Proceedings of the 2014 Winter Simulation Conference
A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, eds.

ENHANCING THE ANALYTIC UTILITY OF THE SYNTHETIC THEATER OPERATIONS RESEARCH MODEL (STORM)

Mary L. McDonald
Stephen C. Upton
Christian N. Seymour
Thomas W. Lucas
Susan M. Sanchez
Paul J. Sanchez

Operations Research Department
Naval Postgraduate School
Monterey, CA 93943, USA

Harrison C. Schramm
Jerry R. Smith

Assessment Division (OPNAV N81)
Navy Headquarters Staff
Washington, DC 20350, USA

EXTENDED ABSTRACT

The Assessment Division, Navy Headquarters Staff (OPNAV N81) uses large-scale simulation to analyze how budgeted capabilities and capacities map to risk in various scenarios. The Navy, along with the Air Force and Marine Corps, use the Synthetic Theater Operations Research Model (STORM) to assess risk in an integrated, campaign setting. Ultimately, analyses performed in STORM inform the decisions made by the Services for future resource planning. STORM is a large, stochastic campaign-level simulation that requires many inputs for a given scenario and generates an enormous amount of output data, which then needs to be turned into an analysis product. We are developing tools and methods that: (1) reduce the amount of manpower and time required to complete STORM output post-processing, (2) determine, in a sequential dynamic manner, a sufficient number of replications to perform, (3) support STORM verification and validation, and (4) boost the speed and precision with which analysts are able to gather insights from a set of simulation runs.

A current impediment to fast and efficient use of STORM is the sheer volume of data it generates. There are many objects and events in a campaign, and STORM output can include a complete trace of the model state over the simulated campaign. Consequently, STORM routinely creates gigabytes of output from a single replication. When multiple replications are required, even more data is generated. Moreover, this output is typically not in a form that can immediately be used for analysis. Thus, some type of post-processing, e.g., filtering and transformation, is required to produce a reduced set of data that is suitable for subsequent analysis. This reduced set may still task computational resources, e.g., memory and disk space, so other techniques may be needed, such as dynamic processing of streamed output, in order to successfully conduct a full analysis of the data. One component of our research is to determine how STORM post-processing can best be improved. The research involves identifying potential data generation and storage efficiencies, automating post-processing tasks, making use of distributed computation where possible, and reducing manpower requirements when using STORM. An additional benefit is the ability to accommodate larger run sizes than are currently feasible.

Since STORM is stochastic, a determination must be made as to the number of independent replications to perform for each set of inputs. Replication allows analysts to better estimate output measures (e.g., blue systems lost), evaluate the variance of responses, and determine the distributions of outcomes (Lucas 2000). As more replications are made, these estimates become more precise. In addition, taking more replications increases our statistical power in detecting alternatives and increases the chances of identifying rare, but perhaps critical, events. The number of runs required depends on the variability of the response and the statistical power desired (Law 2007). Since the variability of the response is usually unknown prior to running the experiments, ideally, the number of runs taken should be determined dynamically.

For the reasons alluded to above, STORM runs are expensive, and current practice is for analysts to perform a predetermined, fixed number of replications for a given set of inputs. Depending on the characteristics of the scenarios being modeled, this may or may not be a sufficient sample size. In scenarios that have low variance and high signal strength, we may need fewer than the predetermined replications. In other scenarios, the ability to take more replications could be of enormous value. One objective of our research is to dynamically (sequentially) calculate appropriate sample sizes for the metrics of interest.

Sometimes the most difficult aspect of gaining insights from a high-dimensional set of output is ‘putting it all together’ to form a coherent narrative that describes: (1) which major entities and platforms initiated key actions, (2) what happened or failed to happen, (3) when and where key combat events occurred, and, probably the most difficult to ascertain, (4) why major events or outcomes occurred or didn’t occur. Of the gigabytes of output that STORM produces, a ‘feature extraction’ process is first needed to determine the functions of the data that are most relevant and meaningful to the campaign analyst. Once key data is acquired from the raw data stream, we experiment with new methods for visualization and analysis of the simulation output data. This process supports the verification and validation process in that it can identify both ‘bugs’ in simulation code as well as unintended defects in the many combat plans that must be created by analysts as part of the scenario development process. Once the scenario has been satisfactorily, analysis and visualization techniques can be used to facilitate a quick understanding of the Who-What-When-Where-Why-How of a simulation event stream, and help analysts assess which scenario variations may be the most fruitful for further study with a focused subset of excursions. In our presentation, we will summarize work conducted to date and show examples of the analysis and visualization techniques that have been developed.

REFERENCES

Kleijnen, J. P. C., S. M. Sanchez, T. W. Lucas, and T. M. Cioppa. 2005. “A User’s Guide to the Brave New World of Designing Simulation Experiments.” *INFORMS Journal on Computing* 17(3): 263-289.

Law, A. M. 2007. *Simulation Modeling and Analysis*. 4th ed. New York, NY: McGraw-Hill.

Lucas, T. W. 2000. “The Stochastic Versus Deterministic Argument for Combat Simulations: Tales of When the Average Won’t Do!” *Military Operations Research: Special Issue on Warfare Analysis and Complexity—the New Sciences*, 5(3): 9-28.

Lucas, T. W., and S. M. Sanchez. 2006. “The Brave New World of Designing Simulation Experiments for Defense and Homeland Security Applications.” In *2006 Joint Statistical Meetings Proceedings*, 1212-1218.

Sanchez, S. M., T. W. Lucas, P. J. Sanchez, C. J. Nannini, and H. Wong. 2012. “Designs for Large-Scale Simulation Experiments with Applications to Defense and Homeland Security.” In *Design and Analysis of Experiments, Volume III,* edited by K. Hinkelmann, 413-441. Hoboken, NJ: Wiley.