Optimization & Analysis for Defense Simulation Models

Jose Ramirez
OptTek Systems, Inc.
Washington, D.C.
ramirez@opttek.com

Benjamin Thengvall
OptTek Systems, Inc.
Boulder, CO
thengvall@opttek.com

ABSTRACT

When performing defense system analysis with simulation models, a great deal of time and effort is expended creating representations of real-world scenarios in Department of Defense (DoD) simulation tools. However, once these models have been created and validated, analysts rarely retrieve all the knowledge and insights that the models may yield. Analysts are limited to simple explorations because they do not have the time and training to perform more complex analyses manually. Additionally, they do not have software integrated with their simulation tools to automate these analyses and retrieve all the knowledge and insights available from their models. Simple, manual explorations are inefficient in their use of computing resources and are often ineffective in providing the best answers to analyst questions. To derive the greatest benefit from a simulation model, an analyst should apply both optimization and statistical analysis techniques. Combining these techniques, and using available simulation optimization and analysis tools, can provide answers to these essential questions and key insights for decision makers. More importantly, the organizational return on investment from simulation studies is increased which builds stakeholder confidence. Tools like these can also be used for model verification and validation.

ABOUT THE AUTHORS

Jose Ramirez is Vice President of Government Services & Principal Investigator at OptTek Systems, Inc. and has a Ph.D. in Operations Management (Operations Research). He served as the Warfighting Analysis Division Chief (Colonel) in the Joint Staff J-8 where he led, managed, and provided analytical guidance to a 40-personnel team of analytical modelers, national defense strategists, and cybersecurity personnel. He has guided campaign level state-of-the-art discrete-event modeling and simulation and data analytics supporting the Chairman of the Joint Chiefs of Staff and Combatant Commanders (INDOPACOM, EUCOM, and NATO). He has strategized and conducted operational assessments of capabilities versus peer adversaries that provide insights on future equipment modernization investment options, and implemented advanced data analytics technology on the DoD's global munitions requirements process.

Benjamin Thengvall is Chief Operating Officer at OptTek Systems, Inc. and has an M.S.E. and Ph.D. in Operations Research and Industrial Engineering. He is an expert in the areas of mathematical modeling, real-time optimization software and services, transportation and scheduling problems, agent-based and discrete-event simulation, and simulation optimization and analysis with over 20 years of experience. He has spent his career providing innovative software solutions to complex real-world problems through mathematical modeling, simulation, and metaheuristic techniques. His experience spans many projects in both the commercial and government spheres.

Optimization & Analysis for Defense Simulation Models

Jose Ramirez
OptTek Systems, Inc.
Washington, D.C.
ramirez@opttek.com

Benjamin Thengvall
OptTek Systems, Inc.
Boulder, CO
thengvall@opttek.com

INTRODUCTION

Simulation and optimization are two powerful methodologies that are both widely used across the Department of Defense (DoD) (see Boginski, et al. (2015); Dirik, et al. (2015); Kannon, et al. (2015); and Hill, et al. (2001)). Simulations are used to understand complex system behavior at multiple levels of fidelity. For example, we can use simulation to perform detailed engineering, design, and testing of individual weapon system components; examine the interaction of components in a single advanced weapon system such as a fighter aircraft or nuclear submarine; and analyze tactical engagements between weapons systems.

Optimization provides a powerful way to determine the best option among many, sometimes infinitely many, options. For example, military analysts use optimization to maximize the amount of fuel and munitions delivered to an area of operations using the least number of ships and cargo aircraft; determine the best allocation of dollars to minimize the risk of failure in a future war; assign military personnel to bases in a way that maximizes personal preferences and professional development; and allocate blue force weapons to red force targets in order to maximize the probability of damage while minimizing collateral damage. When resources are limited, and mission effectiveness is paramount, optimization offers military decision makers keen insights and enables them to make the best choices.

Theoretically, the issue of identifying “best options” falls within the realm of optimization. Until quite recently, however, the methods available for finding optimal decisions have been unable to cope with the complexities and uncertainties posed by many real-world problems, particularly those approached by simulation. In fact, these complexities and uncertainties are the primary reason that simulation is chosen as a basis for studying such problems. Consequently, decision makers have been faced with a “Catch 22”. Many important types of real-world optimization problems can only be treated by the use of a simulation model, but once they are submitted to simulation there are no optimization methods that can adequately organize the search for the best solutions. In short, there has not existed any type of search process capable of effectively integrating simulation and optimization. The same shortcoming is also encountered in settings outside of simulation where complex (realistic) models cannot be analyzed using traditional “closed form” optimization tools like mathematical programming.

Recent developments are changing this picture. Advances in the field of metaheuristics- the domain of optimization that incorporates artificial intelligence and analogs to physical, biological or evolutionary processes – have led to the creation of a new approach that successfully integrates simulation and optimization. As a result, analysts can get the best benefits from their simulation models.

Organizations may fail to take full advantage of their simulation models. Even though large amounts of time and money are invested creating a simulation tool and populating it with validated data, a large part of the valuable knowledge that the model may yield is generally overlooked. Simulation analysts who can access such knowledge are exceedingly valuable to their organization and become highly sought-after resources. Combining optimization and statistical analysis techniques with a simulation model is the key to unlocking this knowledge. Optimization techniques can be used to execute a simulation model many times, varying the input parameter values, to determine the best input values to achieve desired system outputs. The results of these simulation runs can then be explored with statistical techniques to better understand the system modeled by the simulation. Essential optimization and analysis questions that can be answered for simulation models by combining these techniques include:

- Optimization
    - What combinations of input parameters lead to the best and worst performance of the system?

    - What are the best tradeoffs between multiple competing objectives?
- Analysis
    - Which input parameters have the greatest influence on the system being modeled and which have the least?
    - Are there good or bad regions of the input parameter space that can be defined by a subset of input parameters with restricted ranges?
    - Are some areas of the parameter trade space more robust to parameter variation than others?

To derive the greatest benefit from a simulation model, an analyst should apply both optimization and statistical analysis techniques. Combining these techniques can provide answers to these essential questions and key insights for decision makers. More importantly, they increase the organizational return on investment from their analysts and simulation models such as providing a range of force structure capacity (size) options.

In this paper, we first summarize some of the most relevant approaches that have been developed for the purpose of optimizing simulated systems. We then concentrate on the metaheuristic black-box approach that leads the field of practical applications and provide some relevant details of how this approach has been implemented and used in commercial software. As a concrete example, we describe some of the mathematics and logic behind the OptQuest simulation optimization software engine. Lastly, we present some use cases that analysts might encounter and discuss how using a simulation optimization and analysis tool like OptDef integrated with their simulation model can lighten their workload and lead to better study results.

OPTIMIZATION AND STATISTICAL ANALYSIS IN COMMERCIAL SIMULATION PACKAGES

Over the past two decades optimization tools in commercial simulation packages have become widespread and are relatively easy to use, even if not all practitioners exploit them. Commercial simulation packages also have analysis tools that explore the variability uncovered through simulation replications (or Monte Carlo runs) for a single set of input parameters. However, the analysis of all simulation runs resulting from an optimization run is less commonly available, at least in an automated, easy to digest way.

The underlying statistical techniques discussed in this paper are not new. However, in many tools today, to perform variable sensitivity, and good and bad region analysis across simulation runs executed with different combinations of input parameter changes, analysts must use multiple tools, or perform the simulations and then piece together the results of various statistical techniques. Therefore, these types of valuable simulation analyses are done infrequently, and are often performed only by technical consultants and advanced analysts. To perform these analyses, users of discrete event simulation packages export their simulation results and then use specialized statistical tools like JMP or SPSS or write code in languages like R or Python for analysis. Users of spreadsheet-based Monte Carlo simulations have more statistical analysis tools at their disposal, but even for these analysts, gaining insights across all simulation runs is not an automated process.

The critical goals of identifying good and bad regions of a parameter trade space, and of discovering robust solutions, are sometimes pursued by more advanced analysts through generation of a response surface approximation by coupling design of experiments with simulation. This approximate response surface is then explored through various stochastic optimization techniques (Samuelson 2011). Such an approach generally relies on moving from tool to tool for the different steps in the process: generating the design of experiments, executing the simulations, and performing the stochastic optimization. This type of process has the conspicuous shortcoming of frequently oversimplifying complex response surfaces, which can entail a costly loss of valuable insights.

CLASSICAL APPROACHES FOR SIMULATION OPTIMIZATION

Fu (2002) identifies 4 main approaches for optimizing simulations:

- Stochastic approximation (gradient-based approaches)
- (Sequential) response surface methodology
- Random search
- Sample path optimization (also known as stochastic counterpart)

Stochastic approximation algorithms attempt to mimic the gradient search method used in deterministic optimization. The procedures based on this methodology must estimate the gradient of the objective function in order to determine a search direction. Stochastic approximation targets continuous variable problems because of its close relationship with steepest descent gradient search. However, this methodology has been applied to discrete problems (see e.g. Gerencser et al. 1999).

Sequential response surface methodology is based on the principle of building local metamodels. The "local response surface" is used to determine a search strategy (e.g., moving in the estimated gradient direction) and the process is repeated. In other words, the metamodels do not attempt to characterize the response surface for the entire solution space but rather concentrate in the local area that the search is currently exploring.

A random search method moves through the solution space by randomly selecting a point from the neighborhood of the current point. This implies that a neighborhood must be defined as part of developing a random search algorithm. Random search has been applied mainly to discrete problems and its appeal is based on the existence of theoretical convergence proofs. Unfortunately, these theoretical convergence results mean little in practice where it's more important to find high quality solutions within a reasonable length of time than to guarantee convergence to the optimum in an infinite number of steps.

Sample path optimization is a methodology that exploits the knowledge and experience developed for deterministic continuous optimization problems. The idea is to optimize a deterministic function that is based on n random variables, where n is the size of the sample path. In the simulation context, the method of common random numbers is used to provide the same sample path to calculate the response over different values of the input factors. Sample path optimization owes its name to the fact that the estimated optimal solution that it finds is based on a deterministic function built with one sample path obtained with a simulation model. Generally, n needs to be large for the approximating optimization problem to be close to the original optimization problem (Andradottir, 1998).

Leading commercial simulation software employs metaheuristics as the methodology of choice to provide optimization capabilities to their analysts. We explore this approach to simulation optimization in the next section.

SIMULATION OPTIMIZATION APPROACH WITH METAHEURISTICS

Metaheuristics provide a way of considerably improving the performance of simple heuristic procedures. The search strategies proposed by metaheuristic methodologies result in iterative procedures with the ability to explore solution spaces beyond the solution that result from applying a simple heuristic. Genetic algorithms (GAs) and scatter search (SS), for example, are population-based metaheuristics designed to operate on a set of solutions that is maintained from iteration to iteration. On the other hand, metaheuristics such as simulated annealing (SA) and tabu search (TS) typically maintain only one solution by applying mechanisms to transform the current solution into a new one. Metaheuristics have been developed to solve complex optimization problems in many areas, with combinatorial optimization being one of the most fruitful. Very efficient procedures can be achieved by relying on context information, that is, by taking advantage of specific information about the problem. The solution approach may be viewed as the result of adapting metaheuristic strategies to specific optimization problems. In these cases, there is no separation between the solution procedure and the model that represents the complex system.

We can use metaheuristics to create solution procedures that are context independent, that is, procedures capable of tackling several problem classes and that do not use specific information from the problem to customize the search. The original genetic algorithm designs were based on this paradigm, where solutions to all problems were represented as a string of zeros and ones (Katoch, 2021). The advantage of this design is that the same solver can be used to solve a wide variety of problems, because the solver uses strategies to manipulate the string of zeros and ones and a *decoder* is used to translate the string into a solution to the problem under consideration. The obvious disadvantage is that the solutions found by context-independent solvers might be inferior to those of specialized procedures when applying the same amount of computer effort (e.g., search time). We refer to solvers that do not use context information as general-purpose or "black box" optimizers.

Figure 1. shows the black-box approach to simulation optimization favored by procedures based on metaheuristic methodology. In this approach, the metaheuristic optimizer (labeled as "optimization") chooses a set of values for the

input parameters (i.e. factors or decision variables) and uses the responses generated by the simulation model or instance to make decisions regarding the selection of the next trial solution.

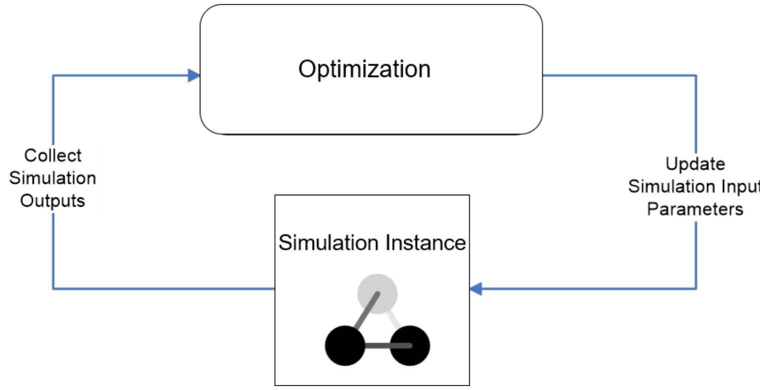

Figure 1: Black box approach to simulation optimization

One of the main design considerations when developing a general-purpose optimizer is the solution representation to be employed. The solution representation is used to establish the communication between the optimizer and the simulation (which is in fact the abstraction of the complex system). As mentioned above, classical genetic algorithms used binary strings to represent solutions. This representation is not particularly convenient in some instances, for instance, when a natural solution representation is a sequence of numbers, as in the case of permutation problems. One of the most flexible representations is an n-dimensional vector, where each component may be a continuous or integer bounded variable. This representation can be used in a wide range of applications, which include all those problems that can be formulated as mathematical programs.

Scatter Search

Scatter search is a population-based metaheuristic for optimization. It has been applied to problems with continuous and discrete variables and with one or multiple objectives. The success of scatter search as an optimization technique is well documented in a constantly growing number of journal articles, book chapters (Glover, Laguna and Martí 2003a, 2003b, 2004; Laguna 2002) and a book (Laguna and Martí 2003). Scatter search consists of five methods:

- Diversification Generation
- Improvement
- Reference Set Update
- Subset Generation
- Solution Combination

The diversification generation method is used to generate a set of diverse solutions that are the basis for initializing the search. The improvement method transforms solutions with the goal of improving quality (typically measured by the objective function value) or feasibility (typically measured by some degree of constraint violation). The reference set update method refers to the process of building and maintaining a set of solutions that are combined in the main iterative loop of any scatter search implementation. The subset generation method produces subsets of reference solutions which become the input to the combination method. The solution combination method uses the output from the subset generation method to create new trial solutions. New trial solutions are the results of combining, typically two but possibly more, reference solutions.

Extensions of the basic SS framework can be created to take advantage of additional metaheuristic search strategies, such as the memory-based structures of tabu search. There are significant differences between classical genetic algorithms and scatter search. While classical GAs rely heavily on randomization and a somewhat limiting operations to create new solutions (e.g., one-point crossover on binary strings), scatter search employs strategic choices and memory along with structured combinations of solutions to create new solutions. Scatter search explicitly encourages the use of additional heuristics to process selected reference points in search for improved solutions. This is especially advantageous in settings where heuristics that exploit the problem structure can either be developed or are already available.

OPTQUEST ENGINE

OptQuest is a commercial simulation optimization engine. Scatter search was the first optimization technology implemented within OptQuest. Over the years, however, OptQuest evolved to include a composite of many technologies both for prediction and for optimization. Engine options and settings may be used to choose technologies and change the behavior of the solution process in order to tackle difficult problems. Since OptQuest is built under the assumption that a computationally expensive black box is used to evaluate the objective function of the optimization problem being solved, prediction models within the engine have the dual purpose of assisting in establishing search directions as well as estimating the value of the objective function before solutions are processed by the objective function evaluator.

Prediction Technologies

The OptQuest Engine includes a *multivariate linear regression* module that is employed for several purposes. Once the optimization process begins and a number of solutions have been evaluated, the multivariate linear regression module is executed with the purpose of evaluating the linearity of the unknown objective function (represented by the black box). If a linear approximation is fairly accurate, then the module can be used to filter out trial solutions that otherwise would be submitted to the black box evaluator and to predict good solutions to evaluate.

OptQuest contains a *neural network* module that can be used to “screen out” trial solutions that are predicted to have inferior objective function values when compared to the best-known objective function value found throughout the search. The neural network can also be employed as a prediction model to help the system accelerate the search by suggesting solutions for evaluation that may be of high quality. The multi-layer neural network is trained during the search process based on a set of already evaluated solutions.

OptQuest Capabilities and Practical Implications

OptQuest replaces manual trial-and-error or simple parametric search schemes with a potent search engine that can pinpoint the best decisions that fall within the domain that the simulation or other evaluation model encompasses. Many defense simulation packages give the decision-maker no help in identifying good alternatives to evaluate. More importantly, they offer no guidance or insight into the nature of alternatives that can yield the best decisions. Common optimization questions defense analysts ask include:

- What is the most effective raid configuration to maximize the number of successful engagements?
- What is the best logistics posture to ensure successful and fast delivery of equipment?
- What is the most cost-effective inventory policy?
- What is the best workforce allocation?
- What is the most productive mission operating schedule?
- How does one minimize cost and maximize specific equipment usage?

The answers to such questions require a painstaking examination of multiple scenarios, where each scenario in turn requires the implementation of an appropriate simulation or evaluation model to determine the consequences for costs, performance, and risks. The critical "missing component" is to disclose which decision scenarios are the ones that should be investigated -- and still more completely, to identify good scenarios automatically by a search process designed to find the best set of decisions. The OptQuest engine automates this search for the best solutions. OptQuest enables the decision maker to specify a variety of important relationships to control the determination of optimal decisions, such as:

- Ranges of key parameters
- Budget limitations
- Asset capacities
- Acceptable minimum and maximum output values

- Limits on resources used
- Links between components or subsystems

OptQuest then determines the strategic options that are investigated under its guidance, and which it successively passes to the simulation or technical model for evaluation. The resulting search isolates scenarios that yield the highest quality outcomes for provided objectives, according to the criteria selected by the decision maker.

USE CASE EXAMPLES

OptDef is a simulation optimization and analysis tool that embeds the OptQuest engine and provides an analyst focused user interface that can be used by defense analysts. OptDef can be used with an appropriate defense simulation model in a variety of real-world applications, such as optimizing air defense configurations, maximizing satellite coverage, perform cyber security vulnerability assessments, and launch and deployment optimization.

- Optimal Blue Response: Figure 2 shows a notional Advanced Framework for Simulation, Integration, and Modeling (AFSIM) scenario. Incoming blue forces attempt to hit all red targets. In this case, an analyst's objective would be to maximize the number of hits while minimizing the number of blue aircraft used. Using the least number of blue forces has the added benefit of saving on cost while still achieving the mission objective. A typical optimization setup would include varying parameters such as weapon type (categorical variable type), number of weapons (integer variable type), and the amount of time each aircraft has on a target (integer variable type). Running the simulation optimization software utilizes the metaheuristic methodologies described earlier to explore the space and find the optimal response (i.e. reach the analyst's objective).

Figure 2. Notional AFSIM scenario

- Maximal Satellite Coverage: An analyst may need to optimize satellite target coverage (i.e. swath). In this case, the objective could be finding the optimal satellite configuration to get the best coverage for high priority targets/areas. An optimization setup could include varying spacecraft orbital parameters and system configuration (e.g. varying orbits, number and type of spacecraft) to reach the objective. The analyst can also specify multiple objectives that would balance the best satellite configuration with cost (perhaps an additional variable would be fuel/energy amount). Utilizing multiple objectives allows the analyst to make data driven decisions that best meet mission needs.
- Cyber Security Optimization and Analysis: An analyst may want to test the limits of their IT system by conducting vulnerability assessments. Unlike the Figure 2 scenario, which has blue forces on the offensive, in the cyber security realm we can focus on a defensive posture. Variable parameters may include number and type of servers that are part of the system architecture. It may also include known speed of response to a detected threat. If there is a cyber-attack on the system, the analyst can optimize the solution to minimize loss of function and duration of effect against it. With simulation optimization, the analyst can explore scenarios in the cyber kill chain that are the most detrimental and identify key components that must be protected at all costs.

CONCLUSION

We have introduced the key concepts associated with the area of optimizing simulations, starting with the approaches that researchers have investigated for many years. We discussed the metaheuristic approach to simulation optimization, which is the approach widely used in commercial applications. Key aspects that are relevant to implementation, in addition to the fundamental strategies underlying the metaheuristic, are the solution representation, the use of metamodels and the formulation of constraints.

We explained the advantage of combining simulation and optimization. The level of performance achieved by the solutions found with optimization far exceeds that to be derived from a manual what-if analysis because of the overwhelmingly large number of possible scenarios that must be considered. We provided an overview of the commercial OptQuest simulation optimization engine as well as examples of defense-related problems that can be solved pairing simulation tools with OptDef.

There is still much to learn and discover about how to optimize simulated systems both from the theoretical and the practical points of view. The rich variety of practical applications and the dramatic gains already achieved by simulation optimization ensure that this area will provide an intensive focus for study and a growing source of practical advances in the future. Simulation optimization software and tools, such as OptQuest and OptDef, can provide great support to the analyst as they explore their data and achieve project/mission objectives.

REFERENCES

Andradottir, S. (1998). A Review of Simulation Optimization Techniques. *Proceedings of the 1998 Winter Simulation Conference*, D.J. Medeiros, E.F. Watson, J.S. Carson and M.S. Manivannan (eds.), 151-158.

Boginski, V., Pasiliao, E. L., & Shen, S. (2015). Special issue on optimization in military applications. *Optimization Letters, 9(8),* 1475-1476.

Dirik, N., Hall, S. N., & Moore, J. T. (2015). Maximizing strike aircraft planning efficiency for a given class of ground targets. *Optimization Letters*, *9*(8), 1729-1748.

Fu, M. (2002). Optimization for Simulation: Theory and Practice. *INFORMS Journal on Computing, 14(3),* 192-215.

L. Gerencser, S. D. Hill and Z. Vago. Optimization over discrete sets via SPSA. *Proceedings of the 38th IEEE Conference on Decision and Control*, Phoenix, AZ, USA, 1999, *2*, 1791-1795.

Glover, F., M. Laguna and R. Martí (2003a) "Scatter Search," in *Advances in Evolutionary Computation: Theory and Applications*, New York, NY, USA: Springer-Verlag.

Glover, F., M. Laguna and R. Martí (2003b) "Scatter Search and Path Relinking: Advances and Applications," in *Handbook of Metaheuristics*, Boston, MA, USA: Kluwer Academic Publishers.

Glover, F., M. Laguna and R. Martí (2004) "New Ideas and Applications of Scatter Search and Path Relinking," in *New Optimization Techniques in Engineering*, Berlin, Germany: Springer.

Hill, R. R., Miller, J. O., & McIntyre, G. A. (2001). Simulation analysis: applications of discrete event simulation modeling to military problems. *Winter Simulation Conference Proceedings*. Arlington, VA, USA: IEEE Computer Society.

Kannon, T. E., Nurre, S. G., Lunday, B. J., & Hill, R. R. (2015). The aircraft routing problem with refueling. *Optimization Letters, 9*(8), 1609-1624.

Katoch, S., Chauhan, S.S. & Kumar, V. (2021). A review on genetic algorithm: past, present, and future. *Multimed Tools Appl (80)*, 8091–8126. https://doi.org/10.1007/s11042-020-10139-6

Laguna, M. and R. Martí (2003) *Scatter Search: Methodology and Implementations in C*, Boston, MA, USA: Kluwer Academic Publishers:

Samuelson, D. 2010. "When Close is Better than Optimal." *ORMS-Today* 37(6):144-152.