Les Cahiers du GERAD ISSN: 0711–2440

Variable neighborhood search

P. Hansen, N. Mladenović,
J. Brimberg, J. A. Moreno Pérez

G–2009–45

September 2009
Revised: July 2018

La collection *Les Cahiers du GERAD* est constituée des travaux de recherche menés par nos membres. La plupart de ces documents de travail a été soumis à des revues avec comité de révision. Lorsqu'un document est accepté et publié, le pdf original est retiré si c'est nécessaire et un lien vers l'article publié est ajouté.

The series *Les Cahiers du GERAD* consists of working papers carried out by our members. Most of these pre-prints have been submitted to peer-reviewed journals. When accepted and published, if necessary, the original pdf is removed and a link to the published article is added.

CITATION ORIGINALE / ORIGINAL CITATION

P. Hansen, N. Mladenović, J. Brimberg, J. A. Moreno Pérez. Variable neighborhood search, Handbook of Metaheuristics, 3rd edition, International Series in Operations Research & Management Sciences, 272, 2019, http://dx.doi.org/10.1007/978-3-319-91086-4.

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2018
– Bibliothèque et Archives Canada, 2018

Legal deposit – Bibliothèque et Archives nationales du Québec, 2018
– Library and Archives Canada, 2018

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Variable neighborhood search

Pierre Hansen [a,b]
Nenad Mladenović [a,c]
Jack Brimberg [a,d]
José A. Moreno Pérez [e]

[a] *GERAD Montréal (Québec), Canada, H3T 2A7*

[b] *HEC Montréal, Montréal (Québec), Canada, H3T 2A7*

[c] *Mathematical Institute, SANU, 11001 Belgrade, Serbia*

[d] *Department of Mathematics and Computer Science, Royal Military College of Canada, Kingston, (Ontario), Canada, K7K 7B4*

[e] *IUDR & Department of Informatics and Systems Engineering, Universidad de La Laguna, 38271 La Laguna, Spain*

pierreh@crt.umontreal.ca
nenad@mi.sanu.ac.rs
jack.brimberg@rmc.ca
jamoreno@ull.es

September 2009
Revised: July 2018
Les Cahiers du GERAD
G–2009–45

Copyright © 2018 GERAD, Hansen, Mladenović, Brimberg, Moreno Pérez

Les textes publiés dans la série des rapports de recherche *Les Cahiers du GERAD* n'engagent que la responsabilité de leurs auteurs. Les auteurs conservent leur droit d'auteur et leurs droits moraux sur leurs publications et les utilisateurs s'engagent à reconnaître et respecter les exigences légales associées à ces droits. Ainsi, les utilisateurs:

- Peuvent télécharger et imprimer une copie de toute publication du portail public aux fins d'étude ou de recherche privée;
- Ne peuvent pas distribuer le matériel ou l'utiliser pour une activité à but lucratif ou pour un gain commercial;
- Peuvent distribuer gratuitement l'URL identifiant la publication.

Si vous pensez que ce document enfreint le droit d'auteur, contactez-nous en fournissant des détails. Nous supprimerons immédiatement l'accès au travail et enquêterons sur votre demande.

The authors are exclusively responsible for the content of their research papers published in the series *Les Cahiers du GERAD*. Copyright and moral rights for the publications are retained by the authors and the users must commit themselves to recognize and abide the legal requirements associated with these rights. Thus, users:

- May download and print one copy of any publication from the public portal for the purpose of private study or research;
- May not further distribute the material or use it for any profit-making activity or commercial gain;
- May freely distribute the URL identifying the publication.

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Abstract: Variable neighborhood search (VNS) is a metaheuristic for solving combinatorial and global optimization problems whose basic idea is a systematic change of neighborhood both within a descent phase to find a local optimum and in a perturbation phase to get out of the corresponding valley. In this chapter we present the basic schemes of VNS and some of its extensions. We then describe recent developments, i.e., formulation space search and variable formulation search. We then present some families of applications in which VNS has proven to be very successful: (i) exact solution of large scale location problems by primal-dual VNS; (ii) generation of solutions to large mixed integer linear programs, by hybridization of VNS and local branching; (iii) generation of solutions to very large mixed integer programs using VNS decomposition and exact solvers (iv) generation of good feasible solutions to continuous nonlinear programs; (v) adaptation of VNS for solving automatic programming problems from the Artificial intelligence field and (vi) exploration of graph theory to find conjectures, refutations and proofs or ideas of proofs.

Acknowledgments: The work of Nenad Mladenović was conducted at the National Research University Higher School of Economics, Nizhni Novgorod, Russia, and supported by RSF grant 14–41–00039. The fourth author is partially funded by Ministerio de Economía y Competitividad (Spanish Government) with FEDER funds, grant TIN2015-70226-R, and by Fundación Cajacanarias, grant 2016TUR19.

1 Introduction

Optimization tools have greatly improved during the last two decades. This is due to several factors: (i) progress in mathematical programming theory and algorithmic design; (ii) rapid improvement in computer performances; (iii) better communication of new ideas and integration in widely used complex softwares. Consequently, many problems long viewed as out of reach are currently solved, sometimes in very moderate computing times. This success, however, has led researchers and practitioners to address much larger instances and more difficult classes of problems. Many of these may again only be solved heuristically. Therefore thousands of papers describing, evaluating and comparing new heuristics appear each year. Keeping abreast of such a large literature is a challenge. Metaheuristics, or general frameworks for building heuristics, are therefore needed in order to organize the study of heuristics. As evidenced by the Handbook, there are many of them. Some desirable properties of metaheuristics [74, 77, 78] are listed in the concluding section of this chapter.

Variable neighborhood search (VNS) is a metaheuristic proposed by some of the present authors some twenty years ago [98]. Earlier work that motivated this approach can be found in [31, 45, 53, 95]. It is based upon the idea of a systematic change of neighborhood both in a descent phase to find a local optimum and in a perturbation phase to get out of the corresponding valley. Originally designed for approximate solution of combinatorial optimization problems, it was extended to address mixed integer programs, nonlinear programs, and recently mixed integer nonlinear programs. In addition VNS has been used as a tool for automated or computer assisted graph theory. This led to the discovery of over 1500 conjectures in that field and the automated proof of more than half of them. This is to be compared with the unassisted proof of about 400 of these conjectures by many different mathematicians.

Applications are rapidly increasing in number and pertain to many fields: location theory, cluster analysis, scheduling, vehicle routing, network design, lot-sizing, artificial intelligence, engineering, pooling problems, biology, phylogeny, reliability, geometry, telecommunication design, etc. References are too numerous to be listed here, but many of them can be found in [79] and special issues of *IMA Journal of Management Mathematics* [93], *European Journal of Operational Research* [78] and *Journal of Heuristics* [104] that are devoted to VNS.

This chapter is organized as follows. In the next section we present the basic schemes of VNS, i.e., variable neighborhood descent (VND), reduced VNS (RVNS), basic VNS (BVNS) and general VNS (GVNS). Two important extensions are presented in Section 3: Skewed VNS and Variable neighborhood decomposition search (VNDS). A further recent development called Formulation Space Search (FSS) is discussed in section 4. The remainder of the paper describes applications of VNS to several classes of large scale and complex optimization problems for which it has proven to be particularly successful. Section 5 is devoted to primal dual VNS (PD-VNS) and its application to location and clustering problems. Finding feasible solutions to large mixed integer linear programs with VNS is discussed in Section 6. Section 7 addresses ways to apply VNS in continuous global optimization. The more difficult case of solving mixed integer nonlinear programming by VNS is considered in Section 8. Applying VNS to graph theory *per se* (and not just to particular optimization problems defined on graphs) is discussed in Section 9. Brief conclusions are drawn in Section 10.

2 Basic schemes

A deterministic optimization problem may be formulated as

$$\min\{f(x)|x \in X, X \subseteq \mathcal{S}\}, \tag{1}$$

where $\mathcal{S}, X, x$ and f denote the *solution space*, the *feasible set*, a *feasible solution* and a real-valued *objective function*, respectively. If $\mathcal{S}$ is a finite but large set, a *combinatorial optimization* problem is defined. If $\mathcal{S} = \mathbb{R}^n$, we refer to *continuous optimization*. A solution $x^* \in X$ is *optimal* if

$$f(x^*) \leq f(x),\ \forall x \in X.$$

An *exact algorithm* for problem (1), if one exists, finds an optimal solution x^*, together with the proof of its optimality, or shows that there is no feasible solution, i.e., $X = \emptyset$, or the solution is unbounded. Moreover, in practice, the time needed to do so should be finite (and not too long). For continuous optimization, it is reasonable to allow for some degree of tolerance, i.e., to stop when sufficient convergence is detected.

Let us denote $\mathcal{N}_k$, $(k = 1, \ldots, k_{max})$, a finite set of pre-selected neighborhood structures, and $\mathcal{N}_k(x)$ the set of solutions in the k^{th} neighborhood of x. Most local search heuristics use only one neighborhood structure, i.e., $k_{max} = 1$. Often successive neighborhoods $\mathcal{N}_k$ are nested and may be induced from one or more metric (or quasi-metric) functions introduced into a solution space $\mathcal{S}$. An *optimal solution* x_{opt} (or global minimum) is a feasible solution where a minimum is reached. We call $x' \in X$ a *local minimum* of (1) with respect to $\mathcal{N}_k$ (w.r.t. $\mathcal{N}_k$ for short), if there is no solution $x \in \mathcal{N}_k(x') \subseteq X$ such that $f(x) < f(x')$. Metaheuristics (based on local search procedures) try to continue the search by other means after finding the first local minimum. VNS is based on three simple facts:

Fact 1 *A local minimum w.r.t. one neighborhood structure is not necessarily so for another;*

Fact 2 *A global minimum is a local minimum w.r.t. all possible neighborhood structures;*

Fact 3 *For many problems, local minima w.r.t. one or several $\mathcal{N}_k$ are relatively close to each other.*

This last observation, which is empirical, implies that a local optimum often provides some information about the global one. For instance, there may be several variables sharing the same values in both solutions. Since these variables usually cannot be identified in advance, one should conduct an organized study of the neighborhoods of a local optimum until a better solution is found.

In order to solve (1) by using several neighborhoods, facts 1 to 3 can be used in three different ways: (i) deterministic; (ii) stochastic; (iii) both deterministic and stochastic.

We first examine in Algorithm 1 the solution move and neighborhood change function that will be used whithin a VNS framework. Function `NeighborhoodChange()` compares the incumbent value $f(x)$ with the new value $f(x')$ obtained from the k^{th} neighborhood (line 1). If an improvement is obtained, the incumbent is updated (line 2) and k is returned to its initial value (line 3). Otherwise, the next neighborhood is considered (line 4).

Algorithm 1 Neighborhood change.

```
    Function NeighborhoodChange (x, x', k)
1   if f(x') < f(x) then
2   |   x ← x' // Make a move
3   |   k ← 1 // Initial neighborhood
    else
4   |   k ← k + 1 // Next neighborhood
    return x, k
```

Below we discuss Variable Neighborhood Descent and Reduced Variable Neighborhood Search and then build upon this to construct the framework for Basic and General Variable Neighborhood Search.

(i) The **Variable Neighborhood Descent** (VND) method (Algorithm 2) performs a change of neighborhoods in a deterministic way. These neighborhoods are denoted as $N_k, k = 1, \ldots, k_{max}$.

Most local search heuristics use one or sometimes two neighborhoods for improving the current solution (i.e., $k_{max} \leq 2$). Note that the final solution should be a local minimum w.r.t. all k_{max} neighborhoods, and thus, a global optimum is more likely to be reached than with a single structure. Beside this *sequential* order of neighborhood structures in VND, one can develop a *nested* strategy. Assume, for example, that $k_{max} = 3$; then a possible nested strategy is: perform VND with Algorithm 2 for the first two neighborhoods from each point x' that belongs to the third one ($x' \in N_3(x)$). Such an approach is successfully applied in [24, 30, 75].

Algorithm 2 Variable neighborhood descent.

Function VND (x, k_{max})
1 $k \leftarrow 1$
2 **repeat**
3 $\quad x' \leftarrow arg\min_{y \in N_k(x)} f(y)$ // Find the best neighbor in $N_k(x)$
4 $\quad x, k \leftarrow$ NeighborhoodChange (x, x', k) // Change neighborhood
until $k = k_{max}$
return x

(ii) The **Reduced VNS** (RVNS) method is obtained when a random point is selected from $\mathcal{N}_k(x)$ and no descent is attempted from this point. Rather, the value of the new point is compared with that of the incumbent and an update takes place in the case of improvement. We also assume that a stopping condition has been chosen such as the maximum CPU time allowed t_{max}, or the maximum number of iterations between two improvements. To simplify the description of the algorithms, we always use t_{max} below. Therefore, RVNS (Algorithm 3) uses two parameters: t_{max} and k_{max}. .

Algorithm 3 Reduced VNS.

Function RVNS(x, k_{max}, t_{max})
1 **repeat**
2 $\quad k \leftarrow 1$
3 $\quad$ **repeat**
4 $\quad\quad x' \leftarrow$ Shake(x, k)
5 $\quad\quad x, k \leftarrow$ NeighborhoodChange (x, x', k)
$\quad$ **until** $k = k_{max}$
6 $\quad t \leftarrow$ CpuTime()
until $t > t_{max}$
return x

The function Shake in line 4 generates a point x' at random from the k^{th} neighborhood of x, i.e., $x' \in \mathcal{N}_k(x)$. It is given in Algorithm 4, where it is assumed that the points from $\mathcal{N}_k(x)$ are numbered as $\{x^1, \ldots, x^{|\mathcal{N}_k(x)|}\}$. Note that a different notation is used for the neighborhood structures in the shake operation, since these are generally different than the ones used in VND.

Algorithm 4 Shaking function.

Function Shake(x, k)
1 $w \leftarrow \lfloor 1 + \text{Rand}(0, 1) \times |\mathcal{N}_k(x)| \rfloor$
2 $x' \leftarrow x^w$
return x'

RVNS is useful for very large instances for which local search is costly. It can be used as well for finding initial solutions for large problems before decomposition. It has been observed that the best value for the parameter k_{max} is often 2 or 3. In addition, a maximum number of iterations between two improvements is typically used as the stopping condition. RVNS is akin to a Monte-Carlo method, but is more systematic (see, e.g., [100] where results obtained by RVNS were 30% better than those of the Monte-Carlo method in solving a continuous min-max problem). When applied to the p-Median problem, RVNS gave equally good solutions as the *Fast Interchange* heuristic of [121] while being 20 to 40 times faster [80].

(iii) The **Basic VNS** (BVNS) method [98] combines deterministic and stochastic changes of neighborhood. The deterministic part is represented by a local search heuristic. It consists in (i) choosing an initial solution x, (ii) finding a direction of descent from x (within a neighborhood $N(x)$) and (iii) moving to the minimum of $f(x)$ within $N(x)$ along that direction. If there is no direction of descent, the heuristic stops; otherwise it is iterated. Usually the steepest descent direction, also referred to as *best improvement*, is used. Also see Algorithm 2, where the best improvement is used in each neighborhood of the VND. This is summarized in Algorithm 5, where we assume that an initial solution x is given. The output consists of a local minimum, also denoted by x, and its value.

Algorithm 5 Best improvement (steepest descent) heuristic.

```
Function BestImprovement(x)
1 repeat
2 |   x' ← x
3 |   x ← arg min_{y∈N(x')} f(y)
  until (f(x) ≥ f(x'))
return x
```

As *Steepest descent* may be time-consuming, an alternative is to use a *first descent* (or *first improvement*) heuristic. Points $x^i \in N(x)$ are then enumerated systematically and a move is made as soon as a direction for descent is found. This is summarized in Algorithm 6.

Algorithm 6 First improvement (first descent) heuristic.

```
Function FirstImprovement(x)
1 repeat
2 |   x' ← x; i ← 0
3 |   repeat
4 |   |   i ← i + 1
5 |   |   x ← arg min{f(x), f(x^i)}, x^i ∈ N(x)
  |   until (f(x) < f(x')) or i = |N(x)|)
  until (f(x) ≥ f(x'))
return x
```

The stochastic phase of BVNS (see Algorithm 7) is represented by the random selection of a point x' from the k^{th} neighborhood of the shake operation. Note that point x' is generated at random in Step 5 in order to avoid cycling, which might occur with a deterministic rule.

Algorithm 7 Basic VNS.

```
Function BVNS(x, k_max, t_max)
1 t ← 0
2 while t < t_max do
3 |   k ← 1
4 |   repeat
5 |   |   x' ← Shake(x, k)                  // Shaking
6 |   |   x'' ← BestImprovement(x')  // Local search
7 |   |   x, k ← NeighborhoodChange(x, x'', k) // Change neighborhood
  |   until k = k_max
8 |   t ← CpuTime()
return x
```

Example. We illustrate the basic steps on a minimum k-cardinality tree instance taken from [86], see Figure 1. The minimum k-cardinality tree problem on graph G (k-card for short) consists of finding a subtree of G with exactly k edges whose sum of weights is minimum.

The steps of BVNS for solving the 4-card problem are illustrated in Figure 2. In Step 0 the objective function value, i.e., the sum of edge weights, is equal to 40; it is indicated in the right bottom corner of the figure. That first solution is a local minimum with respect to the edge-exchange neighborhood structure (one edge in, one out). After shaking, the objective function is 60, and after another local search, we are back to the same solution. Then, in Step 3, we take out 2 edges and add another 2 at random, and after a local search, an improved solution is obtained with a value of 39. Continuing in that way, the optimal solution with an objective function value equal to 36 is obtained in Step 8.

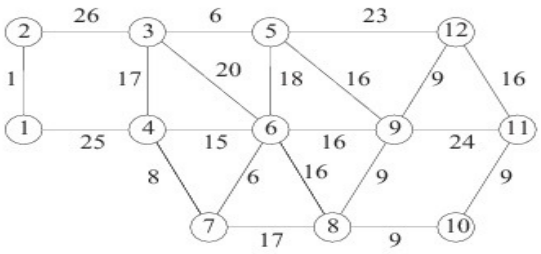


Figure 1: 4-cardinality tree problem.

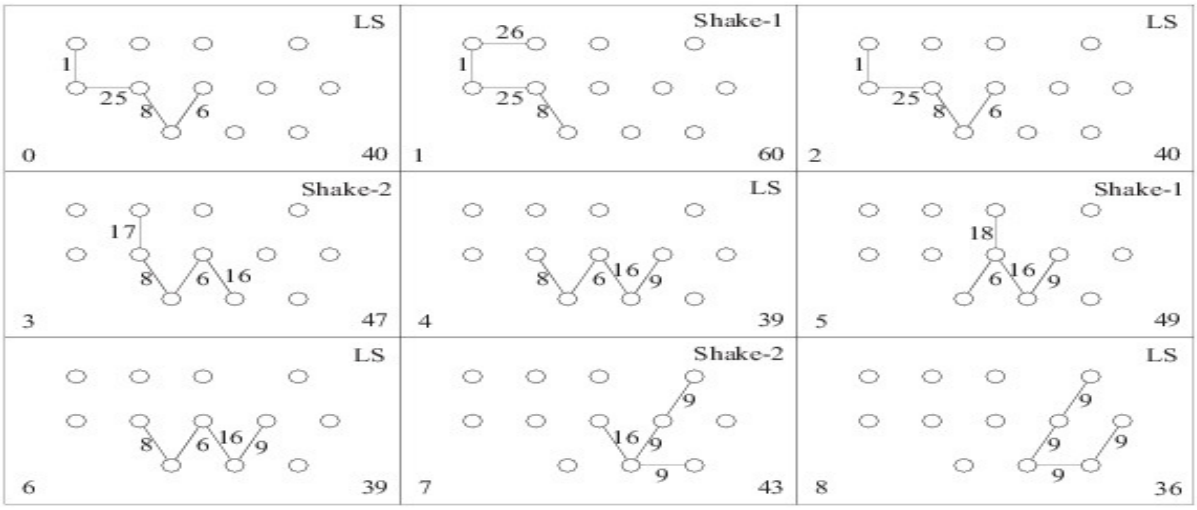


Figure 2: Steps of the Basic VNS for solving 4-card tree problem.

(iv) **General VNS.** Note that the local search step (line 6 in BVNS, algorithm 7) may also be replaced by VND (Algorithm 2). This General VNS (VNS/VND) approach has led to some of the most successful applications reported in the literature (see, e.g., [2, 30, 33, 35, 37, 39, 40, 48, 75, 81, 111, 112]). General VNS (GVNS) is outlined in Algorithm 8 below. Note that neighborhoods $N_1, ..., N_{l_{max}}$ are used in the VND step, while a different series of neighborhoods $N_1, ..., N_{k_{max}}$ apply to the Shake step.

Algorithm 8 General VNS.

Function GVNS $(x, \ell_{max}, k_{max}, t_{max})$
1 **repeat**
2 $k \leftarrow 1$
3 **repeat**
4 $x' \leftarrow$ Shake(x, k)
5 $x'' \leftarrow$ VND(x', ℓ_{max})
6 $x, k \leftarrow$ NeighborhoodChange(x, x'', k)
until $k = k_{max}$
7 $t \leftarrow$ CpuTime()
until $t > t_{max}$
return x

3 Some extensions

(i) The **Skewed VNS** (SVNS) method [69] addresses the problem of exploring valleys far from the incumbent solution. Indeed, once the best solution in a large region has been found it is necessary to go quite far to

obtain an improved one. Solutions drawn at random in far-away neighborhoods may differ substantially from the incumbent, and VNS may then degenerate, to some extent, into a Multistart heuristic (where descents are made iteratively from solutions generated at random, and which is known to be inefficient). So some compensation for distance from the incumbent must be made, and a scheme called Skewed VNS (SVNS) is proposed for that purpose. Its steps are presented in Algorithms 9, 10 and 11. The `KeepBest`(x, x') function (Algorithm 9) in SVNS simply keeps the best of solutions x and x' The `NeighborhoodChangeS` function (Algorithm 10) performs the move and neighborhood change for the SVNS.

Algorithm 9 Keep best solution.

```
Function KeepBest(x, x')
 1 if f(x') < f(x) then
 2     x ← x'
return x
```

Algorithm 10 Neighborhood change for Skewed VNS.

```
Function NeighborhoodChangeS(x, x', k, α)
 1 if f(x') − αρ(x, x') < f(x) then
 2     x ← x'; k ← 1
   else
 3     k ← k + 1
return x,k
```

Algorithm 11 Skewed VNS.

```
Function SVNS (x, k_max, t_max, α)
 1 x_best ← x
 2 repeat
 3     k ← 1
 4     repeat
 5         x' ← Shake(x, k)
 6         x'' ← FirstImprovement(x')
 7         x, k ← NeighborhoodChangeS(x, x'', k, α)
 8         x_best ← KeepBest (x_best, x)
       until k = k_max
 9     x ← x_best
10     t ← CpuTime()
   until t > t_max
return x
```

SVNS makes use of a function $\rho(x, x'')$ to measure the distance between the current solution x and the local optimum x''. The distance function used to define $\mathcal{N}_k$ could also be used for this purpose. The parameter α must be chosen to allow movement to valleys far away from x when $f(x'')$ is larger than $f(x)$ but not too much larger (otherwise one will always leave x). A good value for α is found experimentally in each case. Moreover, in order to avoid frequent moves from x to a close solution, one may take a smaller value for α when $\rho(x, x'')$ is small. More sophisticated choices for selecting a function of $\alpha\rho(x, x'')$ could be made through some learning process.

(ii) The **Variable neighborhood decomposition search** (VNDS) method [80] extends the basic VNS into a two-level VNS scheme based upon decomposition of the problem. It is presented in Algorithm 12, where t_d is an additional parameter that represents the running time allowed for solving decomposed (smaller-sized) problems by Basic VNS (line 5).

For ease of presentation, but without loss of generality, we assume that the solution x represents a set of attributes. In Step 4 we denote by y a set of k solution attributes present in x' but not in x ($y = x' \setminus x$). In Step 5 we find the local optimum y' in the space of y; then we denote with x'' the corresponding solution in the whole space X ($x'' = (x' \setminus y) \cup y'$). We notice that exploiting some *boundary effects* in a new solution can significantly improve solution quality. That is why, in Step 6, the local optimum x''' is found in the whole

Algorithm 12 Variable neighborhood decomposition search.

Function VNDS $(x, k_{max1}, t_{max}, t_d)$
repeat
  $k \leftarrow 1$
  repeat
    $x' \leftarrow$ Shake (x, k); $y \leftarrow x' \setminus x$
    $y' \leftarrow$ BVNS(y, k_{max2}, t_d); $x'' = (x' \setminus y) \cup y'$
    $x''' \leftarrow$ FirstImprovement(x'')
    $x, k \leftarrow$ NeighborhoodChange(x, x''', k)
  until $k = k_{max1}$
until $t > t_{max}$
return x

space X using x'' as an initial solution. If this is time consuming, then at least a few local search iterations should be performed.

VNDS can be viewed as embedding the classical successive approximation scheme (which has been used in combinatorial optimization at least since the sixties, see, e.g., [61]) in the VNS framework. Let us mention here a few applications of VNDS: (i) p-median problem [80]; simple plant location problem [68]; k-cardinality tree problem [119]; 0-1 mixed integer programming problem [88, 64]; design of MBA student teams [46], etc.

4 Changing formulation within VNS

A traditional approach to tackle an optimization problem is to consider a given formulation and search in some way through its feasible set X. Given that the same problem can often be formulated in different ways, it is possible to extend search paradigms to include jumps from one formulation to another. Each formulation should lend itself to some traditional search method, its 'local search' that works totally within this formulation, and yields a final solution when started from some initial solution. Any solution found in one formulation should easily be translatable to its equivalent solution in any other formulation. We may then move from one formulation to another by using the solution resulting from the local search of the former as an initial solution for the local search of the latter. Such a strategy will of course only be useful when local searches in different formulations behave differently. Here we discuss two such possibilities.

4.1 Variable neighborhood-based formulation space search

The idea of changing the formulation of a problem was investigated in [101] using an approach that systematically alternates between different formulations for solving various Circle Packing Problems (CPP). It is shown there that a stationary point for a nonlinear programming formulation of CPP in Cartesian coordinates is not necessarily a stationary point in polar coordinates. A method called *Reformulation Descent* (RD) that alternates between these two formulations until the final solution is stationary with respect to both formulations is suggested. Results obtained were comparable with the best known values, but were achieved about 150 times faster than with an alternative single formulation approach. In this paper, the idea suggested above of *Formulation Space Search* (FSS) is also introduced, using more than two formulations. Some research in that direction has also been reported in [84, 96, 107]. One methodology that uses the variable neighborhood idea when searching through the formulation space is given in Algorithms 13 and 14. Here ϕ (ϕ') denotes a formulation from a given space $\mathcal{F}$, x (x') denotes a solution in the feasible set defined with that formulation, and $\ell \leq \ell_{max}$ is the formulation neighborhood index. Note that Algorithm 14 uses a reduced VNS strategy in the formulation space $\mathcal{F}$. Note also that the `ShakeFormulation()` function must provide a search through the solution space $\mathcal{S}'$ (associated with formulation ϕ') in order to get a new solution x'. Any appropriate method can be used for this purpose.

4.2 Variable formulation search

Many optimization problems in the literature, e.g., min-max problems, demonstrate a flat landscape. It means that, given a formulation of the problem, many neighbors of a solution have the same objective function value.

Algorithm 13 Formulation change.

Function `FormulationChange`$(x, x', \phi, \phi', \ell)$
if $f(\phi', x') < f(\phi, x)$ **then**
  $\phi \leftarrow \phi'$
  $x \leftarrow x'$
  $\ell \leftarrow 1$
else
  $\ell \leftarrow \ell + 1$
return x, ϕ, ℓ

Algorithm 14 Reduced variable neighborhood FSS.

Function `VNFSS`(x, ϕ, ℓ_{max})
repeat
  $\ell \leftarrow 1$ // Initialize formulation in $\mathcal{F}$
  while $\ell \leq \ell_{max}$ **do**
    $x', \phi', \ell \leftarrow$ `ShakeFormulation`(x,x',ϕ,ϕ',ℓ) // $(\phi',x') \in (N_\ell(\phi),\mathcal{N}(x))$ at random
    $x, \phi, \ell \leftarrow$ `FormulationChange`(x,x',ϕ,ϕ',ℓ) // Change formulation
  until *some stopping condition is met*
return x

When this happens, it is difficult to determine which neighborhood solution is more promising to continue the search. To address this drawback, the use of alternative formulations of the problem within VNS is proposed in [99, 103, 106]. In [106] it is named Variable Formulation Search (VFS). It combines a change of neighborhood within the VNS framework, with the use of alternative formulations.

Let us assume that, beside the original formulation and the corresponding objective function $f_0(x)$, there are p other formulations denoted as $f_1(x), .., f_p(x), x \in X$. Note that two formulations are defined as equivalent if the optimal solution of one is the optimal solution of the other, and vice versa. For simplification purposes, we will denote different formulations as different objectives $f_i(x), i = 1, .., p$. The idea of VFS is to add the procedure `Accept`(x, x', p), given in Algorithm 15, in all three basic steps of BVNS: `Shaking`, `LocalSearch` and `NeighborhoodChange`. Clearly, if a better solution is not obtained by any of the $p + 1$ formulations, the move is rejected. The next iteration in the loop of Algorithm 15 will take place only if the objective function values according to all previous formulations are equal.

Algorithm 15 Accept procedure with p secondary formulations.

Logical Function `Accept` (x, x', p)
for $i = 0$ ***to*** p **do**
  if $(f_i(x') < f_i(x))$ **then return** `TRUE`
  if $(f_i(x') > f_i(x))$ **then return** `FALSE`
return `FALSE`

If `Accept` (x, x', p) is included in the `LocalSearch` subroutine of BVNS, then it will not stop the first time a non improved solution is found. In order to stop `LocalSearch` and thus claim that x' is a local minimum, x' should not be improved by any among the p different formulations. Thus, for any particular problem, one needs to design different formulations of the problem considered and decide the order in which they will be used in the `Accept` subroutine. Answers to those two questions are problem specific and sometimes not easy. The `Accept` (x, x', p) subroutine can obviously be added to the `NeighborhoodChange` and `Shaking` steps of BVNS from Algorithm 7 as well.

In [103], three evaluation functions, or acceptance criteria, within the `Neighborhood Change` step are used in solving the *Bandwidth Minimization Problem*. This min-max problem consists of finding permutations of rows and columns of a given square matrix to minimize the maximal distance of the nonzero elements from the main diagonal in the corresponding rows. Solution x may be represented as a labeling of a graph and the move from x to x' as $x \rightarrow x'$. Three criteria are used:

(1) the bandwidth length $f_0(x)$ $(f_0(x') < f_0(x))$;

(2) the total number of critical vertices $f_1(x)$ $(f_1(x') < f_1(x))$, if $f_0(x') = f_0(x)$;

(3) $f_3(x, x') = \rho(x, x') - \alpha$, if $f_0(x') = f_0(x)$ and $f_1(x') = f_1(x)$. Here, we want $f_3(x, x') > 0$, because we assume that x and x' are sufficiently far from one another when $\rho(x, x') > \alpha$, where α is an additional parameter. The idea for a move to an even worse solution, if it is very far, is used within Skewed VNS. However, a move to a solution with the same value is only performed in [103] if its Hamming distance from the incumbent is greater than α.

In [99] a different mathematical programming formulation of the original problem is used as a secondary objective within the `Neighborhood Change` function of VNS. There, two combinatorial optimization problems on a graph are considered: the *Metric Dimension Problem* and *Minimal Doubly Resolving Set Problem*.

A more general VFS approach is given in [106], where the *Cutwidth Graph Minimization Problem* (CWP) is considered. CWP also belongs to the min-max problem family. For a given graph, one needs to find a sequence of nodes such that the maximum cutwidth is minimum. The cutwidth of a graph should be clear from the example provided in Figure 3 for the graph with six vertices and nine edges shown in (a).

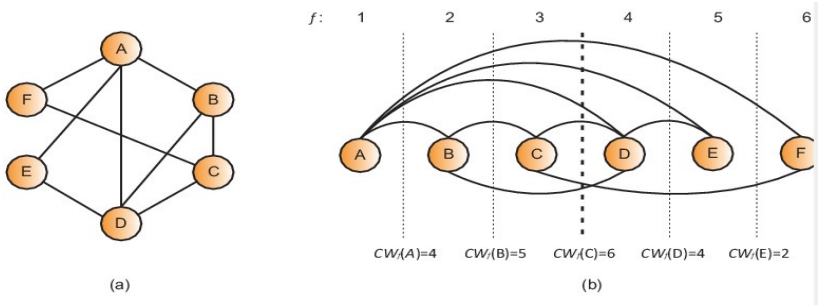


Figure 3: Cutwidth minimization example as in [106].

Figure 3(b) shows an ordering x of the vertices of the graph in (a) with the corresponding cutwidth CW values of each vertex. It is clear that the CW represents the number of cut edges between two consecutive nodes in the solution x. The cutwidth value $f_0(x) = CW(x)$ of the ordering $x = (A, B, C, D, E, F)$ is equal to $f_0(x) = \max\{4, 5, 6, 4, 2\} = 6$. Thus, one needs to find an order x that minimizes the maximum cut-width value over all vertices.

Beside minimizing the bandwidth f_0, two additional formulations, denoted f_1 and f_2 , are used in [106], and implemented within a VND local search. Results are compared among themselves (Table 1) and with a few heuristics from the literature (Table 2), using the following usual data set:

- "*Grid*": This data set consists of 81 matrices constructed as the Cartesian product of two paths. They were originally introduced by Rolim et al. [113]. For this set of instances, the vertices are arranged on a grid of dimension width × height where width and height are selected from the set {3, 6, 9, 12, 15, 18, 21, 24, 27}.
- "*Harwell-Boeing*" (HB): This data set is a subset of the public-domain Matrix Market library[1]. This collection consists of a set of standard test matrices $M = (M_{ij})$ arising from problems in linear systems, least squares, and eigenvalue calculations from a wide variety of scientific and engineering disciplines. Graphs were derived from these matrices by considering an edge (i, j) for every element $M_{ij} \neq 0$. The

[1]Available at `http://math.nist.gov/MatrixMarket/data/Harwell-Boeing/`

data set is formed by the selection of the 87 instances were $n \leq 700$. Their number of vertices ranges from 30 to 700 and the number of edges from 46 to 41686.

Table 1 presents the results obtained with four different VFS variants, after executing them for 30 seconds over each instance. The column 'BVNS' of Table 1 represents a heuristic based on BVNS which makes use only of the original formulation f_0 of the CWP. VFS_1 denotes a BVNS heuristic that uses only one secondary criterion, i.e., f_0 and f_1. VFS_2 is equivalent to the previous one with the difference that now f_2 is considered (instead of f_1). Finally, the fourth column of the table, denoted as VFS_3, combines the original formulation of the CWP with the two alternative ones, in the way presented in Algorithm 15. All algorithms were configured with $k_{max} = 0.1n$ and start from the same random solution.

Table 1: Comparison of alternative formulations within 30 seconds for each test, by average objective values and % deviation from the best known solution. Test are performed on "Grid" and "HB" data sets that contain 81 and 86 instances, respectively.

	BVNS	**VFS_1**	**VFS_2**	**VFS_3**
Avg.	137.31	93.56	91.56	90.75
Dev. (%)	192.44	60.40	49.23	48.22

It appears that significant improvements in solution quality are obtained when at least one secondary formulation is used in case of ties (compare e.g., 192.44% and 60.40% deviations from the best known solutions obtained by BVNS and VFS_1, respectively). An additional improvement is obtained when all three formulations are used in VFS_3.

Comparison of VFS_3 and state-of-the-art heuristics are given in Table 2. There, the stopping condition is increased from 30 seconds to 300 and 600 seconds for the first and the second set of instances, respectively. Besides average values and % deviation, the methods are compared based on the number of wins (the third row) and the total cpu time in seconds. Overall, the best quality results are obtained by VFS in less computing time.

Table 2: Comparison of VFS with the state-of-the-art heuristics over the "Grid" and "HB" data sets, within 300 and 600 seconds respectively.

	81 'grid' test instances				**86 HB instances**			
	GPR [3]	**SA** [44]	**SS** [105]	**VFS** [106]	**GPR** [3]	**SA** [44]	**SS** [105]	**VFS** [106]
Avg.	38.44	16.14	13.00	12.23	364.83	346.21	315.22	314.39
Dev. (%)	201.81	25.42	7.76	3.25	95.13	53.30	3.40	1.77
#Opt.	2	37	44	59	2	8	47	61
CPU t (s)	235.16	216.14	210.07	90.34	557.49	435.40	430.57	128.12

It appears that the best quality results are obtained by VFS in less computing time.

5 Primal-dual VNS

For most modern heuristics, the difference in value between the optimal solution and the obtained approximate solution is not precisely known. Guaranteed performance of the primal heuristic may be determined if a lower bound on the objective function value can be found. To this end, the standard approach is to relax the integrality condition on the primal variables, based on a mathematical programming formulation of the problem. However, when the dimension of the problem is large, even the relaxed problem may be impossible to solve exactly by standard commercial solvers. Therefore, it seems to be a good idea to solve dual relaxed problems heuristically as well. In this way we get guaranteed bounds on the primal heuristic performance. The next difficulty arises if we want to get an exact solution within a branch-and-bound framework since having the approximate value of the relaxed dual does not allow us to branch in an easy way, for example by exploiting complementary slackness conditions. Thus, the exact value of the dual is necessary. A general approach to get both guaranteed bounds and an exact solution is proposed in [68], and referred as Primal-Dual VNS (PD-VNS). It is given in Algorithm 16.

Algorithm 16 Basic PD-VNS.

Function `PD-VNS` (x, k_{max}, t_{max})
1 `BVNS` (x, k_{max}, t_{max}) // Solve primal by VNS
2 `DualFeasible`(x, y) // Find (infeasible) dual such that $f_P = f_D$
3 `DualVNS`(y) // Use VNS do decrease infeasibility
4 `DualExact`(y) // Find exact (relaxed) dual
5 `BandB`(x, y) // Apply branch-and-bound method

In the first stage, a heuristic procedure based on VNS is used to obtain a near optimal solution. In [68] it is shown that VNS with decomposition is a very powerful technique for large-scale simple plant location problems (SPLP) with up to 15 000 facilities and 15 000 users. In the second phase, the objective is to find an exact solution of the relaxed dual problem. Solving the relaxed dual is accomplished in three stages: (i) find an initial dual solution (generally infeasible) using the primal heuristic solution and complementary slackness conditions; (ii) find a feasible solution by applying VNS to the unconstrained nonlinear form of the dual; (iii) solve the dual exactly starting with the found initial feasible solution using a customized "sliding simplex" algorithm that applies "windows" on the dual variables, thus substantially reducing the problem size. On all problems tested, including instances much larger than those previously reported in the literature, the procedure was able to find the exact dual solution in reasonable computing time. In the third and final phase, armed with tight upper and lower bounds obtained from the heuristic primal solution in phase one and the exact dual solution in phase two, respectively, a standard branch-and-bound algorithm is applied to find an optimal solution of the original problem. The lower bounds are updated with the dual sliding simplex method and the upper bounds whenever new integer solutions are obtained at the nodes of the branching tree. In this way it was possible to solve exactly problem instances of sizes up to 7 000 facilities $\times$7 000 users, for uniform fixed costs, and 15 000 facilities $\times$15 000 users, otherwise.

6 VNS for mixed integer linear programming

The Mixed Integer Linear Programming (MILP) problem consists of maximizing or minimizing a linear function, subject to equality or inequality constraints and integrality restrictions on some of the variables. The mixed integer programming problem $(MILP)$ can be expressed as:

$$(MILP) \quad \left[\begin{array}{lll} \min & \sum_{j=1}^{n} c_j x_j & \\ \text{s.t.} & \sum_{j=1}^{n} a_{ij} x_j \geq b_i & \forall i \in M = \{1, 2, \ldots, m\} \\ & x_j \in \{0, 1\} & \forall j \in \mathcal{B} \\ & x_j \geq 0, \texttt{integer} & \forall j \in \mathcal{G} \\ & x_j \geq 0 & \forall j \in \mathcal{C} \end{array} \right.$$

where the set of indices $N = \{1, 2, \ldots, n\}$ is partitioned into three subsets $\mathcal{B}, \mathcal{G}$ and $\mathcal{C}$, corresponding to binary, general integer and continuous variables, respectively.

Numerous combinatorial optimization problems, including a wide range of practical problems in business, engineering and science, can be modeled as MILPs. Several special cases, such as knapsack, set packing, cutting and packing, network design, protein alignment, traveling salesman and other routing problems, are known to be NP-hard [57].

Many commercial solvers such as CPLEX [85] are available for solving MILPs. Methods included in such software packages are usually of the branch-and-bound (B&B) or of branch-and-cut (B&C) types. Basically, those methods enumerate all possible integer values in some order, and prune the search space for the cases where such enumeration cannot improve the current best solution.

6.1 Variable neighborhood branching

The connection between local search based heuristics and exact solvers may be established by introducing the so called *local branching constraints* [52]. By adding just one constraint into (MILP), as explained below, the k^{th} neighborhood of (MILP) is defined. This allows the use of all local search based metaheuristics, such

as Tabu search, Simulating annealing, VNS etc. More precisely, given two solutions x and y of (MILP), the distance between x and y is defined as:

$$\delta(x, y) = \sum_{j \in \mathcal{B}} \mid x_j - y_j \mid.$$

Let X be the solution space of (MILP). The neighborhood structures $\{\mathcal{N}_k \mid k = 1, \ldots, k_{max}\}$ can be defined, knowing the distance $\delta(x, y)$ between any two solutions $x, y \in X$. The set of all solutions in the k^{th} neighborhood of $y \in X$ is denoted as $\mathcal{N}_k(y)$ where

$$\mathcal{N}_k(y) = \{x \in X \mid \delta(x, y) \ \leq \ k\}.$$

For the pure 0-1 MILP given above (MILP with $\mathcal{G} = \emptyset$), $\delta(.,.)$ represents the Hamming distance and $\mathcal{N}_k(y)$ may be expressed by the following *local branching constraint*

$$\delta(x, y) = \sum_{j \in S} (1 - x_j) + \sum_{j \in \mathcal{B} \setminus S} x_j \leq k, \tag{2}$$

where $S = \{j \in \mathcal{B} \mid y_j = 1\}$.

In [81] a general VNS procedure for solving 0-1 MILPs is presented (see Algorithm 17). An exact MILP solver (`MIPSOLVE()` within CPLEX) is used as a black box for finding the best solution in the neighborhood, based on the given formulation (MILP) plus the added local branching constraints. Shaking is performed using the Hamming distance defined above. A detailed description of this VNS branching method is provided in Algorithm 17. The variables and constants used in the algorithm are defined as follows [81]:

- UB - input variable for the CPLEX solver which represents the current upper bound.
- $first$ - logical input variable for CPLEX solver which is `true` if the first solution lower than UB is asked for in the output; if $first$ = `false`, CPLEX returns the best solution found so far.
- TL - maximum time allowed for running CPLEX.
- rhs - right hand side of the local branching constraint; it defines the size of the neighborhood within the inner or VND loop.
- $cont$ - logical variable which indicates if the inner loop continues (`true`) or not (`false`).
- x_opt and f_opt - incumbent solution and corresponding objective function value.
- x_cur, f_cur, k_cur - current solution, objective function value and neighborhood from where the VND local search starts (lines 6-20).
- x_next and f_next - solution and corresponding objective function value obtained by CPLEX in the inner loop.

Algorithm 17 VNS Branching.

```
Function VnsBra(total_time_limit, node_time_limit, k_step, x_opt)
 1  TL := total_time_limit; UB := ∞; first := true
 2  stat := MIPSOLVE(TL, UB, first, x_opt, f_opt)
 3  x_cur:=x_opt; f_cur:=f_opt
 4  while (elapsedtime < total_time_limit) do
 5      cont := true; rhs := 1; first := false
 6      while (cont or elapsedtime < total_time_limit) do
 7          TL = min(node_time_limit, total_time_limit-elapsedtime)
 8          add local br. constr. δ(x, x_cur) ≤ rhs; UB := f_cur
 9          stat := MIPSOLVE(TL, UB, first, x_next, f_next)
10          switch stat do
11              case "opt_sol_found":
12                  reverse last local br. const. into δ(x, x_cur) ≥ rhs + 1
13                  x_cur := x_next; f_cur := f_next; rhs := 1;
14              case "feasible_sol_found":
15                  reverse last local br. constr. into δ(x, x_cur) ≥ 1
16                  x_cur := x_next; f_cur := f_next; rhs := 1;
17              case "proven_infeasible":
18                  remove last local br. constr.; rhs := rhs+1;
19              case "no_feasible_sol_found":
20                  cont := false
21      if f_cur < f_opt then
22          x_opt := x_cur; f_opt := f_cur; k_cur := k_step;
        else
23          k_cur := k_cur+k_step;
24      remove all added constraints; cont := true
25      while cont and (elapsedtime < total_time_limit) do
26          add constraints k_cur ≤ δ(x, x_opt) and δ(x, x_opt) < k_cur + k_step
27          TL := total_time_limit-elapsedtime; UB := ∞; first := true
28          stat := MIPSOLVE(TL, UB, first, x_cur, f_cur)
29          remove last two added constraints; cont =false
30          if stat = "proven_infeasible" or "no_feasible_solution_found" then
31              cont :=true; k_cur := k_cur+k_step
```

In line 2, a commercial MIP solver is run to get an initial feasible solution, i.e., logical variable 'first' is set to value `true`. The outer loop starts from line 4. VND based local search is performed in the inner loop that starts from line 6 and finishes at line 24. There are 4 different outputs from subroutine `MIPSOLVE` provided by variable *stat*. They are coded in lines 11-20. The shaking step also uses the MIP solver. It is presented in the loop that starts at line 25.

6.2 VNDS based heuristics for MILP

It is well known that heuristics and relaxations are useful for providing upper and lower bounds on the optimal value of large and difficult optimization problems. A hybrid approach for solving 0-1 MILPs is presented in this section. A more detailed description may be found in [64]. It combines variable neighborhood decomposition search (VNDS) [80] and a generic MILP solver for upper bounding purposes, and a generic linear programming solver for lower bounding. VNDS is used to define a variable fixing scheme for generating a sequence of smaller subproblems, which are normally easier to solve than the original problem. Different heuristics are derived by choosing different strategies for updating lower and upper bounds, and thus defining different schemes for generating a series of subproblems. We also present in this section a two-level decomposition scheme, in which subproblems created according to the VNDS rules are further divided into smaller subproblems using another criterion, derived from the mathematical formulation of the problem.

6.2.1 VNDS for 0-1 MILPs with Pseudo-cuts

Variable neighborhood decomposition search is a two-level variable neighborhood search scheme for solving optimization problems, based upon the decomposition of the problem (see Algorithm 12). We discuss here an algorithm which solves exactly a sequence of reduced problems obtained from a sequence of linear programming relaxations. The set of reduced problems for each LP relaxation is generated by fixing a certain number of variables according to VNDS rules. That way, two sequences of upper and lower bounds are generated, until an optimal solution of the problem is obtained. Also, after each reduced problem is solved, a pseudo-cut is added to guarantee that this subproblem is not revisited. Furthermore, whenever an improvement in the objective function value occurs, a local search procedure is applied in the whole solution space to attempt a further improvement (the so-called *boundary effect* within VNDS). This procedure is referred to as VNDS-PC, since it employs VNDS to solve 0-1 MILPs, while incorporating pseudo-cuts to reduce the search space [64].

If $J \subseteq \mathcal{B}$, we define the partial distance between x and y, relative to J, as $\delta(J, x, y) = \sum_{j \in J} \mid x_j - y_j \mid$. Obviously we have $\delta(\mathcal{B}, x, y) = \delta(x, y))$. More generally, let $\bar{x}$ be an optimal solution of LP(P), the LP relaxation of the problem P considered (not necessarily MIP feasible), and $J \subseteq B(\bar{x}) = \{j \in N \mid \bar{x}_j \in \{0, 1\}\}$ an arbitrary subset of indices, The partial distance $\delta(J, x, \overline{x})$ can be linearized as follows:

$$\delta(J, x, \overline{x}) = \sum_{j \in J} [x_j(1 - \overline{x}_j) + \overline{x}_j(1 - x_j)].$$

Let X be the solution space of problem P. The neighborhood structures $\{\mathcal{N}_k \mid k = k_{min}, \ldots, k_{max}\}$, $1 \leq k_{min} \leq k_{max} \leq p$, can be defined knowing the distance $\delta(\mathcal{B}, x, y)$ between any two solutions $x, y \in X$. The set of all solutions in the k^{th} neighborhood of $x \in X$ is denoted as $\mathcal{N}_k(x)$, where

$$\mathcal{N}_k(x) = \{y \in X \mid \delta(\mathcal{B}, x, y) \ \leq \ k\}.$$

From the definition of $\mathcal{N}_k(x)$, it follows that $\mathcal{N}_k(x) \subset \mathcal{N}_{k+1}(x)$, for any $k \in \{k_{min}, k_{min} + 1, \ldots, k_{max} - 1\}$, since $\delta(\mathcal{B}, x, y) \ \leq \ k$ implies $\delta(\mathcal{B}, x, y) \ \leq \ k + 1$. It is trivial that, if we completely explore neighborhood $\mathcal{N}_{k+1}(x)$, it is not necessary to explore neighborhood $\mathcal{N}_k(x)$.

Ordering variables w.r.t LP–relaxation. The first variant of VNDS-PC, denoted as `VNDS-PC1`, is considered here for the maximization case. See Algorithm 18 for the pseudo-code of this algorithm which can be easily adjusted for minimization problems. Input parameters for the algorithm are an instance P of the 0-1 MIP problem, a parameter d which defines the number of variables to be released in each iteration and an initial feasible solution x^* of P. The algorithm returns the best solution found until the stopping criterion defined by the variable *proceed*1 is met.

This variant of `VNDS-PC` is based on the following choices. Variables are ordered according to their distances from the corresponding LP relaxation solution values (see lines 4, 6 and 7 in Algorithm 18). More precisely, we compute distances $\delta_j = \mid x_j - \overline{x}_j \mid$ for $j \in \mathcal{B}$, where x_j is a variable value of the current incumbent (feasible) solution and $\overline{x}_j$ a variable value of the LP-relaxation. We then index variables $x_j, j \in \mathcal{B}$, so that $\delta_1 \leq \delta_2 \leq \ldots \leq \delta_p$, $p = \mid \mathcal{B} \mid$. Parameters k_{min}, k_{step} and k_{max} (see line 9 in Algorithm 18) are determined in the following way. Let q be the number of binary variables which have different values in the LP relaxation solution and in the incumbent solution ($q = \mid \{j \in \mathcal{B} \mid \delta_j \neq 0\} \mid$), and let d be a given parameter (whose value is experimentally found) which controls the neighborhood size. Then we set $k_{min} = p - q$, $k_{step} = \lfloor q/d \rfloor$ and $k_{max} = p - k_{step}$. We also allow the value of k to be less then k_{min} (see lines 17 and 18 in Algorithm 18). In other words, we allow the variables which have the same integer value in the incumbent and LP-relaxation solutions to be freed anyway. When $k < k_{min}$, k_{step} is set to (approximately) half the number of remaining fixed variables. Note that the maximum value of parameter k (which is k_{max}) indicates the maximum possible number of fixed variables, which implies the minimum number of free variables and therefore the minimum possible neighborhood size in the VNDS scheme.

If an improvement occurs after solving the subproblem $P(x^*, J_k)$, where x^* is the current incumbent solution (see line 12 in Algorithm 18), we perform a local search on the complete solution, starting from x' (see line 14 in Algorithm 18). The local search applied at this stage is the variable neighborhood descent

Algorithm 18 VNDS for MIPs with pseudo-cuts.

Function `VNDS-PC1`(P, d, x^*)
1 Choose stopping criteria (set $proceed1 = proceed2 =$ `true`)
2 Add objective cut: $LB = c^{\mathrm{T}}x^*$; $P = (P \mid c^{\mathrm{T}}x > LB)$
3 **while** $proceed1$ **do**
4 Find an optimal solution $\overline{x}$ of LP(P)
5 set $UB = c^{\mathrm{T}}\overline{x}$
6 **if** $B(\overline{x}) = \mathcal{B}$ **then**
  break
7 Set $\delta_j = \mid x^*_j - \overline{x}_j \mid$, $j \in \mathcal{B}$
8 Index x_j so that $\delta_j \leq \delta_{j+1}$, $j = 1, \ldots, p-1$, $p = |\mathcal{B}|$
9 Set $q = \mid \{j \in \mathcal{B} \mid \delta_j \neq 0\} \mid$
10 Set $k_{min} = p - q$, $k_{step} = \lfloor q/d \rfloor$, $k_{max} = p - k_{step}$, $k = k_{max}$
11 **while** $proceed2$ **and** $k \geq 0$ **do**
12 $J_k = \{1, \ldots, k\}$; $x' =$ `MIPSOLVE`$(P(x^*, J_k), x^*)$
13 $P = (P \mid \delta(J_k, x^*, x) \geq 1)$
14 **if** $(c^{\mathrm{T}}x' > c^{\mathrm{T}}x^*)$ **then**
15 $x^* =$ `LocalSearch`(P, x'); $LB = c^{\mathrm{T}}x^*$
16 Update objective cut: $P = (P \mid c^{\mathrm{T}}x > LB)$; **break**
  else
17 **if** $(k - k_{step} < k_{min})$ **then**
  $k_{step} = \max\{\lfloor k/2 \rfloor, 1\}$
18 Set $k = k - k_{step}$
19 Update $proceed2$
20 Update $proceed1$
return LB, UB, x^*.

for 0-1 MILPs, as described in [81]. Note that, in Algorithm 18 and in the pseudo-codes that follow, the statement $y =$ `MILPSOLVE`(P, x) denotes a call to a generic MILP solver, for a given 0-1 MILP problem P, starting from a given solution x and returning a new solution y (if P is infeasible, then the value of y remains the same as the one before the call to the MILP solver).

In practice, when used as a heuristic with a time limit as the stopping criterion, `VNDS-PC1` has a good performance. One can observe that, if pseudo-cuts (line 13 in Algorithm 18) and objective cuts (lines 2 and 16) are not added, the algorithm from [88] is obtained, which is a special case of VNDS-PC with a fixed LP relaxation reference solution.

Ordering variables w.r.t. the minimum and maximum distances from the incumbent solution. In the VNDS variant above, the variables in the incumbent integer solution are ordered according to the distances of their values to the values of the current linear relaxation solution. However, it is possible to employ different ordering strategies. For example, in the case of maximization of $c^T x$, consider the following two problems:

$$(LP^-_{x^*}) \left[\begin{array}{ll} & \min \delta(x^*, x) \\ \text{s.t.:} & Ax \leq b \\ & c^{\mathrm{T}}x \geq LB + 1 \\ & x_j \in [0,1]\,,\ j \in \mathcal{B} \\ & x_j \geq 0,\ j \in N \end{array} \right. \qquad (LP^+_{x^*}) \left[\begin{array}{ll} & \max \delta(x^*, x) \\ \text{s.t.:} & Ax \leq b \\ & c^{\mathrm{T}}x \geq LB + 1 \\ & x_j \in [0,1]\,,\ j \in \mathcal{B} \\ & x_j \geq 0,\ j \in N \end{array} \right.$$

where x^* is the best known integer feasible solution and LB is the best lower bound found so far (i.e., $LB = c^{\mathrm{T}}x^*$). Of course, in case of solving $\min c^T x$, the inequality $c^T x \geq LB + 1$ from models $(LP^-_{x^*})$ and $(LP^+_{x^*})$, should be replaced with $c^T x \leq UB - 1$, where the upper bound $UB = c^T x^*$. If $\overline{x}^-$ and $\overline{x}^+$ are optimal solutions of the LP-relaxation problems $LP^-_{x^*}$ and $LP^+_{x^*}$, respectively, then components of x^* could be ordered in ascending order of values $|\overline{x}^-_j - \overline{x}^+_j|$, $j \in \mathcal{B}$. Since both solution vectors $\overline{x}^-$ and $\overline{x}^+$ are real-valued (i.e., from $\mathbb{R}^n$), this ordering technique is expected to be more sensitive than the standard one, i.e., the number of pairs (j, j'), $j, j' \in N, j \neq j'$ for which $|\overline{x}^-_j - \overline{x}^+_j| \neq |\overline{x}^-_{j'} - \overline{x}^+_{j'}|$ is expected to be greater than the

number of pairs (h, h'), $h, h' \in N, h \neq h'$ for which $|x^*_h - \overline{x}_h| \neq |x^*_{h'} - \overline{x}_{h'}|$, where $\overline{x}$ is an optimal solution of the LP relaxation $LP(P)$.

Also, according to the definition of $\overline{x}^-$ and $\overline{x}^+$, it is intuitively more likely the variables x_j, $j \in N$, for which $\overline{x}^-_j = \overline{x}^+_j$, to have that same value $\overline{x}^-_j$ in the final solution, than it is for variables x_j, $j \in N$, for which $x^*_j = \overline{x}_j$ (and $\overline{x}^-_j \neq \overline{x}^+_j$), to have the final value x^*_j. In practice, if $\overline{x}^-_j = \overline{x}^+_j$, $j \in N$, then usually $x^*_j = \overline{x}^-_j$, which justifies the ordering of components of x^* in the described way. However, if we want to keep the number of iterations in one pass of VNDS approximately the same as in the standard ordering, i.e., if we want to use the same value for parameter d, then the subproblems examined will be larger than with the standard ordering, since the value of q will be smaller (see line 8 in Algorithm 19). The pseudo-code of this variant of VNDS-PC, denoted as VNDS-PC2, is provided in Algorithm 19.

Algorithm 19 VNDS for MIPs with pseudo-cuts and another ordering strategy.

Function `VNDS-PC2`(P, d, x^*)
1 Choose stopping criteria (set *proceed*1=*proceed*2=`true`)
2 Add objective cut: $LB = c^{\mathrm{T}}x^*$; $P = (P \mid c^{\mathrm{T}}x > LB)$
3 **while** *proceed*1 **do**
4 Find an optimal solution $\overline{x}$ of LP(P); set $UB = c^{\mathrm{T}}\overline{x}$
5 **if** $(B(\overline{x}) = \mathcal{B})$ **then** **break**
6 Find optimal solutions $\overline{x}^-$ of $LP^-_{x^*}$ and $\overline{x}^+$ of $LP^+_{x^*}$
7 $\delta_j = \mid \overline{x}^-_j - \overline{x}^+_j \mid$, $j = 1, ..., p$; index x_j so that $\delta_j \leq \delta_{j+1}$, $j = 1, ..., p-1$
8 Set $q = \mid \{j \in \mathcal{B} \mid \delta_j \neq 0\} \mid$, $k_{step} = \lfloor q/d \rfloor$, $k = p - k_{step}$
9 **while** *proceed*2 **and** $k \geq 0$ **do**
10 $J_k = \{1, \ldots, k\}$; $x' =$ `MIPSOLVE`$(P(x^*, J_k), x^*)$;
11 **if** $(c^{\mathrm{T}}x' > c^{\mathrm{T}}x^*)$ **then**
12 Update objective cut: $LB = c^{\mathrm{T}}x'$; $P = (P \mid c^{\mathrm{T}}x > LB)$;
13 $x^* =$ `LocalSearch`(P, x'); $LB = c^{\mathrm{T}}x^*$; **break**
else
14 **if** $(k - k_{step} > p - q)$ **then** $k_{step} = \max\{\lfloor k/2 \rfloor, 1\}$
15 Set $k = k - k_{step}$
16 Update *proceed*2
17 $x' =$ `MIPSOLVE`$(P(\overline{x}, B(\overline{x})), x^*)$; $LB = \max\{LB, c^{\mathrm{T}}x'\}$;
18 Add pseudo-cut to P : $P = (P \mid \delta(B(\overline{x}), x, \overline{x}) \geq 1)$;
19 $x' =$ `MIPSOLVE`$(P(\overline{x}^-, B(\overline{x}^-)), x^*)$; $LB = \max\{LB, c^{\mathrm{T}}x'\}$;
20 Add pseudo-cut to P : $P = (P \mid \delta(B(\overline{x}^-), x, \overline{x}^-) \geq 1)$;
21 $x' =$ `MIPSOLVE`$(P(\overline{x}^+, B(\overline{x}^+)), x^*)$; $LB = \max\{LB, c^{\mathrm{T}}x'\}$;
22 Add pseudo-cut to P : $P = (P \mid \delta(B(\overline{x}^+), x, \overline{x}^+) \geq 1)$;
23 Update *proceed*1;
24 **return** LB, UB, x^*.

6.2.2 A double decomposition scheme

In this section we propose the use of a second level decomposition scheme within VNDS for the 0-1 MILP. The 0-1 MILP is tackled by decomposing the problem into several subproblems, where the number of binary variables with value 1 is fixed at a given integer value. Fixing the number of variables with value 1 to a given value $h \in \mathbb{N} \cup \{0\}$ can be achieved by adding the constraint $x_1 + x_2 + \ldots + x_p = h$, or, equivalently, $e^{\mathrm{T}}x = h$, where e is the vector of ones. Solving the 0-1 MILP by tackling separately each of the subproblems P_h for $h \in N$ appears to be an interesting approach for the case of the multidimensional knapsack problem [120], especially because the additional constraint $e^{\mathrm{T}}x = h$ provides tighter upper bounds than the classical LP–relaxation.

Formally, let P_h be the subproblem, obtained from the original problem by adding the hyperplane constraint $e^{\mathrm{T}}x = h$ for $h \in N$, and enriched by an objective cut:

$$(P_h)\left[\begin{array}{ll}\max & c^{\mathrm{T}}x\\ \text{s.t.:} & Ax \leq b\\ & c^{\mathrm{T}}x \geq LB+1\\ & e^{\mathrm{T}}x = h\\ & x \in \{0,1\}^p \times \mathbb{R}_+^{n-p}\end{array}\right.$$

Let h_{min} and h_{max} denote lower and upper bounds on the number of variables with value 1 in an optimal solution of the problem. Then it is obvious that $\nu(P) = \max\{\nu(P_h) \mid h_{min} \leq h \leq h_{max}\}$. Bounds $h_{min} = \left\lceil \nu(LP_0^-) \right\rceil$ and $h_{max} = \left\lfloor \nu(LP_0^+) \right\rfloor$ can be computed by solving the following two problems:

$$(LP_0^-)\left[\begin{array}{ll}\min & e^{\mathrm{T}}x\\ \text{s.t.:} & Ax \leq b\\ & c^{\mathrm{T}}x \geq LB+1\\ & x \in [0,1]^p \times \mathbb{R}_+^{n-p}\end{array}\right. \qquad (LP_0^+)\left[\begin{array}{ll}\max & e^{\mathrm{T}}x\\ \text{s.t.:} & Ax \leq b\\ & c^{\mathrm{T}}x \geq LB+1\\ & x \in [0,1]^p \times \mathbb{R}_+^{n-p}\end{array}\right.$$

We define the order of the hyperplanes at the beginning of the algorithm, and then we explore them one by one, in that order. The ordering can be done according to the objective values of the linear programming relaxations $\mathrm{LP}(P_h)$, $h \in H = \{h_{min}, \ldots, h_{max}\}$. In each hyperplane, `VNDS-PC1` is applied and if there is no improvement, the next hyperplane is explored. We refer to this method as `VNDDS` (short for *Variable Neighborhood Double Decomposition Search*), which corresponds to the pseudo-code in Algorithm 20. This idea is inspired by the approach proposed in [109], where the ordering of the neighborhood structures in Variable Neighborhood Descent is determined dynamically, by solving relaxations of problems. Problems differ in one constraint that defines Hamming distance h ($h \in H = \{h_{min}, .., h_{max}\}$).

Algorithm 20 Two levels of decomposition with hyperplanes ordering.

Function `VNDDS`(P, x^*, d)
1 Solve the LP-relaxation problems LP_0^- and LP_0^+;
  Set $h_{min} = \left\lceil \nu(LP_0^-) \right\rceil$ and $h_{max} = \left\lfloor \nu(LP_0^+) \right\rfloor$;
2 Sort the set of subproblems $\{P_{h_{min}}, \ldots, P_{h_{max}}\}$ so that $\nu(LP(P_h)) \leq \nu(LP(P_{h+1}))$, $h_{min} \leq h < h_{max}$;
3 Find initial integer feasible solution x^*;
4 **for** $(h = h_{min}; h \leq h_{max}; h++)$ **do**
5 $x' =$ `VNDS-PC1`(P_h, d, x^*)
6 **if** $(c^{\mathrm{T}}x' > c^{\mathrm{T}}x^*)$ **then**
  $x^* = x'$
return x^*.

It is important to note that the exact variant of `VNDDS`, i.e., without any limitations regarding the running time or the number of iterations, converges to an optimal solution in a finite number of steps [64].

6.2.3 Comparison

For comparison purposes, five algorithms are ranked according to their objective values for the MIP benchmark instances in MIPLIB [94] and the benchmark instances for the Maximum Knapsack Problem (MKP) in [23]. Tables 3–4 report the average differences between the ranks of every pair of algorithms for the MIPLIP and MKP test sets, respectively.

Table 3: Objective value average rank differences on the MIPLIB set.

ALGORITHM (average rank)	CPLEX (2.14)	VNDS-MIP (1.95)	VNDS-PC1 (2.64)	VNDS-PC2 (3.64)	VNDDS (4.64)
CPLEX (2.14)	0.00	0.18	-0.50	-1.50	-2.50
VNDS-MIP (1.95)	-0.18	0.00	-0.68	-1.68	-2.68
VNDS-PC1 (2.64)	0.50	0.68	0.00	-1.00	-2.00
VNDS-PC2 (3.64)	1.50	1.68	1.00	0.00	-1.00
VNDDS (4.64)	2.50	2.68	2.00	1.00	0.00

Table 4: Objective value average rank differences on the MKP set.

ALGORITHM (average rank)	CPLEX (2.86)	VNDS-MIP (3.09)	VNDS-PC1 (2.09)	VNDS-PC2 (3.23)	VNDDS (3.72)
CPLEX (2.86)	0.00	-0.23	0.77	-0.36	-0.86
VNDS-MIP (3.09)	0.23	0.00	1.00	-0.14	-0.64
VNDS-PC1 (2.09)	-0.77	-1.00	0.00	-1.14	-1.64
VNDS-PC2 (3.23)	0.36	0.14	1.14	0.00	-0.50
VNDDS (3.72)	0.86	0.64	1.64	0.50	0.00

It appears that `VNDS-MIP` outperforms the other four methods on MIPLIB instances, while for the MKP set, the best performance is obtained with the `VNDS-PC1` heuristic.

7 Variable Neighborhood Search for continuous global optimization

The general form of the continuous constrained nonlinear global optimization problem (GOP) is given as follows:

$$(GOP) \quad \left[\begin{array}{lll} \min & f(x) & \\ \text{s.t.} & g_i(x) \leq 0 & \forall i \in \{1,2,\ldots,m\} \\ & h_i(x) = 0 & \forall i \in \{1,2,\ldots,r\} \\ & a_j \leq x_j \leq b_j & \forall j \in \{1,2,\ldots,n\} \end{array} \right.$$

where $x \in R^n$, $f : R^n \to R$, $g_i : R^n \to R$, $i = 1,2,\ldots,m$, and $h_i : R^n \to R$, $i = 1,2,\ldots,r$, are possibly nonlinear continuous functions, and $a, b \in R^n$ are the variable bounds. A box constraint GOP is defined when only the variable bound constraints are present in the model.

GOPs naturally arise in many applications, e.g. in advanced engineering design, data analysis, financial planning, risk management, scientific modeling, etc. Most cases of practical interest are characterized by multiple local optima and, therefore, a search effort of global scope is needed to find the globally optimal solution.

If the feasible set X is convex and objective function f is convex, then (GOP) is relatively easy to solve, i.e., the Karush-Kuhn-Tucker conditions can be applied. However, if X is not a convex set or f is not a convex function, we can have many local optima and the problem may not be solved with classical techniques. For solving (GOP), VNS has been used in two different ways: (i) with neighborhoods induced by using a ℓ_p norm; (ii) without using a ℓ_p norm.

(i) VNS with ℓ_p norm neighborhoods [49, 91, 97, 100]. A natural approach in applying VNS for solving GOPs is to induce neighborhood structures $\mathcal{N}_k(x)$ from the ℓ_p metric given as:

$$\rho(x,y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{1/p}, \quad p \in [1,\infty) \tag{3}$$

and

$$\rho(x,y) = \max_{1 \leq i \leq n} |x_i - y_i|, \quad p \to \infty. \tag{4}$$

The neighborhood $\mathcal{N}_k(x)$ denotes the set of solutions in the k–th neighborhood of x based on the metric ρ. It is defined as

$$\mathcal{N}_k(x) = \{y \in X \mid \rho(x, y) \leq \rho_k\}, \tag{5}$$

or

$$\mathcal{N}_k(x) = \{y \in X \mid \rho_{k-1} < \rho(x, y) \leq \rho_k\}, \tag{6}$$

where ρ_k, known as the radius of $\mathcal{N}_k(x)$, is monotonically increasing with k ($k \geq 2$).

For solving box constraint GOPs, both [49] and [91] use the neighborhoods as defined in (6). The basic differences between the two algorithms reported there are as follows: (1) in the procedure suggested in [91] the ℓ_∞ norm is used, while in [49] the choice of metric is either left to the analyst, or changed automatically in some predefined order; (2) the commercial solver SNOPT [58] is used as a local search procedure within VNS in [91], while in [49], the analyst may choose one out of six different convex minimizers. A VNS based heuristic for solving the generally constrained GOP is suggested in [97]. There, the problem is first transformed into a sequence of box constrained problems within the well known exterior point method:

$$\min_{a \leq x \leq b} F_{\mu,q}(x) = f(x) + \frac{1}{\mu} \sum_{i=1}^{m} (\max\{0, g_i(x)\})^q + \sum_{i=1}^{r} |h_i(x)|^q, \tag{7}$$

where μ and $q \geq 1$ are a positive penalty parameter and penalty exponent, respectively. Algorithm 21 outlines the steps for solving the box constraint subproblem as proposed in [97]:

Algorithm 21 VNS using a ℓ_p norm.

Function `Glob-VNS` (x^*, k_{max}, t_{max})
1 Select the set of neighborhood structures $\mathcal{N}_k$, $k = 1, \ldots, k_{\max}$
2 Select the array of random distributions types and an initial point $x^* \in X$
3 $x \leftarrow x^*$, $f^* \leftarrow f(x)$, $t \leftarrow 0$
4 **while** $t < t_{max}$ **do**
5 $\quad k \leftarrow 1$
6 $\quad$ **repeat**
7 $\quad\quad$ **for** *all distribution types* **do**
8 $\quad\quad\quad y \leftarrow$ `Shake`(x^*, k) // Get $y \in \mathcal{N}_k(x^*)$ at random
9 $\quad\quad\quad y' \leftarrow$ `BestImprovment`(y) // Apply LS to obtain a local minimum y'
10 $\quad\quad\quad$ **if** $f(y') < f^*$ **then**
11 $\quad\quad\quad\quad x^* \leftarrow y'$, $f^* \leftarrow f(y')$, `go to line 5`
12 $\quad\quad k \leftarrow k + 1$
$\quad$ **until** $k = k_{\max}$
13 $\quad t \leftarrow$ `CpuTime`()

The `Glob-VNS` procedure from Algorithm 21 contains the following parameters in addition to k_{max} and t_{max}: (1) *Values of radii* ρ_k, $k = 1, \ldots, k_{\max}$, which may be defined by the user or calculated automatically in the minimizing process; (2) *Geometry* of neighborhood structures $\mathcal{N}_k$, defined by the choice of metric. Usual choices are the ℓ_1, ℓ_2, and ℓ_∞ norms; (3) *Distribution* types used for obtaining random points y from $\mathcal{N}_k$ in the *Shaking* step. A uniform distribution in $\mathcal{N}_k$ is the obvious choice, but other distributions may lead to much better performance on some problems. Different choices of neighborhood structures and random point distributions lead to different VNS-based heuristics.

(ii) VNS without using ℓ_p norm neighborhoods. Two different neighborhoods, $N_1(x)$ and $N_2(x)$, are used in the VNS based heuristic suggested in [118]. In $N_1(x)$, r (a parameter) random directions from the current point x are generated and a one dimensional search along each direction is performed. The best point (out of r) is selected as a new starting solution for the next iteration, if it is better than the current one. If not, as in VND, the search is continued within the next neighborhood $N_2(x)$. The new point in $N_2(x)$ is obtained as follows. The current solution is moved for each x_j ($j = 1, \ldots, n$) by a value Δ_j, taken at random from the interval $(-\alpha, \alpha)$; i.e., $x_j^{(new)} = x_j + \Delta_j$ or $x_j^{(new)} = x_j - \Delta_j$. Points obtained by the plus

or minus sign for each variable define the neighborhood $N_2(x)$. If a relative increase of 1% in the value of $x_j^{(new)}$ produces a better solution than $x^{(new)}$, the + sign is chosen; otherwise the - sign is chosen.

Neighborhoods $N_1(x)$ and $N_2(x)$ are used for designing two algorithms. The first, called VND, iterates over these neighborhoods until there is no improvement in the solution value. In the second variant, a local search is performed with N_2 and k_{max} is set to 2 for the shaking step.

It is interesting to note that computational results reported by all VNS based heuristics were very promising. They usually outperformed other recent approaches from the literature.

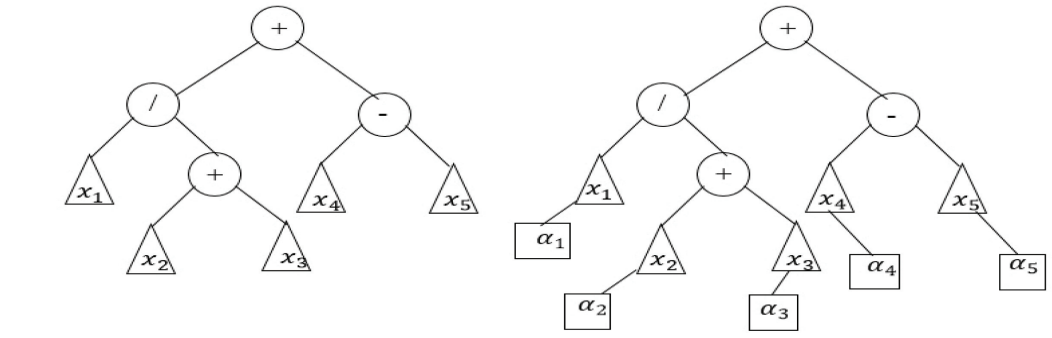
8 Variable neighborhood programming (VNP) - VNS for automatic programming

Building an intelligent machine is an old dream that, thanks to computers, begins to take shape. Automatic programming is an efficient technique that has led to important developments in the field of artificial intelligence. Genetic programming (GP) [87], inspired by the genetic algorithm (GA), is among the few evolutionary algorithms used to evolve a population of programs. The main difference between GP and GA is the representation of a solution. An individual in GA can be a string, while in GP, the individuals are programs. A tree is the usual way to represent a program in GP. For example, assume that the current solution of a problem is the following function:

$$f(x_1, \ldots, x_5) = \frac{x_1}{x_2 + x_3} + x_4 - x_5.$$

Then the code (tree) that calculates f using GP may be represented as in Figure 4a).

Elleuch et al. [50, 51] recently adapted VNS rules for solving automatic programming problems. They first suggested an extended solution representation by adding coefficients to variables. Each terminal node was attached to its own parameter value. These parameters give a weight for each terminal node, with values from the interval [0, 1]. This type of representation allows VNP to examine parameter values and the tree structure in the same iteration, increasing the probability for finding a good solution faster. Let $G = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ denote a parameter set. In Figure 4b) an example of a solution representation in VNP is illustrated.



a) GP solution representation b) VNP solution representation

Figure 4: Current solution representation in automatic programming problem: a) $\frac{x_1}{x_2+x_3} + x_4 - x_5$**; b)** $\frac{\alpha_1 x_1}{\alpha_2 x_2 + \alpha_3 x_3} + \alpha_4 x_4 - \alpha_5 x_5$.

(i) Neighborhood structures . Nine different neighborhood structures are proposed in [50] based on a tree representation. To save space, we will just mention some of them:

- $N_1(T)$ - **Changing a node value operator.** This neighborhood preserves the tree structure and changes only the values of a functional or a terminal node. Each node has a set of allowed values from which one can be chosen. Let x_i be the current solution; then a neighbor x_{i+1} differs from x_i by just a single node. A move within this neighborhood is shown in Figure 5.

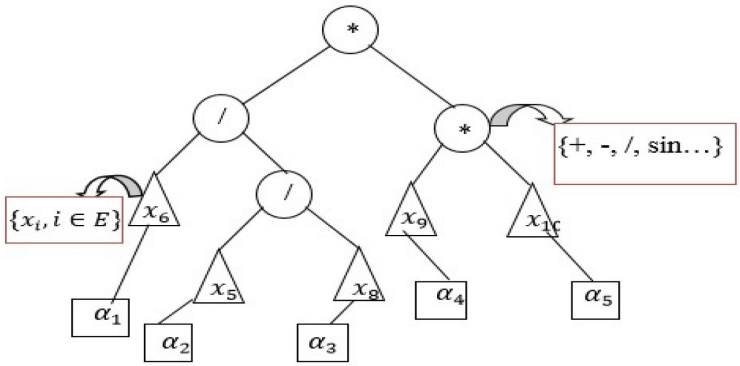


Figure 5: Neighborhood structure N_1: Changing a node value.

- $N_2(T)$ - **Swap operator**. Here, a subtree from the current tree is randomly selected and a new random subtree is generated as shown in Figures 6a1) and 6a2). Then the new subtree replaces the current one (see Figure 6b)). In this move, any constraint related to the maximum tree size should be respected.

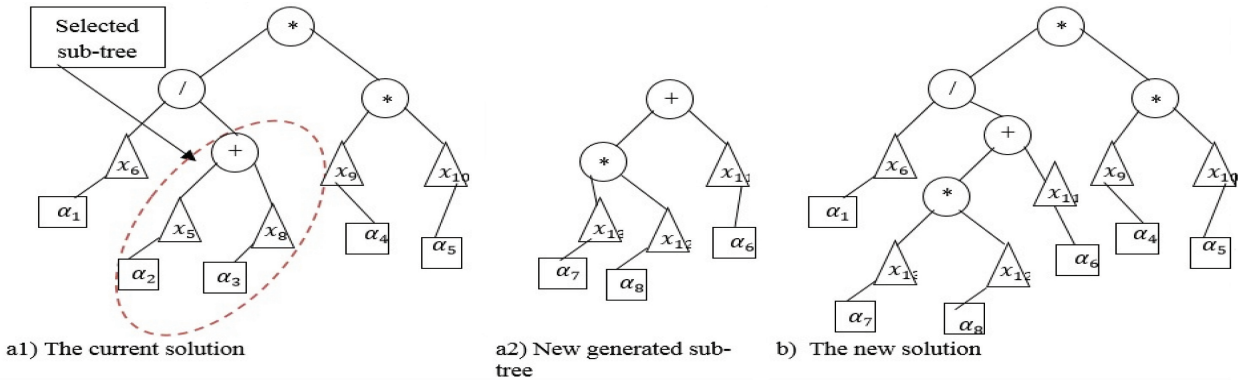


Figure 6: Neighborhood structure N_2: Swap operator.

- $N_3(T)$ - **Changing parameter values.** In the two previous neighborhoods, the tree structure and the node values were considered. In the $N_3(T)$ neighborhood, attention is paid to the parameters. So, the position and value of nodes are kept in order to search the neighbors in the parameter space. Figure 7 illustrates the procedure where the change from one value to another is performed at random.

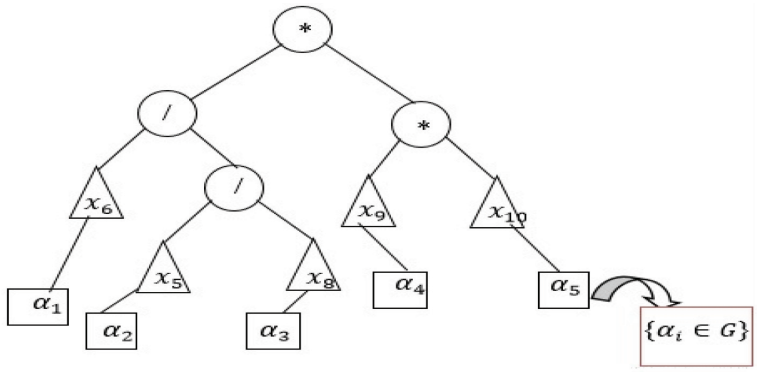


Figure 7: Neighborhood structure N_3: Change parameters.

These neighborhoods may be used in both the local search step ($N_\ell, \ell \in [1, \ell_{max}]$) and in the shaking step ($\mathcal{N}_k, k \in [1, k_{max}]$) of the VNP.

(ii) VNP shaking. The shaking step allows diversification in the search space. The proposed VNP algorithm does not use exactly the same neighborhood structures N_ℓ than the local search. Thus, we denote the neighborhoods used in the shaking phase as $\mathcal{N}_k(T), k = 1, \ldots, k_{max}$. $\mathcal{N}_k(T)$ may be constructed by repeating k times one or more moves from the set $\{N_\ell(T), |\ell = 1, \ldots, \ell_{max}\}$. Consider, for example, the swap operator $N_2(T)$. Let m denote the maximum number of nodes in the tree representation of the solution. We can get a solution from the k^{th} neighborhood of T using the swap operator, where k represents the number of nodes of the new generated sub-tree. If n denotes the number of nodes in the original tree after deleting the old sub-tree, than $n+k \leq m$. The objective of the shaking phase is to provide a good starting point for the local search.

(iii) VNP objective function. The evaluation consists of defining a fitness (or objective) function to assess a solution. This function depends on the problem considered. After running each solution (program) on a training data set, the fitness may be measured by counting the training cases where the returned solution is correct or close to the exact solution.

(iv) An example: Time series forecasting (TSF) problem. Two widely used benchmark data sets of the TSF problem are considered in [50] to study the VNP capabilities: the Mackey-Glass series and the Box-Jenkins set. The parameters for the VNP implementation that were chosen after some preliminary testing are given in Table 5.

Table 5: VNP parameters adjustment for the forecasting problem.

Parameters	Values
The functional set	$F = \{+, *, , pow\}$;
The terminal sets	$\{(x_i, c), i \in [1, ..., m], m =$ number of inputs, $c \in R\}$;
Neighborhood structures	$\{N_1, N_2, N_3\}$;
Minimum tree length	20 nodes;
Maximum tree length	200 nodes;
Maximum number of iterations	50000

The root mean square error (RMSE) is used as the fitness function, as it is normally done in the literature:

$$f(T) = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_t^j - y_{out}^j)^2}$$

where n is the total number of samples, and y_{out}^j and y_t^j are the output of the VNP model and the desired output for sample j, respectively. Next we illustrate with a comparison on a single Box-Jenkins instance.

The gas furnace data for this instance were collected from a combustion process of a methane air mixture [21]. This time series has found a widespread application as a benchmark example for testing prediction algorithms. The data set contains 296 pairs of input-output values. The input $u(t)$ corresponds to the gas flow, and the output $y(t)$ is the CO_2 concentration in the outlet gas. The inputs are $u(t-4)$, and $y(t-1)$, and the output is $y(t)$. In this work, 200 samples are used in the training phase and the remaining samples are used for the testing phase. The performance of the evolved VNP model is evaluated by comparing it with existing approaches. The RMSE achieved by the VNP output model is 0.0038, which is better than the RMSE obtained by other approaches, as shown in Table 6.

Table 6: Comparison of testing error on Box-Jenkins dataset.

Method	Prediction error RMSE
ODE [117]	0.5132
HHMDDE [47]	0.3745
FBBFNT [29]	0.0047
VNP [50]	**0.0038**

9 Discovery science

In all the above applications, VNS is used as an optimization tool. It can also lead to results in "discovery science", i.e., for the development of new theories. This has been done for graph theory in a long series of papers with the common title "Variable neighborhood search for extremal graphs" that report on the development and applications of the AutoGraphiX (AGX) system [5, 39, 40]. This system addresses the following problems:

- Find a graph satisfying given constraints.
- Find optimal or near optimal graphs for an invariant subject to constraints.
- Refute a conjecture.
- Suggest a conjecture (or repair or sharpen one).
- Provide a proof (in simple cases) or suggest an idea of proof.

A basic idea then is to address all of these problems as parametric combinatorial optimization problems on the infinite set of all graphs (or in practice some smaller subset) using a generic heuristic to explore the solution space. This is being accomplished using VNS to find extremal graphs with a given number n of vertices (and possibly also a given number of edges). Extremal graphs may be viewed as a family of graphs that maximize some invariant such as the independence number or chromatic number, possibly subject to constraints. We may also be interested in finding lower and upper bounds on some invariant for a given family of graphs. Once an extremal graph is obtained, VND with many neighborhoods may be used to build other such graphs. Those neighborhoods are defined by modifications of the graphs such as the removal or addition of an edge, rotation of an edge, and so forth. Once a set of extremal graphs, parameterized by their order, is found, their properties are explored with various data mining techniques, leading to conjectures, refutations and simple proofs or ideas of proof.

More recent applications include [35, 37, 56, 62, 71] in chemistry, [7, 40] for finding conjectures, [4, 43] for largest eigenvalues, [25, 72, 73] for extremal values in graphs, independence [6, 11], specialty indexes [18, 19, 20, 83] and others [8, 82, 114, 115]. See [10] for a survey with many further references.

The current list of references in the series "VNS for extremal graphs" corresponds to [4, 5, 6, 7, 8, 9, 11, 12, 13, 18, 19, 20, 25, 35, 37, 39, 40, 43, 56, 62, 71, 72, 73, 82, 83, 114, 115]. Another list of papers, not included in this series, is [10, 14, 15, 16, 17, 36, 38, 63, 65, 66, 67, 70, 116]. Papers in these two lists cover a variety of topics:

(i) **Principles of the approach** [39, 40] and its implementation [5];

(ii) **Applications to spectral graph theory**, e.g., bounds on the index for various families of graphs, graphs maximizing the index subject to some conditions [4, 18, 25, 43, 67];

(iii) **Studies of classical graph parameters**, e.g., independence, chromatic number, clique number, average distance [6, 10, 11, 12, 13, 114, 115];

(iv) **Studies of little known or new parameters of graphs**, e.g., irregularity, proximity and remoteness [14, 72];

(v) **New families of graphs discovered by AGX**, e.g., bags, which are obtained from complete graphs by replacing an edge by a path, and bugs, which are obtained by cutting the paths of a bag [9, 82];

(vi) **Applications to mathematical chemistry**, e.g., study of chemical graph energy, and of the Randić index [19, 20, 37, 56, 62, 63, 65, 66, 71];

(vii) **Results of a systematic study of 20 graph invariants**, which led to almost 1500 new conjectures, more than half of which were proved by AGX and over 300 by various mathematicians [8];

(viii) **Refutation or strengthening of conjectures from the literature** [7, 36, 66];

(ix) **Surveys and discussions about various discovery systems in graph theory**, assessment of the state-of-the-art and the forms of interesting conjectures together with proposals for the design of more powerful systems [38, 70].

10 Conclusions

The general schemes of variable neighborhood search have been presented and discussed. In order to evaluate research development related to VNS, one needs a list of the desirable properties of metaheuristics.

(i) *Simplicity:* the metaheuristic should be based on a simple and clear principle, which should be widely applicable;

(ii) *Precision:* the steps of the metaheuristic should be formulated in precise mathematical terms, independent of possible physical or biological analogies which may have been the initial source of inspiration;

(iii) *Coherence:* all steps of heuristics developed for solving a particular problem should follow naturally from the metaheuristic principles;

(iv) *Effectiveness:* heuristics for particular problems should provide optimal or near-optimal solutions for all known or at least the most realistic instances. Preferably, they should find optimal solutions for most benchmark problems for which such solutions are known;

(v) *Efficiency:* heuristics for particular problems should take a moderate computing time to provide optimal or near-optimal solutions, or comparable or better solutions than the state-of-the-art;

(vi) *Robustness:* the performance of the metaheuristic should be consistent over a variety of instances, i.e., not merely fine-tuned to some training set and not so good elsewhere;

(vii) *User-friendliness:* the metaheuristic should be clearly expressed, easy to understand and, most importantly, easy to use. This implies it should have as few parameters as possible, ideally none;

(viii) *Innovation:* the principle of the metaheuristic and/or the efficiency and effectiveness of the heuristics derived from it should lead to new types of application.

(ix) *Generality:* the metaheuristic should lead to good results for a wide variety of problems;

(x) *Interactivity:* the metaheuristic should allow the user to incorporate his knowledge to improve the resolution process;

(xi) *Multiplicity:* the metaheuristic should be able to produce several near optimal solutions from which the user can choose.

We have tried to show here that VNS possesses to a great extent, all of the above properties. This framework has led to heuristics which are among the very best ones for many problems. Interest in VNS is growing quickly. This is evidenced by the increasing number of papers published each year on this topic. Twenty years ago, only a few; fifteen years ago, about a dozen; ten years ago, about 50, and more than 250 papers in 2016.

Figure 8 shows the parallel increase of the number of papers on VNS and on the other best known metaheuristics. Data are obtained by using the *Scopus search tool*, looking for the terms "Variable Neighborhood Search" (VNS) and "Metaheuristics" (MH). Figure 8 shows the number of times the terms appeared in the abstract of papers in this database. The years used are from 2000 to 2017 but in 2017 only the first 6 months (from January to June) are included. For comparison purposes, the number of papers with MH is divided by 4.

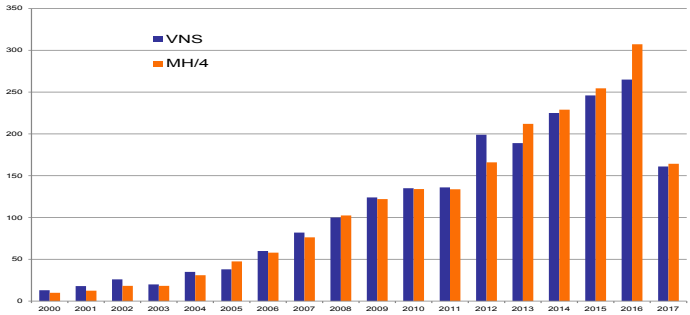


Figure 8: VNS versus MH.

Figure 9 shows the parallel increase of number of papers on VNS and on other most known Metaheuristics. Data are collected again from the *Scopus search tool* to look for the terms Variable Neighborhood Search (VNS), Tabu Search (TS), Genetic Algorithms (GA) and Simulated Annealing (SA). For better illustration, the number of appearances of TS, GA and SA are divided by 3, 50 and 10, respectively.

From the last figure, one can easily see that the relative increase in the number of papers with VNS is larger than the one of other major metaheuristics, especially in the last five years.

In addition, the 18th EURO Mini conference held in Tenerife in November 2005 was entirely devoted to VNS. It led to special issues of the *IMA Journal of Management Mathematics* in 2007 [93], *European Journal of Operational Research* [78] and *Journal of Heuristics* [104] in 2008. After that, VNS conferences took place in Herceg Novi - Montenegro (2012), Djerba - Tunis (2014), Málaga - Spain (2016) and in Ouro Preto - Brazil (2017). Each meeting was covered with before-conference Proceedings (in Electronic notes of Discrete Mathematics) and with at least one post-conference special issue in leading OR journals: Computers and OR, Journal of Global Optimization, IMA JMM, International Transactions of OR.

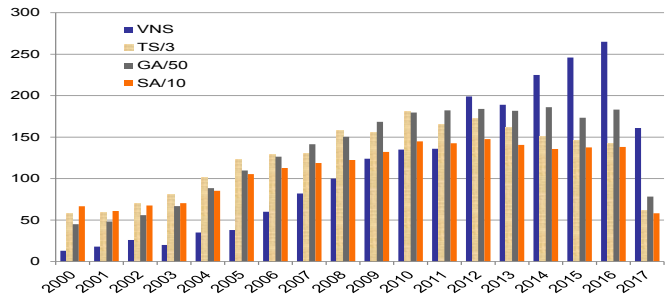


Figure 9: VNS versus other main MHs.

References

[1] Abhishek, K., Leyffer, S., Linderoth, J.: FilMINT: An outer-approximation based solver for nonlinear mixed-integer programs. Technical Report ANL/MCSP1374- 0906, Argonne National Laboratory, 2007

[2] Aloise, D.J., Aloise, D., Rocha, C.T.M., Ribeiro, C.C., Ribeiro, J.C., Moura, L.S.S.: Scheduling work-over rigs for onshore oil production. Discrete Applied Mathematics 154, 695–702 (2006)

[3] Andrade, D. V., Resende, M. G. C.: GRASP with Path-Relinking for Network Migration Scheduling. In *Proceedings of International Network optimization Conference (INOC)*, 2007

[4] Aouchiche, M., Bell, F.K., Cvetković, D., Hansen, P., Rowlinson, P., Simić, S.K., Stevanović, D.: Variable neighborhood search for extremal graphs 16. Some conjectures related to the largest eigenvalue of a graph. European Journal of Operational Research 191, 661–676 (2008)

[5] Aouchiche, M., Bonnefoy, J.M., Fidahoussen, A., Caporossi, G., Hansen, P., Hiesse, L., Lacheré, J., Monhait, A.: Variable neighborhood search for extremal graphs 14. The AutoGraphiX 2 system. In: Liberti L., Maculan N. (eds) Global Optimization: from Theory to Implementation, pp. 281–309. Springer (2005)

[6] Aouchiche, M., Brinkmann, G., Hansen, P.: Variable neighborhood search for extremal graphs 21. Conjectures and results about the independence number. Discrete Applied Mathematics 156, 2530–2542 (2009)

[7] Aouchiche, M., Caporossi, G., Cvetković, D.: Variable neighborhood search for extremal graphs 8. Variations on Graffiti 105. Congressus Numerantium 148, 129–144 (2001)

[8] Aouchiche, M., Caporossi, G., Hansen, P.: Variable Neighborhood search for extremal graphs 20. Automated comparison of graph invariants. MATCH. Communications in Mathematical and Computer Chemistry 58, 365–384 (2007)

[9] Aouchiche, M., Caporossi, G., Hansen, P.: Variable neighborhood search for extremal graphs 27. Families of extremal graphs. Les Cahiers du GERAD G–2007–87, (2007)

[10] Aouchiche, M., Caporossi, G., Hansen, P., Laffay, M.: AutoGraphiX: a survey. Electronic Notes in Discrete Mathematics 22, 515–520 (2005)

[11] Aouchiche, M., Favaron, O., Hansen, P.: Variable neighborhood search for extremal graphs 22. Extending bounds for independence to upper irredundance. Discrete Applied Mathematics 157, 3497–3510 (2009)

[12] Aouchiche, M., Favaron, O., Hansen, P.: Recherche à voisinage variable de graphes extrèmes 26. Nouveaux résultats sur la maille (French). Les Cahiers du GERAD G–2007–55, (2007)

[13] Aouchiche, M., Hansen, P.: Recherche à voisinage variable de graphes extrèmes 13. À propos de la maille (French). RAIRO Oper. Res. 39, 275–293 (2005)

[14] Aouchiche, M., Hansen, P.: Automated results and conjectures on average distance in graphs. In: Bondy A., Fonlupt J., Fouquet J.L., Fournier J.C., Ramírez Alfonsń J.L. (eds) Graph Theory in Paris. Trends in Mathematics. pp. 21–36 (2006)

[15] Aouchiche, M., Hansen, P.: On a conjecture about the Randic index. Discrete Mathematics 307, 262–265 (2007)

[16] Aouchiche, M., Hansen, P.: Bounding average distance using minimum degree. Graph Theory Notes of New York 56, 21–29 (2009)

[17] Aouchiche, M., Hansen, P.: Nordhaus-Gaddum relations for proximity and remoteness in graphs. Computers & Mathematics with Applications 59, 2827–2835 (2010)

[18] Aouchiche, M., Hansen, P., Stevanović, D.: Variable neighborhood search for extremal graphs 17. Further conjectures and results about the index. Discussiones Mathematicae: Graph Theory 29, 15–37 (2009)

[19] Aouchiche, M., Hansen, P., Zheng, M.: Variable neighborhood search for extremal graphs 18. Conjectures and results about the Randic index. MATCH. Communicationa in Mathematical and Computer Chemistry 56, 541–550 (2006)

[20] Aouchiche, M., Hansen, P., Zheng, M.: Variable Neighborhood Search for Extremal Graphs 19. Further Conjectures and Results about the Randic Index. MATCH. Communications in Mathematical and Computer Chemistry 58, 83–102 (2007)

[21] Audet, C., Báchard, V., Le, Digabel, S.: Nonsmooth optimization through mesh adaptive direct search and variable neighborhood search. Journal of Global Optimization 41, 299–318 (2008)

[22] Audet, C., Brimberg, J., Hansen, P., Mladenović, N.: Pooling problem: alternate formulation and solution methods. Management Science 50, 761–776 (2004)

[23] Beasley J. OR-Library: distributing test problems by electronic mail, Journal of the Operational Research Society 41(11) 1069–1072 (1990).

[24] Belacel, N., Hansen, P., Mladenović, N.: Fuzzy J-means: a new heuristic for fuzzy clustering. Pattern Recognition 35, 2193–2200 (2002)

[25] Belhaiza, S., de, Abreu, N., Hansen, P., Oliveira, C.: Variable neighborhood search for extremal graphs 11. Bounds on algebraic connectivity. In: D. Avis, A. Hertz and O. Marcotte (eds.) Graph Theory and Combinatorial Optimization, pp. 1–16. (2007)

[26] Bonami, P., Biegler, L.T., Conn, A.R., Cornuejols, G., Grossmann, I.E., Laird, C.D., Lee, J., Lodi, A., Margot, F., Sawaya, N., Wachter, A.: An algorithmic framework for convex mixed integer nonlinear programs. Discrete Optimization 5, 186–204 (2008)

[27] Bonami, P., Cornuejols, G., Lodi, A., Margot, F.A.: Feasibility pump for mixed integer nonlinear programs. Mathematical Programming 119, 331–352 (2009)

[28] Bonami, P., Lee, J.: BONMIN User's manual. Technical report, IBM Corporation, (2007)

[29] Bouaziz, S., Dhahri, H., Alimi, A.M., Abraham, A.: A hybrid learning algorithm for evolving Flexible Beta Basis Function Neural Tree Model. Neurocomputing 117, 107–117 (2013)

[30] Brimberg, J., Hansen, P., Mladenović, N., Taillard, É.: Improvements and comparison of heuristics for solving the multisource Weber problem. Operations Research 48, 444–460 (2000)

[31] Brimberg, J., Mladenović, N.: A. variable neighborhood algorithm for solving the continuous location-allocation problem. Studies in Locational Analysis 10, 1–12 (1996)

[32] Bussieck, M.R., Drud, A.S., Meeraus, A.: MINLPLib - A collection of test models for mixed-integer nonlinear programming. INFORMS Journal on Computing, 15, 114–119 (2003)

[33] Canuto, S., Resende, M., Ribeiro, C.: Local search with perturbations for the prize-collecting Steiner tree problem in graphs. Networks 31, 201–206 (2001)

[34] Caporossi, G., Alamargot, D., Chesnet, D.: Using the computer to study the dynamics of the handwriting processes. Lecture Notes in Computer Science 3245, 242–254 (2004)

[35] Caporossi, G., Cvetković, D., Gutman, I., Hansen, P.: Variable neighborhood search for extremal graphs 2. Finding graphs with extremal energy. Journal of Chemical Information and Computer Sciences 39, 984–996 (1999)

[36] Caporossi, G., , Dobrynin, A.A., Gutman, I., Hansen, P.: Trees with palindromic Hosoya polynomials. Graph Theory Notes of New York 37, 10–16 (1999)

[37] Caporossi, G., Gutman, I., Hansen, P.: Variable neighborhood search for extremal graphs 4. Chemical trees with extremal connectivity index. Computers & Chemistry 23, 469–477 (1999)

[38] Caporossi, G., Gutman, I., Hansen, P., Pavlović, L., Graphs with maximum connectivity index. Computational Biology and Chemistry 27, 85–90 (2003)

[39] Caporossi, G., Hansen, P.: Variable neighborhood search for extremal graphs 1. The AutoGraphiX system. Discrete Mathematics 212, 29–44 (2000)

[40] Caporossi, G., Hansen, P.: Variable neighborhood search for extremal graphs 5. Three ways to automate finding conjectures. Discrete Mathematics 276, 81–94 (2004)

[41] Carrabs, F., Cordeau, J.-F., Laporte, G.: Variable neighbourhood search for the pickup and delivery traveling salesman problem with LIFO loading. INFORMS Journal on Computing 19, 618–632 (2007)

[42] Carrizosa, E., Martín-Barragán, B., Plastria, F., Romero, Morales, D.: On the selection of the globally optimal prototype subset for nearest-neighbor classification. INFORMS Journal on Computing 19, 470–479 (2007)

[43] Cvetkovic, D., Simic, S., Caporossi, G., Hansen, P.: Variable neighborhood search for extremal graphs 3. On the largest eigenvalue of color-constrained trees. Linear and Multilinear Algebra 49, 143–160 (2001)

[44] Cohoon, J., Sahni, S.: Heuristics for backplane ordering. Journal of VLSI and Computer Systems 2, 37–61 (1987)

[45] Davidon, W.C.: Variable metric algorithm for minimization. Argonne National Laboratory Report ANL-5990 (1959)

[46] Desrosiers J., Mladenović N., Villeneuve D.: Design of balanced MBA student teams. Journal of the Operational Research Society 56, 60–66 (2005)

[47] Dhahri, H., Alimi, A.M., Abraham, A.: Hierarchical multi-dimensional differential evolution for the design of beta basis function neural network, Neurocomputing 97, 131–140 (2012)

[48] Djenic, A., Radojicic, N., Maric, M., Mladenović N.: Parallel VNS for bus terminal location problem. Applied Soft Computing 42, 448–458 (2016)

[49] Dražić, M., Kovacevic-Vujcić, V., Cangalović, M., Mladenović, N.: GLOB - A. new VNS-based software for global optimization In: Liberti L., Maculan N. (eds) Global Optimization: from Theory to Implementation, pp. 135–144, Springer (2006)

[50] Elleuch, S., Jarboui, B., and Mladenović, N.: Variable neighborhood programming - A new automatic programming method in artificial intelligence. GERAD Technical report, G–2016–21 (2016), Montreal, Canada.

[51] Elleuch, S., Jarboui, B., and Mladenović, N. Reduced variable neighborhood programming for the preventive maintenance planning of railway infrastructure, GERAD Technical report, G–2016–92 (2016), Montreal, Canada.

[52] Fischetti, M., Lodi, A.: Local branching. Mathematical Programming 98, 23–47 (2003)

[53] Fletcher, R., Powell, M.J.D.: Rapidly convergent descent method for minimization. The Computer Journal 6, 163–168 (1963)

[54] Fletcher, R., Leyffer, S.: Solving mixed integer nonlinear programs by outer approximation. Mathematical Programming 66, 327-349 (1994)

[55] Fletcher, R, Leyffer, S.: Numerical experience with lower bounds for MIQP branch-and-bound. SIAM Journal of Optimization 8, 604-616 (1998)

[56] Fowler, P.W., Hansen, P., Caporossi, G., Soncini, A.: Variable neighborhood search for extremal graphs 7. Polyenes with maximum HOMO-LUMO gap. Chemical Physics Letters 49, 143–146 (2001)

[57] Garey, M.R., Johnson, D.S.: Computers and intractability: A guide to the theory of NP-completeness. Freeman, New-York (1978)

[58] Gill, P., Murray, W., Saunders, M.A.: SNOPT:An SQP algorithms for largescale constrained optimization. SIAM Journal of Optimization 12, 979–1006 (2002)

[59] Gill, P., Murray, W., Wright, M.: Practical optimization. Academic Press, London (1981)

[60] Glover, F., Kochenberger, G., (eds.) Handbook of Metaheuristics. Kluwer (2003)

[61] Griffith, R.E., Stewart, R.A.: A nonlinear programming technique for the optimization of continuous processing systems. Management Science 7, 379–392 (1961)

[62] Gutman, I., Hansen, P., Mélot, H., Variable neighborhood search for extremal graphs 10. Comparison of irregularity indices for chemical trees. Journal of Chemical Information and Modeling 45, 222–230 (2005)

[63] Gutman, I., , Miljković, O., Caporossi, G., Hansen, P.: Alkanes with small and large Randić connectivity indices. Chemical Physics Letters 306, 366–372 (1999)

[64] Hanafi S, Lazić J, Mladenović N, Wilbaut C, Crévits I.: New variable neighborhood search based 0-1 MIP heuristic. Yugoslav Journal of Operational Research 25, 343-360 (2015)

[65] Hansen, P.: Computers in Graph Theory. Graph Theory Notes of New York XLIII, 20–39 (2002)

[66] Hansen, P.: How far is, should and could be conjecture-making in graph theory an automated process? In: Graph and Discovery, Dimacs Series in Discrete Mathematics and Theoretical Computer Science 69, 189–229 (2005)

[67] Hansen, P., Aouchiche, M., Caporossi, G., Mélot, H, Stevanović, D.: What forms do interesting conjectures have in graph theory? In: Graph and Discovery Dimacs Series in Discrete Mathematics and Theoretical Computer Science 69, 231–251 (2005)

[68] Hansen, P., Brimberg, J., Urošević, D., Mladenović, N.: Primal-dual variable neighborhood search for the simple plant location problem. INFORMS Journal on Computing 19, 552–564 (2007)

[69] Hansen, P., Jaumard, B., Mladenović, N., Parreira, A.: Variable neighborhood search for weighted maximum satisfiability problem. Les Cahiers du GERAD G–2000–62, (2000)

[70] Hansen, P., Mélot, H.: Computers and discovery in algebraic graph theory. Linear Algebra and Applications 356, 211–230 (2002)

[71] Hansen, P., Mélot, H.: Variable neighborhood search for extremal graphs 6. Analyzing bounds for the connectivity index. Journal of Chemical Information and Computer Sciences 43, 1–14 (2003)

[72] Hansen, P., Mélot, H.: Variable neighborhood search for extremal graphs 9. Bounding the irregularity of a graph. Graphs and Discovery 69, 253–264 (2005)

[73] Hansen, P., Mélot, H., Gutman, I.: Variable neighborhood search for extremal graphs 12. A note on the variance of bounded degrees in graphs. MATCH Communications in Mathematical and in Computer Chemistry 54, 221–232. (2005)

[74] Hansen, P., Mladenović, N.: Variable neighborhood search: Principles and applications. European Journal of Operational Research 130, 449–467 (2001)

[75] Hansen, P., Mladenović, N.: J-Means: A. new local search heuristic for minimum sum-of-squares clustering. Pattern Recognition 34, 405–413 (2001)

[76] Hansen, P., Mladenović, N.: Developments of variable neighborhood search. In: Ribeiro C., Han-sen P. (eds.) Essays and surveys in metaheuristics, pp. 415–440, Kluwer (2001)

[77] Hansen, P., Mladenović, N.: Variable neighborhood search. In: Glover F., Kochenberger G. (eds.) Handbook of Metaheuristics, pp. 145–184, Kluwer (2003)

[78] Hansen, P., Mladenović, N., Moreno Pérez, J.A.: Variable neighborhood search. European Journal of Operational Research 191, 593–595. (2008)

[79] Hansen, P., Mladenović, N., Moreno Pérez, J.A.: Variable neighborhood search: methods and applications. 4OR. A Quarterly Journal of Operations Research 6, 319–360 (2008)

[80] Hansen, P., Mladenović, N., Pérez-Brito, D.: Variable neighborhood decomposition search. Journal of Heuristics 7, 335–350 (2001)

[81] Hansen, P., Mladenović, N., Urošević, D.: Variable neighborhood search and local branching. Computers and Operations Research 33, 3034–3045 (2006)

[82] Hansen, P., Stevanović, D.: Variable neighborhood search for extremal graphs 15. On bags and bugs, Discrete Applied Matehmatics 156 986–997 (2005)

[83] Hansen, P., Vukičević, D.: Variable neighborhood search for extremal graphs 23. On the Randic index and the chromatic number. Discrete Mathematics 309 4228–4234 (2009)

[84] Hertz, A., Plumettaz, M., Zufferey, N.: Variable space search for graph coloring. Discrete Applied Mathematics 156, 2551–2560 (2008)

[85] ILOG CPLEX 10.1. User's Manual (2006)

[86] Jornsten, K., Lokketangen, A.: Tabu search for weighted k-cardinality trees. Asia-Pacific Journal of Operational Research 14, 9–26 (1997)

[87] Koza J.R.: Genetic programming: on the programming of computers by means of natural selection. MIT Press Cambridge, (1992)

[88] Lazić, J., Hanafi, S., Mladenović, N., Urošević, D.: Variable neighbourhood decomposition search for 0–1 mixed integer programs. Computers and Operations Research 37) 1055–1067 (2010)

[89] Lejeune, M.A.: A. variable neighborhood decomposition search method for supply chain management planning problems. European Journal of Operational Research 175, 959–976 (2006)

[90] Leyffer, S.: User manual for MINLP_BB Technical report, University of Dundee, UK, (1999)

[91] Liberti, L., Dražić, M.: Variable neighbourhood search for the global optimization of constrained NLPs. In Proceedings of GO Workshop, Almeria, Spain, 2005. (2005)

[92] Liberti, L., Nannicini, G., Mladenović, N.: A good recipe for solving MINLPs. In: V. Maniezzo, T. Stuetze, S. Voss, (eds.) Matheuristics. Hybridizing metaheuristics and mathematical programming, Springer (2008)

[93] Melián, B., Mladenović, N.: Editorial. IMA Journal of Management Mathematics 18, 99–100 (2007)

[94] MIPLIB http://miplib.zib.de/miplib2003/

[95] Mladenović, N.: A variable neighborhood algorithm – a new metaheuristic for combinatorial optimization. Abstracts of papers presented at Optimization Days, Montréal, p. 112, (1995)

[96] Mladenović, N.: Formulation space search – a new approach to optimization (plenary talk). Proceedings of XXXII SYMOPIS'05, pp. 3 (Vuleta J. eds.), Vrnjacka Banja, Serbia (2005)

[97] Mladenović, N., Dražić, M., Kovačevic-Vujčić, V., Čangalović, M.: General variable neighborhood search for the continuous optimization. European Journal of Operational Research 191, 753–770 (2008)

[98] Mladenović, N., Hansen, P.: Variable neighborhood search. Computers and Operations Research 24, 1097–1100 (1997)

[99] Mladenovic N, Kratica J, Kovacevic-Vujcic V, Cangalovic M.: Variable neighborhood search for metric dimension and minimal doubly resolving set problems. European Journal of Operational Research 220, 328–337 (2012)

[100] Mladenović, N., Petrović, J., Kovačević-Vujčić, V., Čangalović, M.: Solving spread spectrum radar polyphase code design problem by tabu search and variable neighborhood search. European Journal of Operational Research 151, 389–399 (2003)

[101] Mladenović, N., Plastria, F., Urošević, D.: Reformulation descent applied to circle packing problems. Computers and Operations Research 32, 2419–2434 (2005)

[102] Mladenović, N., Plastria, F., Urošević, D.: Formulation space search for circle packing problems. Lecture Notes on Computer Science 4638, 212–216 (2007)

[103] Mladenović, N., Urošević, D., Pérez-Brito, D., García-González, C.G.: Variable neighbourhood search for bandwidth reduction. European Journal of Operational Research 200, 14–27 (2010)

[104] Moreno-Vega, J.M., Melián, B.: Introduction to the special issue on variable neighborhood search. Journal of Heuristics 14, 403–404 (2008)

[105] Pantrigo, J.J., Marti, R., Duarte, A., Pardo, E.G.: Scatter search for the cutwidth minimization problem. Annals of Operations Research. 199, 285–304 (2012)

[106] Pardo,E.G., Mladenović, N., Pantrigo, J.J., Duarte, A.: Variable formulation search for the cut-width minimization problem Applied Soft Computing 13, 2242–2252 (2014)

[107] Plastria, F., Mladenović, N., Urošević, D.: Variable neighborhood formulation space search for circle packing. 18th Mini Euro Conference VNS., Tenerife, Spain (2005)

[108] Popper K.: The logic of scientific discovery, London: Hutchinson (1959)

[109] Puchinger, J., and Raidl, G.: Bringing order into the neighborhoods: Relaxation guided variable neighborhood search. Journal of Heuristics 14, 457–472 (2008)

[110] Resende, M. G. C., Martí, R., Gallego, M., Duarte, A.: Grasp and path relinking for the max-min diversity problem. Computers and Operations Research, 37, 498–508 (2010)

[111] Ribeiro, C.C., de, Souza, M.C.: Variable neighborhood search for the degree-constrained minimum spanning tree problem. Discrete Applied Mathematics 118, 43–54 (2002)

[112] Ribeiro, C.C., Uchoa, E., Werneck, R.: A hybrid GRASP with perturbations for the Steiner problem in graphs. INFORMS Journal on Computing 14, 228–246 (2002)

[113] Rolim, J., Sýkora, O., Vrt'o, I.: Optimal cutwidths and bisection widths of 2- and 3-dimensional meshes. In: Graph-Theoretic Concepts in Computer Science, Lecture Notes in Computer Science 1017, 252–264 (1995)

[114] Sedlar, J., Vukicevic, D., Aouchiche, M., Hansen, P.: Variable neighborhood search for extremal graphs 24. Conjectures and results about the clique number. Les Cahiers du GERAD G–2007–33, (2007)

[115] Sedlar, J., Vukicevic, D., Aouchiche, M., Hansen, P.: Variable neighborhood search for extremal graphs 25. Products of connectivity and distance measures. Les Cahiers du GERAD G–2007–47, (2007)

[116] Stevanovic, D., Aouchiche, M., Hansen, P.: On the spectral radius of graphs with a given domination number. Linear Algebra and its Applications 428, 1854–1864 (2008)

[117] Subudhi, B., Jena, D.: A differential evolution based neural network approach to nonlinear system identification. Applied Soft Computing 11, 861–871 (2011)

[118] Toksari, A.D., Güner, E.: Solving the unconstrained optimization problem by a variable neighborhood search. Journal of Mathematical Analysis and Applications 328, 1178–1187 (2007)

[119] Urošević, D., Brimberg, J., Mladenović, N.: Variable neighborhood decomposition search for the edge weighted k-cardinality tree problem. Computers and Operations Research 31, 1205–1213 (2004)

[120] Vimont, Y., Boussier, S., Vasquez, M.: Reduced costs propagation in an efficient implicit enumeration for the 01 multidimensional knapsack problem. Journal of Combinatorial Optimization 15, 165–178 (2008)

[121] Whitaker, R.: A fast algorithm for the greedy interchange of large-scale clustering and median location problems. INFOR 21, 95–108 (1983)

[122] Zhang, C., Lin, Z., Lin, Z.: Variable neighborhood search with permutation distance for QAP. Lecture Notes in Computer Science 3684, 81–88 (2005)