A MONTE CARLO ROLLOUT ALGORITHM FOR STOCK CONTROL

Denise Holfeld* and Axel Simroth**

* Operations Research, Fraunhofer IVI, Germany, Email: denise.holfeld@ivi.fraunhofer.de
** Operations Research, Fraunhofer IVI, Germany, Email: axel.simroth@ivi.fraunhofer.de

Abstract The application of optimization in industrial processes is faced with many challenges. One of the main challenge is the possible inaccuracy of information. In contrast to mathematical optimization theory, information is not completely known a priori. Often information can only be estimated or changes over time. Another challenge is the need of a decision in real time. Both points are relevant for a control of a flexibly designed in-plant block storage. The schedule plan for storages and removals should be able to adapt quickly to changes. In this paper an algorithmic approach is presented which is able to react on dynamic and uncertain changes due to the production process. To this end, optimization algorithms are implemented within a rolling planning process, so it is possible to respond to updated information by adapting the current plan. A novel optimization method is developed to generate cost effective and robust solutions by looking ahead into the future.

Paper type: Research Paper

Published online: 30 October 2013
Vol. 3, No. 4, pp. 279-286

ISSN 2083-4942 (Print)
ISSN 2083-4950 (Online)

Keywords: *optimization under uncertainties, Monte Carlo Rollout, block storage control*

1. INTRODUCTION

An optimal control of storages within a production process is considered. More precisely, it is a block storage for transport wagons which is adapted to a production hall. The wagons are stored in rows, each with only one entry. As a result, wagons behind are blocked by wagons in the front. One storekeeper executes all storage and removal operations between the storage and staging areas. For that purpose, storing positions for necessary storage and removal operations have to be determined, taking into account the uncertainty of the retrieval sequence of production. A position for a storing should be selected, so that future removals will not be blocked. Furthermore, a wagon blocking a future removal should preferably be removed when appropriate. For this task an optimization algorithm was developed and is presented here.

To ensure real-time decisions, a method based on a heuristic in combination with a meta-heuristic to consider uncertainties is used. One try to take into account future developments resulting from a decision, too. Since one cannot simulate all possible developments, a Monte Carlo approach is used that is a statistic method based on a large number of random experiments. The principle of the Monte Carlo Rollout approach used here, is outlined as follows: One has a simple heuristic H to select a decision that is a position for a storage or a removal operation. This heuristic provides a short-sighted solution for the current instance taking into account the current knowledge of information. The Monte Carlo Rollout approach improves the heuristic solution by considering not only current costs, but also a future trend caused by this decision. The application of this method in a manufacturing process is supported by RFID Technology.

2. SETTINGS

2.1. Application in a manufacturing process

An in-plant logistic problem in a fridge and freezer production is considered. More precisely, the storage and the transport of wagons loaded with appliance doors should be optimized. The doors are produced in a foaming facility and are stored on wagons. On one wagon are doors for the fridge part and/or doors for the freezer part of one type of end device. The number of such door combinations on one wagon can vary. A wagon can hold up to 18 door combinations. The doors of all end devices to be produced are manufactured one day ahead. Consequently an intermediate storage of door wagons is necessary. There is only limited space for this storage. Hence, a block storage has to be organized. Several wagons will stand in one row with just one entry. If a wagon is needed before the wagons in front of it, relocations within the storage are necessary. The number of such relocations should be minimized. According to requirements of the final production, door wagons have to be provided on staging areas for all montage stations. There are 18 montage stations which retrieve door type combinations in question for the type of end device they produce at the moment. Which station will produce which type of end device is not known. All stations can choose from a list of production lots. Thus, the sequence of removals is not known a priori.

2.2. Mathematical methods

To optimize a scheduling problem one can use complex algorithms, e.g. branch-and-bound algorithm. However, these are time consuming for extensive problems. Additionally, solutions have to be calculated very often caused by uncertainties. Based on the real time requirement complex algorithms are inapplicable. On this account, a heuristic H is used to determine a solution based on current information. The idea is to improve the solution by a Monte Carlo Rollout (MCRo) approach as meta-heuristic. In this way, the heuristic is combined with a stochastic model for simulating the future uncertainties. This leads to better and more robust solutions for optimization problems under uncertainties.

The MCRo approach combines ideas from Rollout algorithms for combinatorial optimization and the Monte Carlo Tree Search in game theory. The main idea is to create a set of different decisions -- called alternatives -- and to evaluate the behaviour of each alternative in a set of random future developments. Based on the evaluation of the alternatives in the future developments, the best alternative is selected. In the following the Rollout algorithm and the Monte Carlo Tree Search will be briefly described. Afterwards the MCRo method and its application to the scheduling planning problem is presented.

Rollout algorithms (Bertsekas & Castanon, 1999),(Bertsekas, Tsitsiklis & Wui, 1997) can be used for optimization problems that have a sequential structure, i.e. that can be solved by making a sequence of consecutive decision steps with a limited number of alternative decisions in each step. By means of the Rollout method, each alternative decision is evaluated in order to choose the best alternative. The Rollout algorithm iteratively explores all different alternatives in the current decision step. It uses a so-called base heuristic for making decisions in the steps following the current decision. The base heuristic usually is a fast, rather simple but solid heuristic for the problem at hand, that solves the problem in a sequential manner. With the help of this base heuristic the Rollout algorithm gets an evaluation of the alternative at a leaf of the decision tree, namely at that leaf that would be reached if the base heuristic would be applied after choosing the alternative considered. After evaluating all alternatives in the current step, the one that leads to the best results is chosen.

Monte Carlo Tree Search (Chaslot, Bakkes, Szita & Spronck, 2008), (Brügmann, 1993), (Kocsis & Szepesvari, 2006) first was developed for computers to play the board game Go, but in the meantime it is the state-of-the-art technique for a set of single- or multi-person games. Monte Carlo Tree Search is used for problems where no good heuristic was found to evaluate a decision. Instead of using a noisy and possible misleading heuristic evaluation, the alternative decisions are evaluated by means of random games. Often there are game-specific information about the quality of moves which can be used to weight the possible moves.

The MCRo method combines both approaches to handle sequential optimization problems that are afflicted with uncertainties. The evaluation of an alternative decision by solving the problem further with a simple and fast base heuristic is adapted from the Rollout approach. The uncertainties are covered through the random selection of future situations, by means of a random player as in the Monte Carlo Tree Search. So, the optimization problem with uncertainties is modelled as a two-player game. The first player is the decision maker that decides on the base of a simple heuristic. The second player is the random player that creates new future situations by random. The game where both players move consecutively is called MCRo. With a set of different MCRos, an alternative can be proven and evaluated in a set of random future scenarios, and the long-term behaviour and robustness against uncertainties of the alternative could be analysed. The MCRo method is shown schematically in Fig. 1.

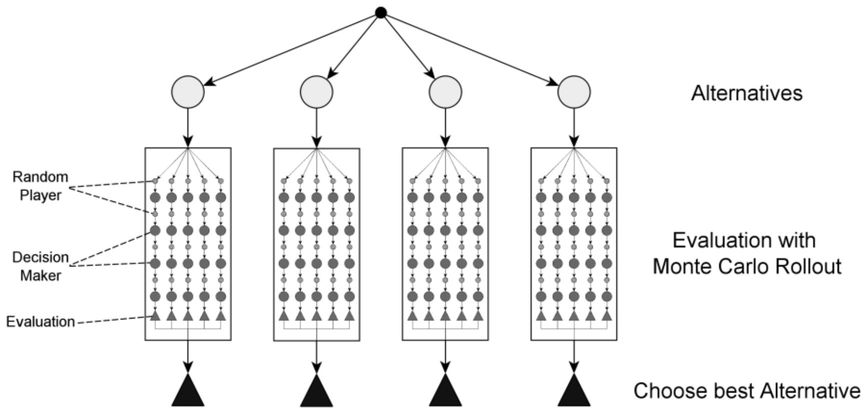

Fig. 1 Schematic structure of the MCRo method

2.3. Application of MCRo to a scheduling problem

The stochastic scheduling problem treated here is a combinational optimization problem in a sequential manner. The decision maker chooses for new storages positions in the stock and appropriate wagons for necessary removals. R is the estimated sequence of retrievals at this time. S^* is the sequence of planned storages and removals over a limited horizon. During the execution of sequence S^* there is a change in the sequence of retrievals *R*. This corresponds to the random player above. The decision maker has to react again.

For a removal, it is assumed, that a matching wagon is selected, which produces the least time expense. For a new storage the base heuristic *H* selects a position in the sense, that the stock is evenly filled. To make a decision where to store a new wagon one should use all known information on this date. Also information about the presence of uncertainty and, associated information therewith. Here, it is known that the sequence of retrievals *R* is estimated and can change over time. *R* results from calculations of time effort with the help of average values. These times may differ from real performance times. A sufficiently large deviation leads to a permutation in *R*. A permutation in *R* is also possible, when a montage station changes its production plan in the short term. A stochastic model is used to simulate possible permutations within *R* and there probabilities of occurrence. This model generates the moves of the "random player".

With the MCRo algorithm a position for a new storage is determined as follows. A limited number of possible positions in the stock $P = \{p_1, p_2, \dots, p_n\}$ are determined. Each of these positions $p_1, p_2, \dots, p_n$ is successively considered and evaluated. For that, it is assumed that the wagon is stored on position p_i. Starting from the resulting stock, *m* so-called MCRos are performed. From here on, the random player and decision maker alternate until a predefined number of future removals (the "depth") were considered. In the random player turn, performance times are generated by the stochastic model. This can lead to a permutation in *R*. Using the base heuristic *H* and the current sequence *R* the decision maker extends the sequence S* to a position for a new storage or removal. All necessary reallocations to fulfill the retrieval sequence in one MCRo will be counted and provide the evaluation of this MCRo. The solution quality of alternative p_i is determined by averaging all *m* MCRo evaluations. After all, the best alternative p^*, (which leads to the smallest average number of relocations) is chosen and the wagon is stored there.

2.4. Use of RFID technology

All transport wagons are equipped with RFID transponders. Wagons loaded with doors from the foaming facility have to pass a RFID gate. There the material numbers are assigned to the transponder of the wagon and the quantity of door combinations on the wagon must be entered. The system returns if a additional module-mounting is necessary or if the wagon can be directly placed on the staging area. On the staging areas are places for new wagons, where the wagon is registered by RFID. If there is a free space in the storage, the management system determine a storage area and the storekeeper has to storage the new wagon on this position. Ideally, this storage operation is combined with a removal operation. Hence, the management system returns beside the storage area for the new storage a position for a removal, too. A removed wagon will be scanned for confirmation and called from stock.

2.5. Results

In preparation for the launch of the storage control several simulations were performed. The storage includes 14 rows with 4 storage areas and 7 rows each, with 5, 6 respectively 8 wagons in one row. New storages from the foaming facility are parked on the staging area, but door wagons corresponding to the retrievals are provided by the storekeeper on this area, too. This staging area can hold up to 34 wagons.

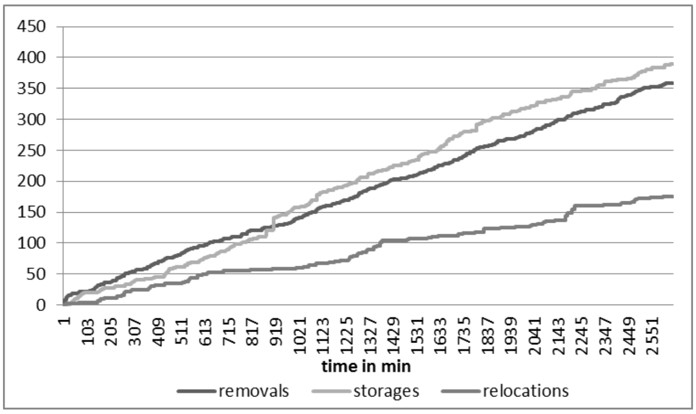

Fig. 2 Removals, storages and necessary relocations over 3 days

For the simulation data from the production over several days were used. A list of all produced lots provides the basis of the simulation. For each assembly station the next production lot is determined randomly. In this way, the dynamic retrieval sequence *R* is generated and thus all removal operations are defined. All storage operations result from the production plan of the foaming facility. The door production works in a 3-shift system, the final assembly in a 2-schift system. Doors produced in the night shift are stored in the early shift. The simulation period is 3 days, so all figures include 6 shifts a 440 minutes. The quantity of removals, storages and necessary relocations over the simulation period is shown in Fig. 2.

Based on these data, two different base heuristics were tested with different degrees of dynamic. Here, the degree of dynamic is reflected in the number of permutations in *R*. With heuristic *H1* the closest emptiest row is selected for a new storage. This is a simple and very quick heuristic without foresight. In contrast, the second heuristic *H2* requires more effort. With foresight to the next planned removals, corresponding rows are excluded for a new storage. Fig. 3 shows the results for both heuristics with and without meta-heuristic. Without MCRo the results of the

H2 are better than the results of *H1*, but deteriorate by the increasing uncertainty. This is due to the foresight. With an increasing degree of dynamic the foresight is inaccurate and *H2* takes the wrong decision. The results of both heuristics were improved by the use of MCRo as meta-heuristic. Better results are achieved by using the simple heuristic *H1* in combination with the MCRo.

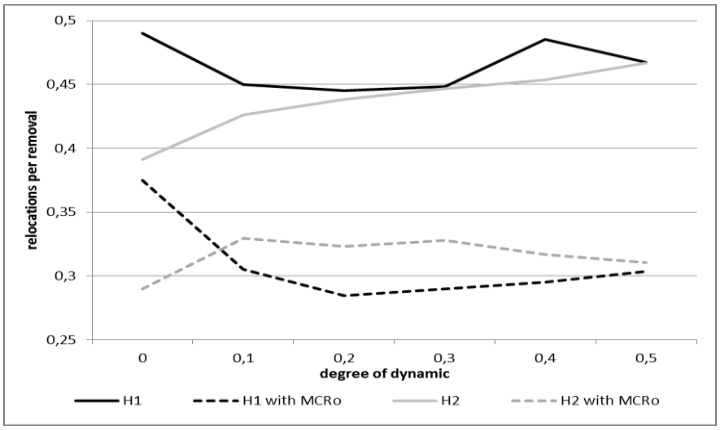

Fig. 3 Comparison of two heuristics with different degrees of dynamic

The number of relocations in Fig. 2 results from the use of *H1* in combination with the meta-heuristic. Fig. 4 shows the improvement of *H1* by using MCRo over the time. For illustration the number of relocations per removal is presented and compared. The use of the MCRo as meta-heuristic leads to an improvement of up to 50%.

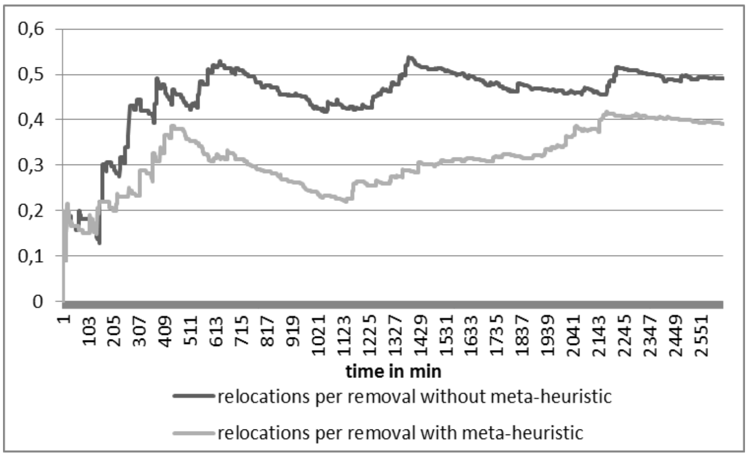

Fig. 4 Reduction of relocations by use of the meta-heuristic

3. CONCLUSION

A fast and robust optimization algorithm was developed to control a block storage with up to 8 items in one row. A simple and fast heuristic is used to ensure real-time applications. To generate robust solutions with respect to uncertainties in processes a Monte Carlo Rollout approach is used as meta-heuristic. It has been shown, that the expense of removal operations can be drastically reduced using the MCRo approach.

ACKNOWLEDGEMENTS

This work was supported by the German Federal Ministry of Education and Research (BMBF) in the funding program "KMU-innovativ: IKT" under the Project Number 01IS10024B ("EZOLAT").

REFERENCES

Bertsekas, D.P.; Castanon, D. A., (1999), "Rollout Algorithms for Stochastic Scheduling Problems", Journal of Heuristics, No. 5, pp. 89-108.

Chaslot, G.; Bakkes, S.; Szita, I.; Spronck, P., (2008), "Monte-Carlo Tree Search: A New Framework for Game AI", Proceedings of the Fourth Artificial Intelligence and Interactive Digital Entertainment Conference, pp. 216-217.

Bertsekas, D.P.; Tsitsiklis J.N.; Wui C., (1997), "Rollout Algorithms for Combinatorial Optimization", Journal of Heuristics, Vol. 3, pp. 245–262.

Kocsis, L.; Szepesvari, C., (2006), "Bandit based Monte-Carlo Planning", ECML-06. Number 4212 in LNCS, Springer, pp. 282-293.

Brügman B., (1993), "Monte Carlo Go".

BIOGRAPHICAL NOTES

Denise Holfeld is a researcher in Operation Research. She studied mathematics from 2002 to 2008 and completed with a degree in Business Mathematics. Subsequently, she worked in the research group Scientific Computing at the Technical University Dresden. Since 2010 she is member of the Operations Research group of the Fraunhofer Institute for Transportation and Infrastructure Systems IVI, where she is concerned with the formulation of mathematical optimization models and the design and implementation of solution approaches and algorithms. Her research interests are in Discrete Mathematics and Combinatorial Optimisation.