Rig Routing with Possible Returns and Stochastic Drilling Times*

Pavel Borisovsky[0000-0003-2673-7644], Anton Eremeev[0000-0001-5289-7874], Yulia Kovalenko[0000-0003-4791-7011], and Lidia Zaozerskaya[0000-0002-0287-3759]

Sobolev Institute of Mathematics, Novosibirsk, Russia
borisovski@mail.ru, julia.kovalenko.ya@yandex.ru

Abstract. We consider a real-world vehicle routing problem with time windows, arising in drilling rigs routing and well servicing on a set of sites with different geographical locations. Each site includes a predetermined number of wells which must be processed within a given time window. The same rig can visit a site several times, but the overall number of site visits by rigs is bounded from above. Each well is drilled by one rig without preemptions. It is required to find the routes of the rigs, minimizing the total traveling distance. We also consider a stochastic generalization of the problem, where the drilling times are supposed to be random variables with known discrete distributions. New mixed-integer linear programming models are formulated and tested experimentally. A randomized greedy algorithm is proposed for approximate solving the problem in stochastic formulation, if the number of possible realizations of drilling times is so high that existing MIP solvers are not suitable.

Keywords: Vehicle routing problem · multiple visits · stochastic duration

1 Introduction

The area of exploration or production of gas and oil raises a number of optimization problems for managing the drilling rigs activities that include drilling and traveling between wells. The widely used approach to modeling and solution of such problems is based on the Mixed Integer Linear Programming (MILP). One of the earliest studies in this direction [4], considers a rather complex and detailed problem of scheduling the drilling and other tasks for several offshore oil production platforms to maximize the total profit. For the rigs, only the number of moves are counted, but not the travel distances. In [14], a simpler model optimizing the drilling durations and travel times is proposed. It also considers the possibility of rigs outsourcing and compatibilities of rigs and wells. No individual time windows for each well are given, but rather a common deadline for the whole project.

* The final authenticated publication is available online at https://doi.org/10.1007/978-3-030-77876-7_4

The problem of our interest was introduced in [7]. In this problem, several rigs travel between a set of sites, and each site has a certain number of wells to drill. This problem was classified as the Split Delivery Vehicle Routing Problem with Time Windows (SDVRPTW). A MIP model based on the classical VRP model with time windows (see e.g. [15]) was proposed, and solutions found by a commercial MIP solver were compared to those from Variable Neighborhood Search metaheuristic [7], [8]. In our paper, we consider a generalization of this problem, in which it is allowed to re-visit the same site by the same rig. This feature can be beneficial from the real-life production perspective, but makes the problem more complicated.

In the mentioned above papers, all the necessary data are supposed to be deterministic. Here we extend the study to the case of uncertain drilling durations, assuming that they are random variables with known discrete distributions. Note that a similar assumption is considered in [1] for the scheduling of a set of offshore oil rigs given that the drilling time is a random variable with a known distribution. The authors propose a Monte-Carlo approach, in which the samples of drilling times are simulated and then are used as input data in the GRASP heuristic. As a result, a set of approximate solutions is built and its properties are investigated statistically.

Among the classic problems, the closest one to our formulation seems to be the Split Delivery Vehicle Routing Problem with Time Windows and Uncertain Service Times. There are many papers devoted to some particular aspects of this problems, but we are not aware of any research on the case combining all the indicated problem settings. The problem with random travel and service times originated from [9]. In [3], this case is extended with the time windows constraints. Many papers deal with the robust approach, in which the probability distribution of uncertain parameters are not given, and the solution to be found must be suitable for all possible realizations of uncertain data. Among these papers we can mention [13], [10], in which the service times vary within some convex set. A comprehensive survey of stochastic and robust solution of different VRP type problems can be found in [11].

The SDVRPTW with possibility to re-visit sites has a similarity with production scheduling problems, if one considers rigs as machines, wells as product orders, and sites as orders of the same type. The distances between the sites correspond to setup or changeover times, which should be minimized, while all products should be produced within the given time windows. The production scheduling problems of such kind were successfully solved using time-decomposition techniques and MIP-formulations based on the event points approach (see e.g. [2] [5], [12]).

In our work, we aim at the following three main goals: (i) to compare two different approaches to defining a MIP formulation of the problem, the one based on the classical VRP model with time windows (as in [7], [15] etc.) and the one based on the event points approach [5], (ii) to extend the deterministic problem formulation to a stochastic optimization problem where the drilling time at each site is a random value with a known discrete distribution, testing the MIP-solving techniques for finding exact and approximate solutions to this problem, (iii) to develop a heuristic capable of solving approximately the stochastic optimization problems of higher dimension, compared to MIP-solvers.

In order to reach the second goal, we apply a quantille optimization approach from [6]. This approach is more general than required for our stochastic problem formulation, since in our case the objective function of any fixed solution (which consists of a set of rig routes and the assignment of drilling tasks to rigs) is the total traveling length and does not depend on the random variables. The latter ones only influence on the feasibility of a solution with respect to time window constraints. In the stochastic formulation, we assume that a threshold $\alpha \in (0, 1]$ is given and it is required that the obtained solution should satisfy all time window constraints with a probability not less than α.

2 Deterministic Problem Formulation

We have a set of sites $I = \{i_1, \ldots, i_{|I|}\}$ which must be served by a set of vehicles $U = \{u_1, \ldots, u_{|U|}\}$ (drilling rigs). Each site $i \in I$ is characterized by the total number of planned wells n_i and the time window $(a_i, b_i]$, in which all wells should be drilled. A vehicle can visit site $i \in I$ several times, but the total number of visits of i by all vehicles is bounded by $m_i \leq n_i$. Each well is drilled by one rig without preemptions. Drilling a well of site $i \in I$ by vehicle $u \in U$ requires d_{ui} time units. A subset I_u of sites, that can be served by vehicle $u \in U$, is given. Each vehicle u is initially located at an individual depot id_u. The traveling time between sites i and j for vehicle u is denoted by s_{uij}. It is required to find rigs routes between sites and assignments of wells to rigs minimizing the total traveling time.

In this section, we propose two models for the considered problem. The first one is based on the event point approach and the second one uses the classic approach from VRP theory. Before that, we provide an example which indicates that there are instances where the same rig visits a site several times in any optimal solution.

2.1 Illustrative Example

Consider an instance with 6 sites (see Fig. 1) and the following input data. The number of wells at the sites with odd indices is 5, the number of wells at the sites with even indices is 8. Time windows:

$a_1 = 20$, $b_1 = 30$, $a_2 = 10$, $b_2 = 40$, $a_3 = 30$, $b_3 = 40$,
$a_4 = 20$, $b_4 = 50$, $a_5 = 40$, $b_5 = 50$, $a_6 = 30$, $b_6 = 60$.

There are 3 drilling rigs. The durations of wells drilling at all sites do not depend on the assignment of vehicles to wells and all equal to 2 (i.e. $d_{ui} = 2$). In this example, we suppose that the rigs are identical, i.e., the traveling time between sites i and j is the same for all vehicles (s_{uij} do not depend on u). The distances between pairs of sites are indicated for each edge, if a direct transportation is possible. Direct transportation is prohibited between all other pairs of sites (i.e. $s_{uij} = \infty$). Vertex id corresponds to the initial location of the rigs here (a common depot).

Odd-numbered sites have narrow time windows. For each pair of time windows $(a_i, b_i]$ and $(a_{i+1}, b_{i+1}]$ for $i = 1, 2, 3$, the following condition holds: $(a_i, b_i] \subset (a_{i+1}, b_{i+1}]$, i.e. the i-th window is contained in the $i+1$-st, dividing it into three parts with durations equal to 10. Thus, if returns of the rigs to previously visited sites are prohibited, then moving from site $i+1$ to site i is impossible. The optimal solution with $f = 24$ is uniquely determined up to the assignment of rigs to the routes. It is shown in Fig. 1 on the left. The route of each rig has a unique marking.

If returns to the previously visited sites are allowed, then for each pair of sites i, $i+1$ for $i = 1, 2, 3$, the drilling rig can perform part of the work (drill 4 wells) on site $i+1$, then move to site i, do all the work there, return to the site $i+1$ to process the remaining wells there. The optimal solution with $f = 21$ is shown in Fig. 1 on the right. The value of the objective function is smaller by 3, compared to the case where returns are prohibited.

Based on this example, it is easy to build a family of problems with $6k$ sites and $3k$ machines, with the values of objective function $21k$ (if returns are allowed) or $24k$ (if returns are prohibited) for $k \in N$.

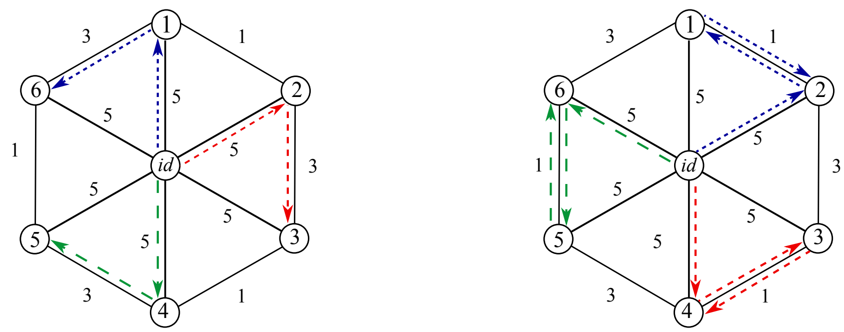

Fig. 1. Optimal solutions in the case of single visits (left) and multiple visits (right).

2.2 MIP Model Based on Event Points

The set of event points for each vehicle u is defined as $K_u = \{1, 2, \ldots, k_u^{\max}\}$, where $k_u^{\max} \leq \sum_{i \in I_u} m_i$. Let U_i denote the subset of rigs suitable for site $i \in I$, i.e. $U_i = \{u \in U : \ i \in I_u\}$. Introduce the variables:

$x_{uik} \in \{0, 1\}$ such that $x_{uik} = 1$ iff vehicle u visits site i in event point k;
$y_{uik} \in \mathbb{Z}^+$ is the number of wells of site i drilled by vehicle u in event point k;
$t_{uk}^s \geq 0$ is the starting time of works for vehicle u in event point k;
$t_{uk}^f \geq 0$ is the completion time of works for vehicle u in event point k;
$t_{uk}^w \geq 0$ is the traveling time and waiting time between sites in event points $k-1$

and k.
$t_{uk} \geq 0$ is the traveling time between sites in event points $k-1$ and k.

Then the set of feasible solutions is defined as follows:

$$1 \leq \sum_{u \in U_i} \sum_{k \in K_u} x_{uik} \leq m_i, \ i \in I, \tag{1}$$

$$\sum_{i \in I_u} x_{uik} \leq 1, \ u \in U, \ k \in K_u, \tag{2}$$

$$\sum_{i \in I_u} x_{u,i,k-1} \geq \sum_{i \in I_u} x_{uik}, \ u \in U, k \in K_u, k > 2, \tag{3}$$

$$x_{u,id_u,1} = 1, \ x_{u,id_u,k} = 0, \ x_{u,i,1} = 0, \ u \in U, \ i \in I_u, \ k \in K_u, k > 1, \tag{4}$$

$$\sum_{u \in U_i} \sum_{k \in K_u} y_{uik} = n_i, \ i \in I, \tag{5}$$

$$y_{uik} \geq x_{uik}, \ i \in I, \ u \in U_i, \ k \in K_u, \tag{6}$$

$$y_{uik} \leq n_i x_{uik}, \ i \in I, \ u \in U_i, \ k \in K_u, \tag{7}$$

$$t^w_{uk} \geq \sum_{i \in I_u \cup \{id_u\}} s_{uij} x_{u,i,k-1} - s_{\max}(1 - x_{ujk}), \tag{8}$$

$$u \in U, \ j \in I_u, \ k \in K_u, \ k > 1,$$

$$\sum_{1<k' \leq k} \sum_{i \in I_u} d_{ui} y_{u,i,k'} + \sum_{1<k' \leq k} t^w_{u,k'} \leq \sum_{i \in I_u} b_i x_{uik} + b_{\max}(1 - \sum_{i \in I_u} x_{uik}), \tag{9}$$

$$u \in U, \ k \in K_u, \ k > 1,$$

$$\sum_{1<k'<k} \sum_{i \in I_u} d_{ui} y_{u,i,k'} + \sum_{1<k' \leq k} t^w_{u,k'} \geq \sum_{i \in I_u} a_i x_{uik}, \ u \in U, \ k \in K_u, \ k > 1. \tag{10}$$

Here $b_{\max} = \max_{i \in I} b_i$, $s_{\max} = \max_{i,j \in I, \ u \in U_i \cap U_j} s_{uij}$. Inequality (1) provides the upper bound on the number of visits of a site. Constraint (2) implies that in any event point on rig u at most one site may be served. Constraints (3) ensure continuous usage of event points, i.e. if an event point is used for visiting some site, then the previous one is used as well. The initial positions of vehicles are given by constraints (4). Conditions (5) guarantee that all wells of site i will be drilled. If a site i is not served by rig u in the event point k (i.e. $x_{uik} = 0$) then the number of drilled wells should be zero – this is ensured by inequality (7). Constraint (6) indicates that at least one well must be drilled if a rig visits site i. Conditions (9) and (10) ensure rigs routes feasibility with respect to time windows. The traveling time plus waiting time between event points $k-1$ and k is calculated in (8).

We also can modify constraint (1) for the case of the upper bounds on the number visits m'_i for each rig instead of all rigs:

$$\sum_{k \in K_u} x_{uik} \leq m'_i, \ i \in I, \ u \in U_i. \tag{11}$$

The optimization criterion for the presented model is formulated in the following form: minimize

$$f = \sum_{u \in U} \sum_{k \in K_u,\ k>1} t_{uk}, \tag{12}$$

$$t_{uk} \geq \sum_{i \in I_u \cup \{id_u\}} s_{uij} x_{u,i,k-1} - s_{\max}(1 - x_{ujk}), \tag{13}$$

$$u \in U,\ j \in I_u,\ k \in K_u,\ k > 1.$$

The traveling time between event points $k-1$ and k is calculated in (13), and the objective function (12) summarizes the traveling times between all event points.

Using additional variables $t^s_{uk} \geq 0$ and $t^f_{uk} \geq 0$, we can rewrite constraints (9) and (10) in the equivalent form

$$t^f_{uk} \geq t^s_{uk} + \sum_{i \in I_u} d_{ui} y_{uik},\ u \in U,\ k \in K_u, \tag{14}$$

$$t^s_{uk} \geq t^f_{u,k-1} + t_{uk} - b_{\max}(1 - \sum_{i \in I_u} x_{uik}),\ u \in U,\ k \in K_u, k > 1, \tag{15}$$

$$t^f_{uk} \leq \sum_{i \in I_u} b_i x_{uik},\ u \in U,\ k \in K_u, \tag{16}$$

$$t^s_{uk} \geq \sum_{i \in I_u} a_i x_{uik},\ u \in U,\ k \in K_u. \tag{17}$$

Our preliminary computational experiment shows that model (1)-(7), (12)-(17) is more appropriate for commercial solvers (CPLEX, GUROBI) than model (1)-(10), (12)-(13). The model contains $\sum\limits_{u \in U} \left(\sum\limits_{i \in I_u} m_i \right) \cdot |I_u|$ Boolean variables as well as integer variables.

3 Stochastic Model

In this section we consider a stochastic version of the problem, and construct mixed integer linear programming model similar to model (1)-(7), (12)-(17).

Suppose that drilling times d_{ui} of wells on sites are discrete random variables with values d_{uih} and probabilities p_{ih}, $h = 1, \ldots, v_i$, $\sum_{h=1}^{v_i} p_{ih} = 1$. Here we assume that these probabilities do not depend on u, in other words each outcome h defines the whole vector $(d_{u_1,ih}, ..., d_{u_{|U|},ih})$. Considering all possible combinations of drilling times at sites, we form the total set of possible scenarios SC with cardinality $v = \prod_{i \in I} v_i$. Let $d_{u,i,sc}$ denote the drilling time of a well of site i by rig u in accordance with scenario sc, and p_{sc} be the probability of scenario sc. Now our goal will be to define the rig routes and assign the number of wells drilling to each rig in each visit to a site, minimizing the traveling distance, s.t.

the probability of satisfying all time windows constraints is not less than a given threshold level α.

Given some values of all Boolean variables $\mathbf{x} = (x_{uik})$ and integer variables $\mathbf{y} = (y_{uik})$ from MIP problem (1)–(7), (12)-(17) and given a specific realization sc of the random scenario, one can define a function $Q(\mathbf{x}, \mathbf{y}, sc)$ to be 0 if the system of constraints (14)–(17) is consistent for the fixed $\mathbf{x}, \mathbf{y}, sc$, and define $Q(\mathbf{x}, \mathbf{y}, sc) = 1$ otherwise. Using this function, the problem from Section 2 for a single fixed scenario sc may be defined as an optimization problem w.r.t. two vectors of variables $\mathbf{x}$, $\mathbf{y}$ and a vector $\mathbf{t} = (t_{uk})$, asking to minimize the objective function (12), subject to the set of constraints (1)–(7), (13), and $Q(\mathbf{x}, \mathbf{y}, sc) \le 0$.

Let us denote the system of constraints (1)–(7), (13) on variables $\mathbf{x}$, $\mathbf{y}$, $\mathbf{t}$ by $(\mathbf{x}, \mathbf{y}, \mathbf{t})\mathbf{R} \le \mathbf{r}$, where matrix $\mathbf{R}$ and a row-vector $\mathbf{r}$ are defined appropriately on the basis of the input data. Then the stochastic optimization problem mentioned above may be formulated as

$$\min_{\mathbf{x},\mathbf{y},\mathbf{t}} \sum_{u \in U} \sum_{k \in K_u,\ k>1} t_{uk},$$

$$(\mathbf{x}, \mathbf{y}, \mathbf{t})\mathbf{R} \le \mathbf{r},$$

$$\Pr\{Q(\mathbf{x}, \mathbf{y}, sc) \le 0\} \ge \alpha.$$

Let $\mathbf{w} = (w_1, \ldots, w_{|SC|})$ be a vector of scenario indicators. Application of Theorem 1 from [6] shows that the stochastic optimization problem is equivalent to the following deterministic MIP problem.

$$\min_{\mathbf{x},\mathbf{y},\mathbf{t},\mathbf{w}} \sum_{u \in U} \sum_{k \in K_u,\ k>1} t_{uk},$$

$$(\mathbf{x}, \mathbf{y}, \mathbf{t})\mathbf{R} \le \mathbf{r},$$

$$Q(\mathbf{x}, \mathbf{y}, sc) \le 1 - w_{sc}, \quad sc \in SC,$$

$$\sum_{sc \in SC} w_{sc} p_{sc} \ge \alpha.$$

Here the confidence set of level α is formed by the Boolean variables w_{sc} such that $w_{sc} = 1$ if scenario sc belongs to the confidence set, and $w_{sc} = 0$ otherwise.

An equivalent of the constraint $Q(\mathbf{x}, \mathbf{y}, sc) \le 1 - w_{sc}$ with variables for starting times and completion times may be written as follows.

$$t^f_{u,k,sc} \ge t^s_{u,k,sc} + \sum_{i \in I_u} d_{u,i,sc} y_{uik},\ u \in U,\ k \in K_u,\ sc \in SC, \tag{18}$$

$$t^s_{u,k,sc} \ge t^f_{u,k-1,sc} + t_{uk} - b_{\max}(1 - \sum_{i \in I_u} x_{uik}), \tag{19}$$

$$u \in U,\ k \in K_u, k > 1,\ sc \in SC,$$

$$t^f_{u,k,sc} \le \sum_{i \in I_u} b_i x_{uik} + b_{\max}(1 - w_{sc}),\ u \in U,\ k \in K_u,\ sc \in SC, \tag{20}$$

$$t^s_{u,k,sc} \ge \sum_{i \in I_u} a_i x_{uik} - b_{\max}(1 - w_{sc}),\ u \in U,\ k \in K_u,\ sc \in SC, \tag{21}$$

3.1 Illustrative Example

Consider an instance with two rigs and one site including two wells. The site has a time window $(0, 4]$, and the drilling time may be 2 or 3 with some probability. The traveling time from rig deports to the site is 1. It is easy to see that if wells are drilled for 2 time units then only one rig visits the site and the objective is 1, but if the wells are drilled for 3 time units then both rigs visit the site and the objective is 2.

4 MIP Model Based on VRP Approach

The proposed VRP-based model is similar to the model from [7], but allows to visit a site by the same rig several times.

For each site $i \in I$ we create m_i copies and introduce a new set of sites I'. All copies of the same original site have identical set of wells. Denote by I'_i all copies of the original site i, $I' = \cup_{i \in I} I'_i$. Traveling times between site copies from I'_i are equal to zero, traveling times between copies of different sites are equal to the traveling times between these sites. Introduce a dummy site f_s corresponding to starting and completion point of the rout of each rig, set $I'_f := I' \cup \{f_s\}$. Put traveling times $s_{u,f_s,i'} := s_{u,id_u,i}$ and $s_{u,i',f_s} := 0$ for $i' \in I'_i,\ i \in I$. All rigs are suitable for the dummy site f_s, i.e. $U_{f_s} = U$. Set $I'_u := \cup_{i \in I_u} I'_i \cup \{f_s\}$ for all $u \in U$ and $U_{i'} = U_i$ for all $i' \in I'_i,\ i \in I$.

Introduce Boolean variables $x_{ui'j'}$ such that $x_{ui'j'} = 1$ if rig u visits site-copy i' and travels to site-copy j', and $x_{ui'j'} = 0$ otherwise. Let Real variables $t^s_{ui'}$ defines the starting time of works for vehicle u on site-copy i', and integer variables $y_{ui'}$ counts the number of wells of site-copy i' drilled by vehicle u. We formulate the following mixed integer linear programming model: minimize

$$f = \sum_{u \in U} \sum_{i' \in I'_f} \sum_{j' \in I'_f} s_{ui'j'} x_{ui'j'}, \tag{22}$$

$$\sum_{j' \in I'_f} x_{ui'j'} = \sum_{j' \in I'_f} x_{uj'i'},\ u \in U,\ i' \in I'_u \setminus \{f_s\}, \tag{23}$$

$$\sum_{u \in U_{j'}} \sum_{i' \in I'_u} x_{ui'j'} \le 1,\ j' \in I', \tag{24}$$

$$\sum_{i' \in I'_i} \sum_{u \in U_i} y_{ui'} = n_i,\ i \in I, \tag{25}$$

$$y_{ui'} \ge \sum_{j' \in I'_u} x_{uj'i'},\ i' \in I',\ u \in U_{i'}, \tag{26}$$

$$y_{ui'} \le n_i \sum_{j' \in I'_u} x_{uj'i'},\ i' \in I',\ u \in U_{i'}, \tag{27}$$

$$t^s_{ui'} + y_{ui'} d_{ui} + s_{u,i',j'} \le t^s_{u,j'} + b_{\max}(1 - x_{ui'j'}), \tag{28}$$

$$i' \neq j' \in I',\ i:\ i' \in I'_i,\ u \in U_{i'} \cap U_{j'},$$

$$s_{u,f_s,j'} \le t^s_{uj'} + b_{\max}(1 - x_{u,f_s,j'}),\ j' \in I',\ u \in U_{j'}, \tag{29}$$

$$t^s_{ui'} \ge \sum_{j' \in I'_u} a_i x_{uj'i'},\ i \in I,\ i' \in I'_i,\ u \in U_i, \tag{30}$$

$$t^s_{ui'} + y_{ui'} d_{ui} \le \sum_{j' \in I'_u} b_i x_{uj'i'},\ i \in I,\ i' \in I'_i,\ u \in U_i, \tag{31}$$

$$\sum_{i' \in I'_u} x_{u,f_s,i'} = 1,\ u \in U, \tag{32}$$

$$\sum_{i' \in I'_u} x_{u,i',f_s} = 1,\ u \in U. \tag{33}$$

Constraints (23) guarantee that each site-copy has exactly one predecessor and one successor in the route. Inequalities (24) indicate that each rig visits each site-copy at most ones. Conditions (25)–(27) ensure that the required number of wells are drilled at each site, and each well is drilled by one rig. Constraints (28)–(29) set the starting times of the works on sites for rigs. Inequalities (30)–(31) ensure feasibility of rig routes with respect to time windows. Conditions (32)–(33) indicate that each rig starts and completes its rout in depot.

The model contains $\sum\limits_{u \in U} \left(\sum\limits_{i \in I_u} m_i \right)^2$ Boolean variables and $\sum\limits_{u \in U} \sum\limits_{i \in I_u} m_i$ integer variables.

4.1 Stochastic Version

In the stochastic version, as in Section 3, we introduce binary variables w_{sc} equipped with the constraint (**??**), add the scenario index to variables $t^s_{u,i}$ and replace constraints (28)-(31) by the following scenarios-based conditions:

$$t^s_{u,i',sc} + y_{ui'} d_{u,i,sc} + s_{ui'j'} \le t^s_{u,j',sc} + b_{\max}(2 - x_{ui'j'} - w_{sc}), \tag{34}$$

$$i' \neq j' \in I',\ i:\ i' \in I'_i,\ u \in U_{i'} \cap U_{j'},\ sc \in SC,$$

$$s_{u,f_s,j'} \le t^s_{uj'sc} + b_{\max}(2 - x_{u,f_s,j'} - w_{sc}), j' \in I',\ u \in U_{j'},\ sc \in SC, \tag{35}$$

$$t^s_{u,i',sc} \ge \sum_{j' \in I'_u} a_i x_{uj'i'} - b_{\max}(1 - w_{sc}), i \in I,\ i' \in I'_i,\ u \in U_i,\ sc \in SC, \tag{36}$$

$$t^s_{u,i',sc} + y_{ui'} d_{u,i,sc} \le \sum_{j' \in I'_u} b_i x_{uj'i'} + b_{\max}(1 - w_{sc}), \tag{37}$$

$$i \in I,\ i' \in I'_i,\ u \in U_i,\ sc \in SC.$$

5 Greedy Algorithm for Stochastic Optimization

The number of binary variables w_{sc} grows exponentially in the number of sites with uncertain drilling time. Therefore, a straightforward application of a MIP solver allows to solve only small-sized instances. In order to treat larger instances, we propose a simple randomized greedy heuristic.

Recall that each scenario sc is some realization of random drilling times and it is represented as an $|I|$-dimensional vector $(sc_1, \ldots, sc_{|I|})$. Let us say that scenario sc dominates scenario sc' iff $sc_i \geq sc'_i$ for all i. Clearly, in this case if there is a solution to the considered stochastic problem with $w_{sc} = 1$ one may always set $w_{sc'} = 1$ in this solution without violation of its feasibility or worsening the objective function value. For any subset of scenarios $S \subset SC$ define $D(S) \subset SC$ as the set of all scenarios that are dominated by at least one element of S (note that each scenario dominates itself, so $S \subseteq D(S)$). For a subset S consider a stochastic optimization problem in which the constraint $\sum_{sc} p_{sc} \cdot w_{sc} \geq \alpha$ is excluded, all variables $w_{sc}, sc \in S$ are fixed to one, and all other w_{sc} are fixed to zero. Denote this problem as $P(S)$.

With these notations, we may reformulate our stochastic problem as follows: Find a subset of scenarios S such that the total probability of $D(S)$ is not less than α and an optimal value of the objective function of problem $P(S)$ is minimal. The proposed greedy algorithm is aimed at finding such a subset S and its outline is given below.

Algorithm 1 Randomized Greedy Algorithm

1: Set $S := \emptyset, p := 0$.

2: Repeat until $p \geq \alpha$ or the running time exceeds the given limit.

- 2.1 Choose scenarios $sc^1, ..., sc^r \in SC$ uniformly at random.
- 2.2 Solve r stochastic optimization problems $P(S \cup \{sc^1\}), ...P(S \cup \{sc^r\})$ and let $f^1, ..., f^r$ be the objective function values for the obtained solutions.
- 2.3 Choose sc^j with the minimal value f^j, add it to S, and remove all dominated scenarios: $S := S \cup \{sc^j\}, SC := SC \setminus D(\{sc^j\})$.
- 2.4 Update the current value of p as the total probability of $D(S)$.

The solution of stochastic optimization problems at Step 2.2 can be done in parallel. Due to the random nature of the algorithm, it is reasonable to run it several times and choose the best result. One positive feature of the algorithm is that it produces a sequence of solutions with increasing probability values p and the corresponding values of f, which gives a better understanding of the problem structure to a decision maker (this will be illustrated in Section 6.3).

6 Implementation and the Computer Experiments

In the experiments, we used a server with two AMD EPYC 7502 processors (each one has 32 cores, hyper-threading mode on), OS Ubuntu 20.04. MILP solver Gurobi (version 9.0.3) was applied to solve MIP problems[1] coded in GAMS.

6.1 Testing Deterministic Models

First, we tested the event-point-based and the VRP-based models on instances with 50 sites from [8], where the results of Gurobi were presented for the VRP-based model version with no returns. Two versions of the models are investigated, when two and when three visits of sites are allowed for rigs. Computational experiment with the same parameters as in [8] did not improve the objective values. We believe that this is due to the structure of the instances from [8], where time windows for objects are uniformly distributed during the planing horizon and have lengths less than or equal to the total drilling time of wells on the objects. The event-point-based model demonstrated slightly worse results than VRP-based model.

Second, we compared the two deterministic models on a family D of problems D_k, $k \in N$, constructed on the basis of the example from Subsection 2.1. Problem D_k consists of k subproblems with the structure as shown in Fig. 1 and the same initial data. It has $6k$ sites and $3k$ drilling rigs. Each subproblem has the set of sites $G_v = \{6v-5, 6v-4, ..., 6v\}$, the set of rigs $U_v = \{3v-2,\ 3v-1, 3v\}$ and the point of the initial location of rigs id_v for $v = 1, .., k$. Rigs of U_v can serve all sites from G_v, as well as the first site from the set with the next index (i.e. G_{v+1}). In the distance matrices, we put $s_{u,i,6v+1} = s_{u,6v+1,i} = i + 10$ for $v = 1, ..., k-1$, $i \in G_v$, $u \in U_v$. For all v and $i \in G_v$, $u \in U_v$, put $s_{u,id_v,i} = 5$. Direct transportation is prohibited between all other pairs of sites. For the forbidden movements of drilling rigs, we will assign sufficiently large values as distances. In the experiment, we set this value to 265.

For the VRP-based model, the Gurobi solver was used with the parameters Presolve = 2, GomoryPasses = 0, Method = 0, MinRelNodes = 10627, ImproveStartTime = 8640. For MIP model based on event points we did not find any parameters settings better than the default ones, so the default settings were used. The results for $k = 1, 2, .., 6$ are shown in Table 1. In the case of one visit, these instances required little solving time in both models, although the VRP-based model required less time for most instances. In the case of two visits, as k grows, the problems become more difficult for both models. For all instances, solutions with the optimal value of the objective function were found, but using the VRP-based model, the solver failed to prove the obtained solutions optimality in 10 hours of CPU time for $k \geq 4$. The EP-based model yields the best results for this series.

[1] The choice of this solver was based on a preliminary experiment, which indicated that on the MIP instances considered here Gurobi has an advantage to other solvers available to us (e.g. it was approximately twice as fast in comparison with CPLEX 12.10.0.0).

Table 1. Comparison of models on series D.

k	$\|I\|$	$\|U\|$	at most 1 visit			at most 2 visits		
			Obj	Time		Obj	Time	
				EP-based	VRP-based		EP-based	VRP-based
1	6	3	24	0,89	**0,21**	21	2,50	**1,82**
2	12	6	48	7,14	**1,55**	42	149,36	**114,93**
3	18	9	72	77,94	**7,41**	63	**576,40**	10502,38
4	24	12	96	328,89	**11,71**	84	**703,38**	> 36000
5	30	15	120	97,46	**7,95**	105	**5888,08**	> 36000
6	36	18	144	**40,11**	88,36	126	**10695,31**	> 36000

6.2 Testing Stochastic Models

The experiments were done on subinstances of the instances from [8]. In two series $S1$ and $S1'$ (10 instances in each series) we take the first seven or the first ten sites, and all given six rigs. Sites have from 5 to 30 wells. Note that instances from $S1'$ are characterized by shorter time windows than instances from $S1$.

For five random sites of each instance we suppose that drilling time can take two values 2 or 3, and the probability of value 2 is generated randomly from the interval $[0.75, 0.85]$. Drilling times are equal to 2 for the rest of the sites. For the sake of simplicity, we assume that $d_{u,i,sc}$ do not depend on u, i.e. they are the same for any fixed pair i and sc.

We test event-point-based and VRP-based models for threshold levels $\alpha = 0.5; 0.6; 0.7; 0.8; 0.9; 0.99$ in two versions, when one and when two visits of sites are allowed for rigs. The results for two instances with seven sites of series $S1$ are presented in Table 2 (the full results for all instances are available at https://gitlab.com/YuliaKovalenko-gl/stochastic-vrp-problem.git) The running time of Gurobi is greater on series $S1'$ than on series $S1$ due to the structure of the instances in this series.

In most of the instances the optimal stochastic solution at level $\alpha = 0.8$ has a lower traveling distance, compared to the worst-case scenario, where all drilling times are equal to 3. In the instances presented in Table 2, the version allowing up to two visits yields a solution with a lower objective traveling distance, compared to the version with no returns. As we can see the running time of Gurobi has no specific tendency as a function of threshold α. In the two-visit-version, none of the considered models clearly dominates the other one in terms of running time of Gurobi. The VRP-based model demonstrates better results in a majority of the cases with no returns.

6.3 Evaluation of the greedy algorithm

For testing the greedy algorithms, two instances from previous section, namely 7_2 and 7_3, were taken and four larger problems were generated on the basis of the instances S1.1, S1.2, S1.3, S1.4 from [8]. The original instances contain 50

Table 2. Comparison of models on series with 7 sites.

α	at most 1 visit Obj	Time EP-based	Time VRP-based	at most 2 visits Obj	Time EP-based	Time VRP-based
Instance 7_2						
0.5	10	34,86	0,63	**9**	22,09	3,12
0,6	10	24,54	6,70	10	64,34	41,14
0,7	10	29,38	6,07	10	60,62	29,09
0,8	11	74,47	9,65	**10**	62,67	63,48
0,9	13	410,16	30,88	**11**	211,69	727,08
0,99	13	133,62	23,92	13	315,74	381,81
Instance 7_3						
0,5	16	20,65	4,35	16	29,59	51,6
0,6	16	4,33	3,07	16	13,42	27,08
0,7	17	8,64	4,64	17	22,56	113,11
0,8	19	62,14	9,95	**18**	24,28	143,03
0,9	21	21,88	17,35	**20**	56,21	159,91
0,99	22	13,47	4,06	**21**	7,03	51,46

sites and 6 rigs. Here only the first 12 sites are extracted. All the travel times are kept unchanged, but the drilling times $d_{u,i}$ now take value 1 with probability 0.8 and value 3 with probability 0.2 for each site. The total number of scenarios is then $2^{12} = 4096$. The MIP model for the first instance has 547980 columns, 7216 discrete-columns, and 3858896 rows. The straightforward application of Gurobi with $\alpha = 0.7$ could not find a feasible solution in five hours.

In the implementation of the greedy algorithm, at Step 2.1, the number of considered scenarios is $r = 3$, and in case of large instances they are chosen at random among the scenarios, in which $d_{u,i,sc}$ have value 3 for more than five sites, otherwise the probability of $D(\{sc\})$ is negligibly small. At Step 2.2, three problems are solved in parallel by Gurobi, each process is allowed to use up to four CPU cores, and the solving time of one problem is limited by 180 seconds. The algorithm stops when it reaches the level $\alpha = 0.95$, or when the running time exceeds the overall time limit, which was set to one hour. Although the algorithm may work with both EP and VRP based models, the VRP case was chosen, because it showed better performance in earlier tests with no returns.

For each problem instance, five independent runs of the greedy algorithm were made and the best results were collected and summarized in Table 3. As before, the smaller problems were solved in two variants: with at most one or at most two visits of each site (this is marked with "1v" or "2v" in the table). Column "optimistic" shows the objective function values of the solution with the best realization of the drilling times, i.e. the scenario sc with all $d_{u,i,sc} = 1$; similarly, column "pessimistic" corresponds to the worst-case scenario with all $d_{u,i,sc} = 3$. The other columns show the best results provided by the algorithm after reaching the given probability threshold. The cells, for which no feasible solutions were obtained are marked with "–". For example, let us fix $\alpha = 0.8$, then for instance S2.1 there exists a solution with the cost $f = 35$ that is valid with probability at least α. For the smaller problems, the obtained solutions are quite close to the optimal ones (results known to be optimal are marked by "*", compare to Table 2). For problems S2.1...S2.4, which can not be straightforwardly solved by the MIP solver in practically acceptable time, the greedy algorithm still yields reasonable solutions.

Table 3. Results of the greedy algorithm

Instance	optimistic	p=0.6	p=0.7	p=0.8	p=0.9	p=0.95	pessimistic
7_2(1v)	10	10*	10*	11*	13*	13*	13
7_2(2v)	9	10*	10*	11	11*	13	13
7_3(1v)	16	16*	18	19*	21*	22	22
7_3(2v)	15	16*	18	19	20*	21	21
S2_1(1v)	13	29	33	35	44	–	–
S2_2(1v)	10	21	22	29	31	31	31
S2_3(1v)	16	39	39	42	43	45	45
S2_4(1v)	12	35	40	42	–	–	–

7 Conclusions

In this paper, we have studied a generalization of the drilling rig routing problem suggested by I. Kulachenko and P. Kononova, allowing to re-visit the same site by the same rig and assuming that at some sites the drilling durations are random variables with known discrete distributions. We have compared two different approaches to defining a MIP formulation of the problem, a one based on the classical VRP model with time windows and a one based on the event points approach. Also, we have carried out a computational experiment, comparing the performance of Gurobi solver on these MIP models and found out that in different cases either one of the models has an advantage. To solve approximately the stochastic optimization problems of higher dimension, we have developed a randomized greedy heuristic, which demonstrated promising results.

Acknowledgement

The research is supported by Russian Science Foundation grant 21-41-09017. A server of Sobolev Institute of Mathematics, Omsk Branch is used for computing.

References

1. Bassi, H.V., Ferreira Filho, V.J.M., Bahiense, L.: Planning and scheduling a fleet of rigs using simulation–optimization. Comp. & Ind. Eng. **63**(4), 1074–1088 (2012)

2. Borisovsky, P.A., Eremeev, A.V., Kallrath, J.: Multi-product continuous plant scheduling: combination of decomposition, genetic algorithm, and constructive heuristic. International Journal of Production Research **58**(9), 2677–2695 (2020)
3. Errico, F., Desaulniers,G., Gendreau, M. Rei,W. Rousseau, L.-M.: The vehicle routing problem with hard time windows and stochastic service times. EURO Journal on Transportation and Logistics **7**(3), 223–251 (2018)
4. Iyer, R.R., Grossmann, I.E., Vasantharajan, S., Cullick, A.S.: Optimal Planning and Scheduling of Offshore Oil Field Infrastructure Investment and Operations. Ind. Eng. Chem. Res. **37**(4), 1380–1397 (1998)
5. Ierapetritou, M.G., Floudas, C.A.: Effective continuous-time formulation for short-term scheduling: I. Multipurpose batch process. Ind. Eng. Chem. Res. **37**(11) 4341–4359 (1998)
6. Kibzun, A.I., Naumov, A.V., Norkin, V.I.: On reducing a quantile optimization problem with discrete distribution to a mixed integer programming problem. Automation and Remote Control **74**, 951–967 (2013)
7. Kulachenko, I., Kononova, P.: A Matheuristic for the Drilling Rig Routing Problem. In: Kononov, A. et al. (eds.): MOTOR 2020, LNCS, vol. 12095, pp. 343–358. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-49988-4_24
8. Kulachenko, I., Kononova, P.: A hybrid algorithm for drilling rig routing problem. Diskretny Analys i Issledovanie Operacii. To appear. (2021) (in Russian)
9. Laporte,G., Louveaux, F., Mercure, H.: An integer L-shaped algorithm for the capacitated vehicle routing problem with stochastic demands. Transportation Science **26**(3), 161–170 (1992)
10. Nasri, M., Metrane, A., Hafidi, I., Jamali, A.: A robust approach for solving a vehicle routing problem with time windows with uncertain service and travel times. International Journal of Industrial Engineering Computations **11**, 1–16 (2020)
11. Oyola, J., Arntzen, H., and Woodruff, D.L.: The stochastic vehicle routing problem, a literature review, part I: models. EURO Journal on Transportation and Logistics **7**(3) 193–221 (2018)
12. Shaik, M.A., Floudas, C.A., Kallrath, J., Pitz, H.J.: Production scheduling of a large-scale industrial continuous plant: short-term and medium-term scheduling. Comp. Chem. Engng. **33**, 670–686 (2009)
13. Souyris, S., Cristián E. Cortés, C.E., Fernando Ordóñez, F., Andres Weintraub, A.: A robust optimization approach to dispatching technicians under stochastic service times. Optimization Letters **7**, 1549–1568 (2013)
14. Tavallali, M. S., Zare, M.: Planning the Drilling Rig Activities – Routing and Allocation. Computer Aided Chemical Engineering 43, 1219–1224 (2018)
15. Cordeau, J.-F., Desaulniers, G., Desrosiers J., Solomon, M.M., Soumis, F. VRP with time windows. In: The vehicle routing problem / edited by P. Toth, D. Vigo, SIAM Monographs on Discrete Mathematics and Applications, SIAM, pp. 157–193 (2002)