynamic update of Pareto Front (LSDPF), is based on existing methods previously cited. However, some improvement techniques have been added. The paper is organized as follows. Section 2 presents the LSDPF algorithm with several exploration strategies. Some numerical results are given in section 3 in order to test the different exploration strategies and a parameter of the algorithm. Then computational experiments compare our proposed method to the bLSET algorithm developed by (Raith, 2010) and the Pulse algorithm developed by (Duque et al., 2015). Finally section 4 concludes the paper and gives some future work directions.

2 LABEL SETTING ALGORITHM WITH DYNAMIC UPDATING OF PARETO FRONT (LSDPF)

The LSDPF algorithm is based on a two-phase method introduced by (Ulunngu and Teghem, 1995). The first phase aims to compute dominated solutions by solving several Mono-Objective Shortest Path Problems. The second phase aims to get only non-dominated solutions and is based on classic label-setting algorithm which has been introduced in (Martins, 1984). The specificity of the proposed algorithm is that the final Pareto front can dynamically evolve at each iteration of the label-setting algorithm, and not necessary when the target node is reached.

2.1 First Phase

As explained shortly before, the first phase consists in running a set of Mono-Objective Shortest Path Problems (using a simple adaptation of (Dijkstra, 1959)), and is composed by two steps.

The first step provides an upper bound of a given criterion by solving a Mono-Objective Shortest Path Problem from $s$ to $t$, minimizing the other criterion. These upper bounds allow to reduce the exploration of the graph for the next step.

The second step consists in solving reverse one-to-all Mono-Objective Shortest Path Problems from $t$, where each objective function is defined by a linear combination of both criteria:

$$\alpha \sum_{(i,j) \in P} c_{ij}^1 + (1 - \alpha) \sum_{(i,j) \in P} c_{ij}^2$$

We note $A$ the set of all $\alpha$ values tested. After each one-to-all resolution, each traversed node is characterized by a pair of values $(d^\alpha_i, s^\alpha_i)$, representing a feasible path from $i$ to $t$ with a distance equal to $d^\alpha_i$ and an insecurity equal to $s^\alpha_i$. This search is stopped at a node $i$ when both $d^\alpha_i$ and $s^\alpha_i$ are greater than upper bounds found in the first step. Only non dominated pairs - in Pareto sense - of values $(d^\alpha_i, s^\alpha_i)$ are saved.

Two specific cases are important: $\alpha = 1$ and $\alpha = 0$. Both allows to determine the lower and the upper bound for each criterion at each node.
• For $\alpha = 1$, is obtained at each node $i \in V$ the lower bound of the distance criterion defined by $(LB_i^1 = d_i^1)$, which is associated to an upper bound of the insecurity criterion defined by $(UB_i^1 = s_i^1)$.

• For $\alpha = 0$, is obtained at each node $i \in V$ the lower bound of the insecurity criterion defined by $(LB_i^0 = s_i^0)$, which is associated to an upper bound of the distance criterion defined by $(UB_i^0 = d_i^0)$.

Finally, the set of pairs $(d_i^0, s_i^0)$ is used to initialize the Pareto front. Additional one-to-one Mono-Objective Shortest Path from $s$ to $t$ can be done (i.e. using more values of $\alpha$), in order to extend the initial Pareto Front.

### 2.2 Second Phase

The second phase is based on a label-setting algorithm with additional improvements related to the Bi-Objective Shortest Path problem. These improvements allow to reduce the number of labels created and thus finding the final Pareto front faster. The proposed method (LSDPF) is presented in the algorithm 1. Let us define:

• $s$ and $t$ respectively the source and target nodes.

• $i$ and $j$ the indexes of nodes and $u$, $v$, $w$ and $x$ the indexes of labels.

• $l_u$ a label, defined by a tuple $(i, vd, vs)$ where:
  - $i$ is the node associated to the label.
  - $vd$ and $vs$ represent respectively the values of distance and insecurity of a feasible path from $s$ to $i$.

• $LN_i$ the set of labels at node $i$.

• $Q$ contains unexplored labels.

• $l_v \prec_p l_u$ means that the label $l_v$ is strictly dominated by $l_u$ in the Pareto sense.

The final Pareto front contained all non-dominated solutions is obtained at node $t$: it corresponds to the set of labels $LN_t$.

The main steps of the LSDPF algorithm are:

• Line 1: the FIRSTPHASE function represents the first phase of the proposed method explained in the section 2.1. It allows to determine all pairs $(d_i^0, s_i^0)$. The Pareto front is initialized in line 4.

• Line 7: the NEXTLABEL function defines the order to explore the label, which affects the evolution of the Pareto front and thus the convergence of the algorithm. Three strategies explained later have been tested.

• Line 12: the current label is explored if and only if it is not dominated by another label existing on the same node.

---

**Algorithm 1** Label Setting algorithm with Dynamic update of Pareto Front (LSDPF)

1: **FIRSTPHASE**
2: $LN_i \leftarrow \emptyset$, $\forall i = \{1, \ldots , n\}$
3: $LN_s \leftarrow \{(s,0,0)\}$
4: $LN_t \leftarrow \{(t,0,0)\}$
5: $Q \leftarrow \{(s,0,0)\}$
6: while $Q \neq \emptyset$ do
7:   $(i, vd, vs) \leftarrow$ NEXTLABEL($Q$)
8:   $Q \leftarrow Q \setminus (i, vd, vs)$
9:   for all $(i, j) \in A$ do
10:     $(j, vd', vs') \leftarrow (j, vd + c_{ij}, vs + c_{ij}^2)$
11:     $I_u \leftarrow (j, vd', vs')$
12:     if $\exists l_v \in LN_j/l_v \prec_p I_u$ then
13:       $\exists l_v \in LN_j/l_v \prec_p (j, vd' + LB_j^0, vs'+ LB_j^0)$
14:       if $j \neq t$ then
15:         $Q \leftarrow Q \cup I_u$
16:       end if
17:       $LN_j \leftarrow LN_j \setminus I_v$
18:       UPDATELABELS($Q$, $LN_j$, $I_u)$
19:       DPF($LN_t$, $(j, vd', vs')$)
20:     end if
21:   end if
22: end for
23: end while
24: return $LN_t$

---

**Algorithm 2** UPDATELABELS

1: **function UPDATE**($Q$, $LN_j$, $I_v$)
2: for all $l_v \in LN_j/l_v \prec_p I_v$ do
3:   $LN_j \leftarrow LN_j \setminus I_v$
4:   if $l_v \in Q$ then
5:     $Q \leftarrow Q \setminus I_v$
6:   end if
7: end for
8: end function

---

**Algorithm 3** DPF

1: **function DPF**($LN_t$, $(j, vd', vs')$)
2: for all $\alpha \in A$ do
3:   $I_v \leftarrow (t, vd' + d_i^\alpha, vs' + s_i^\alpha)$
4:   if $\exists I_u \in LN_i/l_v \prec_p I_v$ then
5:     $LN_i \leftarrow LN_i \setminus I_v$
6:   end if
7:  for all $l_x \in LN_i/l_x \prec_p I_x$ do
8:  end for
9: end if
10: end for
11: end function
2.3 Exploration Strategies

The exploration strategy determines in which order the labels will be explored (goal of the NEXTLABEL function in the algorithm). From a strategy to another one, the number of explored labels can be different as well as the computation time of the LSDPF algorithm. The selected strategy is thus important, three strategies are presented and have been tested.

2.3.1 Strategy 1

The first strategy is simple. It consists in selecting the label that contains the minimum \( vd' \) value among all labels present in \( Q \). If several labels have the same minimum value \( vd' \), the label with the smallest value of \( v' \) is selected among these labels. The graphic evolution of the set of explored nodes/labels can be represented by a circle around the source node, which become increasingly larger. Only distance criterion as first criterion has been tested (and not the insecurity one). Indeed, the range of distance values is bigger than the range of insecurity values in cycling context. Considering the distance criterion first, the search will be guided to more interesting labels.

2.3.2 Strategy 2

The second strategy aims to explore the most promising label first. From the set of labels presents in \( Q \), the next label to treat is the label with the smallest value \( (vd' + LB_1^j) \). In case of equalities, the label with the smallest value \( (vd' + LB_2^j) \) is selected. As previously, only distance criterion is tested as first criterion for the three same reason. The representation graph of explored labels is close to an ellipsoid around the source and target nodes.

2.3.3 Strategy 3

The last strategy uses a linear combination of \( vd' + LB_1^j \) and \( vs' + LB_2^j \). Unlike other strategies, this one will take into account both criteria to guide the exploration. Thus the next label presents in \( Q \) to explore is the label minimizing the function:

\[
\beta (vd' + LB_1^j) + (1 - \beta) (vs' + LB_2^j)
\]

3 COMPUTATIONAL EXPERIMENTS

The LSDPF algorithm has been implemented in C++ language using the library Boost [http://www.boost.org](http://www.boost.org). Computational experiments have been performed on an Intel Core i7 vPro processor quad-core with 8GB of RAM and using Linux system. First, numerical results are given for the three strategies and several values of \( A \) on real data sets taking into account the study case (Cycling). Next, the efficiency of the LSDPF algorithm with the best exploration strategy is compared to efficient benchmarks algorithms on instances of the literature.

3.1 Comparison of exploration strategies

To compare the different exploration strategies presented in section 2.3, we used real roads networks that come from the cooperative OpenStreetMap project.
The same interpretation can be done for all the graphs: the second strategy seems to be better than the first one as we might expect but also than third one, which can be surprising. The second strategy improves by approximately 20% the execution time of the first strategy and between 40% and 90% execution times of the third strategy. Results of the third strategy are not as good as expected. The linear combination of both criteria doesn’t seems guide the exploration towards good solutions quickly.

The table 2 presents the average number of explored labels for each strategy and each graph. The parameter $\beta$ is equal to 0.5. The first column indicates the graph name, columns from 2 to 4 shows the average execution time of each strategy, and last column shows the average number of non-dominated solutions on the Pareto front.

Table 2: Average number of explored labels for each strategy

| Graph | Str. 1 | Str. 2 | Str. 3 | \( |A| \) |
|-------|--------|--------|--------|--------|
| Paris | 194.8  | 158.6  | 403.6  | 2      |
| Berlin| 206.2  | 150.8  | 246    | 2      |
| SF    | 87040.8 | 71815.2 | 715942.4 | 1280  |

The result of the comparison between the strategies 1 and 2 is that when the number of explored label decreases, the execution time decrease. However the same interpretation cannot be done for the difference found between the strategies 1 and 3 and between the strategies 2 and 3. Even if the strategy 3 explores less labels, its average of execution times are bigger than others, meaning the linear combination of both criteria cannot be used to quickly guide the exploration towards good solutions.

### 3.2 Comparison of several $|A|$ values

Only two strategies are used for the numerical experiments of several $|A|$ values. Three cases have been tested:
- $A = \{0; 1\} (|A| = 2)$
- $A = \{0; 0.5\} (|A| = 3)$
- $A = \{0; 0.25; 0.5; 0.75; 1\} (|A| = 5)$

Table 5 presents the average execution time (in ms) of the first phase for each case. Table 3 shows the average execution time (in ms) of the second phase for each case and each strategy (strategy 1 noted S1 strategy 2 noted S2). The total average execution time (in ms) of LSDPF algorithm is presented in table 4. Finally, the number of explored labels (in thousands) of LSDPF algorithm is given in table 6.

Table 3: Average execution time of second phase (ms) for several values of $|A|$ several strategies

| Graph | $|A| = 2$ | $|A| = 3$ | $|A| = 5$ |
|-------|----------|----------|----------|
|       | S1 | S2 | S1 | S2 | S1 | S2 |
| Paris | 165 | 129 | 154 | 163 | 142 | 156 |
| Berlin| 146 | 91  | 139 | 113 | 110 | 114 |
| SF    | 86844 | 71666 | 64430 | 75884 | 67118 | 82026 |

Table 4: Average execution time of LSDPF (ms) for several values of $|A|$ and several strategies

| Graph | $|A| = 2$ | $|A| = 3$ | $|A| = 5$ |
|-------|----------|----------|----------|
|       | S1 | S2 | S1 | S2 | S1 | S2 |
| Paris | 195 | 159 | 193 | 201 | 207 | 221 |
| Berlin| 206 | 151 | 223 | 197 | 257 | 261 |
| SF    | 87041 | 71862 | 64641 | 76094 | 67488 | 82396 |
From previous tables 5 to 6, it seems not interesting to use more than 2 $\alpha$ values whatever the strategy 1 or 2. The fact that the execution time increase when the number of $\alpha$ values increases in a given strategy can be explained by the additional consumption of CPU time needed by DPF function. Moreover, the degradation of execution time of the strategy 2 is more significant in comparison with the strategy 1.

### 3.3 Efficiency of LSDPF

In order to compare our results with benchmarks algorithms solving the BSP - the bounded Label Setting algorithm (bLSET) (Raith, 2010) and the pulse algorithm (Duque et al., 2015) - computational experiments have been performed on instances from the 9th DIMACS challenge. Three graphs have been used. The first graph corresponds to New York City. This graph contains 264,346 nodes and 733,846 arcs. The second corresponds to another representation of the San Francisco Bay Area and contains 321,270 nodes and 800,172 arcs. The third corresponds to the state of Florida and contains 1,070,376 nodes and 2,712,798 arcs. Criteria considered for these graphs correspond to the physical distance and transit time between nodes. To fairly compare algorithms, the same pairs of source and target nodes considered in (Duque et al., 2015) have been used for each instance. As (Duque et al., 2015), the 30 instances of each graph has been clustered into the same three equal sized groups, denoted S (small), M (medium) and L (large), based on the number of non-dominated solutions found on the Pareto front.

The table 7 compares the average execution time of the bLSET algorithm, the Pulse algorithm, and LSDPF algorithm. All times are in milliseconds. Column 3, 4 and 5 presents the average execution time of the bLSET algorithm, Pulse algorithm and LSDPF algorithm. Values followed by + means that at least one of the instances could not be solved in less than one hour and so the search has been stopped.

Table 7: Averaging times (ms) for the bLSET, the Pulse algorithm and the LSDPF on 9th DIMACS challenge graphs

<table>
<thead>
<tr>
<th>Cluster</th>
<th>bLSET</th>
<th>Pulse</th>
<th>LSDPF</th>
</tr>
</thead>
<tbody>
<tr>
<td>NY-S</td>
<td>34.10</td>
<td>62.30</td>
<td>320</td>
</tr>
<tr>
<td>NY-M</td>
<td>147.40</td>
<td>301.160</td>
<td>523.20</td>
</tr>
<tr>
<td>NY-L</td>
<td>422.70</td>
<td>881.260</td>
<td>1367.660*</td>
</tr>
<tr>
<td>BAY-S</td>
<td>8.80</td>
<td>67.80</td>
<td>160</td>
</tr>
<tr>
<td>BAY-M</td>
<td>49.90</td>
<td>552.40</td>
<td>5700</td>
</tr>
<tr>
<td>BAY-L</td>
<td>171.80</td>
<td>317.430</td>
<td>1055.50</td>
</tr>
<tr>
<td>FLA-S</td>
<td>14.70</td>
<td>330.120</td>
<td>350</td>
</tr>
<tr>
<td>FLA-M</td>
<td>94.10</td>
<td>566.150*</td>
<td>3479.10</td>
</tr>
<tr>
<td>FLA-L</td>
<td>552.30</td>
<td>262.7430*</td>
<td>888.590*</td>
</tr>
</tbody>
</table>

Table 7 shows that execution times of bLSET are average bigger than other algorithms. Regarding the Pulse and LSDPF algorithms, the Pulse algorithm is particularly efficient on instances with a small number of solutions unlike other methods. For these instances, LSDPF takes up to 7 times more to find the entire Pareto front but execution times remain reasonable since they are on average lower than 3 seconds. Nevertheless, we can see that LSDPF has the same average order execution times to solve instances whatever the number of solutions which is not the case of the two other algorithms where it appears to have an exponential time progression. Moreover they may not find the entire Pareto front in the case of larger instances (see table 8). To conclude, LSDPF is clearly the faster method to find the whole Pareto front for medium and larger instances.

Table 8: Number of instances unsolved in less than an hour for bLSET, Pulse algorithm and the LSDPF.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>bLSET</th>
<th>Pulse</th>
<th>LSDPF</th>
</tr>
</thead>
<tbody>
<tr>
<td>NY-L</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>FLA-M</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FLA-L</td>
<td>6</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Finally, the criterion used for the shortest paths search on these instances are distance and time, that are not completely independents criteria. The Pareto front thus contains fewer non-dominated solutions that it is possible to have using more conflicting criteria, such as distance and insecurity for example.
4 CONCLUSION

In this paper, we consider a classical Bi-objective Shortest Path Problem (BSP) which considers both distance and insecurity inspired from a cycling context. We proposed an exact algorithm based on label-setting methods and called LSDPF, in order to find all non-dominated solutions of the problem. One of the specificities of LSDPF algorithm is to dynamically update the whole Pareto front during the search allowing to reduce the number of explored labels. Different strategies have been proposed and tested. Numerical experiments on real data sets and on instances of the literature were conducted. Comparison with recent benchmarks algorithms to solve BSP (the bounded Label Setting algorithm (blSET) (Raith, 2010) and the pulse algorithm (Duque et al., 2015)) shows that our method outperform the benchmarks algorithms.

The first perspective of this work is to generalize our methods to several conflictual criteria on same type of instances. Another promising research direction is to use some graph contractions in the case of multi-criteria.

REFERENCES


