A New Mathematical Model For the Minimum Linear Arrangement Problem

Mahdi Moeini[1,2], Serigne Gueye[1] and Sophie Michel Loyal[3]
[1]*Laboratoire Informatique d'Avignon (LIA), Université d'Avignon et des Pays de Vaucluse, LIA-CERI 339, chemin des Meinajaries, Agroparc BP 91228, 84911, Avignon, France*
[2]*Centre de Recherche en Informatique de Lens (CRIL-CNRS UMR 8188), Université d'Artois, Rue Jean Souvraz,SP 18, F 62307 Lens Cedex, France*
[3]*Laboratoire de Mathématique Appliquée du Havre (LMAH), Université du Havre, 25 rue Philippe Lebon, 76058, Le Havre, France*

Keywords: Integer Programming, Minimum Linear Arrangement Problem, Valid Inequality.

Abstract: This paper addresses a classical combinatorial optimization problem called the Minimum Linear Arrangement (MinLA) Problem. The MinLA problem has numerous applications in different domains of science and engineering. It is known to be NP-hard for general graphs. The objective of this paper is to introduce a new mathematical model and associated theoretical results, including novel rank inequalities. Preliminary computational experiments are reported on some benchmark instances.

1 INTRODUCTION

The Minimum Linear Arrangement (MinLA) is a challenging problem in combinatorial optimization. For a given graph, the MinLA consists in arranging the nodes of the graph on a line in such a way to minimize the sum of the distance between the adjacent nodes. In the literature, the MinLA is known under different names such as the *optimal linear ordering*, the *edge sum problem*, the *minimum-1-sum* (see e.g., (Amaral et al., 2008), (Horton, 1997), (Caprara and Gonzalez, 2005), (Caprara et al., 2010), (Petit, 1999), (Hungerlaender and Rendl, 2012), (Schwarz, 2010)).

In addition to theoretical interests, the MinLA has many practical applications. A non-exhaustive list includes the *design of VLSI layouts*, the *graph drawing*, the *single machine job scheduling, etc.* ((Petit, 1999), (Hungerlaender and Rendl, 2012), (Schwarz, 2010)). The MinLA can be solved efficiently for some particular kinds of graphs. For example, there are polynomial time algorithms for the MinLA on trees, outerplanar graphs, and certain Halin graphs (Caprara et al., 2010). But in general, the MinLA is NP-hard (Garey et al., 1976). Because of the hardness of finding optimal values, lower bounding and heuristic algorithms are usually applied to get good approximations ((Caprara et al., 2010), (Petit, 1999), (Schwarz, 2010)).

In this paper, we present a new 0-1 linear programming problem and investigate the associated new polyhedra for which valid inequalities are provided using lifting techniques or introducing some rank inequalities. The lower bound corresponding to the optimal value of the relaxed problem is tested on standard benchmarks. The results are modest and show that many facets of the corresponding new polyhedra has to be discovered.

The structure of the paper is as follows. In Section 2, the Minimum Linear Arrangement (MinLA) problem is reviewed and the new formulation of the MinLA is given. Section 3 is devoted to theoretical results on new rank inequalities. The computational experiments are reported in section 4 and the last section includes some conclusions.

2 THE MINIMUM LINEAR ARRANGEMENT (MinLA) PROBLEM

The present section includes the basic definitions and models of the MinLA problem.

2.1 Definitions and Preliminaries

Let $G = (V,E)$ be an undirected (connected) graph, where V (with $|V| = n$) is the set of nodes and E de-

Moeini M., Gueye S. and Michel Loyal S..
A New Mathematical Model For the Minimum Linear Arrangement Problem.
DOI: 10.5220/0004827800570062
In *Proceedings of the 3rd International Conference on Operations Research and Enterprise Systems* (ICORES-2014), pages 57-62
ISBN: 978-989-758-017-8

notes the set of edges. A layout is defined as an one-to-one function $\varphi : V \longrightarrow [1..n]$. The Minimum Linear Arrangement problem (MinLA) is the combinatorial optimization problem consisting in finding a layout φ minimizing the following sum

$$\sum_{uv \in E} |\varphi(u) - \varphi(v)|.$$

(see e.g., (Amaral et al., 2008), (Horton, 1997), (Caprara and Gonzalez, 2005), (Caprara et al., 2010), (Petit, 1999), (Hungerlaender and Rendl, 2012), (Schwarz, 2010)).

There are various mathematical formulations for the MinLA. An overview of the different models can be found in (Petit, 1999) and (Schwarz, 2010). A basic formulation of MinLA is obtained by considering the following variables. Let us define the variables x_{ik} as:

$$x_{ik} = \begin{cases} 1 & \text{if the label } k \text{ is assigned to the node } i, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

With these variables, the MinLA can be formulated as follows

MinLA-Quadratic:

$$\min \sum_{(i,j) \in E} \sum_k \sum_l |k-l| x_{ik} x_{jl} \tag{2}$$

s.t:

$$\sum_{i=1}^{n} x_{ik} = 1 \quad : k \in \{1, \cdots, n\}, \tag{3}$$

$$\sum_{k=1}^{n} x_{ik} = 1 \quad : i \in \{1, \cdots, n\}, \tag{4}$$

$$x_{ik} \in \{0,1\} \quad : i,k \in \{1, \cdots, n\}, \tag{5}$$

This is a 0-1 quadratic programming problem for which different reformulation has been proposed (see (Schwarz, 2010) for an overview of different formulations of MinLA).

We propose a non-standard linearization of MinLA working as follows:
Let us define

$$f_{kl} := \sum_{(i,j) \in E} x_{ik} x_{jl} \in \{0,1\}. \tag{6}$$

If we apply this definition on (2)-(5), i.e., **MinLA-Quadratic**, the formulation of the MinLA becomes:

A New Linearization for MinLA:

$$\min \sum_{k,l} |k-l| f_{kl} \tag{7}$$

s.t:

$$\sum_{i=1}^{n} x_{ik} = 1 \quad : k \in \{1, \cdots, n\}, \tag{8}$$

$$\sum_{k=1}^{n} x_{ik} = 1 \quad : i \in \{1, \cdots, n\}, \tag{9}$$

$$f_{kl} \geq (x_{ik} + x_{jl} - 1) : \forall (i,j) \in E, \forall k,l \in \{1, \cdots, n\}, \tag{10}$$

$$x_{ik} \in \{0,1\} \quad : i,k \in \{1, \cdots, n\}, \tag{11}$$

$$f_{kl} \in \{0,1\} \quad : \forall k,l \in \{1, \cdots, n\}. \tag{12}$$

Unfortunately, the formulation (7)-(12) has two drawbacks. The first one is due to the large number of constraints (10), that is $O(n^4)$, and the second one to the fact that the resulting bound is very poor. Indeed, one may observe that the solution

$$f_{kl} = 0 \quad (\text{for } k,l = 1, \cdots, n),$$

$$x_{ik} = \frac{1}{n} \quad (\text{for } i,k = 1, \cdots, n),$$

is feasible, with 0 as the objective value. Hence, it is necessary to strengthen the relaxation by introducing valid inequalities (see (Amaral et al., 2008), (Amaral, 2009), (Amaral and Letchford, 2011), (Horton, 1997), (Caprara et al., 2010), (Schwarz, 2010)). We show in Section 3, how to deal with these two aspects.

Notice that f_{kl} can be seen as a binary flow between the locations k and l. Indeed, f_{kl} equal 1 if the two entities located in k and l are linked by an edge in G, and 0 otherwise. As a consequence, one may see that the graph whose the adjacency matrix is represented by $f = \{f_{kl}\}$ is isomorphic to G. Hence we transform the initial problem in a new one consisting in finding an optimal isomorphic graph for G.

3 VALID INEQUALITIES

In order to reduce the number of constraints (10), we show in the following theorem how we can produce $O(n^3)$ equivalent inequalities using lifting techniques (Nemhauser and Wolsey, 1998).

Theorem 1: If we define A as the *adjacency matrix* of the (connected) graph $G = (V,E)$, the following inequalities are valid for (7)-(12):

(i) For $k,l,i_0 \in \{1,2,\cdots,n\}$:

$$f_{kl} \geq \sum_{j=1}^{n} A_{i_0 j} x_{jl} + \alpha(1 - x_{i_0,k}), \tag{13}$$

where

$$\alpha := \min(A_{i'j'} - A_{i_0 j'}),$$

such that $1 \leq i' \neq i_0 \leq n$ and $1 \leq j' \leq n$.

(ii) For $k, l, i_0 \in \{1, 2, \cdots, n\}$:

$$f_{kl} \geq \sum_{j=1}^{n} A_{i_0 j} x_{jl} + \sum_{k'=1, k' \neq k}^{n} \alpha_{k'} x_{i_0, k'}, \qquad (14)$$

where

$$\alpha_{k'} := \min(A_{i' j} - A_{i_0 j}),$$

such that $i' \neq i_0, j$ and $j \neq i_0$ and $k' \neq k$.

Proof: **(i)** Remember the definition of f_{kl}

$$f_{kl} := \sum_{(i,j) \in E} x_{ik} x_{jl} \in \{0, 1\}$$

For constant values of k and l, and setting $i \longleftarrow i_0$, we can do a lifting by the assumption of $x_{i_0,k} := 1$. Under these assumptions, we have $f_{kl} \geq \sum_{j=1}^{n} A_{i_0 j} x_{jl}$. Our objective consists in finding an $\alpha \in \mathbb{R}$, such that the above inequality remains true even for $x_{i_0,k} \neq 1$. Thus we have to solve the following problem

$$\alpha := \min \frac{1}{(1 - x_{i_0,k})} (f_{kl} - \sum_{j=1}^{n} A_{i_0 j} x_{jl}),$$

such that $x_{i_0,k} \neq 1$ and $k, l, i_0 \in \{1, 2, \cdots, n\}$. The corresponding optimal value is given by:

$$\alpha := \min(A_{i' j'} - A_{i_0 j'}),$$

where, $1 \leq i' \neq i_0 \leq n$ and $1 \leq j' \leq n$.

(ii) This part is similar to the part (i) with the difference of the assumption on $x_{i_0,k'} = 0$ for all $k' \neq k$.

We know that

$$f_{kl} \geq \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} x_{ik} x_{jl},$$

where $k < l$. We are looking for the real values $\alpha_{k'}$ such that

$$f_{kl} \geq \sum_{j=1}^{n} A_{i_0,j} x_{jl} + \sum_{k' \neq k, k'=1}^{n} \alpha_{k'} x_{i_0 k'},$$

particularly, we are looking for $\alpha_{k'}$ that minimizes $f_{kl} - \sum_{j=1}^{n} A_{i_0,j} x_{jl}$ under the constraint $x_{i_0,k'} = 1$. There are two cases:

Case (1) $k' = l$ and $k' \neq k$: in this case $\alpha_{k'} = \min_{j \neq i_0} A_{i_0,j}$.

Case (2) $k' \neq k, l$: in this case $\alpha_{k'} = \min_{j \neq i_0, i' \neq i_0, i' \neq j} (A_{i',j} - A_{i_0,j})$.

By taking the minimum of these cases, we obtain the result of the theorem. ■

Now, we focus on the second drawback of the formulation (7)-(12). In order to find better lower bounds for the continuous relaxation of (7)-(12), we introduce some sets of valid inequalities.

The first set of valid inequalities is related to the connectivity of G. We note that $\sum_l f_{kl}$ is equal to the degree of any node having the label k, hence the following inequalities are valid for any connected graph G:

$$\sum_{l=1}^{n} f_{kl} \geq (\text{minimum vertex degree of the graph } G), \qquad (15)$$

$$\sum_{l=1}^{n} f_{kl} \leq (\text{maximum vertex degree of the graph } G), \qquad (16)$$

$$\sum_k \sum_l f_{k,l} = 2|E|, \qquad (17)$$

where, $|E|$ is the total number of the edges.

A stronger version of these inequalities is given through the following theorem:

Theorem 2: For any connected graph G the following equalities are valid for (7)-(12):

$$\sum_k f_{kl} = \sum_{j=1}^{n} (\text{degree of node } j)\, x_{jl} \quad : \quad l \in \{1, \cdots, n\} \qquad (18)$$

Proof: for any value of l, we have

$$\sum_k f_{kl} = \sum_k \left(\sum_{(i,j) \in E} x_{ik} x_{jl} \right) = \sum_{(i,j) \in E} \left(\sum_k x_{ik} \right) x_{jl}$$

Since $\sum_k x_{ik} = 1$, hence $\sum_k f_{kl} = \sum_{(i,j) \in E} x_{jl}$. If we define A as the *adjacency matrix* of the graph G, then

$$\sum_k f_{kl} = \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} x_{jl} = \sum_{j=1}^{n} \sum_{i=1}^{n} A_{ij} x_{jl} = \sum_{j=1}^{n} x_{jl} \left(\sum_{i=1}^{n} A_{ij} \right)$$

Since *the degree of the node* (j) is equal to $\sum_{i=1}^{n} A_{ij}$, consequently,

$$\sum_k f_{kl} = \sum_{j=1}^{n} (\text{degree of node } j)\, x_{jl}$$

where l belongs to $\{1, \cdots, n\}$. ■

A well known set of valid inequalities for different formulations of the MinLA consists in using triangular inequalities. The following theorem is based on this idea.

Theorem 3: The following inequalities are valid for (7)-(12):

- If $k, l, m \in \{1, \cdots, n\}$, then

$$|k - l| f_{kl} \leq |k - m| + |m - l| : \forall k, l, m, \qquad (19)$$

$$|k - l| f_{kl} \leq |k - m| + |m - l| f_{ml} : \forall k < l < m, \qquad (20)$$

• For all $k,l \in \{1,\cdots,n\}$ and for all $(i,j) \in E$, *if degree of the node i is equal to* 1, then

$$x_{ik} - x_{jl} + f_{kl} \leq 1. \quad (21)$$

Proof: All of these inequalities can be easily obtained by using the fact that $|k-l| \leq |k-m| + |m-l|$ (for all triple k,l,m) and $x_{ik}, f_{kl} \in \{0,1\}$ (for all $i,k,l \in \{1,\cdots,n\}$). ■

Definition: For a given graph G, the chromatic number of G is denoted by $\chi(G)$ and is equal to the minimum number of colors that one needs to color the nodes of G in such a way that all adjacent nodes of G get different colors.
The following theorem introduces new valid inequalities that involve the chromatic number of a given graph G.

Theorem 4 (Chromatic Inequalities): For any graph G having the chromatic number $\chi(G)$ and for all triple k,l,s, the following inequalities are valid for (7)-(12):

$$f_{k,l} + f_{l,s} + f_{k,s} \leq \chi(G). \quad (22)$$

Proof: First of all, we note that the left hand side of (22) is at most equal to 3. If $\chi(G) \geq 3$, we are done; otherwise, if $\chi(G) = 2$, then there is (at least) one node among the nodes having the labels k,l,s that is not connected to one of the others. Without loss of generality, let us suppose that the node having the label k is not connected to the node with the label s; this means that there will be no flow between the nodes k and s. Consequently, the nodes k and s can have the same colors and hence $f_{k,s} = 0$, which completes the proof. ■

The results of the theorem 4 can be useful when $\chi(G) = 2$. This corresponds to some special graphs such as trees and the graphs without any cycles of odd lengths. This is rather restricting in using the inequalities (22). Notice that finding the chromatic number of a given graph is an NP-hard problem. Nevertheless, tight upper bound of $\chi(G)$ may be found by using heuristic algorithms.

Definition: For a given graph G, *matching* is a subset of edges in which, no two edges are adjacent to a same node. A *maximum matching* of G is defined as a matching with the maximum cardinality. The matching number of G, noted by $\nu(G)$, is the size of a maximum matching.

As in the case of chromatic number, some valid inequalities depending on the matching number may be introduced.

Theorem 5 (Matching Inequalities): For any graph G having the matching number $\nu(G)$ and for all k and "$0 \leq i$ and $k+2i+1 \leq n$"; the following inequality is valid for (7)-(12):

$$\sum_{i=1}^{i:k+2i+1 \leq n} f_{k+2i,k+2i+1} \leq \nu(G). \quad (23)$$

Proof: For an arbitrary (connected) path in G, the left hand side of (23) is (at most) the number of the edges in the path such that any couple of edges is separated by at least one edge. By noting this fact, one can conclude that the number of these edges cannot exceed the matching number of G. ■

4 COMPUTATIONAL EXPERIMENTS AND NUMERICAL RESULTS

In this section, we present the preliminary results that we have obtained by applying the valid inequalities of the previous section.

The model has been coded in C++ and has been solved with IBM CPLEX 12.2 in an Intel Core 2 Duo of 3 GHz and 3.25 GB of RAM. The experiments have been carried out on some benchmark instances already used in (Caprara et al., 2010), (Caprara et al., 2011), and (Schwarz, 2010). Table 1 reports some characteristics of the instances. In this table, for each graph, the number of nodes (n), of edges (m), and of triangles (t) are reported. The absence of triangles can be useful for chromatic inequalities.

The results are reported in Tables 2 and 3. In Table 2, we denote by "Optimal" the known optimal values in the literature (see (Caprara et al., 2010), (Caprara et al., 2011), and (Schwarz, 2010)). The column "Optimal" corresponds to optimal values obtained through exact algorithms, such as Branch-and-Cut procedures. Concerning our experiments, "LP(2-3)" is used to denote the optimal value of the relaxed linear program under the valid inequalities of the theorems 2 and 3. Table 3 contains more results on the smaller sized instances. More precisely, Table 3 presents the optimal value of the relaxed linear program under the valid inequalities of the theorems 1-3 (denoted by *LP(1-3)*) that are compared to the results of (*LP(2-3)*). The CPU time (in seconds) of each case is shown in a side column (i.e., *cpu*). There is a time limit of 1200 seconds on CPLEX.

The number of the constraints (10) is huge. Hence, in our experiments, the constraints (10) have not been used. We just considered the remaining constraints of the model (7)-(12) as well as some of the rank inequalities. This concerns the valid inequalities that have been introduced in the theorems 1, 2,

Table 1: The characteristics of the benchmark instances.

Name	n	m	t
*bcspwr*01	39	46	2
*bcspwr*02	49	59	3
*bcspwr*03	118	179	23
*bcspwr*04	274	669	582
*can*_24	24	68	60
*can*_61	61	248	396
*can*_62	62	78	2
*can*_73	73	152	32
*can*_96	96	336	320
*can*_144	144	576	912
*can*_161	161	608	592
*can*_187	187	652	620
*can*_229	229	774	690
*curtis*54	54	124	78
*dwt*_59	59	104	30
*dwt*_66	66	127	62
*dwt*_72	72	75	0
*dwt*_87	87	227	147
*dwt*_162	162	510	464
*dwt*_209	209	767	707
*dwt*_221	221	704	608
*dwt*_245	245	608	374
lshp − 265	265	744	480
*ibm*32	32	90	28
*will*57	57	127	94

and 3. The valid inequalities of the theorems 4 and 5 have been also tested but we have observed that, unfortunately, they do not significantly help to improve the bound. That is why, for the sake of conciseness, we report only the results of the other theorems with more impact.

According to the numerical results that are reported in Tables 2 and 3, we get the following observations:

- the relaxed LP model under the rank inequalities of the theorems 2 and 3 (i.e., *LP(2-3)*) gives (in most of the cases) integer values for the variables f_{kl}.
- the rank inequalities of the theorem 2 are particularly efficient in approximating the integral envelope of the model (7)-(12).
- inclusion of the lifting inequalities (i.e., theorem 1) has a huge influence on the size (consequently, on the computational time) of the problem but contributes slightly in improvement of the lower bounds. Due to this fact, in our experiments, we included only 1500 lifting inequalities in the relaxed LP model.

Table 2: The results of the experiments under the valid inequalities of the theorems 2 and 3 (*LP(2-3)*).

Name	Optimal	cpu	LP(2-3)	cpu
*bcspwr*01	106	0.7	54	0.031
*bcspwr*02	161	1.8	70	0.047
*bcspwr*03	662	189.5	241	0.359
*bcspwr*04	3696	limit	1189	3.295
*can*_24	210	2.8	138	0.015
*can*_61	1137	538	650	0.094
*can*_62	210	4.2	95	0.093
*can*_73	1083	limit	241	0.109
*can*_96	2105	27786	780	0.188
*can*_144	2873	1710.6	1460	0.484
*can*_161	5657	limit	1478	0.577
*can*_187	3827	limit	1496	0.781
*can*_229	7461	limit	1732	1.515
*curtis*54	454	54.5	214	0.062
*dwt*_59	289	27.4	150	0.078
*dwt*_66	192	1.7	189	0.093
*dwt*_72	167	21.2	79	0.124
*dwt*_87	932	1901.4	424	0.172
*dwt*_162	2281	limit	1078	0.578
*dwt*_209	5905	limit	1824	1.421
*dwt*_221	3603	limit	1500	1.702
*dwt*_245	3422	limit	1093	2.233
lshp − 265	5497	limit	1441	1.827
*ibm*32	485	250.5	178	0.031
*will*57	335	30.5	214	0.062

5 CONCLUSIONS

In this paper, we presented a new mathematical program for solving the Minimum Linear Arrangement (MinLA) problem. This formulation has been followed by introduction of some new valid inequalities. We presented some preliminary numerical results showing that, except for *dwt*_66, more investigations are necessary in order to have a better description of the associated polyhedra. In any case, due to the fact that the formulation imply only $O(n^2)$ additional variables, the computational time is very small. As a perspective, additional polyhedral analysis will be done to improve the bound and to develop a branch-and-cut algorithm. The works in these directions are currently in progress.

ACKNOWLEDGEMENTS

This work was financially supported by the region of Haute-Normandie (France) and the European Union.

Table 3: The results of the experiments by taking into account the lifting inequalities.

Name	Optimal	cpu	LP(2-3)	cpu	LP(1-3)	cpu
can_24	210	2.8	138	0.015	150	434.156
curtis54	454	54.5	214	0.062	253	1200.203
dwt_59	289	27.4	150	0.078	154	10.218
dwt_66	192	1.7	189	0.093	191	1.922
will57	335	30.5	214	0.062	263	1200.250

REFERENCES

Amaral, A. (2009). A new lower bound for the single row facility layout problem. In *Discrete Applied Mathematics, 157, pp. 183-190*.

Amaral, A., Caprara, A., Letchford, A., and Gonzalez, J. (2008). A new lower bound for the minimum linear arrangement of a graph. In *Electronic Notes in Discrete Mathematics, 30, pp. 87-92*.

Amaral, A. and Letchford, A. (2011). A polyhedral approach to the single row facility layout problem. In *Technical Report*.

Caprara, A. and Gonzalez, J. (2005). Laying out sparse graphs with provably minimum bandwidth. In *INFORMS Journal on Computing, Vol. 17, No. 3, pp. 356-373*.

Caprara, A., Letchford, A., and Gonzalez, J. (2010). Decorous lower bounds for minimum linear arrangement. In *accepted to INFORMS Journal on Computing*.

Caprara, A., Oswald, M., Reinelt, G., Schwarz, R., and Traversi, E. (2011). Optimal linear arrangements using betweeness variables. In *Math Programming Computation, Vol. 3, pp. 261-280*.

Garey, M., Johnson, D., and Stockmeyer, L. (1976). Some simplified np-complete graph problems. In *Theoretical Computer Science, No. 1, pp. 237-267*.

Horton, S. (1997). *The Optimal Linear Arrangement Problem: Algorithms and Approximation*. PhD Thesis, Georgia Institute of Technology.

Hungerlaender, P. and Rendl, F. (2012). A computational study and survey of methods for the single-row facility layout problem. In *Technical Report*.

Nemhauser, G. and Wolsey, L. (1998). *Integer and Combinatorial Optimization*. Wiley.

Petit, J. (1999). Experiments on the minimum linear arrangement problem. In *Technical Report*.

Schwarz, R. (2010). *A branch-and-cut algorithm with betweenness variables for the Linear Arrangement Problem*. PhD Thesis, Universitaet Heidelberg.