Test and Evaluation of Complex Systems of Systems

Ann G. Arnold and William F. Kujawa

Introduction

The engineering of complex systems of systems (SOS) has been an area of attention recently. The SOS are typically comprised of several independently procured systems, where system couplings are interdependent and the combined operation of the systems represents fulfilling the overall objective. These large, complex SOS are rarely developed under a single architecture, resulting from a strategic development decision. Component systems are developed one by one, and the overall SOS evolves over time as various visions of how each individual system meets the overall requirements evolves. Although each system is designed and developed to meet a specific functional or performance need, its requirements and design most likely did not develop in response to concerns for the SOS objectives. The methodology presented addresses the identification of SOS requirements and the decomposition of these requirements into Measures of Effectiveness (MOEs), which are quantifiable and measurable, and Measures of Performance (MOPs) for each component system.

The methodology to be presented here was developed for assessing system-of-systems performance for the Joint Countermine (JCM) Advanced Concept Technology Demonstration (ACTD), which was comprised of 13 pre-production (novel) systems designed to support the overall military operation. The vision of the JCM ACTD was to address the current capability gaps in the critical sea-to-land transition zone necessary to enable seamless movement during an amphibious operation and subsequent movement to contact with the enemy. The technologies and systems were selected because they were judged to 1) fulfill a critical countermine need, and 2) they were technologically mature enough so that their acquisition could be accelerated were they judged by the warfighter to provide exceptional military utility. The approach that was utilized in creating a framework for a complex system-of-systems assessment blended three key elements:

- Data analysis and analyst observation of live systems play in a representative environment

- Structured collection of users' impressions regarding the suitability, logistics, reliability of the systems, and

- Modeling and simulation to extrapolate results beyond conditions achievable in live exercises.

The assessment approach has a significant impact throughout the planning process, impacting the exercise design, data collection, instrumentation, and the modeling and simulation component.

One reason for a system-of-systems test is to observe the interactions among the legacy systems and the novel systems. Two systems operating together might demonstrate more capability than would be expected if each were tested separately. On the other hand, two systems that perform satisfactorily in isolation might interfere with each other, or be redundant, when operated together.

A three-tier approach is suggested to develop quantitative measures for this effort. The high-level system-of-systems requirements for overall mission success are first defined. These are then decomposed into quantifiable and measurable Measures of Effectiveness (MOEs). The MOEs relate to each function of the overall mission and should be doctrinally sound and intuitively clear. Measures of Performance (MOPs) are then defined for each of the component systems. Figure 1 illustrates the three levels of quantifiable measures for the Joint Countermine ACTD.

Two important points need to be made regarding these measures. First, they are readily calculated with data that are easily collected. Secondly, there were no predefined thresholds accompanying any of the MOEs or MOPs. For military exercises, success often depends on making an acquisition decision based on properly characterized performance, leading to an understanding of how a system will enhance the system utility.

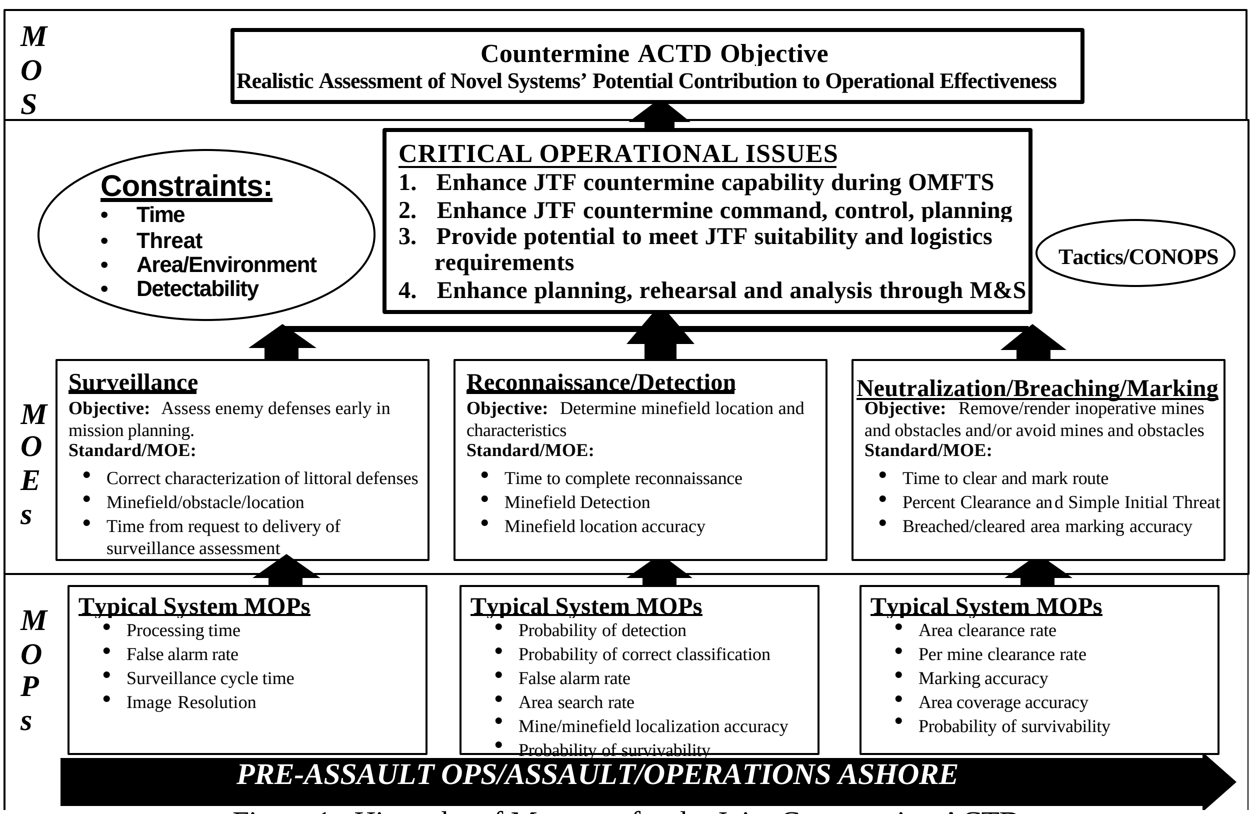

Figure 1. Hierarchy of Measures for the Joint Countermine ACTD

The overall test design is the first step in producing a context for the evaluation of the utility of these systems. The exercise design concept must:

- Naturally motivate the use of the system
- Provide the maximum opportunity to demonstrate measurable utility of each system to the top-level MOEs and system-of-systems requirements
- Demonstrate synergy of the new technologies with existing systems
- Present a significant but fair challenge to each system.

The next step is to further divide the exercise into functional subphases that are amenable to analysis. This provides self-contained operations that can be simulated if necessary as well as executed in the test to produce meaningful results. The results then support the users' ultimate evaluation of the capability of the system.

An important part of this entire process is to define the data requirements necessary to calculate the system MOPs and subphase MOEs. The high-level system requirements that are applicable to the overall mission generate several MOEs that are applicable to each subphase of the exercise. Each MOE generates several MOPs, applicable to each system participating in each subphase. Each MOP (and MOE) generates particular data requirements. All data requirements must be feasible from an engineering standpoint. Instrumentation of the participating systems is necessary to record the required data. If connectivity among the systems is required to maintain an accurate and timely picture of the mission, some of the more complex and costly instrumentation issues may be solved. In addition, it may be important in some situations to establish ground truth for various phases of the exercise; for example, the performance of some systems may depend on environmental factors.

In addition to the quantitative data, another important source of data on the performance of the participating systems is the opinions and impressions of the operators, those running the operation, and special observers and subject matter experts. Special observers and data collectors should be stationed with each participating system and at the central administration to conduct interviews and administer questionnaires.

General Analysis Approach

The analysis methodology relies on comparing the difference between real, observed performance of the systems in an exercise scenario, with the expected performance based on projections of system performance parameters via analytic calculation with support from simulations as appropriate. This technique compensates for not having a large body of representative test performance data. Instead, consistency is tested between one or two samples of observed performance and the expected performance based on the best available information.

In order to provide a basis for subsequent evaluation of the results, it is necessary to understand as much as possible about the performance of the individual systems before the exercise takes place. This understanding can come from tests conducted by the system developers or through modeling and simulation.

Prior knowledge of system performance allows the evaluator to predict how the outcome of mission scenarios would benefit from the contribution of each system in the system of systems. After the exercise, the observed performance will be compared against the predictions. Two outcomes can result:

- The observed performance during the demonstration is consistent with expectations based on modeling and simulation or other analyses, thereby indicating that the contributions of the system to the mission are understood.

- The observed performance is inconsistent with expectations, in which case, additional analysis must be accomplished to determine whether the demonstrated performance is anomalous. If not, then the effectiveness of the system in the overall mission must be reassessed before attempting to form an estimate of its utility, which may still be a significant enhancement over current capabilities.

This approach assumes that there are sufficient performance data on all participating systems to predict the success of specific concepts of operation in the scenario. The process by which baseline information is factored into the evaluation process is illustrated in Figure 2.

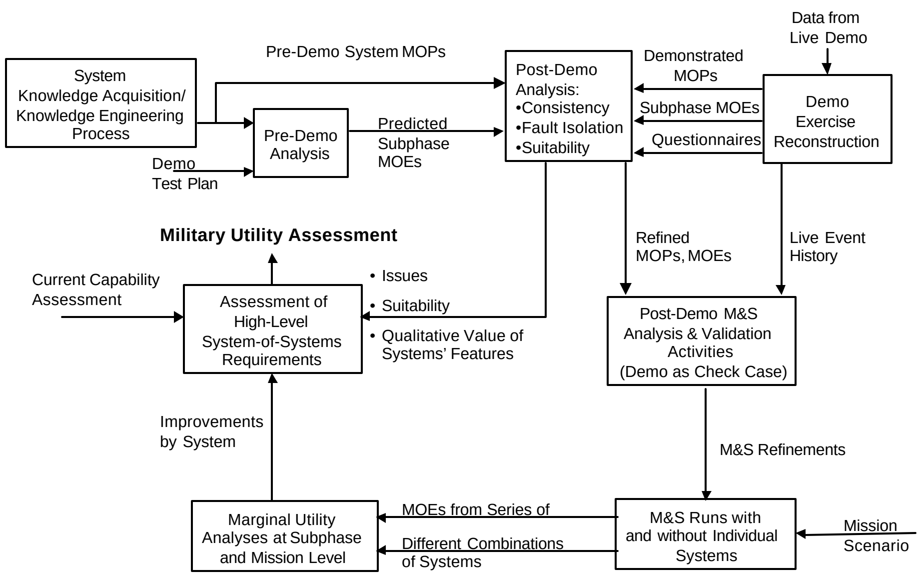

Figure 2. Assessment of Military Utility for the Joint Countermine ACTD

For some systems, considerable data are available to estimate system MOPs without further testing and analysis. For these systems, the utility of the exercise is to provide an estimate of their contribution to the MOEs and overall system-of-systems requirements in context of the mission. For other systems, there is a need to collect more data, run simulations, or otherwise improve the understanding of its relevant performance factors. The decisions and processes summarized in Figure 2 are intended to produce refined estimates of each system's MOPs and MOE contributions in related exercises so that the best possible estimates of predicted MOEs and mission-level results can be compared to observed results during the exercises, thereby contributing to an assessment of their utility.

Prior to any large-scale system-of-systems demonstration, estimates of system MOPs are generally based primarily on engineering specifications, laboratory tests, and field demonstrations. MOEs involving one or more systems may be determined from these MOPs, simple probabilistic models, and planning tools that incorporate the best estimates of system MOPs into characterizations of small segments of the large-scale demonstration. Assessment of the top-level requirements, however, requires the use of complex model and simulation systems to examine the relative effectiveness of the many permutations.

After the demonstration is conducted, the model can be applied to the exercise scenario as actually executed using observed conditions to determine whether the observed results can be considered to have been drawn from the population of expected outcomes based on simulation. The observed data are then used to refine the system models. The outputs of these models and simulations provide a basis for evaluating the utility of the novel systems.

Two aspects of system performance are illustrated in Figure 2. One aspect is functional performance. This type of performance is the one most amenable to a quantitative analysis by comparing demonstration results to expectations from models and simulations. The other aspect of performance is related to Reliability, Availability, Maintainability, and Durability (RAMD), suitability for field operations, and other factors not easily measured in the narrow confines of one or two large-scale exercises. The observations of participants and exercise monitors will provide a source for qualitative assessment of these factors. Adjustments to assessments of the utility based on knowledgeable predictions of the effect of future modifications on system performance can be based on the models developed via the process shown. Figure 2 allows for various reasons for unexpected performance, including problems with the predictions, unexpected changes in the scenario or environment, and/or system malfunction.

Conclusion

In order to determine whether the goals and objectives of a mission involving multiple systems are met, it is important that a robust assessment be conducted. The analysis methodology presented here and developed for complex military exercises creates the framework for providing the information needed to support investment decisions on a broad spectrum of individual technologies. The exercise data, when combined with simulations and other test data, provided a realistic assessment of the performance of the systems operating together in a representative scenario.

Acknowledgement:

This effort was performed for CAPT R. T. Schnoor and COL T. J. Singleton (Ret), Code 32CM, of the Office of Naval Research for the Joint Countermine Advanced Concept Technology Demonstration Program.