Building Trustworthy Systems: Guided State Estimation as a Feasible Approach for Interpretation, Decision and Action Based on Sensor Data

John James*, *Member, IEEE* Frank Mabry

***Abstract* – Sensor management plays a key role in control of critical infrastructure systems. This paper describes an approach for improving capabilities for *interpretation, decision, and action* based on sensor data through application of an *intermediate level of aggregation*. Improvements in complex system understanding are needed now at the interface between *human understanding* of system state and *machine understanding* of system state. The human understanding of the state of the system (situation understanding) must be achieved under ever more demanding time constraints. As expectations increase for faster, more-informed (better) decisions by humans at the supervisory-control level, improvements are needed for providing support for interpreting sensor data to understand current system behaviors and make informed human decisions on actions needed to cause future system behaviors to comply with some planned sequence of events or patterns of behavior. Likewise, as the number and capabilities of networked sensors increase, improvements are needed in enabling autonomous control systems at local levels to understand current system behaviors and make informed machine decisions on actions needed to cause future system behaviors to comply with some planned sequence of events or patterns of behavior. The paper discusses achieving an intermediate level of aggregation: (1) as a scientific basis for understanding complex system behaviors, (2) as an effective tool for creation of technologies based an intermediate level of aggregation and (3) as a basis for education of leaders who must make decisions based on understanding of the current system state.**

Index terms – trustworthy systems, sensor fusion, data mining, decision support

I. Introduction

Sensor management plays a key role in control of critical infrastructure systems. *This paper describes an approach for improving capabilities for* interpretation, decision, and action *based on sensor data through application of an* intermediate level of aggregation.

The idea of a guided state estimation as an approach for supporting human-in-the-loop situation understanding and decision processes has a firm foundation in the military community. Military operations are distributed in time and space and executed at multiple echelons that are used to partition operations into multiple levels which operate at progressively slower time scales over progressively larger spatial scales. Automated sensors and automated sharing of sensor data has been achieved in the 4th Infantry Division of the U. S. Army and will be gradually achieved in all Army units. Synchronization matrices are used at many echelons to summarize interrelationships between loosely coupled (and tightly coupled) components of an operation and also as a useful tool for monitoring execution of a planned sequence of events requiring cooperation among the components.

This paper discusses some initial experiments in building information systems suitable for construction of intermediate levels of aggregation of operational information. Section II describes how this information is obtained, analyzed and displayed, Section III describes how data transformations are used to maintain trust in these intermediate levels of aggregation, Section IV discusses implications of this work, and Section V summarizes the paper.

II. Estimating Systems State Using Intermediate Levels of Aggregation

Say there is some metric for determining degree of attainment of system Purpose:

- Completely attained
- More than Adequately Attained
- Adequately Attained
- Less than Adequately Attained
- Minimally attained.

Then, to the degree that measures are available to indicate *closeness to achieving system purpose* and also that measures are available for estimating the relative *contribution* that elements in a knowledge *value chain* make to achieve the system purpose, then a cost-based allocation of resources can be made to protect, in priority, those assets which contribute the most to completion of enterprise purpose.

Figure 1 provides an overview of the Area of Operations used at the United States Military Academy for Operation Highland Warrior. Figure 2 is an extract from the North West corner of Figure 1 and represents an area appropriate for training a platoon-sized (about 30 soldiers comprised of four squads) element. A variety of platoon-level tasks are tested during Operation Highland Warrior. These include Movement to Contact, Hasty Defense, and

*** This work was partially supported by an endowment establishing the Adam Chair in Information Technology. The views expressed herein are those of the authors and do not purport to reflect the position of the United States Military Academy, the Department of the Army, or the Department of Defense.**

Deliberate Defense. The U. S. Army has taken each of the higher-level unit tasks and broken them down into detailed sub-tasks, conditions under which the sub-tasks must be performed, and standards to which the tasks must be completed.

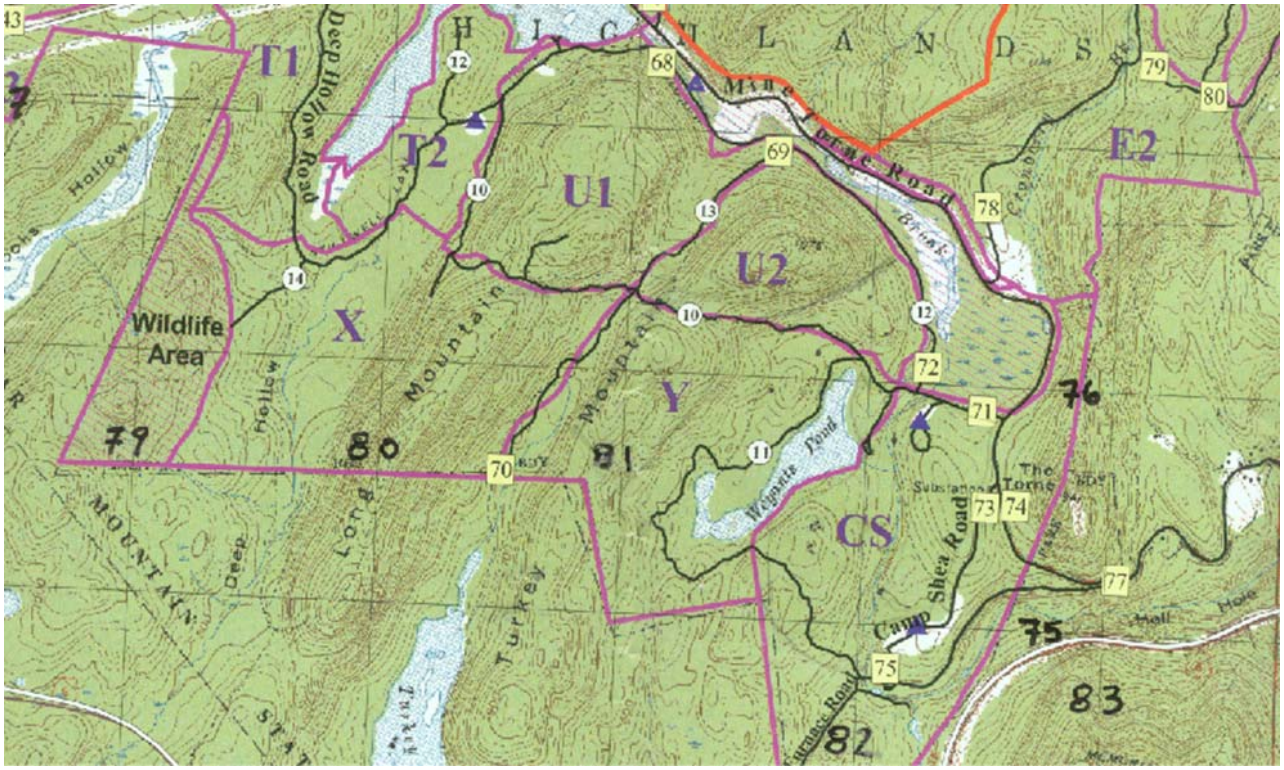

Figure 1. The Area of Operations for Operation Highland Warrior

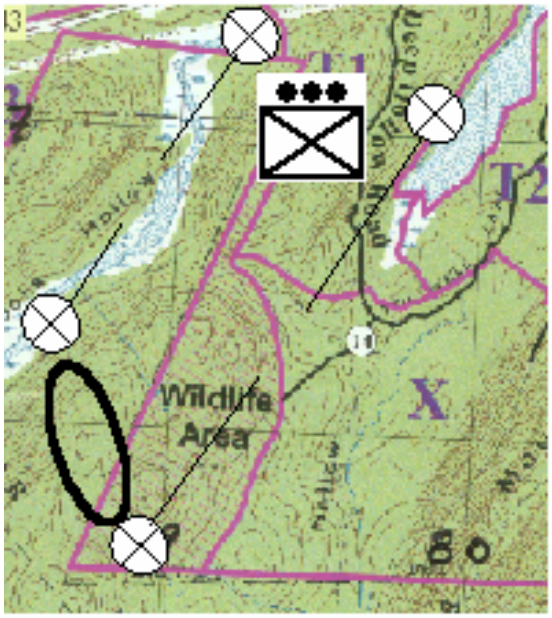

Figure 2. Platoon-Sized Area of Operations

Our initial effort has focused on building a capability for capturing data concerning platoon-level operations that can then be aggregated to successively higher echelons (i.e. company, battalion, brigade, and division echelons respectively).

The Personal Digital Assistant (PDA) display of squad and platoon Task-Condition-Standard information for a deliberate defense is shown in Figure 3. During Operation Highland Warrior, the platoons and squads are evaluated by experienced officers concerning their ability to execute the required tasks to the required standards under the expected conditions.

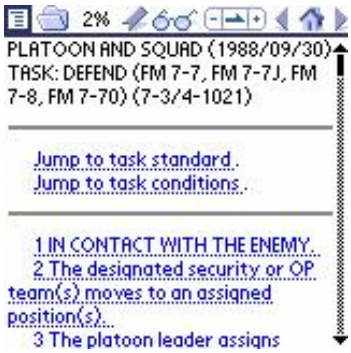

Figure 3. Defensive Tasks, Conditions and Standards

In order to estimate the *state* of the platoon in completing a task we have created a data structure that is amenable to automatic collection and dissemination. The eXtensible Markup Language (XML) representation of a contact report is given in Figure 4.

```
<contact>
      <unit>1plt </unit>
      <time>09:36 </time>
      <loc>824755 </loc>
      <ENsize>0   </ENsize>
      <blKIA>2    </blKIA>
      <blWIA>1    </blWIA>
      <rdKIA>0    </rdKIA>
      <rdWIA>0    </rdWIA>
      <fires>no   </fires>
      <comments></comments>
</contact>
```

Figure 4. A contact report

The PDA screen summarizing a set of contact reports (giving the time of the contact, the unit reporting the contact, the *location* of the unit, the *size* of the enemy, the *blue* killed in action (KIA), and the *red* KIA) is shown in Figure 5.

Figure 5. Display of a set of contacts

The companies (about 100 soldiers) of an Army Battalion (about 500 soldiers) use available fixed and mobile command, control, and communications equipment to create voice and data networks. The newer data networks support automatic sharing of **situational awareness** data as well as implementation of required analog and digital communication networks. One top-level partitioning of information system components is into two sets: one set for those sub-systems associated with *administration and logistics* and one set for those sub-systems associated with *force-level control* (command and control).

As the time for commencing an operation draws close, those Battalion-level information assets that allow commanders and staffs to understand the current locations and activities of friendly and enemy forces (i.e. the *intelligence estimation* assets of force-level control) will have a relatively high priority.

Once an operation begins, those Battalion-level information systems that enable *force level control* functions will have a relatively high priority. The force-level control functions are those that position the company (15 tanks) and platoon (four tanks) elements for application of deadly force as well as those systems that coordinate requests for supporting fire. Deadly force is applied by the combat-crew (tank) level and by supporting fire elements (mortars, artillery, aircraft, …). The Army uses a synchronization matrix to summarize the activities required by different force structure elements during different phases of an operation. The *synchronization matrix* provides a means for constructing metrics to estimate whether subordinate units of a given unit have met time and spatial constraints for achieving a ***commander's intent***. Thus, by phase and unit by echelon, we can estimate if goals are being: completely attained, more than adequately attained, adequately attained, less than adequately attained, or minimally attained.

The joint force information presented in different contexts to different individuals should address the needs of the user. This is particularly true in the case of engagement decisions where the different views of the common operational picture should reflect the fact that engagement decisions are made primarily by officers in the Air Force and Navy and primarily by combat weapons crews in the Army and Marine Corps. Estimates of the relative importance of different information system elements will require ***on-line identification of system state*** since the information system architecture (like the force structure it supports) will change as an operation proceeds. Changes will occur at the network level, at the middleware level, and at the application level.

The information system supporting joint force operations should explicitly support building and maintaining a *belief support* system that provides an estimate of the

trustworthiness of the information on (1) friendly forces, (2) enemy forces, and (3) unit activity. Given that during an operation the only *system invariant* may be the commander's intent, the belief support system should be based upon maintaining the highest degree of trust in that information most critical to achieving the commander's intent. A reasonable place to start is in maintaining trustworthiness in the data enabling estimation of unit state as it relates to the *synchronization matrix.*

The Department of Defense is engaged in a massive effort to transform the services to exploit the leverage possible with more effective use of information system technologies. Joint operational doctrine is being developed to provide concepts for applying that leverage to win battles. It may be that the idea of the *synchronization matrix* will be updated by some other mechanism to achieve joint force integration, perhaps some form of "just-in-time" collaboration to dynamically readjust unit missions. However, it will remain the case that some form of allocation of combat functions to combat units over time will be used to organize effort to have the joint force operate in harmony to meet the commander's intent. Whatever the approach for allocation of unit to functions over time, the supporting information system needs to be structured to establish and maintain an estimate of the trustworthiness of the information being presented. Thus, it is expected that some form of *intermediate level aggregation* will be consistent with execution of any operation, whether a synchronization matrix is explicitly prepared or not.

We are convinced that some variation of the ***hybrid automaton*** ideas of hybrid control theory will be needed for modeling the system *state* since this approach is the most general and also provides a rigorous way of restricting the set of hybrid trajectories to a collection of discrete and continuous variables. The general approach is mathematically rigorous and, at some point, may support automatic *generation* of system of systems solutions. However, we will concentrate on the *constructive* assembly of components of known models into progressively more complex systems of systems and adaptive control of the (well-understood) composed system. This approach supports the *verification and validation* [8] of our system of systems for autonomous combat agents since a necessary step in the *composition process* for our systems will be the satisfaction of *independence of components* constraints except where we explicitly allow feedback loops.

Thus the basic agent in our ***modeling and simulation framework*** is a ***hybrid automaton*** [2] that is a collection $H = (X, V, Init, f, Inv, R)$ where

X is a finite collection of state variables. We assume $X = (X_D \cup X_C)$ with X_D countable and $X_C \in \Re^n$;

V is a finite collection of input variables. We assume $V = (V_D \cup V_C)$ with V_D countable and $V_C \in \Re^n$;

$Init \subseteq X$ is a set of initial states;

$f: X \times V \to X_C$ is a vector field, assumed to be globally Lipschitz in X_C and continuous in V;

$Inv \subseteq X \times V$ is an invariant set;

$R: X \times V \to 2^X$ is a reset relation.

We refer to $x \in X$ as the state of H and to $v \in V$ as the input of H. Associated with this model are rigorous definitions of continuous and discrete states and associated models of continuous and discrete and hybrid (combination of continuous and discrete) behaviors consisting of continuous, discrete and hybrid trajectories from a set of initial states to a set of final states. The *Common Operational Picture* problem can then be characterized as the problem of accurately estimating the current system (friendly and enemy) state. Likewise, the *Command and Control* problem can then characterized as the problem of determining the combination of (friendly) inputs that will cause the hybrid automaton to move from a current state (a collection of state variables) to a desired future state, given the set of environmental and enemy inputs over time. Optimal control theory (and other technologies) can be used to apply some notions of reward and cost functions to assist in choosing among a set of possible future states.

III. Using Data Transformations to Achieve and Maintain Trust in Intermediate Levels of Aggregations

So how does one achieve and maintain "trust" in aggregations of the ever growing amount of data into intermediate levels of representation. We have been conducting experiments in building small decision support systems that enable evaluators of cadet Summer training exercises to use Personal Digital Assistants (PDAs) to record evaluations of small unit operations.

The choices for the eXtensible Markup Language (XML) data structures used for the small decision support systems and the data flow supported by the systems are governed by constraints chosen to foster *trust* that the data is relevant and believable:

1. Data collected supports the evaluation of individual, combat-crew, and unit activities,

2. Evaluations are achieved by estimating how the combat *Task*s have been completed under specified *Conditions* and to the specified *Standards*.

3. Evaluations are conducted by officers who have experience in executing the relevant *Task*s under specified *Conditions* and to the specified *Standards*.

4. The data representation and transformation(s) support maintenance of a complete record of the sequential actions taken to achieve an evaluation.

5. The data representation and transformation(s) support creation of a complete explanation of the sequential actions taken to achieve an evaluation.

6. The data representation and transformation(s) support creation of a "post-facto" proof that the data from beginning to end has been recorded in an "un-reversable" form.

These rules ensure that two conditions for credibility in defending conclusions (developed earlier during work on forensics for criminal investigations) are met:

1. personal credibility of individuals involved (honesty, expertise, experience) is achieved, and

2. testimony can be presented in a convincing, and **understandable** manner.

Achieving the solution relies on a number of developments associated with the representational revolution that XML has engendered. All substantive investigative operations in the developed system that transform data represented in XML are portrayed in XSL-T. XSL-T is itself an XML vocabulary and is used in the system to self-analyze the representation of the solution investigation system and produce from the recorded series of investigative steps an explanation of the investigation. All data imported into the system are represented in XML and their initial version is stored and includes a non-reversible "digest-key." The digest key used also includes the use of the value of the transactional digest key (which is encoded starting with the initial command execution of the system and incorporates a user's secret input and the workstation's machine-id and current time and date). Each transactional step uses the previous transaction's digest key and any new data and/or submitted data. All interactions are via web-page forms. A standalone, single user server is invoked for each web action. The web server composes an XML version of the input from the web browser and then initiates the form's indicated action (the action sequence is itself originally represented in XML and processed to develop the run time command support). Each command executed, adds to the transactional store any additional input data brought into the investigation (such as the information extracted from an email system of a suspect). Because the number of commands implemented is small (two or three) development of post-facto explanation support has been minimized. The explanation of atomic transformation steps can be done by referring to the step wise input transactions recorded by the web-server.

Having followed these guidelines allows an application environment to be established that provides a very interesting benefit: With little or no development overhead the system can provide a multi-leveled explanation of how any result has been derived. User actions, initial data, incorporated data and system function can be explained in detail. Once a form of presentation has been explained at a "micro-level" it can be subsequently elided over in "macro-level" explanations (with immediate descent into detail upon request).

This work is extensible to include the automatic update of position, location, and activity information using the data available over the automated information systems deployed with U. S. Army units. The combination of human estimates with machine- generated data can then be used to generate unit-task-completion-estimates and associated "trustworthiness" estimates based upon the associated confidence in the human and machine estimates of system state. We expect to continue this work to have cadet projects which will extend the current results to include the estimates of task completion and aggregation of platoon-state (i.e. status of a platoon in accomplishing the platoon leader's intent to defend) to company state (i.e. status of a company in accomplishing the company commander's intent to defend).

IV. Implications

As the complexity of networked systems continues to expand, scientists and engineers need better science and improved technologies for understanding and controlling complex system behaviors. Specifically, improvements are needed now at the interface between *human understanding* of **system state** and *machine understanding* of **system state**. The human understanding of the state of the system (situation understanding) must be achieved under ever more stringent time constraints. As expectations increase for faster, more-informed (better) decisions by humans at the supervisory-control level, improvements are needed for providing support for ***interpreting*** sensor data to understand current system behaviors and make informed ***human decisions*** on ***actions*** needed to cause future system behaviors to comply with some ***planned sequence of events*** or ***patterns of behavior.*** Likewise, as the number and capabilities of networked sensors increase, improvements are needed in enabling ***autonomous control systems*** at local levels to understand current system behaviors and make informed ***machine decisions*** on ***actions*** needed to cause future system behaviors to comply with some ***planned sequence of events*** or ***patterns of behavior.*** This need for increased scientific results and more effective technologies applies to all of the emerging large-

scale critical infrastructure protection systems (financial, power, telecommunications, government, security).

Trustworthiness of complex systems: Achieving an intermediate level of aggregation enables incremental construction of higher-level system abstractions: (1) as a scientific basis for understanding complex system behaviors, (2) as an effective tool for creation of technologies based an intermediate level of aggregation and (3) as a basis for education of leaders who must make decisions based on *understanding* of the current **system state**. The cadet projects are extensions of ongoing data mining efforts for understanding the *state* of command and control networks supporting military operations as well as the higher-level *state* of the execution of an operation. The need is to have better tools for estimating the *state* of command and control networks as well as ensuring the *trustworthiness* of the characterization of the *state* of the operation.

V. SUMMARY

We have discussed modeling military systems as representative of modeling other complex systems. The approach discussed for modeling IA components of military systems rests upon the notion that the system at hand is intended to achieve some useful *purpose* and that a *system of systems* approach provides a feasible methodology for composing the system as an aggregation of sub-systems. Many subsystem processes have continuous process models while higher system models are usually discrete. Composition of components requires consideration of interaction of subsystems, especially when feedback loops are present. We have described how achieving an intermediate level of aggregation enables incremental construction of higher-level system abstractions (i.e. an intermediate representation of achieving the commander's intent).

VI. REFERENCES

[1] Mettala, E. G., James, J. R., Coleman, N., Gallagher, E. J., Harris, R. L., Smith, J. G., and Graham, M. "Domain-Specific Software Architectures: Government Needs and Expectations." Presented at the IEEE Symposium on Computer-Aided Control System Design, Napa, CA 17-19 March 1992.

[2] John Lygeros, George Pappas and Shankar Sastry "An Introduction to Hybrid System Modeling, Analysis and Control" Preprints of the First Nonlinear Control Network Pedagogical School, pages 307-329, Athens, Greece, 1999.

[3] M. Kezunovic, " Data Integration and Information Exchange for Enhanced Control and Protection of Power Systems" Proceedings of the 36th Hawaii Conference on System Science, Kona, Hawaii, January, 2003.

[4] R. G. Little, " Toward more Robust Infrastructure: Observations on Improving Resilience and Reliability of Critical Systems" Proceedings of the 36th Hawaii Conference on System Science, Kona, Hawaii, January, 2003.

[5] Linda Wills, Suresh K. Kannan, Sam Sander, Murat Guler, Bonnie Heck, J.V.R Prasad, Daniel P. Schrage, and George Vachtsevanos. "An Open Platform for Reconfigurable Control." IEEE Controls Systems Magazine, 21(3), June 2001.

[6] H. Schafer, D. J. Ragsdale, J. R. Surdu, and C. A. Carver, The IWAR Range: A Laboratory for Undergraduate Information Assurance Education, " *The Journal of Computing in Small Colleges,* vol. 16, pp. 223-232, 1 November 2001.

[7] D. Welch, D. Ragsdale, and W. Schepens, "Training for Information Warfare," *IEEE Computer*, 2002.

[8] John James and Dave Barton "A Framework for Verification and Validation of Integrated and Adaptive Control Systems" Proceedings, 11th IEEE International Symposium on CACSD, Anchorage, Alaska, September, 2000.

[9] Meystel, Alexander M. and Albus, Jim S. *Intelligent Systems Architecture, Design, and Control,* Wiley Series on Intelligent Systems, John Wiley and Sons, 2001.