Commentary on the paper *"Multi-Objective Scheduling for NASA's Future Deep Space Network Array"* by Mark D. Johnston

Commentator: Roman Barták

Charles University, Faculty of Mathematics and Physics
Malostranské nám. 2/25, 118 00 Praha 1, Czech Republic
roman.bartak@mff.cuni.cz

Paper Summary

The paper addresses a problem of scheduling communication requests for future NASA Deep Space Array-Based Network (DSAN) – a terrestrial network of antenna sites that support communication for missions from high earth orbit to the outer planets. The main difference of this network from the existing Deep Space Network is using a large amount of smaller antennas. Consequently, users are not requesting a particular antenna but rather a communication capacity that can be satisfied by different sets of antennas. This gives more flexibility to the network but also increases complexity of the scheduling problem by adding a resource allocation part. Another complexity is handling many different objectives arising mainly from the competing requests from users.

The paper gives some details on this future real-life problem with the focus on narrative description of scheduling objectives, both from the mission (user) perspective and from the system perspective. The main part of the paper is dedicated to a solving technology. First, the multi-objective nature of the problem is highlighted and evolutionary algorithms are presented as a good solving technology for this type of problems. Two particular algorithms, NSGA II and GDE3, are described in more details as good candidates (GDE3 is then used in experiments). Finally, a particular formulation of the DSAN problem using a vector of real-life decision variables is presented. The paper is concluded by two experimental studies. The first experiment with two missions in contention shows convergence of the method to a set of solutions at the Pareto-optimal frontier. The second experiment with extrapolated 2015 mission data focuses on scalability of the method and it shows runtimes dependent on population size and problem granularity.

The paper is mostly clearly written, unless the reader does not go into technical details, because the problem specification remains mostly at a narrative level without a formal description. In the following, I will try to highlight some particular questions and propose some solutions from a more academic point of view that may help other researchers to look at the problem.

Problem Specification

The paper studies a future real-life resource allocation problem. Both these characteristics, future and real-life, make a more formal description very complicated. Real-life problems are frequently too complex to describe them formally in a limited space and future problems are not known in full details. From this perspective, it is clear why the author did not attempt to formalize the problem now and stayed with a narrative description. Nevertheless, a more precise description would be most valuable for the research community that can contribute to solving the problem. As I understand from the paper, the main complexity of the problem is specification of objectives while the resource allocation core is more or less known. A small difficulty is that it is not clear if there is a difference between a hard constraint and objective. The author attempts to model the constraints, like a limited capacity of antennas, as the highest level objective. The reason is that sometimes there is no solution satisfying all hard constraints and by treating the constraints as objectives, it is possible to identify these bottlenecks. On the other side, there is a difficulty that obtained solutions might not be applicable because the constraints are violated. I understand why author went in this direction (everything is just objective) because this is a typical model for evolutionary algorithms where it is complicated to guarantee satisfaction of hard constraints. Nevertheless, from the formal view, there should be a difference between the constraint and the objective. For example, I cannot imagine that maximal available antennas are modelled as objective, because it is surely not easy to build one more antenna so relaxing this constraint is not possible.

I shall now try to propose some basic formalisation of the problem. First, I would propose to define some notions before going into further description. It seems that pass is the smallest "object" we work it. Pass is an interval with a function describing minimal request for antennas (and possible other resources) in each time point and a delta describing the maximal number of additional antennas over this interval that may be useful for communication. When allocating the pass, it is possible to use a shorter interval, but it is necessary to stay within given bounds and the minimal number of allocated antennas in the scheduled time points must be guaranteed. These seem to be the only hard constraints in the paper. There might be other constraints about the pass, like its minimal duration. Several passes are grouped into contacts so it is possible to specify constraints between passes like the minimal and maximal distance between two subsequent passes in a contact or the minimal and maximal sum of durations of scheduled passes in the contact. Some of these constraints tend to be assumed as objectives, but again, it should be clearly specified what a hard constraint is. Finally, the problem is treated as a multi-criteria optimisation problem, but already some combination of objectives is done. Namely, the objectives related to a single user seem to be combined so it would be useful to explicitly describe this combination function.

The main reason why I argue for a more formal description of the problem is the possibility that other researchers may contribute directly to solving this particular problem rather than providing only "general" technology. The formal problem description may be simplified and accompanied by example problems (existing data) so it can be used for benchmarking. Something similar happened with other real-life problems like RLFAP (radio link frequency assignment problem) where the actual data are confidential, but still the core of problem is formally described and published.

Solving Technology

As I already mentioned, the paper contains a formulation of the problem useful for evolutionary algorithms. Namely, each pass is modelled using three real-valued variables in the interval [0,1] that describe shift of start and end times of the interval in respect to given bounds and possible over-allocation of antennas. This model naturally captures the hard constraints mentioned above (each pass is scheduled within the given interval and the minimal requested number of antennas is allocated), but all other "constraints" must be treated as objectives. The obtained vector of decision variables also naturally forms a solution candidate in the population used in evolutionary algorithms. There are still some unanswered questions, namely how exactly the crossover and mutation operations are realised over these vectors. This is in particular relevant to GDE3 algorithm that somehow combines the vectors and some guarantee that the vector remains a solution candidate should be included (the values must stay in the interval [0,1]). Perhaps some normalisation can be applied. Another open question is whether other hard constraints, like the maximal number of available antennas, can be encoded in the representation so the evolutionary approach can focus merely on optimisation. Such problems can be encoded in such a way that a population member is not directly the solution candidate, but a "procedure" how to obtain a feasible solution. I am not sure though whether this technique is applicable for this particular problem.

There is a clear advantage of using evolutionary algorithms – providing a set of candidate solutions. Having more alternative solutions is useful so the user can select among them. However, if the solutions are not similar (and the algorithms tend to obtain a greater diversity) then I can hardly imagine the negotiation process between more users because when many parties are involved in the decision process then chances to find a solution are going down. Hence providing a set of solutions at once is probably not that big advantage here and it might be useful to explore solving techniques proving a single solution too. For example, constraint satisfaction provides a rich formalism for expressing both hard and soft constraints and it has the advantage of giving only the solutions satisfying the hard constraints. There also exist iterative techniques that keep a partial good solution satisfying hard constraints while not-yet scheduled requests are added iteratively to this solution (Müller, 2005). The process is repeated until all/most requests are satisfied. A nice feature is that the solving process can be visualised and the solver can work in interaction with the user.

At the end, I would like to briefly discuss a feature that seems very important for the DSAN problem – problem dynamicity. The studied problem is naturally dynamic because the user requests are changing (urgent requests when anomalies occur) and, for example, bad weather may modify the requested number of antennas. As I understand, some predicted changes may be included among the objectives so, for example, more antennas are allocated for the pass if bad weather is expected. This way, a more robust schedule is constructed. Unexpected events, like urgent requests for communication from users are handled by re-scheduling. The current approach assumes that the last obtained population will be used as a start population with a modified set of objectives (what if the number or requests also changes?). By this way we can obtain a new set of solutions but the question is how far these solutions are from the original solution. Assume that the original solution has been published and it is being implemented while an urgent request arrives. It is important to integrate this request but without changing dramatically the current solution (like re-allocating gates at the airport after a delayed flight). The users will not be happy and sometimes even cannot accept, if already agreed allocation is changed. Hence similarity to the previous solution should probably be included among objectives when such dynamic situations occur. Minimal perturbation problem (Barták et al., 2003) is an attempt to formalise dynamic optimisation problems in the context of constraint satisfaction.

References

Barták, R.; Müller, T.; and Rudová H. 2004. A New Approach to Modelling and Solving Minimal Perturbation Problems. In *Recent Advances in Constraints, 2003*. Springer-Verlag LNAI 3010, 223-249.

Müller, T. 2005. *Constraint-Based Timetabling*, PhD thesis, Charles University.