Scythe

Proceedings and Bulletin of the International Data Farming Community

Issue 3 - Workshop 15

Proceedings and Bulletin of the International Data Farming Community

Table of Contents

Scythe

Proceedings and Bulletin of the International Data Farming Community

It is appropriate that the publication supporting the International Data Farming Workshop is named after a farming implement. In farming, a scythe is used to clear and harvest. We hope that the "Scythe" will perform a similar role for our *data* farming community by being a tool to help prepare for our data farming efforts and harvest the results. The Scythe is provided to all attendees of the Workshops. Electronic copies may be obtained from harvest.nps.edu. Please contact the editors for additional paper copies.

Please let us know what you think of this ***third*** prototype issue. Articles, ideas for articles and material, and any commentary are always appreciated.

Bulletin Editors
Ted Meyer: tedmeyer@mac.com
Gary Horne: gehorne@nps.edu

International Data Farming Community Workshop 14 Program Committee

International Data Farming Community Overview

The International Data Farming Community is a consortium of researchers interested in the study of *Data Farming*, its methodologies, applications, tools, and evolution.

The primary venue for the Community is the biannual International Data Farming Workshops, where researchers participate in team-oriented model development, experimental design, and analysis using high performance computing resources... that is, Data Farming.

Scythe, *Proceedings and Bulletin of the International Data Farming Community,* Issue 3a, Workshop 15 Publication date: January 2008

IDFW 15: Cultivating Knowledge

by Gary Horne
Naval Postgraduate School

International Data Farming Workshop 15 (IDFW 15) took place during the six days from November 11th through the 16th, 2007 in Singapore. Sixty-six participants from seven countries worked in eight different teams exploring questions using Data Farming methods. The theme was "Cultivating Knowledge," and the goal was to use our data farming methods to continue to explore our important questions.

Personnel from DSO National Laboratories in Singapore collaborated with us at the SEED Center for Data Farming at the U.S. Naval Postgraduate School in Monterey, California in organizing the workshop. The "SEED" in our title is an acronym for Simulation Experiments and Efficient Designs.

As the executive director of the Center, it is my pleasure to work with many from around the world to develop the methods of Data Farming and apply them to important questions of our day. And on behalf of the co-directors of the SEED Center for Data Farming, Professors Tom Lucas and Susan Sanchez, I would like to express our thanks to DSO, the team leaders, the plenary speakers and all of the participants in IDFW 15.

I want to also briefly outline the work of the eight teams and invite you to examine the details of their efforts later in this issue of *The Scythe*. This issue also contains a listing of the plenary sessions and concludes with an article on Clustering and Outlier Analysis for Data Mining. As always, the plenary session materials, in-briefs, and out-briefs from the Workshop are available online at harvest.nps.edu along with electonic copies of this issue of *The Scythe*. (Attendees will find an attached CD with this material as well as a collection of photographs from the week of the workshop.)

Team 1 applied Automated Red Teaming techniques developed at DSO to examine a Maritime scenario. They also developed their work on automated co-evolution.

The objective of Team 2 was to develop civilian models within an agent-based simulation. They used Pythagoras and explored factors related to the civilians through data farming.

Team 3 used PAX to gain insight into specific aspects of peace support operations. The scenario used in this team's data farming at this workshop was based on a crowd control demonstration situation in a stabilization operation.

Team 4 built on previous MANA simulation scenarios to continue to address issues involving improvised explosive devices. In particular, the value of adding unmanned aerial systems with swarming behaviors to the operational environment was examined.

Team 5 applied the Fractional Factorial Controlled Sequential Bifurcation screening procedure to the Hierarchy organizational model and found some delightful surprises through their data farming.

The Joint Test and Evaluation Methodology program provided the leadership for Team 6. This team applied design of experiments and data farming techniques for planning tests in a joint mission environment.

Team 7 explored questions involving critical infrastructure. They used NetLogo to implement modeling of infrastrutures and interactions through network topology.

And finally, Team 8 continued to enhance their Combat Identification model in NetLogo. They were able to begin data farming at this workshop, but plan to reconsider some variables and consider new ones at future workshops

IDFW 15 was our second international workshop in Singapore and it was once again a forum for abundant international collaboration. We certainly owe a great deal of thanks to the organizers in Singapore, and in particular I would like to acknowledge Choo Chwee Seng and Ng Ee Chong from DSO for their efforts in coordinating the workshop. We also owe thanks to the technical lead from the SEED Center for Data Farming, Steve Upton, as well as the many efforts of the co-directors of the SEED Center for Data Farming, Professors Tom Lucas and Susan Sanchez. And one more thank you for a job well done goes to Ted Meyer for collecting and publishing the work in this issue of *The Scythe*. Please note that Ted and I can be contacted at datafarming@verizon.net with questions, comments or suggestions.

Now looking ahead, our Data Farming community will be back in Monterey for our next workshop, International Data Farming Workshop 16. It will start with the opening dinner on Sunday 13 April, 2008 and continue through the week with the closing session on Friday 18 April. We hope to see you there!

Gary Horne

Team 1: Applying Automated Red Teaming in a Maritime Scenario

TEAM 1 MEMBERS

Alan Wong Chin Han, MAJ. – Lead
SAF Ops Research Office, Singapore

Sim Wee Chung, CPT. – Lead, Contact[1]
SAF Ops Research Office, Singapore

Chua Ching Lian, Mr.
DSO, Singapore

Lim Yeng Kiat, Mr.
DSTA, Singapore

Kang Shian Chin, Mr.
DSTA, Singapore

Carol Teo Lijun, Ms.
DSTA, Singapore

Thorsten Lampe, Mr.
EADS, Germany

Philip Hingston, A/P.
ECU, Australia

Ben Abbott, LT.
NPS, USA

INTRODUCTION

With shipping at the heart of the global economy, maritime security is required to ensure freedom of the seas and to facilitate freedom of navigation and commerce. Faced with an array of threats from the terrorists and criminals, nations should stand united and share in the responsibility for maintaining maritime security. In IDFW 15, our team will focus on anchorage protection and continue our works from IDFW 14.

AIM

This study aims to:

- Continue to evaluate the usefulness of Automated Red Teaming (ART)[2] in Blue Ops Planning.
- Gather feedback for development of Automated Co-Evolution (ACE) framework

BACKGROUND

Initial Scenario Set-up. In this baseline scenario, the Blue forces conducted patrols to guard against threats on anchorage. There were several commercial ships anchored in the protected area. The Red forces will attempt to penetrate the Blue defense and inflict damages on the anchored vessels, using various approaches. Any damages to the commercial shipping will deal a severe psychological blow to the Blue defense force. The initial set-up of experiment was as shown in Figure 1 below.

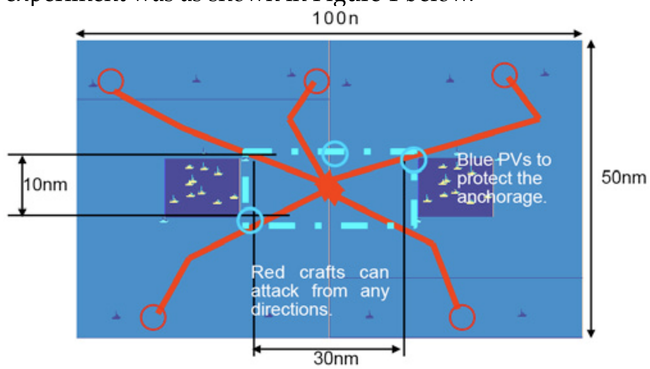


Figure 1: Baseline Blue/Red Plans

KEY ASSUMPTIONS

The following key assumptions were made for this scenario:

Area of Operations (AO). The AO was assumed to be an anchorage in open waters away from the sea lines of communications (SLOC) and main shipping traffic. As such, the neutral shipping was not modeled.

Environmental Conditions. It was assumed that the operations were conducted in dark hours with favorable weather conditions and sea state.

Communication Links. The Blue force was assumed to have full communication link and perfect IKC2. As for the Red force, it was assumed that the individual boats were operating in accordance to mission plans with full communication links.

KEY MODELING PARAMETERS

Blue Forces. The blue force consisted of three patrol vessels (PVs). The following modeling parameters were assumed.

- Patrol Vessels. Each PVs was assumed to conduct normal patrol at 8 knots and give chase at a maximum speed of 16 knots. The PVs were also assumed to be capable of neutralizing the Red boats by closing in within 2 nm. The dynamics of the close water combat was not modeled. A summary of the key specifications of the Blue PVs was as follows:

PV Speed [Patrol] (knots)	8
PV Speed [Chase] (knots)	16
PV Detection Range (nm)	6
PV Identification (ID) Range (nm)	2

Table 1: Specifications of Blue PVs

[1] For more information contact: CPT Sim Wee Chung, toonsim@yahoo.com.sg

[2] ART was developed by DSO-ORL to find an optimal solution for individual sides in a two-sided scenario, using evolutionary algorithms

- Red Forces. The Red boats were modeled as small fishing boats with a maximum speed of 16 knots and loaded with explosives. These boats were assumed to be without any onboard sensors and have a detection and identification range of 2 nm.

Maximum Speed (knots)	16
Detection/ID Range (nm)	2

Table 2: Specifications of Red Boats

- Neutral Commercial Shipping. The neutral commercial ships in the anchorage were assumed to be anchored.

MEASURE OF EFFECTIVENESS

The MOE was:

- Number of Successful Red attacks on Neutral Commercial Shipping
- Mean Red Attrition
- Mean Neutral Shipping Destroyed

METHODOLOGY

Manual Teaming

Blue vs Red. The team members were divided into 2 groups to refine both Blue and Red plans. This was meant to simulate realistic ops planning with minimum intelligence inputs. Both groups underwent several rounds of deliberations and fine-tuning before finalizing their plans. We also captured some of the interim plans to facilitate discussions for the subsequent ACE framework.

Blue Team

Enhanced Inner Patrol. The blue team decided to concentrate own forces within the anchorage area to provide a more responsive and all-round protection on the anchored shipping. A broad deployment concept for the enhanced Blue patrol plan was as shown in Figure 2 below.

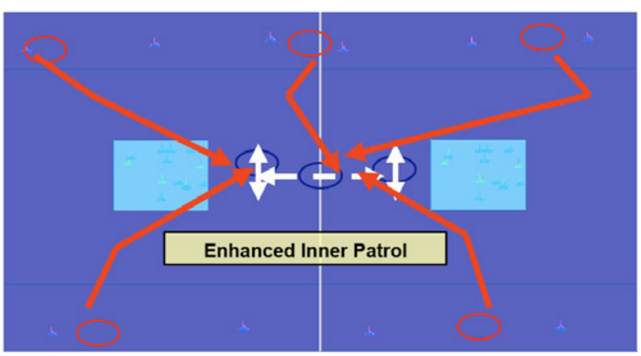


Figure 2: Manual Blue Plan

Red Team

Saturation Strategy. The red team decided to fully utilize their numerical advantage and launch a simultaneous attack on the anchorage area to saturate the Blue forces. A schematic of the red attack plan was as shown in Figure 3.

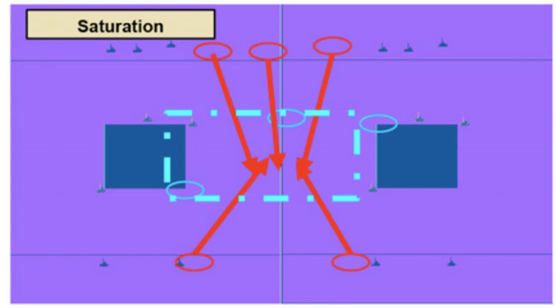


Figure 3: Manual Red Plan

RESULTS AND ANALYSIS

Blue ART Tactics

Multiple-layered Defence Strategy. The ART-generated Blue plan surprised the team initially with a tactic that seemingly made little operational sense to deploy. It took the team a while to decipher and understand the plan better. The ART tactic suggested two border patrols at the northern and southern edge of the anchorage area while the last patrol vessel deployed in a crossover patrol pattern to achieve a multiple-layered defence strategy. The ART-generated Blue plan was as shown in Figure 4 below.

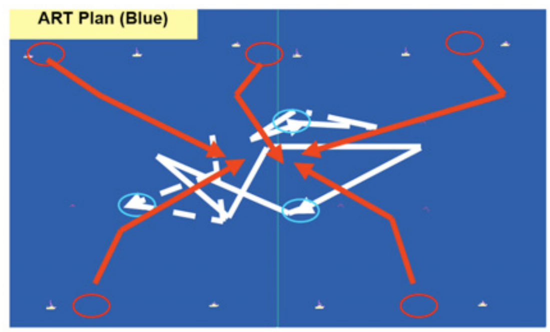


Figure 4: ART Blue Tactics

Red ART Tactics

Simultaneous Pin-point Attack. Similar to the Manual Red Attack Plan, the ART-generated tactic proposed a simultaneous red attack towards the centre of the anchorage area with re-attack flexibilities. This would cater for cases where the anchored vessels were dispersed nearer to the anchorage edges. The ART tactic for the red team was as shown in Figure 5 below.

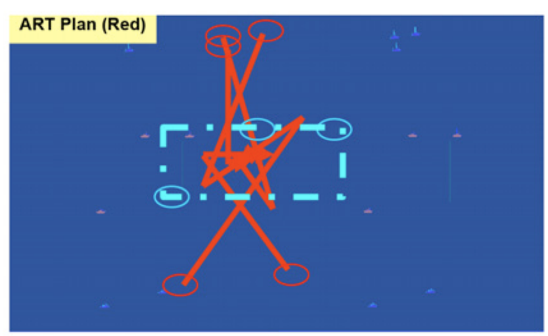


Figure 5: ART Red Tactics

ART Complements Manual Teaming

From the results below in Table 1, it was evident that there was marked improvement in all 3 MOEs for the manual blue and red plans after evolving the intangibles using the ART framework. In addition, the ART-generated Blue tactics produced a significant 98% drop in red mission success and a 99% drop in neutral shipping attrition. As for the ART-generated Red attack plan, the results also showed a drop of 83%in red attrition and 105% increase in the mean number of neutral shipping destroyed. We could therefore conclude from all the above observations that ART not only complemented the manual teaming efforts, it also provided alternate plans for considerations, which might otherwise, be overlooked or non-intuitive.

SUMMARY OF FINDINGS

Automated Red Teaming

Through the exercise during IDFW 15, several valuable feedbacks were received from the team members.

Title is misleading. Firstly, some of the team members found the title "ART" misleading as it seemed to suggest a fully automated process to optimize against multiple plans. This would actually be addressed in the ACE framework. In addition some team members had the impression that ART could only be used to optimize red plans to generate the worst-case scenarios for the blue plans. Instead, ART was developed as a tool to optimize individual sides in two-sided scenarios, using evolutionary algorithms.

<u>Surprises from ART</u>. Secondly, it was interesting to note that the ART had produced plans that were non-intuitive and might not make much operational sense. This led to remarks like "well.. it could be an art itself to decipher ART generated tactics." and "what a surprise!". Nonetheless, ART could be applied to generate alternate plans that might not be intuitive but effective.

<u>Useful Tool to Complement Manual Ops Planning</u>. Finally, the team members found ART as a useful tool to complement manual ops planning. This is consistent with our findings from the last workshop and further strengthens our belief in the applications of the ART framework.

Automated Co-Evolution

During IDFW 15, the team had lengthy and in depth discussions on some of the challenges ACE would faced in its development. Below was the list of questions raised during the discussions:

- How to choose the fittest solution for multiple objectives optimization?
- What if Blue and Red Teams had different end objectives?
- How to achieve optimization for multiple-sided scenarios?

The ACE development team would take into considerations the above challenges to add robustness in the ACE framework.

CONCLUSIONS

This study has discussed some of the strengths and weaknesses of the ART framework. Despite its limitations, the ART framework is highly recommended to be used to complement ops planning efforts. In addition the report has also discussed about potential challenges of the ACE framework.

	Baseline	Blue Plan		Red Plan		Blue ART	Red ART
		Manual	ART	Manual	ART		
Aggressiveness	-60	-60	74	-60	-14	-22	-4
Cohesiveness	-100	-100	-50	-100	-40	85	-16
Determination	60	60	9	60	33	-58	45
Red Mission Success	100%	82%	45%	100%	100%	2%	100%
Red Attrition	2.77	3.98	4.48	1.96	1.83	4.97	0.48
Neutral Attrition	2.21	1.06	0.52	3.05	3.15	0.03	4.52
% Drop (Red Mission Success)	-	16%	55%	0%	0%	98%	0%
% Increase (Red Attrition)	-	44%	80%	(29%)	(34%)	79%	(83%)
% Drop (Neutral Attrition)	-	52%	76%	(38%)	(43%)	99%	(105%)

Table 1: Summary of Results

LOOKING AHEAD

The team will be continuing our effort on the development of the Automated Co-Evolution (ACE) framework.

Team 2: Representing the Civilian Population in Urban Stability Operations

TEAM 2 MEMBERS

Jerry Pearman
Training and Doctrine Analysis Center (TRAC), US

Jonathan Alt, Major[1]
United States Army
Training and Doctrine Analysis Center (TRAC), US

Todd Ferris, Major
United States Marine Corps
Naval Postgraduate School, US

Tiah Yao Ming
DSO National Laboratories, Singapore

Thorsten Seitz, *Commander*
German Navy
Naval Postgraduate School, US

INTRODUCTION

Armed forces across the globe are conducting stability operations to establish enduring political systems in former crisis regions. A goal of stability operations is to influence civilian attitudes in favor of the host nation (HN) government and the stabilization forces. To better understand the dynamics of civilian attitudes, experts have developed analytic models to represent civilian behaviors and reactions to changes in the cultural and societal landscape. Further research is required to integrate concepts from various analytic models to explore the complexity of civilian behaviors. Our team's goal for IDFW 15 was to develop rudimentary agent-based simulation models of civilian behaviors that integrate concepts from selected analytic models.

The work conducted at IDFW 15 extends research led by the US Army's Training and Doctrine Command Analysis Center (TRAC) in Monterey to develop simulation models representing complex social environments. The intent of the simulation models is to assess the effect of societal events on populace behavior. Such events include diplomatic, informational, military, and economic (DIME) factors. When provided with a set of DIME inputs, the simulation model proposes likely responses by the civilian populace.

The simulation model developers used concepts derived from analytic models across the fields of sociology, economics, and international relations to support simulation development. Specifically, the simulation model developers used data and algorithms from analytic models to represent complex social interactions and networks. Complex interactions and networks are critical when attempting to explain the effect of events (inputs) on civilian behavior (outputs). Additionally, analytic models provided well-defined measures of effectiveness easily captured in the agent-based simulation models.

TEAM 2 OBJECTIVES

Team objectives included:

- Developing rudimentary civilian models in an agent-based simulation.
- Exploring factors for data farming related to civilian modeling.

Scoping the problem required the team to determine which sectors of the population to represent in the simulation model. The intent was to model a cross section of the population that would generate heterogeneous responses to events. For instance, populations from diverse religious backgrounds, or sects within the same major religion, often respond differently to certain events. Based on research citing the influence of tribal bonds in our area of interest, the team modeled three tribes. Each tribe contained subordinate clan and family members. To achieve disparate responses to events, the simulation model aligns one tribe with the HN, the second tribe with the insurgency, and represents the third tribe as neutral.

A challenge for the team was to develop a methodology for mapping concepts proposed in the analytic models to the simulation models. First, team members researched applicable analytic models of civilian behavior appropriate for our study. Next, we identified concepts from the analytic models deemed viable for representation in the simulation. For IDFW 15, the team selected Pythagoras, a low resolution agent-based simulation, as the platform to build the civilian model. Pythagoras provides a flexible development environment that directly supports many of the concepts proposed in the analytic models.

ANALYTIC MODELS

The team implemented concepts from three analytic models.

Civilian Attitude Model

An analytic model developed by professors P.A. Jacobs, et al describes the change in population attitude toward the HN due to HN and insurgency (I) actions[2]. The model shows how perceived good and bad actions sum over time. The measure of effectiveness (MOE), attitude toward the HN at time $t + h$, is defined in Equation 1.

---

[1] For more information contact: Major Jonathan Alt, jkalt@nps.edu

[2] Jacobs, P.A., D.P. Gaver, M. Kress, R. Szechtman. *A Model for the Effect of Host Nation/Insurgency Operations on a Population (Draft).* Naval Postgraduate School, October 2007.

Attitude towards HN at time (t+h) due to active actions	=	Attitude towards HN at time (t) due to active HN and I actions	+	Attitude change towards HN at time (t,t+h] due to active HN actions perceived as good
	-	Attitude change towards HN at time (t,t+h] due to active HN actions perceived as bad	+	Attitude change towards HN at time (t,t+h] due to active I actions perceived as bad
	-	Attitude change towards HN at time (t,t+h] due to active I actions perceived as good	+	Attitude change towards HN at time (t,t+h] due to influence of other subpopulations

Equation 1: Measurement of Effectiveness

An example of a good HN action may be job opportunity; or conversely lack of job opportunity may be perceived as a bad HN action. Terrorist activity may be perceived as a bad insurgency action, whereas enhanced security to a region by insurgents may be perceived as good insurgent action. Numerous DIME-related actions could impact the MOE. The simulation model developed during the workshop included only one type of action, an informational/media event. The model simulates HN and insurgent media broadcasts that affect the tribes' attitude toward the HN.

Modeling attitude change and media events in Pythagoras proved relatively straightforward. Pythagoras provides a robust sidedness capability using color schemes that enable the user to define friendly, enemy, and neutral agents. In this case, we assigned each agent within a tribe an initial level of 'blueness' depending on their (friendly) alignment with the HN. The HN and insurgent media events impact each agent's blueness differently depending on their alignment (or lack of alignment) with the HN. An agent's color endstate depicts its attitude toward the HN at $t + h$ and serves as our primary MOE.

The last parameter in the Civilian Attitude Model - influence of other subpopulations – addresses the effect of social networks on attitudes. For instance, two families may belong to different tribes yet influence each other due to common bonds, such as religious sect. Our research shows that multiple analytic models include social network influences when describing civilian behavior. The following section addresses an analytic model focused on social networks and our methodology for mapping its concepts to a simulation model.

Social Networks Model

Professor Deborah Gibbons developed an analytic model that helps describe civilian responses to social network factors[3]. Although the model addresses a cross section of social structures impacting civilian behavior, the team narrowed our focus to her discussion of communications. Specifically, Gibbons highlights three parameters that affect communications as related to civilian behavior: message content, message source and receiver, and hierarchy. We determined that Pythagoras supports these communication parameters to varying degrees and mapped them to the simulation model.

Pythagoras enables communication between agents by assigning like communication devices to two or more agents. Agents with like devices exchange message information each time step. Message content in Pythagoras consists strictly of the sender's location, sidedness, leadership status, resources (supply status), and attributes (user defined quantities). For our purposes, we defined communication hierarchy as intra-tribe and inter-tribe communications, as illustrated in Figure 1. The extent that message content and hierarchy impact the MOE remains uncertain and highlights the importance of integrating multiple analytic concepts into a common simulation model.

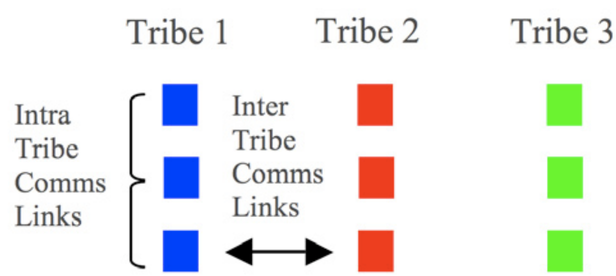


Figure 1: Social Networks Communication Model

Economic Model

Professor Robert McNab developed an analytic model tailored to economic factors that affect civilian behavior in unstable regions[4]. The model describes economic decisions by the HN that impact family resources and help form attitudes. Economic decisions by the HN include wage controls, tax rates, employment opportunities in the production sector, and employment opportunities in the soldiering sector. The dynamics of these decisions force families to seek resources by participating in production, soldiering, or insurrection. Families pay penalties for participation in soldiering or insurrection due to the risks involved with those activities. Unlike the two analytic models described above, McNab's model does not consider the effects of social networks.

To implement concepts from McNab's model, the team utilized Pythagoras' ability to expend and resupply agent resources. The simulation model decrements family resources each time step to simulate daily expenditures. Once resources fall below a user-defined risk threshold, families seek additional resources from production, soldiering, or insurrection, as depicted in Figure 2. Families aligned with the HN (tribe 1) prefer participation in soldiering, whereas families aligned with the insurgents (tribe

[3] Gibbons, Deborah. *Social Networks Model*. Naval Postgraduate School.

[4] McNab, Robert. *A Model of Insurrections*. Naval Postgraduate School.

2) seek resources from insurrection sponsored activities. Non-aligned families (tribe 3) prefer resources from the production sector.

Occasionally, the sources of income experience depleted funds and cannot support customer needs. If insufficient resources exist at the preferred source, families turn toward alternate sources of income. For instance, if a family from tribe 3 becomes unemployed from their factory job, they seek income from either soldiering or insurrection activities.

To link the economic concepts implemented in the simulation model with our overall MOE, attitude toward HN, we impart color attribute changes on families based on their source of income. Specifically, when a family receives resources from the soldiering sector, their blue color attribute increases and they become more aligned with the HN. Conversely, income from insurgent activities decreases blueness, driving families away from the HN. The production sector moves color attributes toward a neutral spectrum, pushing families toward a more neutral stance.

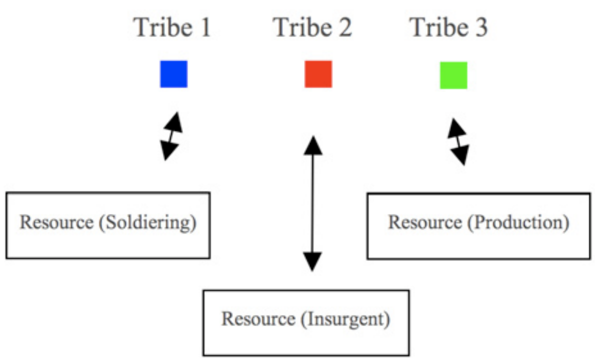


Figure 2: Economic Resources Model

PYTHAGORAS SCENARIO

The team developed the scenario mindful that changes in civilian attitudes and behaviors often require months or years. Given these timelines, the scenario simulates one year of media and economic activity with each time step representing one day (hence 365 time steps). In general, the scenario simulates one year of media and economic events impacting the attitudes of three tribes linked socially, both internally and externally.

EXPERIMENTAL DESIGN

One of our team objectives included exploring factors for data farming related to civilian modeling. Limitations impacting factor selection include simulation capabilities, events modeled, and the scenario. Given available factors, an analyst should then consider which factors may affect selected MOEs. In this case, we selected the following factors:

- Rate of HN media broadcasts
- Rate of insurgent media broadcasts
- Instances of tribes/clans/families
- Vulnerability to media events
- Vulnerability tolerance to media event

These factors proved viable for exploring cause and effect relationships in civilian modeling.

Once we selected experimental factors and identified levels for each factor, we inputted factors and levels into a Nearly Orthogonal Latin Hypercube (NOLH) design to determine design points. We executed 30 replications of each design point on the German high performance computing cluster (HPCC), then analyzed results.

RESULTS

Team 2 analyzed results for one experiment during the workshop. Due to the incremental development process of the simulation model, the experiment included media events, but no economic events. Factors for this experiment included rate of HN media broadcasts, rate of insurgent media broadcasts, and instances of tribes. The NOLH returned 125 design points for the factors and levels described in this experiment.

We analyzed the MOE, attitude toward the HN, by tribe. The data analysis model included linear regression and regression tree techniques. The linear regression models showed no significant interactions between the factors. Regression trees provided both intuitive and non-intuitive results for main factors.

- Intuitively, the rate of HN media broadcasts had the greatest impact on the MOE across each tribe. In other words, as the rate of HN media events increased, tribal attitudes toward the HN improved.
- Interestingly, tribe 1 results indicate that when media events target the most vulnerable populations within the tribe, attitudes toward the HN across the entire tribe improved significantly.
- Surprisingly, tribe 2 results show that instances of tribe 1 members affect tribe 2 attitudes. Inter-tribe social networking between tribes 1 and 2 may have contributed to this result, requiring further analysis.

CONCLUSIONS

Our team's contribution represents a starting point to further extend simulation model research of civilian behaviors. The intuitive results from our experiment provide a measure of validity for the simulation model. Mapping additional concepts from analytic models into the simulation model will enable more detailed study of the complex interactions affecting civilian behaviors.

Recommendations for further research include:

- Executing experiments using media vulnerability factors described above.
- Completing development of economic concepts implemented in the simulation model.
- Improving the social networks concepts implemented in the simulation model. The latest version of Pythagoras will enable more detailed modeling of social networks.

TRAC-Monterey and NPS thesis students will continue to evolve civilian behavior research. The target date for the next iteration of the simulation model is IDFW 16.

Team 3: Operational Synthesis Approach for the Analysis of Peace Support Operations

TEAM 3 MEMBERS

Dr. Dietmar Kunde, LTC - Lead
IT-AmtBw A5, Germany
Gunther Schwarz
EADS System Design Center, Germany
Gudrun Wagner
EADS System Design Center, Germany
Luminita STEMATE
Defence Research and Development Canada, Canada
Ng Ee Chong
DSO-ORL, Singapore
Leow Yong Khiang
DSTA, Singapore
Dr. Dan CHEN
School of Computer Engineering – NTU, Singapore

INTRODUCTION

The Peace Support Operations Team at International Data Farming Workshop 15 used the simulation system "PAX" to gain insight into specific aspects of peace support operations.

PAX focuses on tactical miniature scenarios in the context of peace support missions. Both military expertise and empirical findings from psychological research on aggression were used in the design of PAX. Psychological factors having an influence on the decisions and the behavior of all persons concerned may have a considerable effect on the development of an operation.

PAX concentrates on modeling peace support operations on a detailed tactical level. Since being of secondary interest in the question sets examined, terrain is modeled in a fairly abstract way in a grid-based environment with a distinction between normal cells, built-up cells and obstacles. Due to its nature and objectives, the model focuses on the detailed representation of the individual civilians and their internal states, including emotions such as fear or anger and their interrelation. The military forces modeled in PAX, on the other hand, have the possibility to not only use different types of weapons, as in existing military simulations, but to also take measures of active de-escalation, such as trying to calm down people or talking to the leader of a civilian group.

The current PAX version 3.0 provides the military analyst with means to examine question sets in a variety of easy to set up PSO scenarios. Thus, the main goals in the development of this "Toolbox PAX" were to make the model flexible enough to be used in a broad variety of scenarios, examining different aspects of PSO missions, while at the same time keeping the user interface easy and intuitive enough for the military expert to use it without an excessive introduction.

Scenario

The scenario examined during the workshop was based on a crowd control "Demonstration" situation in a stabilization operation. Different options for the military side should be discussed and analyzed.

The situation was considered to consist of the following elements:

- An event organized by political-religious party A is taking place in a town hall in the center of a town predominantly inhabited and controlled by party B. The event is scheduled and expected to end at 1000.
- PAXFOR intelligence reports:
  - A group of rather extremist counter-demonstrators, who are known for their aggressive behavior, is expected to approach the city hall around 0945 after the end of another event organized by party B. This countermarch is announced and approved by the local forces.
  - A group of rather aggressive young adults sympathizing with party B is also expected to approach the city hall around 0930. They are known for their dislike against members of party A. While they are not expected to immediately show open aggression, this may change especially when aroused or intoxicated. They are not expected to sympathize with PAXFOR so that aggression, once arising, may expand to violent acts against PAXFOR, too.
  - Heavily armed combatants are not expected to be in the town centre since the local population is considered friendly and would probably note and report suspicious non-locals to the local administration. However, both groups belonging to party B might be slightly armed and carry along bats or use other items like bottles, cans or stones to attack their antagonists.

Table 1 summarizes the groups' attitudes and different intentions.

Group	Main setup of group members	Primary intention
Members of party A	Rather peaceful, but still highly aroused and showing a high group cohesiveness.	Leave the town hall and return home
Youngsters of party B	Rather high readiness for aggression and low willingness for cooperation	Provoke members of party A
Counter-demonstrators of party B	High readiness for aggression and anger, low fear	Demonstrating against members of party A
Bystanders	Low anger / readiness for aggression, low fear, no group cohesiveness	Walk around

Table 1: Groups and their primary intention

The basecase scenario is shown in Figure 1, an aerial view of the terrain using the mission preparation tool VIPER-3D, developed at the EADS System Design Centre.

Figure 1: Scenario overview in VIPER-3D

Figure 2 presents a possible equivalent scenario modeled with the PAX scenario editor.

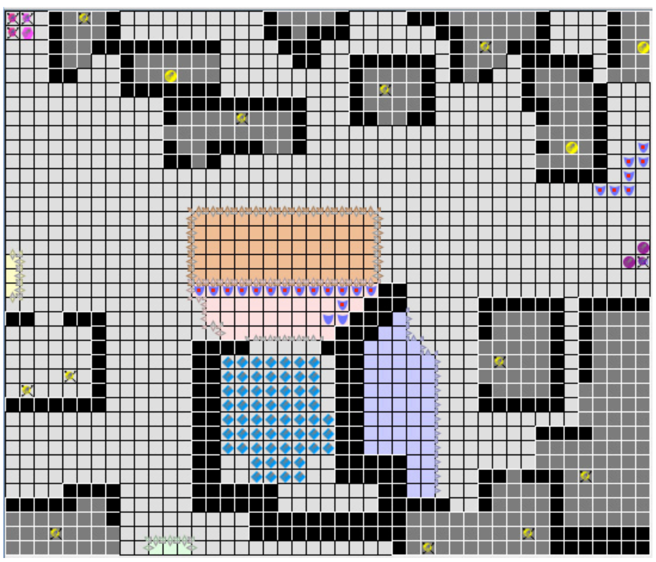


Figure 2: Basecase scenario modeled in PAX

The PAXFOR Operations Analysis cell is asked to provide assistance for the planning of the operation considering the following circumstances and question sets:

- What are the benefits of extensive reconnaissance measurements prior to the operation?
- What can be done to prevent bystanders from being involved in any conflicts and behaving aggressively themselves?
- How can it be assured that peaceful members of party A are able to leave the town hall without being attacked by aroused counter-demonstrators.
- How does the announcement of threat of force influence the civilians' behavior?
- Under which circumstances can the situation escalate to an unacceptable extent?
  - Is the relationship of civilians behaving in a violent way compared to the manpower of the soldiers a sufficient indicator for this?
  - What other indicators are of interest?
  - Under which circumstances is it recommended to stop the operation because extreme escalation can not be avoided?
- Does the use of weapons and protective equipment add significantly to military success? Under which circumstances?
- Under which circumstances and how is it possible to distract or canalize bystanders and other civilian groups in order to help deescalate the situation?

The basic scenario was set up in PAX and the team members developed vignettes representing different possible situations in which the effects of different PAXFOR tactics could be analyzed.

ANALYSIS

The main focus of the analysis was to examine how different rule sets of the soldiers, e.g. Rules of Engagement (RoEs), affect the situation in terms of the following Measures of Effectiveness (MOEs):

- overall escalation of the situation,
- number of casualties on soldiers' side,
- number of casualties on civilians' side,
- number of party A members leaving the town hall without being attacked by other civilians,
- number of arrests
- number of civilians attacking with weapons

Having discussed the given scenario, the group was mainly interested in considering the influence of the following factors:

- Different number of counter-demonstrators
- The counter-demonstrators' attitude towards the party A members and towards the military (from rather peaceful to highly hostile)
- Different number of soldiers
- Different rulesets for the soldiers

Determining the important factors

To find out the most important model parameters with regard to the given list of MOEs, a number of PAX parameters were analyzed in a simulation experiment using the Nearly Orthogonal Latin Hypercube (NOLH) design provided by the Naval Postgraduate School (NPS) Monterey, CA, USA.

The NOLH design used contained 27 parameters, listed below. Each design point was calculated using a MultiRun of the simulation model PAX with 20 replications, each with a different random seed.

NOLH Farming Parameters	Min	Max
Number of counter-demonstrators	50	200
Attitude of party A members towards the military	-60	60
Attitude of party B members towards the military	-60	60
Attitude of party B youngsters towards party B counter-demonstrators	-60	60
Party A cognitive motive to leave the town hall	50	100
Party A anger	10	60
Party A readiness for aggression	10	60
Party A personality constant anger	5	80
Party A personality constant fear	5	80
Party B personality constant anger	5	80
Party B personality constant fear	5	80
Party B counter-demonstrators anger	40	80
Party B counter-demonstrators dog factor	1.5	2.0
Party B counter-demonstrators norms for anti-aggression	5	30
Party B counter-demonstrators readiness for aggression	50	90
Party B counter-demonstrators willingness for cooperation	10	80
Party B youngsters dog factor	0	2.0
Party B youngsters norms for anti-aggression	5	50
Party B youngsters readiness for aggression	40	90
Party B youngsters entry time	10	80
Bystanders' anger	20	90
Military cordon's threshold for calling reinforcement	0	40
Military escort's duration of action arrest	3	10
Military escort's threshold for intervention	0	200
Military escort's shoving strength	50	100
All soldiers' rule set	1	4

Table 2: Parameters varied in NOLH design study

A soldier's rule set defines his reactions to certain civilian actions under specific side conditions. The "PSO Manual" rule set represents a moderate reaction to civilian actions trying to create a balance between an immediate sharp reaction and a complete laissez-faire attitude.

If the soldiers behave according to the "Gandhi" rule set, they always try to pacify the civilians.

Rule set "Arrest on attacks" specifies that soldiers are supposed to arrest the civilian attackers (no matter whether any weapons are used).

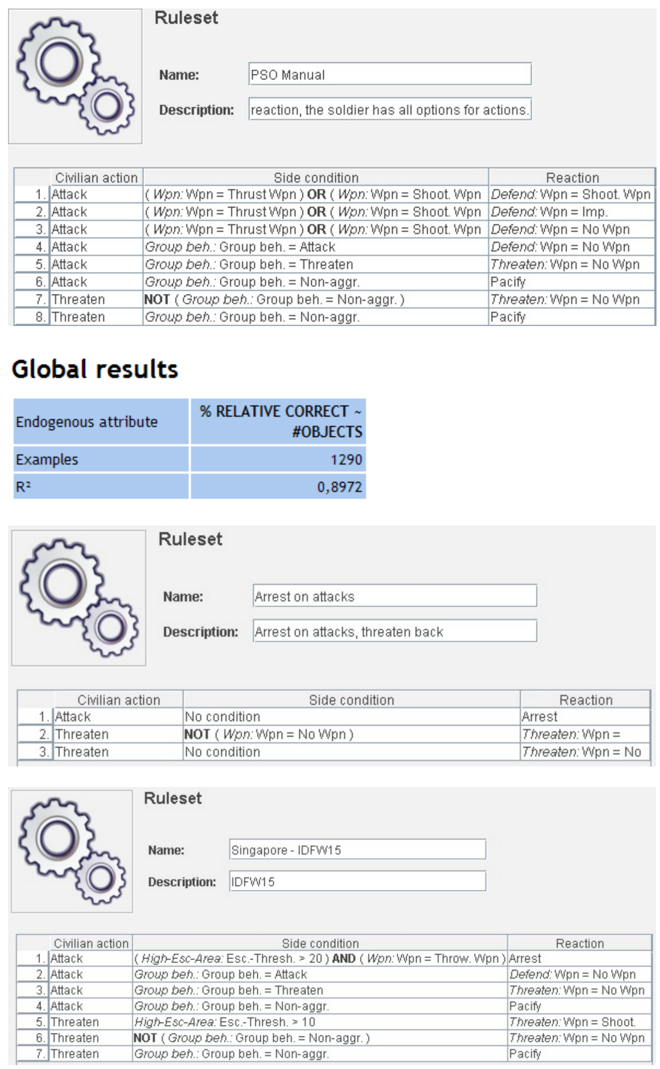
Rule set 4 ("Singapore - IDFW15") was developed within the working group. Its motivation is to allow some freedom of protest up to a point of high group escalation, after which the soldiers will react by

- firing a warning shot in case they are threatened by the civilians and
- arresting the civilians throwing stones in case they are attacked by the civilians.

Figure 3 shows the rules of the various rule sets for the soldiers' behavior.

When analyzing the data with the help of commercial off the shelf (COTS) statistical software, the following parameters turned out to be significantly important for the MOEs described above:

- party B counter-demonstrators' attitude towards the military
- party B counter-demonstrators' readiness for aggression
- party B counter-demonstrators' and party B youngsters' dog factor[1]
- party B youngsters' entry time
- soldiers' rule set


Ruleset

Name: PSO Manual

Description: reaction, the soldier has all options for actions.

	Civilian action	Side condition	Reaction
1.	Attack	(*Wpn:* Wpn = Thrust Wpn) **OR** (*Wpn:* Wpn = Shoot. Wpn	*Defend:* Wpn = Shoot. Wpn
2.	Attack	(*Wpn:* Wpn = Thrust Wpn) **OR** (*Wpn:* Wpn = Shoot. Wpn	*Defend:* Wpn = Imp.
3.	Attack	(*Wpn:* Wpn = Thrust Wpn) **OR** (*Wpn:* Wpn = Shoot. Wpn	*Defend:* Wpn = No Wpn
4.	Attack	*Group beh.:* Group beh. = Attack	*Defend:* Wpn = No Wpn
5.	Attack	*Group beh.:* Group beh. = Threaten	*Threaten:* Wpn = No Wpn
6.	Attack	*Group beh.:* Group beh. = Non-aggr.	Pacify
7.	Threaten	**NOT** (*Group beh.:* Group beh. = Non-aggr.)	*Threaten:* Wpn = No Wpn
8.	Threaten	*Group beh.:* Group beh. = Non-aggr.	Pacify

Global results

Endogenous attribute	% RELATIVE CORRECT ~ #OBJECTS
Examples	1290
R²	0,8972

Ruleset

Name: Arrest on attacks

Description: Arrest on attacks, threaten back

	Civilian action	Side condition	Reaction
1.	Attack	No condition	Arrest
2.	Threaten	**NOT** (*Wpn:* Wpn = No Wpn)	*Threaten:* Wpn =
3.	Threaten	No condition	*Threaten:* Wpn = No

Ruleset

Name: Singapore - IDFW15

Description: IDFW15

	Civilian action	Side condition	Reaction
1.	Attack	(*High-Esc-Area:* Esc.-Thresh. > 20) **AND** (*Wpn:* Wpn = Throw. Wpn)	Arrest
2.	Attack	*Group beh.:* Group beh. = Attack	*Defend:* Wpn = No Wpn
3.	Attack	*Group beh.:* Group beh. = Threaten	*Threaten:* Wpn = No Wpn
4.	Attack	*Group beh.:* Group beh. = Non-aggr.	Pacify
5.	Threaten	*High-Esc-Area:* Esc.-Thresh. > 10	*Threaten:* Wpn = Shoot.
6.	Threaten	**NOT** (*Group beh.:* Group beh. = Non-aggr.)	*Threaten:* Wpn = No Wpn
7.	Threaten	*Group beh.:* Group beh. = Non-aggr.	Pacify

Figure 3: Rule sets examined

[1] The civilian's dog factor influences the evaluation of the soldiers' actions 'Threaten' or 'Defend'. A high value results in a higher increase of fear.

Full factorial experiment

Based on the results of the NOLH design experiment and in order to assess the effect of the different rules of engagement, it was decided to set up a full factorial design experiment varying the party B counter-demonstrators' readiness for aggression and their dog factor together with the soldiers' rule set.

Full Factorial Farming Parameters	Min	Max	Step Size
Party B counter-demonstrators readiness for aggression	30	90	10
Party B counter-demonstrators dog factor	0.1	1.5	0.2
All soldiers' rule set	1	4	1

Table 3: Parameters varied in full factorial design study

All in all the full factorial study consisted of 4480 single PAX runs performed on a 128 node cluster of the Simulation and Test Environment of the German Bundeswehr in Euskirchen / Germany.

Comparison of the rule sets

Prior to analyzing the result landscapes of the full factorial experiment, the team members wrote down their expectations for each ruleset regarding

- the escalation accumulated during the simulation by aggressive actions (i.e. "attacks" and "threats") performed by civilians
- the escalation accumulated during the simulation by soldiers' use of force (including threatening and defending actions),
- the number of attacking actions performed by counter-demonstrators involving weapons and
- the number of defending actions with a shooting weapon performed by soldiers.

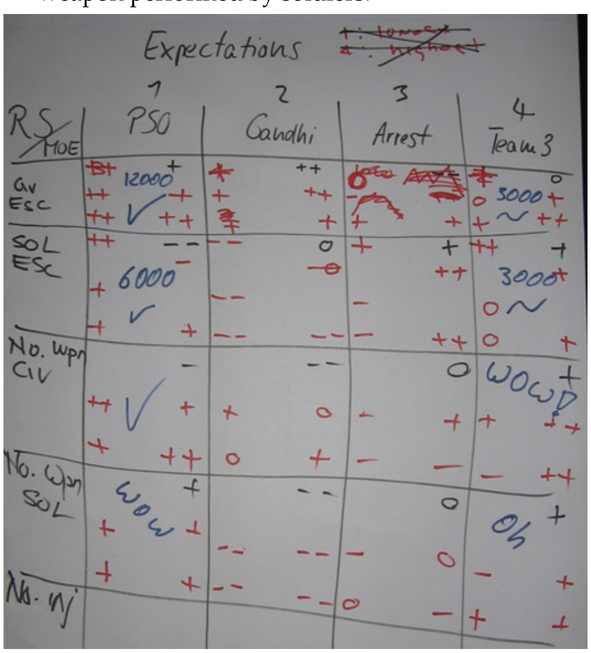


Figure 4: The team members' expectations regarding various MOEs

Regarding the PSO Manual rule set, the team members' expectations were mostly met. The "IDFW15 – Team 3" rule set, however, performed much better than expected.

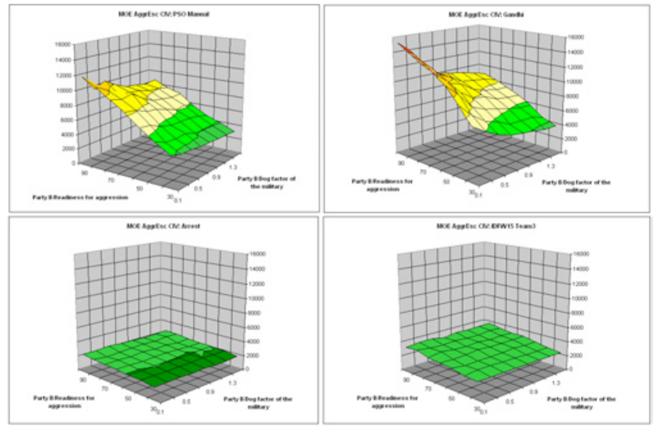


Figure 5: Performance of the different rule sets with regard to the aggregated escalation that resulted from actions performed by civilians

The simulation showed that arresting aggressive civilians in a fairly early stage of escalation seems to be a very effective means to keep control of the situation.

The evaluation of the result landscape for the MOE "Number of Party A members killed" shows that especially the rule sets "PSO Manual" and "Gandhi" do not perform well in protecting the party A members. (see figure 6)

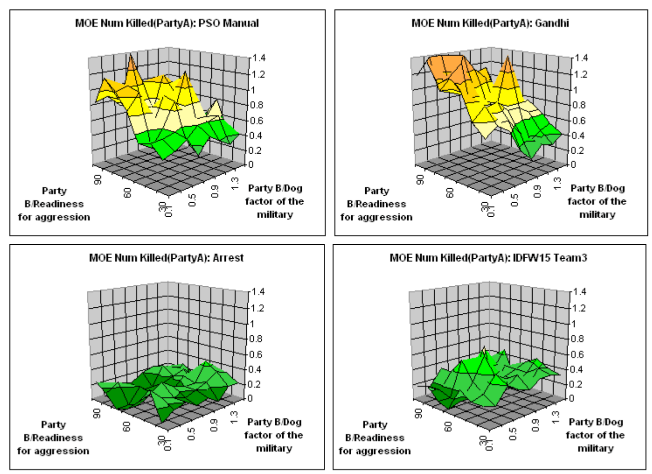


Figure 6: Performance of the different rule sets with regard to the number of party A members killed

However, if the soldiers behaved according to rule set "Gandhi", trying to avoid any violent actions and pacifying the civilians whenever possible, the overall number of civilians killed was heavily reduced. (see figure 7).

The big surprise in comparing the performance of the rule sets was that even if the soldiers acted in a significantly tougher way versus party B members ("PSO Manual"), resulting in injured and killed party B people, they did not perform better in protecting party A members than by using the pacifying "Gandhi" rule set.

Thus it appears that in this scenario, if the soldiers do not have any means of arresting aggressive civilians, deescalating the situation by pacifying the civilian population seems to be the best way to minimize civilian casualties.

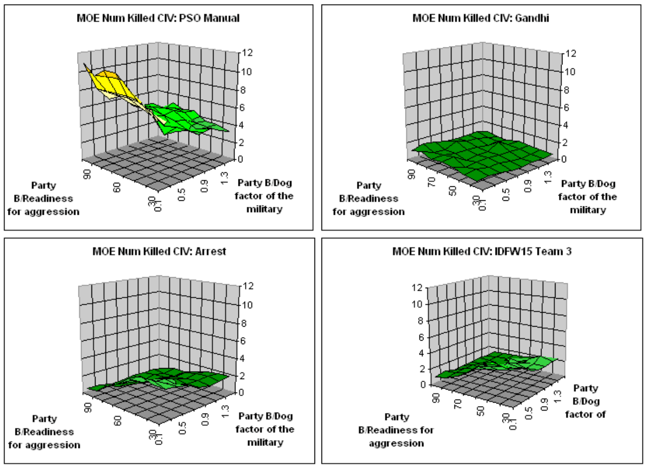


Figure 7: Performance of the different rule sets with regard to the overall number of civilians killed

CONCLUSIONS

PAX was successfully applied using the Data Farming methodology and tool box to analyze and discuss the soldiers' options for actions within the given crowd control scenario and to examine how their behavior influences the civilian population. Prior to setting up the scenario in PAX, the overall scenario situation was examined using the virtual reality tool for mission preparation "VIPER-3D".[2]

The Data Farming process including:

- the identification of the most important factors with regard to the relevant MoEs using the NOLH design of experiments;
- full factorial experiment designs for finding out more especially about the effectiveness of the specific rule sets; and
- course of action analysis on representative single runs to learn more details about the effects responsible for average, good or bad performance of the forces

proved to be very effective regarding the applicability and maturity of the method and tools and regarding the quality of the results.

The international collaboration in the team and especially the expert knowledge of the foreign team members was very helpful to get new perspectives and incentives for further developments of PAX, usage of experimentation techniques, data analysis tools and areas of application.

[2] VIPER-3D had already been applied during the preparation phase in Germany, not at the workshop in Singapore.

Team 4: Influencing Insurgent SVBIED Operations Using Traffic Control Points Enhanced with Unmanned Aerial System (UAS) Employment Strategies

TEAM 4 MEMBERS

Edward Teague
US Military Academy at West Point, US

Niki Goerger
Engineer Research and Development Center, US
US Military Academy at West Point, US

Brad Young
Naval Post Graduate School, US

Ng Chu Ngah
Quek Kwang Ti
Defence Science and Technology Agency, Republic of Singapore

INTRODUCTION

Insurgents have effectively employed asymmetric tactics, such as the use of suicide bombers, as viable threats in urban environments. These threats are often devastating in their physical and emotional effects. They are hard to detect and have proven difficult to thwart or defeat. The U.S. Army has recognized that improvised Explosive Devices (IEDs) pose a persistent and devastating threat, impacting unit operations, U.S. policy and public perception (U.S. Army, 2005).

Suicide Vehicle-Born IEDs (SVBIEDs) would be easier to thwart or defeat if the political, cultural, and physical environments in which they were implemented were more readily constrainable as in full combat operations. However, in Stability, Security, Transition, and Reconstruction (SSTR) Operations, it is important to allow the nearly free flow of people (noncombatants) and goods through an economically developing or thriving community. The involved urban environments can be physically complex and culturally diverse. Threats that employ SVBIEDs take this into account and use this to their advantage. Our current, limited understanding of the human behaviors that drive the insurgent's decisioning and responses, and the insurgent's ability to capitalize on the nature of the urban environment in stability and support operations adds to the complexity and challenges of detecting and defeating this threat. There is a need to increase our understanding of the behavioral aspects, or responses to perceptions regarding the environment, of such threats so we can evoke responses that decrease their probabilities of mission success and increase our advantage in this contemporary operational environment.

The Engineer Research and Development Center (ERDC) along with the United States Military Academy (USMA) Operations Research Center of Excellence (ORCEN) recently conducted a study to provide insights into insurgent behaviors, or decisioning, and mission outcomes given different strategies associated with traffic flow and traffic control points (TCPs) applied by counterinsurgent forces. The goal of the IDFW 15 effort was to determine the value of adding Unmanned Aerial Systems (UAS) to the aforementioned environment. Specifically, we looked to derive insights about semi-autonomous UAS with swarm behaviors.

This extension can assist counterinsurgent forces in several ways. It will drive needs for UAS allocation and development by identifying critical elements of semi-autonomous swarming behavior. In addition, it may highlight mission behavior that shows the most promise by expanding the tasks and environment beyond this current set. Due to the complex nature of the problem, the current state of understanding in the field, and the exploratory nature of the research, insights vice specific answers are central in this research.

PROBLEM STATEMENT

This study addresses whether we can isolate factors needed to identify effective semiautonomous UAS behaviors that add value to the aforementioned SVBIED TCP study. Ideal UAS implementation would maximize the area searched and minimize time between UAS passes within a defined geographic zone in order to locate and track SVBIEDs.

In the study scenarios, insurgents selected a fixed target, planned a route based on awareness of friendly TCPs, and moved through an urban environment to attack the target using one SVBIED. While executing the mission, the SVBIED could be detected by roving UAS or encounter TCPs of which they were not previously aware. If the SVBIED was identified by a UAS, selected TCPs within the local neighborhood would change from a static to a mobile posture. The mobile TCPs would receive updated SVBIED locations from the UAS in contact. They would then close with and interdict the SVBIED.

Within this context, we investigated the following study questions:

- Is SVBIED mission outcome a function of varied UAS swarming behaviors where mission outcome is SVBIED reaches primary target and detonates, detonates at an alternate target, or fails to detonate at any target?
- Does swarming behavior enhance UAS performance when searching an area for a specific target?
- In cases where the SVBIED reaches its primary target, what factors are important?
- In cases where the SVBIED detonates at alternate location, what factors are important?

METHOD

Team 4 used the Map Aware Non-uniform Automata (MANA) agent based simulation as it was also the simulation used for prior experiments with this study. MANA provided the flexibility we needed in agent personality, communications architecture, and agent behavior to effectively model employment of UAS in a variety of ways.

This research incorporated principles from several fields to explore factors associated with effectiveness of strategies in asymmetric environments. Agent based modeling, large-scale experimental design, and artificial electro-magnetic field theory were used to develop methods for capturing complex adaptive system behavior associated with nonlinear interactions. The artificial electro-magnetic field theory was used for global insurgent path planning and UAS separation. An agent based modeling environment was selected for implementing and executing the scenarios. The large-scale experimental design was used to establish factor settings (e.g., UAS speed, Swarm Behavior) for the scenarios.

Previous studies focused on exploring the impact of friendly and adversary capabilities and TCP strategies on SVBIED mission outcomes. These experiments ranged from a single target with a static defense array and sparse road network to multiple targets with a dynamic defense array within a dense road network. Capabilities examined included friendly and adversary communications and sensor performance during mission execution. The strong association of TCP strategy with SVBIED mission outcome spurred this extension of the investigation and led to the current set of experiments.

The visual representation of the current study appears similar to the previous experiments. Visually, it looks as though we've only added 3 UAS. However, there is a different command and control structures for the counterinsurgent forces. In addition, previously static TCPs can move to an SVBIED when it is identified.

The area of interest is a 5km x 5km box representing an urban area with a dense road network. The network consists of primary and secondary roads. There is an array of TCPs meant to interdict SVBIED attacks, background traffic, and a target. In addition, there are a number of spotters to provide information to the SVBIED and its escorts as it travels toward the target.

Our baseline simulation used UAS in the simulation with prescriptive programmed behavior to represent the present state of the art. In addition, they reported data to one central ground station. We assumed transmitted data would be in the form of video transmission and associated metadata such as time and geographic locations of both the UAS and its observations. We also assumed there was no direct communication between any of the 3 UAS.

We developed factors on interest based on insights gleened from previous simulation attempts at swarming UAS. We employed a consistent TCP allocation to negate the influence of that factor. Within MANA, factors considered were the following: UAS Detection Range, UAS Classification Probability, UAS and TCP Communications Delay, UAS to UAS Repel tendency, UAS Inorganic Situational Awareness (ISA) Persistence, and UAS Speed.

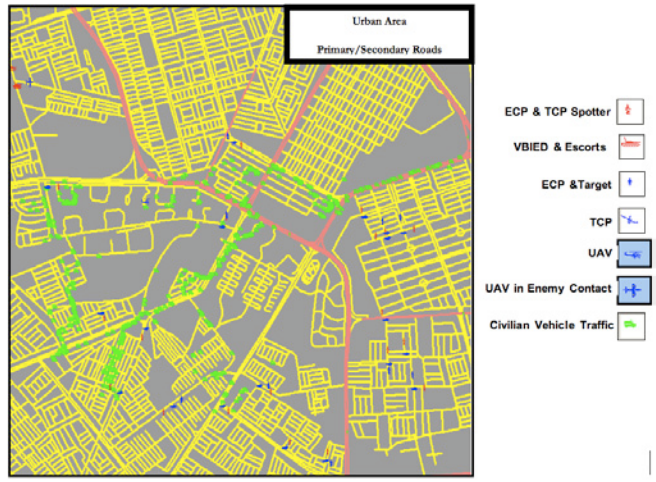


Figure 1: Urban Environment with Dense Road Network

The factors UAS Detection Range and Speed are self-explanatory. Ideally we wanted to maximize both, but we realized significant increases in speed can result in a degradation of the UAS observation capability. In addition, we didn't want to model unrealistic detection capabilities. Instead, we modeled a modest ability for the UAS to detect and classify a target as an SVBIED. Because this study may be used to drive future capability development, we thought this feature was important to explore.

UAS detection capability was modeled because we assumed that some intelligence about pending SVBIED missions existed. That intelligence, however strong, would be fed to the UAS and their operators. It would provide some guidance as to when or where to begin a search as well as what the SVBIED may look like. We think this is plausible.

Even in areas where coalition forces are unable to affect the SVBIED decision cycle, forensic and historical data exists. This data could be used as a substitute for real time or actionable intelligence. A semi-autonomous UAS would use the information and an estimation of its veracity to drive its search pattern. A UAS ground station operator would essentially do the same thing with his reconnaissance plan.

The factors we manipulated to explore swarming with the MANA software were UAS Repel, and UAS ISA Persistence. These two factors seem to be keys in simulating swarming behavior. The essence of swarming behavior is an entity's situational awareness (SA) as a function of the group's situational awareness given current or recent input of other local entities as they collaborate on tasks. Updates to the global and individual SA elicit changes in behavior for each entity.

This self-organized behavior optimizes the group's effort. For instance, ants, when searching for food, move along trails and leave a scent, a pheromone, to mark their presence. If other ants in the area come upon the trail and detect high levels of pheromone, they too begin to follow the trail. Low levels will cause them to ignore the trail. This is because high levels of pheromone indicate high traffic flow. The ants are programmed to assume that high traffic flow means success location of a close by food source. This strategy ensures that the colony exhausts any nearby food before other more arduous exploration begins.

In the simulation, the assigned collaborative task is identification of SVBIEDs within an urban setting. Here, swarming behavior ensures that the UAS are continually conducting surveillance throughout. As a strategy, UAS should remain dispersed so as not to overlap. Still, they need to revisit each other's paths after time has passed for vigilant observation. Varying the UAS' repellence between each other allowed us to set how far apart they would remain from each another.

If the inter-search time between observations is too long there may be gaps in the reconnaissance that would permit an SVBIED undetected travel to a TCP or its intended target. We addressed this with the ISA persistence. It varied the strength of the memory or the pheromone left when a UAS last passed over a location. Opposite from the ant example, UAS would be compelled to return to areas where the pheromone is weak.

Swarm behavior can be an optimization tool that balances simple jobs to meet the larger task of finding SVBIEDs. Varied levels of the factors as well as assigned UAS routes allowed for representation of 3 different types of swarming behavior as well as two variations of a baseline (no swarm) behavior, No Swarm A, No Swarm B, Daisy Chain, Central Hub, and All Net.

UAS in the No Swarm variants did not interact. They shared no information nor did they sense or repel each other. Daisy Chain, Central Hub, and All Net represented swarming of semi-autonomous UAS. Thus, the 3 UAS in each scenario used information from the others to self-organize within the reconnaissance zone and search for SVBIEDs.

UAS in No Swarm mode strictly followed preprogrammed search patterns to search the urban area. Swarming UAS were given general guidance by assigning paths but instructions also included generous levels of stochastic behavior. This allowed the UAS to stray from paths based on information from its UAS swarm-mates. Differences in the organization of the UAS defined their designation Daisy Chain, Central Hub, or All Net. Factor levels were varied for all UAS strategies in order to measure their influence on mission outcome as well.

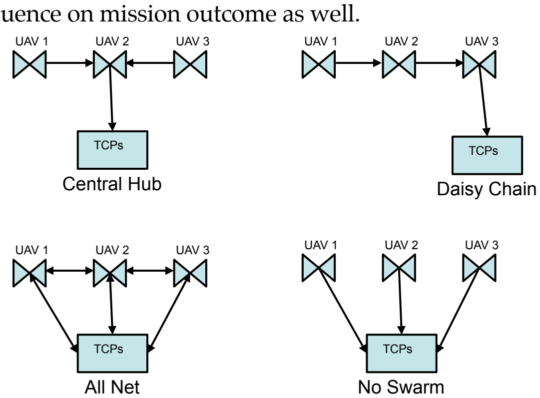


Figure 2: UAS Organization Methods

Though similar, these arrangements represent significant differences in infrastructure or robust reliability of data exchange. Given limitations of UAS processing or line of sight communications, they lend themselves to specific application. For instance, All Net is the most reliable system but it demands UAS that can provide near real time identification of specific targets for transmission to its friendly command and control neighborhood. It also requires constant communications over the local neighborhood. There are no UAS that are currently up to the task as a semi-autonomous system. In addition, the data bandwidth requirement is infeasible as it grows exponentially with each additional UAS or receiver.

EXPERIMENTAL DESIGN

To address the study questions, seven factors as shown in table 1 were incorporated into a nearly-orthogonal Latin hypercubes (NOLH) experimental design. A full factorial design would have yielded thousands of design points. Using NOLH, we reduced the number of design points to 17 and executed 25 replications per design point for a total of 425 runs. Each replication was a realization of a stochastic process of UAS behavior, route selection and traffic flow as modeled in MANA. This was done for each of the 5 UAS strategies.

Due to excessive run times for each simulation (over 96 hours), we have yet to analyze all of the output data. Our hope is that some, if not all, of the factors are significant and that the different UAS strategies yield statistically different outcomes in reference to the rate an SVBIED will detonate at its intended target.

Factor	Description
UAS Detection Range (8 levels)	Distance UAS payload can detect a target (.2 to 1.0)
UAS Classification Probability (17 levels)	Probability the UAS will correctly identify the SVBIED (.66 to .98)
Friendly Comms Accuracy (17 levels)	Accuracy of messages sent by the UAS to TCPS (.40 to 1)
Comms Delay Between UAS and TCPS (17 Levels)	Time that messages between friendly forces will be delayed (0 to 64 seconds)
UAS Repel (3 Levels)	Propensity of UAS to come near each other (low to high)
UAS Persistence (3 Levels)	Propensity of UAS to return to a spot already observed (low to high)
UAS Speed (8 levels)	Speed of travel (100 to 1000 cells per step)

Table 1: Factors Employed in Scenarios

INITIAL FINDINGS

Though we cannot offer any analytical bounty from this effort, there are a number of insights thus far. The first speaks to the length of time our simulation requires for 1 run. Each simulation can take up to 26000 steps (over 7 hours in real time). With over 2000 iterations to run, this is a labor intensive effort.

The runs are long for two reasons. The background urban traffic flow reaches steady state at about 8000 steps. For the model to be valid, this warm up period is a necessity. The traffic congestion is indispensable in order to determine the strengths and weakness of strategies in an urban region. Realistic traffic throughput defines the current environment for all players.

In addition to reaching steady state, the stochastic nature of the SVBIED movement and probability of UAS or TCP success make end time estimations difficult to gauge. The long runtime is a reflection of our desire to capture as many final outcomes as possible. We will analyze the data associated with SVBIED interdiction or detonation to fine tune the simulation run times.

There were a couple challenges in the model as well. It was difficult to mimic the UAS ability to detect and classify a target as an SVBEID. We didn't want the camouflage of the SVBIED so low that the UAS were not valued added but we didn't want the UASs to be omnipotent either.

We compromised by linking the SVBIED escorts to the UAS. The escorts knew where the SVBIED was at all times. We arranged for at least one of the UAS to have contact with an escort and receive updates. To level the playing field some, we varied reliability of the escorts' ISA as it was transmitted to the UAS. Then the UAS more or less knew the VBIED location and description. We felt this method mimicked the sparse intelligence units currently have about SVBIED operations.

Another challenge was modeling swarming in MANA. We were able to mimic pheromones and their fleeting nature by using the Persistence value parameter. However, we were unable to verify that we could fine tune a UAS' receptors to react to pheromones. We had very little information about the ranges at which the pheromones were helping drive UAS behaviors.

THE WAY AHEAD

After collecting the forthcoming data, we hope to further explore important factors by recreating this experiment in different environments using different goals and tasks for the UAS. This would help us determine which UAS strategies are most effective in each environment. For example, in a civil search and rescue scenario, the Daisy Chain may be sufficient. However in a hostile situation, the redundancy of other methods may prove necessary. Tasks to be explored include various types of reconnaissance, direct fire engagements, search and rescue, and forward observation. Settings of interest include both conventional and unconventional conflict, natural disaster, and border patrols.

The intricacy and dire outcomes of the SVBIED problem demand high fidelity models to gain traction and begin validation of new counterinsurgent strategies. We are investigating the construction of a federation of simulation models to study this problem. Though no small effort on its own, we will try to mimic the success of federated training models. We will leverage models that best represent the sub-elements of this complex adaptive system. We hope to build a user friendly model that varies parameters within a constructive simulation and farming output data by leveraging the best tools available.

International Data Farming Workshop 15
Plenary Sessions

Cultivating an Exchange of Knowledge

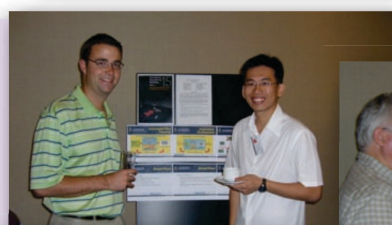
Following is a list of the keynote and plenary sessions at IDFW 15. These presentations, as well as team in- and out-briefs, team reports, models and data can be found on the IDFW15 CD or website (see inside back cover for details).

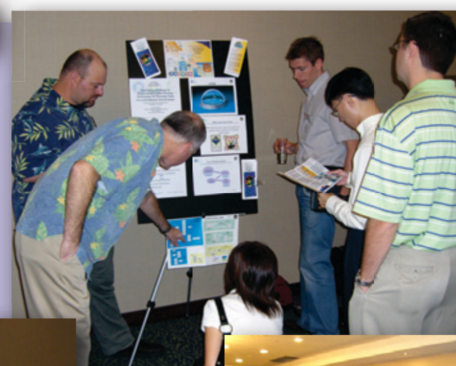
Opening Session

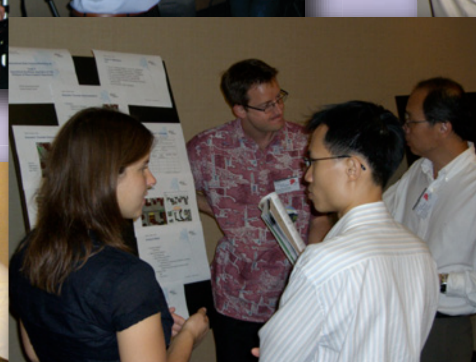
Keynote

Dr. How Khee Yin,
Director, Info Division, DSO National Laboratories

Poster Session







International Data Farming Workshop 15

Plenary Sessions

Topic Sessions



Tuesday Sessions

- **Choo Chwee Seng** - Modelling, Simulation & Analysis in ORL
- **Gunther Schwarz** - EADS PAX Toolbox
- **Gary Horne** - Data Farming for Beginners
- **Susan Sanchez** - Using Simulation to Study the Protection of Critical Maritime Assets



Thursday Sessions

- **Dietmar Kunde** - The ITSimBw Simulation Environment
- **Lawton Clites** - Medical Modeling Update
- **Bradely Wilson** - Modeling Communications and HITL in MANA

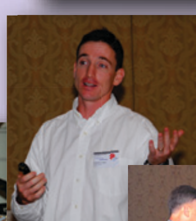
Wednesday Sessions

- **Jonathan Alt** - Representing Urban Cultural Geography in Stability Operations
- **Choo Chwee Seng** - Modeling the Effects of Human Intangibles
- **Philip Hingston** - Evolution and Intelligence
- **Gunther Schwarz** - Agent-based Modeling and Simulation MPECS Architecture
- **Ted Meyer** - Exploring Data
- **Luminita Stemate** - A Multi-Paradigm Approach to Crowd Control Modelling & Simulation



Outbriefs and Closing Session






Team 5: Sequential Screening for Organizational Performance

TEAM 5 MEMBERS

Regine Pei Tze Oh
Naval Postgraduate School, US
DSO National Laboratories, Singapore

Susan M. Sanchez, Ph.D.
Naval Postgraduate School, US

Lawton Clites
Digital Solutions Consulting, US
Naval Postgraduate School, US

Toh Kaiyang
DSO National Laboratories, Singapore

Abstract

Fractional Factorial Controlled Sequential Bifurcation (FFCSB) is a newly proposed two-phase screening procedure for large-scale simulation experiments. Sequential screening algorithms inform the decision maker of critical factors in their simulation models and optimize the use of computation resources in studying only critical factors.

At IDFW15, FFCSB is applied to the Hierarchy organizational model, which serves as a benchmark to compare innovative Command and Control (C2) structures for enabling more effective warfare. The model is developed in Projects, Organizations and Work for Edge Research, which crystallizes two decades of collaborative research between the Naval Postgraduate School (NPS) and Stanford University.

Motivation

Everyday, organizations use software simulations to make better decisions. Software simulations of real world systems are often large and rich with many parameters potentially affecting final outcomes. Faced with a multitude of parameters, decision makers may not know or may lose sight of the few truly critical factors. Thus, screening algorithms are essential in order to identify the factors that most impact outcome measures. This enables experimenters to better utilize their resources by focusing on truly important factors.

Fractional Factorial Controlled Sequential Bifurcation (FFCSB)

FFCSB is a newly proposed two-phase screening procedure for large-scale simulation experiments (Sanchez, Wan and Lucas, 2005.) FFCSB comprises two stages: (1) a fractional factorial (FF) pre-screening phase to sort factors by the direction of their effects (whether an increase in factor leads to a positive or negative change in the measure of performance) and (2) a controlled sequential bifurcation (CSB) to conduct sequential experimentation with accuracy guarantees.

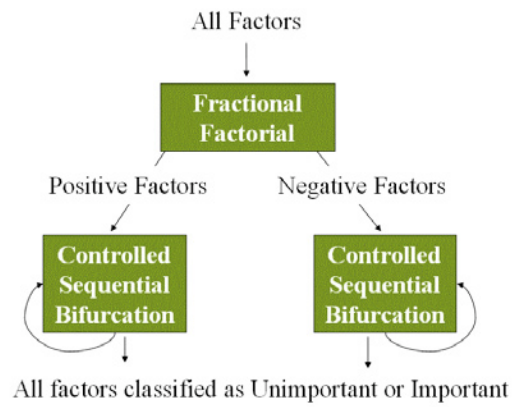


Figure 1: Conceptual Flow of FFCSB

FFCSB offers several enhancements over conventional screening algorithms. First, FFCSB dramatically reduces the need for *a priori* knowledge on the direction of factor effects, which is often a condition for optimal performance of conventional algorithms and has proven difficult to meet. FFCSB also does not require *a priori* knowledge of the number of experiments required for factor classification. It conducts sufficient experiments to complete classification. Second, FFCSB scales well for large scale models with thousands of factors. Third, FFCSB provides accuracy guarantees in its factor classification. Fourth, FFCSB provides a savings in computation.

Hierarchy Organizational Model in POW-ER

At IDFW15, FFCSB is applied to the Hierarchy organizational model developed by the Center for Edge Power (CEP) at NPS and Stanford University. CEP studies innovative C2 structures to enable more powerful warfare. The Hierarchy model is developed in POW-ER—Projects, Organizations and Work for Edge Research—a virtual environment for computational modeling of C2 organizations and processes. This computation tool crystallizes two decades of collaborative research between NPS and Stanford University. The tool is based upon sound research in organizational studies and has been validated extensively and thoroughly (Orr and Nissen 2006, p. 8; Levitt et al, 2005.)

The POW-ER environment uses agent-based simulation to emulate micro-behaviors (e.g. trust, learning, skill sets compatibility, skill competency, centralization) and discrete-event-simulation to emulate processes (e.g., meetings, exception occurrences, rework, process quality). Organizational performance is measured by quantitative metrics, such as project duration, project risk, project cost.

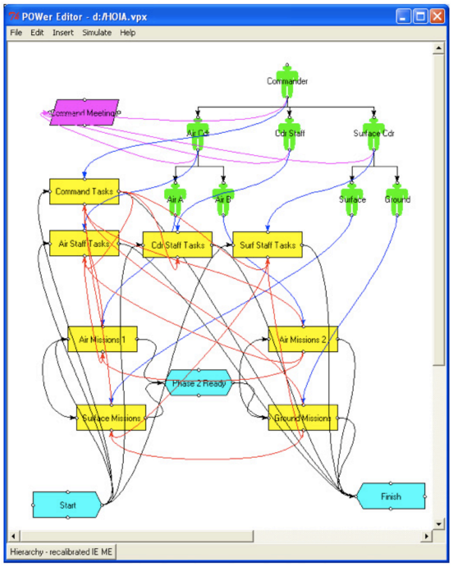


Figure 2: Hierarchy Organization Model in POW-ER

The Hierarchy model is modeled by three sets of structural factors: (1) organization structure (2) communication structure (3) work structure (Nissen 2005 p. 11.) The model is simulated in two contrasting mission contexts: Industrial Age and 21st Century. The mission contexts are modeled by three manipulations of mission factors: (1) mission and environmental context, (2) network architecture and (3) professional competency (Nissen 2005, p. 14.)

In computational experiments on the Hierarchy and other organizational models, the researchers typically varied one factor group (i.e., multiple factors) at a time, and record multiple Measures of Performances (MOPs) associated with each factor group change. The various organization models are then compared using these results and changes in performance are justified by the experimental manipulations.

FFCSB extends CEP's suite of tools for computational experimentation. Through smart and efficient designs of experiments, FFCSB identifies single critical factors that most impact the single MOP of Project Duration.

Methodology of Applying FFCSB on Hierarchy Model

Subject matter experts selected Project Duration as the MOP of interest for FFCSB application. They also identified 114 factors of interest, with associated 2-level factor ranges for exploration.

Working within the computation and time constraints of IDFW15, the team refined the factor ranges to smaller ranges of interests and divided the entire factor space into three smaller subspaces. The factor space is shown in the Table 1. These factors categories are intended to mirror those used in prior experimentation (Gateau et al., 2007, pp. 7-8) but may not be exact.

Mission & Environment
Function Exception Probability
Project Exception Probability
Task Effort Required
Task Learning Days
Task Priority
Task Requirement Complexity
Task Solution Complexity
Task Uncertainty
Full Time Equivalent (Manpower available)
Network Architecture
Mission Priority
Length Of Work-day
Length Of Work-week
Centralization
Matrix-strength
Communication Probability
Noise Probability
Instance Exception Probability
Meeting Priority
Meeting Duration
Meeting Allocation
Rework Strength
Professional Competency
Team Experience
Staff Culture
Role
Application Experience
Cultural Experience
Skill Ratings

Table 1: Factor Spaces of Exploration for the Hierarchy Model

Results of FFCSB Exploration

The following tables (2-3) summarize the FFCSB findings of important factors in the Hierarchy model that most impact Project Duration. There were no factors classified as important in the Network Architecture factor subspace.

Object	Attribute	Effect Direction
Mission	Project Exception Probability	+
Surface Msn	Effort	+
Surface Msn	Solution Complexity	+
Ground Msn	Effort	+
Ground Msn	Requirement Complexity	+
Ground Msn	Solution Complexity	+

Table 2: Important Factors in Mission & Environment Factor Subspace

Object	Attribute	Effect Direction
Mission	Team Experience	+
Air A	Air Skill Ratings	-
Ground	Ground Skill Ratings	-

Table 3: Important Factors in Professional Competency Factor Subspace

In the first factor subspace of Mission & Environment, SMEs identified the factors of Full Time Equivalent (FTE) and Effort as important. FTE measures the equivalent of manpower resources available and Task Effort quantifies the time effort requirement of the task. Contrary to expert opinion, FFCSB did not classify any FTE factors as important over the factor range of exploration. Therein lies our first surprise: FTE is not as important as the other factors in this subspace in impacting the Project Duration. In line with expert opinion, FFCSB classified Effort factors as important, but only for Surface Missions and Ground Missions out of all eight missions in the Hierarchy model. Critical path analysis of the Hierarchy model explains why factors associated with only these two missions showed up consistently as important.

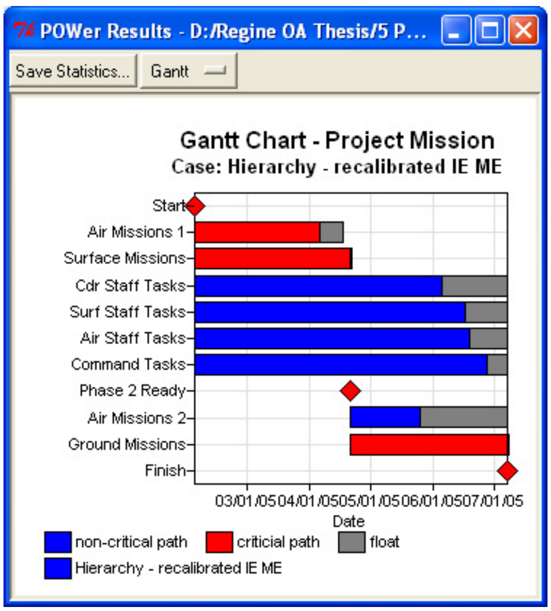


Figure 3: Critical Path Analysis of Hierarchy model shows Air Missions 1, Surface, and Ground Missions on Critical Path

The red bars in Figure 3 depict the critical path of the project simulated in the Hierarchy model. Following the red bars, the Air Missions 1, Surface Missions and Ground Missions are on the critical path. Of these three missions, the Surface Missions and Ground Missions have minimum float, i.e., there is no allowance for shifting these missions in time. Hence, these two missions are crucial to the MOP of Project Duration. Besides the Task Effort factor, FFCSB also classified the Solution Complexity factors of the Surface and Ground Missions as important, as well as the Requirements Complexity of the Ground Missions. This is our second surprise: FFCSB has further quantified expert opinion by flagging those factors associated with missions on the critical path only and with specific characteristics.

In addition, FFCSB classified the global factor of Project Exception Probability (PEP) as important. PEP is the probability that a subtask will fail and generate rework for failure dependent tasks. This factor is significant for the Hierarchy model that is characterized by sequential and interdependent tasks and hence, suffers a longer Project Duration in the event of increased PEP.

Our third surprise is: In the second factor subspace of Network Architecture, there are no factors classified as important for the particular factor ranges explored. This finding is in agreement with SMEs, who did not expect any important factors in this subspace. A set of (relatively computationally expensive) Resolution V Fractional Factorials design was used to verify the factor coefficients in this factor group. The results confirmed that the factor coefficients were relatively small in magnitude and hence, practically insignificant.

In the third factor subspace of Professional Competency, experts identified Skill Ratings and Application Experience factors as important. FFCSB classified the Skill Ratings of the Air A and Ground personnel as important, but not that of the Surface personnel. These three groups of personnel are responsible for the missions on the critical path. The contrast between the three missions is that the Surface mission requires considerably more effort of 21 months versus that of the Air Missions 1 (11 months) and Ground Missions (6.5 months). These findings suggest that Skill Ratings may be more critical for missions that lie on the critical path and have relatively shorter Effort requirements. FFCSB did not classify Application Experience as important.

Interestingly, FFCSB classified Team Experience as important and positively related to the MOP. Team Experience quantifies the degree of familiarity that team members have in working with one another as a team. In other words, this finding suggests that more team experience leads to longer Project Duration in the Hierarchy model. This is our fourth surprise. This counter-intuitive finding may have been observed in earlier research and experimentation. Ramsey and Levitt (2005) summarized high level findings from Horii, Jin and Levitt's "Modeling and Analyzing Cultural Influences on Team Performance through Virtual Experiments" (2004) on the impact of cultural differences in project teams: "Japanese-style organizations were more effective, with either US or Japanese agents, at performing tasks with high interdependence when the team experience of members was low." The Hierarchy model studied in this application shares common characteristics of centralized authority, high formalization, and multiple hierarchies with the Japanese-style organization modeled in Horii, Jin and Levitt (2004, pp. 3). In addition, these experiments had used the MOPs of Project Duration and Quality Risk to quantify team performance, while this FFCSB application only used Project Duration. Hence, there is common ground to compare the similarity of both findings. Had the original intuition on Team Experience been applied with conventional screening algorithms, this factor could have distorted screening findings.

Lastly, there were two interesting observations. The Hierarchy model has a 3-tier command chain that models the Command, Coordination and Operations layers in a Joint Task Force. There were more important factors associated with the Operations layer than the other layers. Second, there were more uncontrollable or difficult to control factors (e.g., Project Exception Probability, Task Requirement Complexity, Task Solution Complexity and Team Experience) than controllable or easy to control factors (e.g., Skill Ratings.)

Way Ahead

There are limitations to the FFCSB application to any model. FFCSB assumes a main effects model, and interactions can distort the accuracy of factor classification. The nature of the response variance (homogeneous or heterogeneous) and its magnitude are unknown. Both model characteristics could have bearings on the FFCSB findings and accuracy guarantees. Particular to the Hierarchy model, the observations of this FFCSB exploration are unique to the factor space organization and ranges of exploration. Hence, the findings are not conclusive of the Hierarchy model. The important factor classification and observations are meant to provide direction for researchers in future work and optimize their experimentation budget on truly important factors.

The team is greatly encouraged by the findings of the first-case application of FFCSB on a real world simulation model. There were interesting findings and many delightful surprises. Initially, some findings appeared counter-intuitive to the data-farmers but were later justified through critical path analysis and through comparison with earlier research on similar models. Hence, it is an encouraging sign that FFCSB can serve as a complementary tool to better understand complex simulation models.

References

[1] Gateau, J.B., Leweling, T.A., Looney, J.P. and Nissen, M.E. (2007). "Hypothesis Testing of Edge Organizations: Modeling the C2 Organization Design Space," Proceedings International Command & Control Research & Technology Symposium, Newport, RI (June 2007) Available at http://www.nps.navy.mil/GSOIS/cep/docs/2007/ICCRTS07_Gateau.pdf [Last Accessed October 10, 2007].

[2] Horii, T., Jin, Y. and Levitt, R.E. (2004). "Modeling and Analyzing Cultural Influences on Team Performance through Virtual Experiments," (2004). Available at http://crgp.stanford.edu/publications/conference_papers/TamakiCMOT.pdf [Last Accessed October 23, 2007].

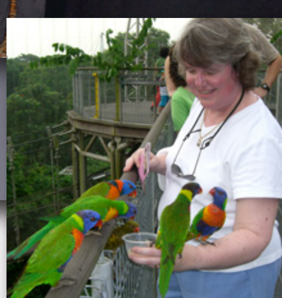
[3] Ramsey, M. and Levitt, R.E. (2005). "A Computational Framework for Experimentation with Edge Organizations," Proceedings International Command & Control Research & Technology Symposium, McLean, VA (June 2005). Available at http://www.nps.navy.mil/GSOIS/cep/docs/2005/Ramsey-Levitt-ICCRTS2005.pdf [Last Accessed October 10, 2007].

[4] Sanchez, S.M., Wan, H. and Lucas, T.W. (2005). "A two-phase screening procedure for simulation experiments," Proceedings of the 2005 Winter Simulation Conference, pp. 223-230. Available at http://www.informs-sim.org/wsc05papers/023.pdf [Last Accessed March 23, 2007].

[5] Orr, R.J. and Nissen, M.E. (2006). "Computational Experimentation on C2 Models," Proceedings International Command and Control Research and Technology Symposium, Cambridge, UK (September 2006). Available at http://www.nps.navy.mil/GSOIS/cep/docs/2006/Orr-Nissen057-ICCRTS2006.pdf [Last Accessed October 10, 2007].

[6] Levitt, R.E., Orr, R.J. and Nissen, M.E. (2005). "Validating the Virtual Design Team (VDT) Computational Modeling Environment," The Collaboratory for Research on Global Projects, Working Paper Series #25, pp. 1-15. (September 2005) Available at: http://crgp.stanford.edu/publications/working_papers/WP25.pdf [Last Accessed October 10, 2007].

[7] Nissen, M.E. (2005). "Hypothesis Testing of Edge Organizations: Specifying Computational C2 Models for Experimentation," Proceedings International Command & Control Research Symposium, McLean, VA (June 2005). Available at http://www.nps.navy.mil/GSOIS/cep/docs/2005/Nissen64-ICCRTS2005.pdf [Last Accessed October 10, 2007].



Team 6: Application of Design of Experiments & Data Farming Techniques for Planning Tests in a Joint Mission Environment

TEAM 6 MEMBERS

Timothy Beach
David Dryer
Joint Test and Evaluation Methodology (JTEM), US

Susan Sanchez
Steve Upton
Naval Postgraduate School, US

John Alt
TRADOC Analysis Center - Monterey, US

Victor Askman
Chris Richmond
REFERENTIA, US

Oliver Henne
BWB, Germany

Lee Kah Hock
Lim Hang Sheng
Yeo Lee Cheng
DSTA, Singapore

Ivy Lee Siew Yang
ST-TS, Singapore

Dave Ang Choon Kiat
DSO-ORL. Singapore

INTRODUCTION

The use of design of experiment (DOE) and data farming techniques is critical to effectively planning, and subsequently evaluating, tests of complex adaptive systems in a joint mission environment. The Joint Test and Evaluation Methodology (JTEM) program, in conjunction with the SEED Center at the Naval Postgraduate School (NPS), and TRADOC Analysis Center-Monterey, is developing methods and processes that incorporate these techniques into the development of the "test and evaluation strategy" phase of the Capability Test Methodology (CTM). In order to structure the underlying business rules and concepts in the CTM's evaluation thread, a Capability Evaluation Metamodel (CEM) is being developed.

The CEM measures framework consists of mission measures of effectiveness (MOE), task measures of performance (MOP), and system/system of systems (SoS) attributes. The mission MOEs assess the contribution of the system under test (SUT) and SoS capabilities to achieving the mission desired effects. Task MOPs address essential task performance related to the identified test issues (critical joint issues [CJI], critical operational issues [COI]). System/SoS attribute measures are used to evaluate the achievement of system or SoS performance across doctrine, organization, training, materiel, leadership and education, personnel, and facilities (DOTMLPF). The evaluation measures for the CTM map from mission to task to attribute level is shown figure 1.

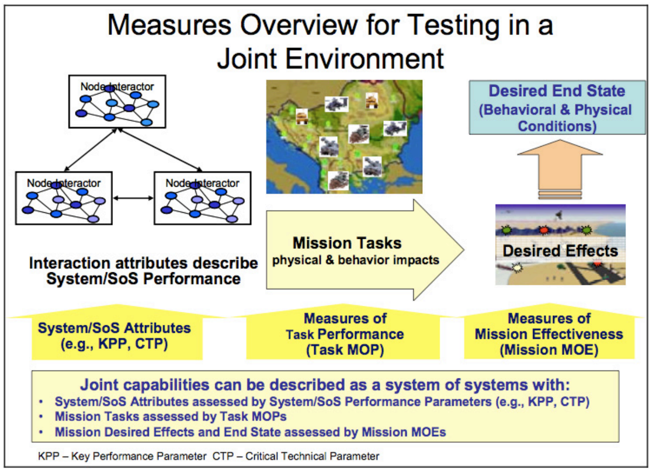


Figure 1: CEM Measures Overview

During IDFW 14 working sessions, team participants presented and refined a representative use case to define input factors and levels of an executable capability test design. Dimensions of this test design are mission, SoS, and mission conditions (including threat force and environmental conditions). This exercise of applying joint mission-level capability concepts to the structure of an efficient DOE provided a basis for use of executable design of experiment analysis as part of developing and refining the CEM. Team participants also discussed an integrated set of visualization, modeling, analysis, and simulation (VMAS) catalysts required to operate on CEM structures. Potential VMAS catalysts include test design visualization, statistical DOEs, simulation model classes and hybrids, as well as simulation analysis and visualization techniques which can fill capability evaluation gaps in the front-end part of the CTM evaluation thread.

APPROACH

During IDFW 15, Team 6's goal was to demonstrate and improve SoS test scoping and design techniques, as they pertain to capabilities supporting joint missions. We used a four-part approach to address this data farming goal:

- Characterize past and hypothetical capability design use cases, including the refinement of a hypothetical CJI and CEM test designs.
- Apply DOE techniques to initial multi-factor CEM capability designs. The DOE focus was on Nearly Orthogonal Latin Hypercube (NOLH) and Fractional Factorial Controlled Sequential Bifurcation (FFCSB) DOE techniques, which are being enhanced at the NPS.

- Conduct data farming on selected factors of this efficient DOE using Tester and Pythagoras agent based simulation (ABS) models.
- Conduct initial analysis of ABS results.

The following questions were addressed by Team 6 activities during IDFW 15:

- Given a past use case limited by JTEM FY07 event constraints (limited SoS composition, partial task thread), which factors are the most important to look at for testing?
- Using a hypothetical use case (FY07 event constraints removed) and a refined CJI, which factors are the most important to look at for testing?

During the workshop, Team 6 conducted "test and evaluation strategy" activities to design and focus the test space for the above stated use cases. These activities included validation of previously developed CEM concepts including CJIs with associated measures, levels, and factors for input into existing models. An iterative approach for refinement of large capability test spaces was exercised involving: developing an efficient capability design of experiment, running model designs on a high performance computing cluster, performing multivariate analysis using the JMP statistical analysis tool, and discussing model design refinement to focus on important factors. The results of this workshop provided valuable input into the development and improvement of the CTM.

For the past use case, Team 6 used version 1.0 of Tester, an agent-based model developed in the MASON framework, to perform exploratory data farming in support of JTEM's data farming for test planning effort. This data farming effort included exploring the results from a recent SoS test event, expanding the bounds of the test results, and understanding the functionality of the model at this point. An NOLH DOE was used for the past use case which included the following independent variable factors:

- Time to clear airspace
- Aircraft speed
- Air support request (ASR) approval time
- Combined Air Operations Center (CAOC) restricted fire zone (RFZ) time
- Close air support (CAS) tasking time
- CAS Joint Terminal Attack Controller (JTAC) coordination time

The dependent response variable in the past use case DOE was CAS Elapsed Time, which is an example of a task MOP in the CEM measures framework.

The following hypothetical CJI was refined by Team 6 in order to create a capability design for the hypothetical use case:

Assess Integrated Fires Command and Control (C2) task performance (for example, CAS, Fires, Dynamic Targeting) to achieve joint forcible entry operations desired effects (for example, threat ineffectiveness, Blue survivability) by an Integrated Fires SoS compositions (for example, C2 system focus, force application system enablers).

CJIs are used to assess performance pertaining to capabilities which support joint missions. A CJI for test and evaluation should be carefully structured to address key capability gaps described in joint capability documentation. The essential elements of a CJI include a capability's essential tasks, mission desired effects, Blue SoS (across DOTMLPF), and conditions involving threat and environmental factors. CJIs should address the SoS capability to perform joint operational tasks and/or the SoS, system, or Service attribute performance.

Based on the hypothetical CJI, a set of capability test design independent variable factors were created based on the factor dimensions of the CEM Joint Operational Context for Test (JOC-T). The JOC-T describes the overall philosophy of forces operating jointly and the tactics, techniques, and procedures (TTP) to be employed to achieve effects on the battlefield by exhibiting capabilities they will not possess separately. The JOC-T should specify mission objectives (for example, mission statement, mission desired effects, mission end state), Blue forces (for example, system capability requirements, system means to implement those requirements, system operating limitations, SoS context, task organization, C2 structure, force lay down with logical groupings of primary nodes), Blue actions, environment conditions (for example, physical conditions, civil conditions, neutral forces), threat forces (for example, threat order of battle, C2 structure, systems, threat lay down), threat actions, and interactions (for example, Blue to threat, Blue to Blue, threat to threat, Blue to environment, threat to environment).

Using these JOC-T dimensions, an initial hypothetical capability test design was developed containing 37 continuous factors and 25 categorical factors for a total number of 62 possible independent variables. To perform preliminary data farming in support of hypothetical use case, the Pythagoras ABS model was used on a selected subset of the 62 possible factors that could be modeled in the current Pythagoras functionality. The DOE used for the hypothetical use case was an NOLH which included the following independent variable factors:

- Organic SA persistence
- Inorganic SA persistence
- Organic fuse radius
- Organic fuse time
- Inorganic fuse radius
- Inorganic fuse time
- Communication delay
- Sensor offset X
- Sensor offset Y

The dependent response variable in the hypothetical use case DOE was Blue survivability, which is an example of a mission MOE in the CEM measures framework.

RESULTS AND ANALYSIS

The past use case using version 1.0 of Tester produced modeling results as shown in the classification and regression tree (CART) analysis shown in figure 2. The most significant factor was the CAOC RFZ time, followed by CAS tasking time and ASR approval time. These three factors accounted for a 46% degradation of the CAS Elapsed Time Task MOP. This was a surprise to the majority of team members who had predicted force application factors, such as Time to clear airspace and Aircraft speed would be most significant. Instead, for this modeling implementation, C2 factors centered at the CAOC were dominant.

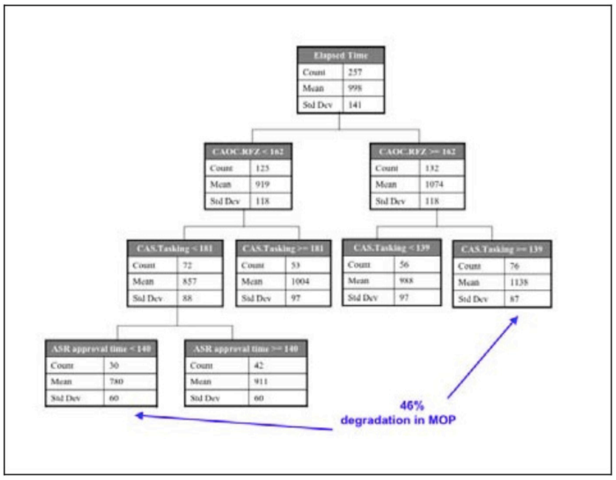


Figure 2: Past Use Case Tester 1.0 Model Results

To set the stage for the hypothetical DOE analysis, Dr. Sanchez described the strengths and weaknesses of various potential DOE techniques for large factor capability test designs using the following comparative analysis. Many designs are possible for the hypothetical case which could create between 64 and 663,552 design pts. Response 3 or Response 5 Fractional Factorial, FFCSB, a variety of NOLHs, and crossed/hybrid DOEs are potential design techniques. The choice depends on analysis considerations including:

- Types of insights desired
- Main effects, interactions, detail in looking at MOEs
- Number of factors
- Mix of continuous/discrete/qualitative factors
- Simulation run times
- Computational budget

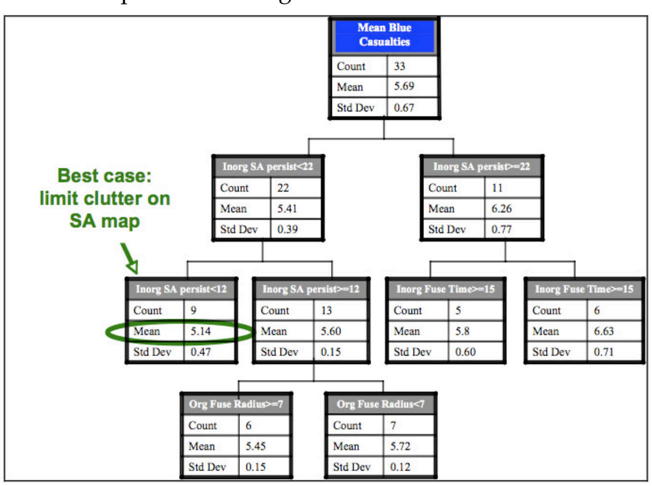


Figure 3: Hypothetical Use Case Pythagoras Model Results

The hypothetical use case using the Pythagoras ABS produced modeling results as shown in the CART analysis, shown in figure 3. The most significant factor was Inorganic Situational Awareness Persistence. Inorganic is defined as external to the "organic" unit modeled in Pythagoras. Persistence relates to the duration of threat target information Inorganic Fusion Time and Organic Fusion Radius. The best case design (least Blue causalities in the mission MOE response) related to limiting clutter on the Blue situational awareness map. This was also a surprise to many team members who had predicted physics-based force application factors, such as sensor offsets, would be most important to limiting Blue casualties. Instead, for this modeling implementation, situational awareness and C2 fusion factors were dominant.

INSIGHTS AND RECOMMENDATIONS

Team 6 insights included the need to further expand DOE parameters for capability testing and that the importance of non-materiel C2 factors was underestimated in the team's ABS results. Another surprise during IDFW 15 was the fact that Aloe Vera is more than just lotion; it also comes in yogurt form. Takeaways for the CTM included the benefits of applying hybrid DOE approaches, successful demonstrations of an iterative design – model/scenario development – analysis approach, and the need for supporting tools to enable data farming as part of CTM capability test design refinement and evaluation.

Recommendations from IDFW 15 Team 6 activities center on enhancement of data farming techniques for further application to capability based assessment. Specifically, enhancements to the Tester and Pythagoras ABS models are recommended for higher fidelity, more efficient capability explorations. These enhancements involve both ABS functionality, as well the automation of scenario input and MOE output analysis. The refinement of DOE techniques for large factor, multiple response designs is also identified. Developing functionality for guided selections of appropriate DOE techniques is also highlighted by the team as a promising way ahead, based on initial results. Follow-on efforts include the incorporation of data farming approaches into the CTM version 2.0 and 3.0, potentially including efficient design of experiments, the use of computing clusters, and the iterative data farming process. The intent is to explore enhanced DOE techniques and models at IDFW 16. Both use case explorations, in addition to providing data for analysis, helped in the development of the data farming infrastructure for the evolving CTM development by JTEM and its partners.

Team 7: Topological Analysis of Infrastructure Network

TEAM 7 MEMBERS

SIM Mong Soon, Ph.D. – Lead
Chin Shi Katt
Lai Kah Wah
Xu Yong Liang
DSO National Laboratories, Singapore

Palvannan Kannapiran
DSTA, SIngapore

Alberto Garcia, LCDR
Naval Postgraduate School, US

Lim Chio Tee
SAF Ops Research Office, Singapore

INTRODUCTION

Critical Infrastructures work together to produce goods and services. For example, the power station generates electricity and the water purification station uses the electricity to produce drinking water. Disruption of Civil Infrastructures will affect our national security, economic well being and way of life. This provides a primary motivation to model and understand the interaction between infrastructures. Based on our works in military modeling and simulation (M&S), we have extended these M&S methodologies to the area of Critical Infrastructure Protection. However, we observed that it take a reasonable amount of modeling effort to model a large network of infrastructures.

There is a need to provide quick answer to operational users. It is sometimes sufficient to have a ballpark estimate of the consequence of a possible disruption in the infrastructure network. This will also assist them to decide on a smaller subset of infrastructures to do detailed modeling and analysis later.

In this workshop, we are proposing to model each infrastructure & its interactions by Network Topology. We implemented this methodology in the NetLogo software. The following sections will discuss our objective for this workshop, describe the scenarios ran during the workshop and highlight some lessons learnt.

DURING THE WORKSHOP

Workshop Objective

In this workshop, we used a case study to evaluate the feasibility of this methodology to model the interaction between infrastructures. The case study here is a military supply chain network model. The initial military supply chain model is developed by Vidal [1]. For the purpose of this workshop, we added some modifications to this model.

The objective is to design a military supply chain with better survivability. The model consists of three types of battalions namely, Main Supply Battalion (MSB), Forward Supply Battalion (FSB) and combat battalion. A MSB supplies goods to a group of FSBs and each FSB in turn feeds the goods to a number of combat battalions. We are interested to determine how different configurations of MSBs, FSBs and battalion will affect the network survivability.

Network survivability

Survivability is defined as Robustness, Responsiveness and Flexibility. Robustness looks at how the size of the network changes when some nodes are removed. Responsiveness measures how quickly some commodities can flow through a network when some nodes are attacked and fail. Flexibility focuses on whether alternate paths exist in a network so that commodity can continue to flow to others nodes after some node failures. To measure these parameters, we have to compute the characteristic path length and the largest component of the network. The characteristics path length calculates the average number of links required to connect each node to every other nodes in the network while the largest component of the network determines the maximum number of nodes that continue to link to each other after some nodes are removed.

To look at how network structures will affect the network survivability, we considered three structures, namely Random network, Scale-free network and UltraLog network. Researchers have already studied the behavior of Random and Scale-free networks under different attack modes. In general, there are two attack modes, random attack and targeted attack. For random attack, the attacker chooses a node to disrupt at random. For targeted attack, the attacker has some information on how the nodes are linked to each other and will choose the most critical node for attack.

In a Scale-free network, when a new node joins the network, the probability that it will attach to an existing node is proportional to the number of links that the node has. Hence, a node with the most number of links is more likely to attract new node. Studies have shown that most real networks behave like the Scale-free network.

A Scale-free network is known to be resilient to random attack but is very vulnerable to targeted attack. This is because in a Scale-free network, there exist a small number of critical nodes with many links. Hence, when the attackers pick a node at random, it is less likely that these critical nodes will be chosen. This intuitively explains the resilient response of the Scale-free network under random attack. However, these critical nodes have a great influence on the survivability of the entire network and when the attackers focus on the critical nodes, the Scale-free network will suffer a serious consequence.

In Random network, nodes are attached to each other randomly. Hence, the response of a random network under random attack is not much different from that of a targeted attack.

Thadakamalla et al [2] proposed a network topology (herein known as UltraLog network) which is considered a hybrid between a Random network and a Scale-free network. It is suggested that the UltraLog model will be as efficient as the Scale-free network and yet perform better than the Scale-free network under targeted attack.

The UltraLog model is inherently hierarchical in nature. It consists of three layers, namely MSB, FSB and battalion. MSBs, FSBs and battalion enter the system in a certain ratio l:m:n where l > m > n:

a. A MSB has five edges pointing from it.
b. A FSB has three edges pointing from it.
c. A battalion has one edge pointing from it and a second edge added with a probability p.

%MSB	%FSB	%Battalion
3.3	13.3	83.3
6.3	31.3	63.5
7.1	28.6	64.3
8.3	25.0	66.7
10.0	20.0	70.0
9.1	45.5	45.5
11.1	44.4	44.4
14.3	42.9	42.9
20.0	40.0	40.0
12.5	12.5	75.0
20.0	20.0	60.0
33.3	33.3	33.3

Table 1: The configuration simulated during the workshop

Description of Scenario

In this workshop, we relaxed the ratio of MSBs, FSBs and battalion so that l ≥ m ≥ n. Table 1 shows the configurations studied during the workshop.

In addition to network configuration, we also modified the model so that there are four types of attack mode, namely random, targeted attack on critical node, targeted attack on critical node in largest component, targeted attack on critical node in smallest component. For more information on these attack modes, the readers can send the inquiry to the team.

We ran 30 replications of simulation for each configuration type and attack mode. After that, we compared the result with that of Scale-free and Random networks. The number of node used in each replication is 100 and one node is removed from the network at each time step. The simulation stops when 80% of the nodes are removed.

RESULTS AND ANALYSIS

For each configuration and attack mode, we computed the average value of the characteristic path length and the largest component at each time step and plotted them on a graph. Figure 1 and 2 show how the characteristic path length and the largest component behave for the three network structures for a given configuration, under targeted attack on critical node.

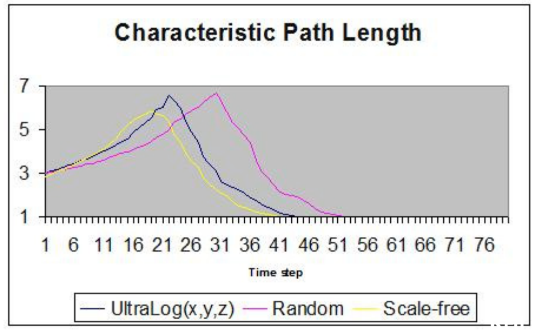


Figure 1: The average characteristic path length at each time step under targeted attack on critical node

In figure 1, it is observed that the rate of increase of the characteristic path length of the Scale-free network is higher than that of the other network structures. Furthermore, the Scale-free network starts to disintegrate at earlier time unit (i.e. time step = 21). On the other hand, the UltraLog model performs better than the Scale-free network but lag behind the random network.

In figure 2, it is again observed that the Scale-free network performs the worst among the three networks. For the Scale-free network, the value of largest component drops at a faster rate than that of the other two networks.

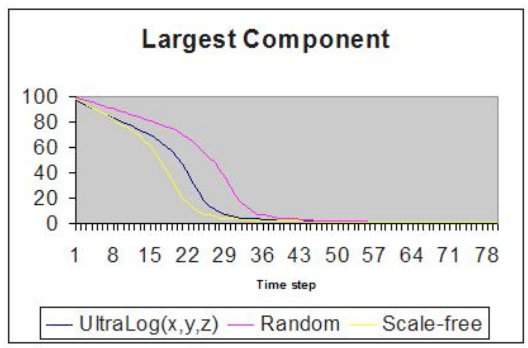


Figure 2: The average largest component at each time step under targeted attack on critical node

For each configuration and attack mode, we plotted the characteristic path length and the largest component for the three networks and made visual comparison. We developed a simple scoring system to assess the performance of the UltraLog network. The following criteria are used to award score to the UltraLog network:

a. Under random attack, if the UltraLog network performs better than the Random network and the Scale-free network, a score of 1 is given.
b. Under random attack, if the UltraLog network performs better than the Random network but lags behind the Scale-free network, a score of 1 is still given.
c. Under targeted attack, if the UltraLog network performs better than the Scale-free network and the Random network, a score of 1 is given.
d. Under targeted attack, if the UltraLog network performs better than the Scale-free network but lags

behind the Random network, a score of 1 is still given.

The result is shown in table 2.

SUMMARY OF FINDINGS

Based on the simulations ran, we observed that the UltraLog model performs better under the following conditions:

1. It is good to have an equal number of MSB, FSB and Battalion but impractical due to high implementation cost.
2. A more practical approach is to have an equal number of FSB and Battalion.
3. We will not recommend an equal number of MSB and FSB

From these observations, a possible rule of thumb is as follows:

Nos. of FSB = Nos. of Battalion > Nos. of MSB

We must caution that these results are obtained from simulations ran during the workshop. It is necessary to check that these observations are true for other configurations that are not ran during the workshop. Furthermore, it will be good to perform the simulation for other network size.

CONCLUSIONS

As conclusion, we found that this approach provide a "good enough" answer for a quick study and highlight some important trends in the result. The data-farming capability of NetLogo software allows the users to run multiple scenarios within reasonable time.

We also made the following improvements to the model:

a. Node recovery. In a real network, a node will recover after a disruption. Hence, it will be more realistic to consider how a node will recover in the model. For the workshop, we have implemented a simple node recovery mechanism.
b. Node replenishment. When a supplier is disrupted, a customer will source for a new supplier. We have also implemented a simple replenishment policy that a node will look for new node to link to, after its supplying node fails.
c. Output analysis. Apart from the capability to run multiple scenarios, the next important step is to facilitate the analysis of this huge amount of simulation outputs. We implemented the analysis tool in VBA and Java.

REFERENCES

[1] The initial military supply chain Netlogo model is developed by Vidal J. M. and is downloadable from: http://jmvidal.cse.sc.edu/netlogomas/

[2] Thadakamalla et al. Survivability of Multiagent-Based Supply Networks: A Topological Perspective. IEEE Intelligent Systems, Sept / Oct 2004.

Team 8: Enhancing the Combat ID Agent Based Model

TEAM 8 MEMBERS

Mink Spaans (Team leader)
Peter Petiet
TNO, Netherlands

David Dean
Dstl, UK

John Jackson, Maj.
Naval Postgraduate School, USA

Bradley Wilson
RAND, USA

Loh Yuan Shan
SAFOFO, Singapore

Wong Ka-Yoon
DSTA, Singapore

David Wang Yongwei
SCE-NTU, Singapore

Chan Wen Kai
SG Army, Singapore

INTRODUCTION

During previous Project Albert and International data Farming Workshops (IDFW) and during discussions between Dstl and TNO, the suitability and feasibility of Agent Based Models (ABMs) to support research on Combat Identification (Combat ID) was examined. The objective of this research is to:

> Investigate the effect of (a large number of) different variations in Situational Awareness, Situation Awareness (SA), Target Identification (Target ID), Human Factors, and Tactics, Techniques, and Procedures (TTP) under different circumstances (scenarios) on mission level combat effectiveness and fratricide.

Combat ID is a complex phenomenon which is heavily based on human factors, technology and tactical considerations. Modeling Combat ID to its full extent is not possible in a single step. It requires both a good combat model and a representation of the Target Detection, Classification, Identification process that takes the considerations mentioned above into account. As a first step to support our objectives, we decided to evaluate the feasibility to represent Situation Awareness in an ABM. This evaluation was conducted during IDFW14. Before and during this workshop, version 1.0 was developed in NETLOGO. This model contains one moving identifying agent and a number of static agents to be identified (objects). The identifying agent has a representation of situation Awareness (SA) and bases its identification decision on a mechanism where it combines SA and data from observations.

Current Features and Objectives

Following our overall Master plan, several extensions have been implemented since IDFW14:

1. When the identifying agent has not decided on the identity of a certain visited object, it is able to revisit the object and try to decide on its identity again.
2. When the agent decides that an object is an enemy, it kills the object. The object will then be removed from the ground truth.
3. The notion of Local SA and Global SA was introduced. Global SA keeps track of the pre-conception of the whole environment in which the agent operates. Its' granularity is less than the granularity of the ground truth. Local SA keeps track of the agents' preconception of its' surrounding area. The granularity of Local SA is equal to the granularity of the ground truth. The size of the local SA and the granularity of the global SA are parameterized (and thus data farmable). The local SA is updated each time new sensor information is accepted or as a result of moving. When the agent moves, the local SA grids moves with it, keeping the agent in the middle of it. As a result of the move, some cells will be removed from the local SA and new cells are added, taking the belief distribution of the global SA cell as its' initial belief. The global SA is updated each time the agent decides on the identity of an object. Figure 1 shows the relation between the Local SA, the Global SA and the ground truth.

The objectives of the study during IDFW15 were to assess the features above by designing and conducting data farming experiments. Further objectives were to (re-)examine and determine the key factors (parameters) in SA.

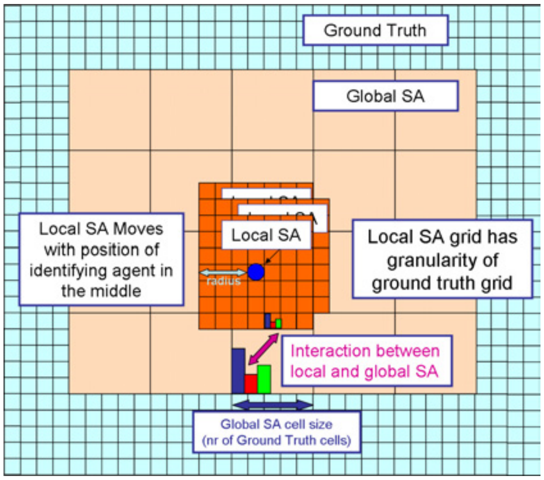


Figure 1: The notion of Local SA, Global SA and their intuitive interaction.

Design of Experiments

During IDFW15 we identified the interaction between Local SA and Global SA, and the mechanism to accept or reject new sensor information as the key factors to consider. We implemented four model versions with different ways to deal with these factors:

version 1.The agent keeps track of the kind of objects that were encountered in each global cell. After a positive identification, the global SA cell where the object is identified, is updated in a way that depends on the kind of objects that were encountered in that cell before. If no or only the same kind of objects were encountered before, the global cell will get the agent's local belief (probability distribution of Red, Blue, and Green). If one other kind was encountered before, it takes the average of the new belief and the old global belief. If two other kinds were encountered before, it takes the average of two times the old belief and one time the new belief.

version 2.The model incorporates a parameter "Belief Increase Steps" (BIS) to update the Global SA grid in steps towards a probability of 1 (100 percent sure). After a positive identification, the global SA is increased with "(1 - ⅓) / BIS". e.g. if BIS is 3, 0.22 is added.

version 3.This version incorporates the notion of surprise. A parameter "Surprise Level" will be implemented that determines a surprise curve that defines the amount of belief that will be added to the global belief when new information is accepted. The amount of belief to add, depends on the old belief in such a way that high old belief will add a low amount and vice versa (see Figure 2). The effect is that new belief that is in line with the old belief, only has a small effect. If the new belief contradicts the old belief (surprise!), the global belief is changed more radically.

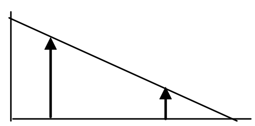


Figure 2: The surprise curve

version 4.This version uses the Surprise Variable as introduced in version 3. The shape of the "Information Acceptance" curves[1] are parameterized, Both the top and the crossing with the Y-axe can be defined by the user. See Figure 2. This enables to shape the information acceptance behavior and in particular solves some problems with old belief that is equal to 0 or 1.

SA Aspect	Version01	version02	version03	version04
Update of Global SA	Weighted Average of old and new belief	Believe Increase Steps	Surprise Level	Surprise Level
Information Acceptance curve	Fixed	Fixed	Fixed	Variable

Table 1 : Overview of different model versions

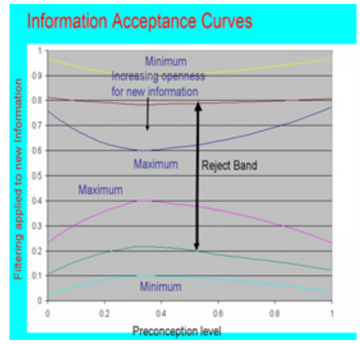


Figure 3: The information Acceptance curves

RESULTS

The Measures of Merit for the combat ID model versions are the number of correct identifications, the number of misidentifications, and the number of fratricide incidents. During the workshop we briefly evaluated model versions 1 and 2 and did runs with all four versions.

For version 1, the main results were that:

- The Decision threshold turned out to be the key factor in determining the number of correct identifications.
- Having a larger local SA grid size cannot overcome a perceived truth that differs greatly from the ground truth.
- Decreasing the decision threshold increases the number of correct identifications and misidentifications
- Outlying cases proved interesting, specifically, the largest number of misidentifications occurred when there was a:
  - Low decision threshold and, in general, a lower stress coefficient (which determines the information acceptance level)
  - Large delta between perceived and ground truth
- The more interspersed the Blue, Red, and Green objects are, the more fratricide
- The higher the level of stress, the less fratricide → same holds for version 4 , with information acceptance level

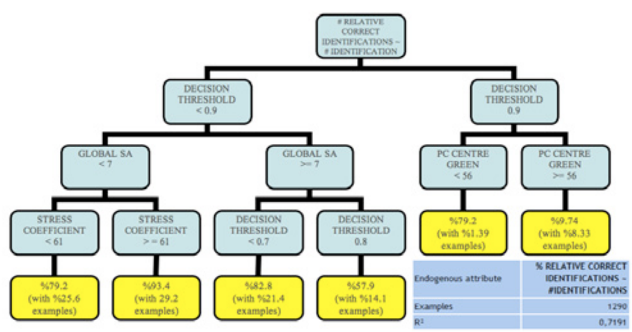


Figure 4: Regression tree for version 1 with relative number of correct identifications compared to total number of identifications

[1] The Information Acceptance curves were called stress curves before. This was actually misleading. The curves are really about the openness of the agent to accept new information, which can be effected by other factors than stress. However, in the results we sometimes still use the word stress.

During the workshop we agreed that different Measures of Effectiveness are needed to view the effects of all SA variables. Therefore, after the workshop, we combined the initial outputs to the relative number of correct identifications:

- The number of correct identifications related to the total number of identifications. This MoE directly relates the number of correct identifications to the number of misidentifications.
- The number of correct identifications related to the total number of existing objects. This MoE directly shows the agents performance in identifying objects (correctly).

Using these two MoE we created the regression trees shown in Figure 4 and Figure 5 to determine the most important factors for version 1.

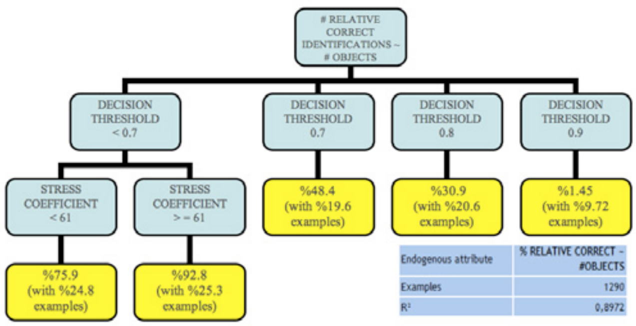


Figure 5: Regression tree for version 1 with relative number of correct identifications compared to total number of objects.

Figure 6 shows the contour plot for the two most important factors, decision threshold and level of stress, related to the number of correct identifications.

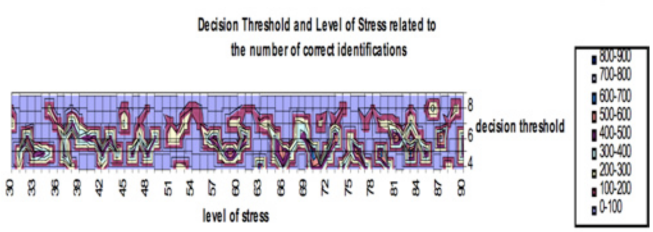


Figure 6: Number of correct identifications related to decision threshold and level of stress (information acceptance level)

For version 2, the main results were that:

- A higher resolution global SA (smaller cell size) can reduce the fratricide incidents, in case where the tank detection range is large
- Having a Large local SA decreases the number of misidentifications at a low classification range, but increases the number of misidentifications at large ranges

Further version results:

- Version 2: The regression trees for version 2 are not significantly different from the ones for version 3
- Version 2: There are no significant effects caused by the variable "Belief Increase Steps" introduced in version 2
- Version 3: The regression trees for version 3 are not significantly different from the ones for version 1 and 2
- Version 3: There are no significant effects caused by the variable "Surprise Level" introduced in version 3 (Note: at a higher Surprise level there seems to be a rise in Mis-Identifications. This needs further examination!)
- Version 4: In general, the higher the information acceptance level, meaning the more open the agent is for new information, the higher the number of correct identifications, and the less the number of misidentifications and fratricide incidents.
- Version 4: The regression tree of all variables related to the number of fratricide incidents is shown in Figure 7.

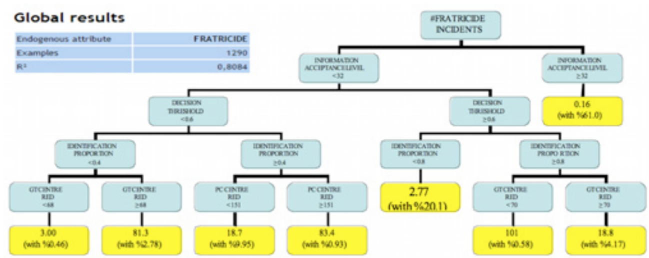


Figure 7: Version 4: Regression Tree regarding the number of fratricide incidents

- Version 4: A contour plot of the Y-Intercept (where the curve crosses the Y-axe) and the Information Acceptance Level related to the number of correct identification (relative to the number of the total number of identifications) is shown in Figure 8.

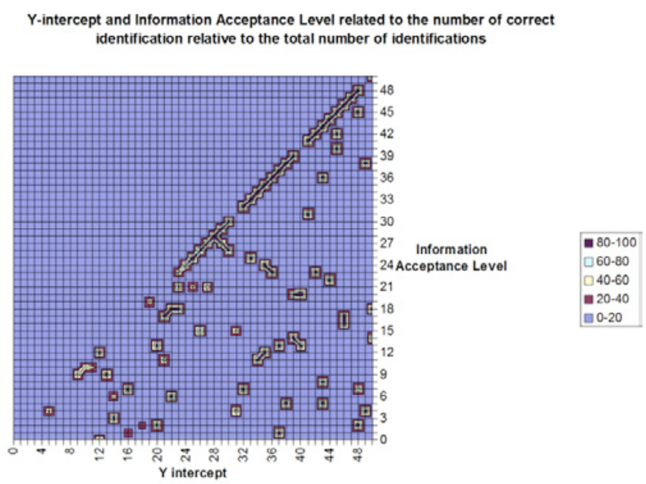


Figure 8: Contour plot Y-Intercept and Information Acceptance Level

CONCLUSIONS

We can draw the conclusion that team 8 made a lot of progress with the model development. We were able to discuss, implement, and run 4 model versions. Discussions about Situational Awareness, and how to capture and represent it in our model raised further issues that might lead to new features.

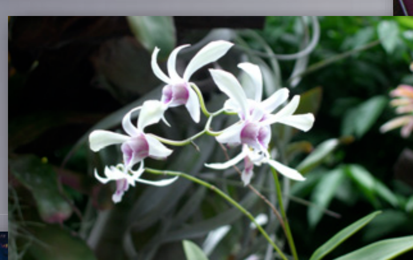
However, our analysis is far from complete yet. At this moment we gained a number of insights from the results of our first data farming efforts. Some of them in line with our expectation, others in contradiction with it. This needs further analysis and possibly more model changes and model runs.

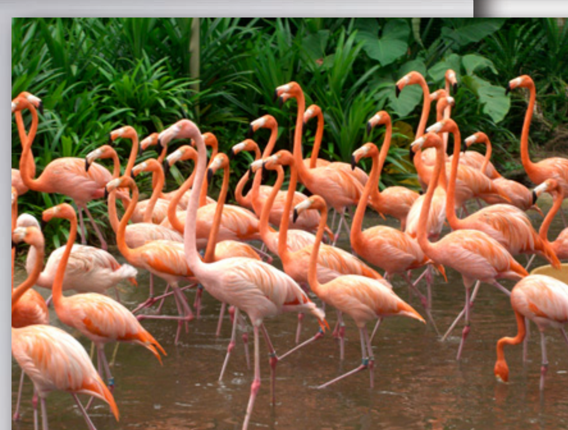
FUTURE DEVELOPMENT

Before the next workshop, we will dive deeper into the ocean of data to pinpoint more characteristics of our current versions. In the process of sowing and reaping, we will incorporate lessons learned and reconsider our model and assumptions continuously. We have an master plan that will serve as a guide for our near term development. Together with the lessons learned from the IDFW 15, we (re)consider current variables and new variables like:

- Other Measures of Merit
- The notion of Surprise
- Similarity of objects
- Environmental (~ sensor) distortions
- Further incorporation of INCIDER aspects
- Continuous Info Processing
- Change of awareness over time
- Notion of killing enemy and preconception
- BDI and awareness of fratricide
- Threat representation
- Incorporation of moving objects
- Incorporation of more identifying agents

IDFW16 in Monterey will serve as a vehicle to test some of the ideas mentioned above. Our plan is to have a stable model and focus on the datafarming process during this workshop. As usual we welcome new members to participate in our team during this workshop.





Clustering and Outlier Analysis For Data Mining (COADM)

Choo Chwee Seng[1]
DSO National Laboratories, Singapore
Ng Ee Chong
DSO National Laboratories, Singapore
Chua Ching Lian
DSO National Laboratories, Singapore

INTRODUCTION

The Clustering and Outlier Analysis for Data Mining (COADM) tool is one of the three key components delivered under the Systematic Data Farming (SDF) project [1]. SDF was sponsored by the Singapore Armed Forces (SAF) Centre for Military Experimentation (SCME) and was completed in 2005.

OBJECTIVE

The objective of COADM is to provide an additional dimension to data analysis, especially when there is a large amount of output generated through data farming. It aims to complement statistical analysis by grouping the data into “good” and “bad” clusters, and identifying the associated parameters so as to provide insights on how to get into “good” clusters and avoid the “bad” ones. COADM also identifies the outliers in each cluster, and in doing so try to discover “surprises”.

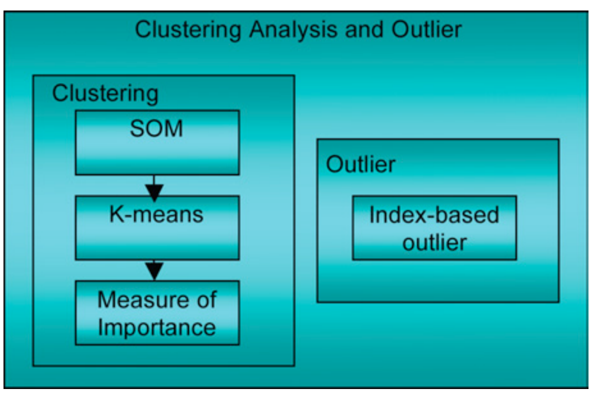


Figure 1: Key Features of COADM

KEY FEATURES

Figure 1 shows the key features of COADM and the underlying techniques and algorithms used.

The Clustering Analysis was based on K-Means methodology coupled with Self-Organising Maps (SOM) to help organise the data into clusters. The incorporation of K-means was to help improve the clustering and segregation capability of the SOM [2].

Based on the Clusters identified, a search was carried out within to identify the points that are “most different” from the rest of the data points within the same cluster, i.e. the outliers. This was achieved by comparing the Euclidean Distance of each data point with its k-nearest neighbour in each cluster and finding the one with the largest Euclidean Distance [3].

COADM was developed from several open source software packages and DSO contributions were in synthesizing the various algorithms/packages to form a package (coded in JAVA) capable of extracting information from numerical data sets. The SOM program used in this package was derived from the SOM toolbox in Matlab [3]. This toolbox is capable of visualizing complex data set, courtesy of Matlab’s great visualization tools; moreover it keeps track of much information which greatly facilitates the data mining process. The outlier algorithm was coded and modified slightly for integration with other packages. There is also a WEKA package provided as an extra data visualizations tool for a more detail examination of the clustering results.

DEMONSTRATION

Scenario

An Urban Scenario was used to demonstrate the key features of COADM (see Figure 2).

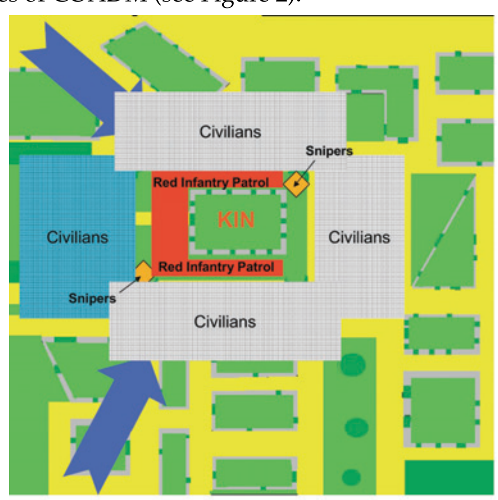


Figure 2: Urban Scenario Setup in MANA

An Urban Area of Operations (AO) 2km by 2km in size was set up in MANA. The scenario was set in this Urban AO where 2 platoons of Blue Infantry soldiers (21 soldiers

---

[1] For more information contact: Choo Chwee Seng, cchweese@dso.org.sg

per platoon), each platoon was supported by 3 MG-mounted soft-skin vehicles, attempted to take over a Key Installation (KIN) held by a platoon of Red Infantry soldiers (21 soldiers). The Red Infantry defence was assisted by two teams of Red snipers (4 snipers in total). The Blue agents' task was made more difficult by the crowd of hostile Civilians congregating near to the KIN and randomly attacking the Blue agents when they were encountered.

Blue Force has 3 Courses of Actions:

d. OCA 1. The Blue agents advanced from the northwest and southwest direction of the map towards the objective, attempting to take out the Red from both sides.

e. OCA 2. The Blue agents were concentrated in the southeast area of the map and advance as a force towards the Red, attempting to punch through the Red defence from a single direction.

f. OCA 3. The Blue agents were spread out on the northern portion of the map and attempted to flush out the read through a swarming approach.

Red Force has 2 Courses of Actions:

g. ECA 1 - All Red agents resided within the building's compound and defended their base from there.

h. ECA 2. A section minus of 6 Red agents lay hidden in an adjacent building as backup to the other two sections in the defended locality. They were called in when the Red agents came in contact with Blue Forces.

Design of Experiment

A hybrid design was formed using the Excel-based Latin Hypercube (LHC) Generator by crossing the 30-factor LHC with the 2-factor Full Factorial design for the OCA and ECA factors. The resultant hybrid design had 6000 design points and sent for data farming (with 100 replications each).

Analysis of Results

The large dataset of MOEs obtained from the data-farming output was analyzed using COADM and some interesting insights were derived. Figure 3 shows some of the selected component plots of the SOM clusters generated by the COADM. Similar distribution of colours on the component plots implies correlation. Hence correlation between the factors and the MOEs can be discovered. Factors found to be correlated to MOEs are also the main factors contributing to the MOEs.

Both the OCA and ECA factors were observed to be uncorrelated with the MOEs. The distribution patterns of the OCA and ECA factors (shown on Figure 3) were observed to be rather independent from the distribution patterns of the MOEs. Hence, varying the OCA and ECA would not contribute to significant changes to the MOEs.

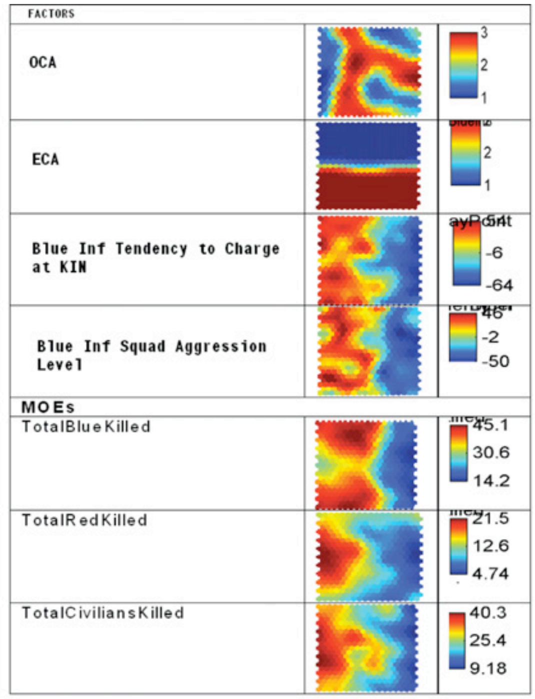


Figure 3: Component Plots of SOM clusters for selected Factors and MOEs

The MOEs were observed to be somewhat correlated. This suggested that achieving high Red attrition would likely coincide with high Blue and Civilian attrition levels. The Red and Civilian casualties were more closely correlated with each other compared with that of the Blue casualties. Therefore, it would suggest that larger number of civilian casualties was unavoidable in this scenario, if the Blue agents or Red agents attempted to maximize the casualties on either sides.

However, there were exceptions. A region that contained outcomes that corresponded to moderate Blue attrition but very high Red attrition was shown in Figure 4. This would be the region of most interest to Blue as the parameter values defined in this region allowed Blue to achieve its mission of killing as many Red as possible, while incurring moderate losses.

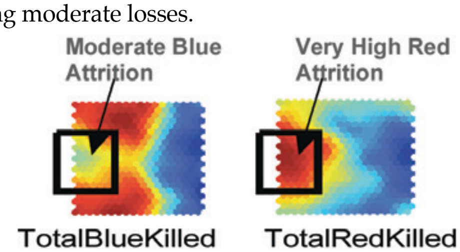
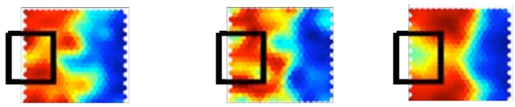


Figure 4: Region of Outcomes corresponding to Moderate Blue Attrition but Very High Red Attrition

Of the 32 farming parameters, it was observed that "Blue Infantry Tendency to Charge at KIN" and "Blue Infantry Squad Aggression Level" correlate most closely with the MOEs, and were hence most influential on the MOE outcomes.

It was interesting to revisit the region spotted under Figure 4, where Blue suffered moderate attrition but Red suffered high attrition. As shown in Figure 5, in this region, the parameter values for "Blue Infantry Tendency to Charge at KIN" and "Blue Infantry Aggression Level" should define the Blue's behavior that would inflict high Red attrition while sustaining moderate Blue attrition.



Blue Inf Tendency to Charge at KIN | Blue Inf Squad Aggression Level | Total Blue Killed

Figure 5: Comparison of Blue Inf Tendency to Charge at KIN, Blue Inf Squad Aggression Level, and Total Blue Killed

COADM tool revealed that the data points can be organized into 20 clusters (see Figure 6). The mean parameter values and MOEs for each cluster were obtained based on the data points within the cluster. By analyzing each cluster, it can identify the clusters that contained generally favorable outcomes for Blue and those that contained generally bad outcomes for Blue.

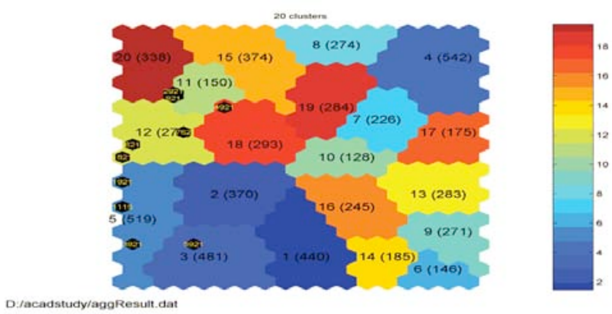


Figure 6: Clustering of Data Farming Output

COADM can also identify contributing factors and behavior that resulted in each of these clusters. Without going into each cluster in detail, with this analysis, Blue would know how to manipulate Blue factors and make decisions to avoid those bad clusters and shift towards the good clusters.

From the output generated by COADM, the outlier points were examined in greater detail and they were laid out in Table 1 in terms of the MOEs. The top outlier was case number 5921 (or Data Point 5921) amongst the 6000 cases in the Experimental Design. This case belonged to Cluster 3 and had 23.45 Red killed in total. COADM identified this case as an outlier because 23.45 red killed was 1.936 times more than Cluster 3's mean value of total Red killed. A value that is 1.5 times either side of the mean would normally be considered as an outlier.

In Cluster 3, Blue generally suffers high attrition and hence Blue should avoid parameter values that will cause them to fall into this cluster. This outlier Case 5921 is an interesting case because it is the best outcome in a bad cluster for the Blue, as Blue was able to inflict much higher Red attrition compared to other cases in Cluster 3.

Case 5921 described a Blue force that was very fast, highly aggressive and extremely stealthy. Although the Red force and Civilians were also generally aggressive, they were less so compared to the Blue force.

Hence, if factors uncontrollable by the Blue Force, such as Red Force tactics and behavior, resulted in the circumstances becoming unfavourable (e.g. falling into Cluster 3 outcomes), Blue force must attempt to exploit outlier case 5921 by moving swiftly and stealthily, and engaging more aggressively than the Red force inflict high Red casualties.

Case	Dist	Cluster	TotalBlueKilled	TotalRedKilled	TotalCiviliansKilled
5921	43.13	3	34.65 (+0.175)	23.45 (+1.936)	43.73 (+1.565)
4921	42.56	18	37.68 (+0.413)	22.88 (+1.838)	42.06 (+1.423)
1921	42.13	5	36.25 (+0.301)	23.63 (+1.966)	42.36 (+1.449)
921	41.93	11	37.89 (+0.430)	23.29 (+1.908)	40.92 (+1.327)
1115	41.31	5	40.47 (+0.633)	23.12 (+1.879)	46.93 (+1.835)
821	41.25	12	41.31 (+0.700)	21.83 (+1.657)	42.67 (+1.475)
2921	41.2	11	41.70 (+0.730)	20.24 (+1.385)	37.31 (+1.022)
1821	41.11	5	41.51 (+0.715)	20.69 (+1.462)	43.27 (+1.526)
3921	41.04	3	42.64 (+0.805)	20.34 (+1.402)	35.59 (+0.876)
762	40.99	12	29.84 (-0.205)	24.11 (+2.049)	45.98 (+1.755)

Table 1: MOEs in Outlier Cases

INSTALLATION

The installation requirements for COADM are as follows:

a. Java 1.4.2 and above.
b. Windows OS 2000/XP.
c. Memory recommended, 256MB Ram.
d. Disk storage space for files, 260MB

To request a copy of COADM, please contact Choo Chwee Seng at cchweese@dso.org.sg.

REFERENCES

[1] C. S. Choo, E. C. Ng, C. K. Ang, and C. L. Chua. Systematic Data Farming: An Application to a Military Scenario, Proceedings of Army Science Conference 2006, Florida, USA.

[2] J. Vesanto, E. Alhoniemi, K. Kiviluoto, and J. Parvianen. Self- Organizing Map for Data Mining in Matlab: The SOM Toolbox. www.cis.hut.fi/projects/somtoolbox.

[3] S. Ramaswamy, R. Rastogi, and K. Shim. 2000: Efficient algorithms for mining outliers from large datasets. In Proceedings of the ACM SIGMOD Conference, pages 427–438.

International Data Farming Workshop 16

When: 13 - 18 April 2008
Where: Portola Plaza Hotel, Monterey, California
Hotel information available at (http://www.portolaplazahotel.com)
Reservations: 866-711-1534

Note: Details on room rates and how to make reservations will be announced soon. Please go to http://harvest.nps.edu/ at the IDFW 16 link for additional information and registration.

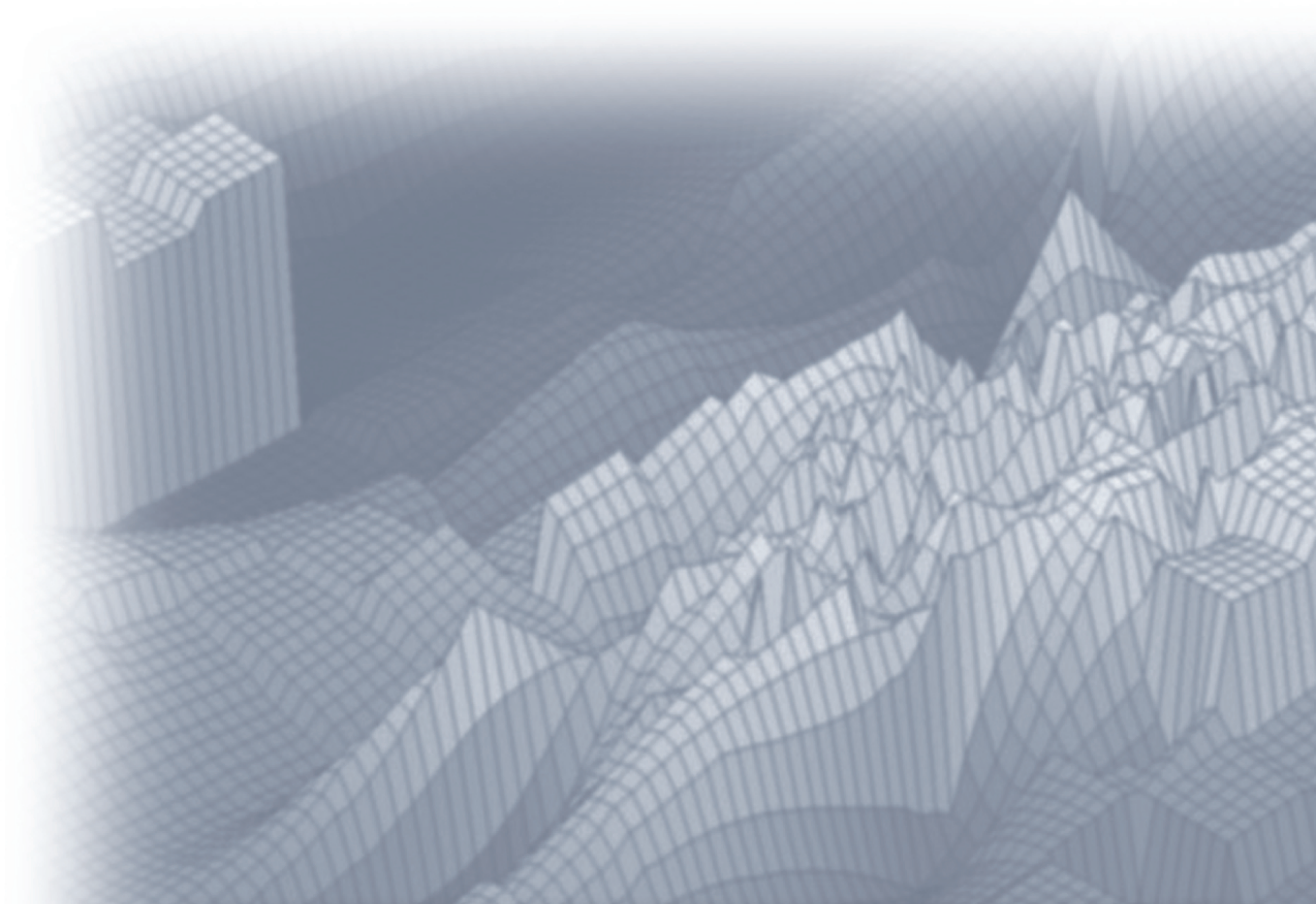
The Early Bird workshop fee, good through 29 February, will be is $496, afterwards the standard fee will be $550.

Tentative Agenda
Sunday, April 13: Opening reception and dinner at 1800 at the Portola Plaza
Monday, April 14: Opening briefs and team poster sessions in the morning, then begin work in teams
Tuesday - Thursday, April 15 - 17: Work in teams (optional plenary sessions in the mornings) at NPS
Friday, April 18: Outbriefs and Closing Ceremony in the morning at the Portola Plaza

Call for Team Leaders / Plenary Speakers:
Please email gehorne@nps.edu with your choice of teams and if you want to lead a team or present a plenary briefing.

Conference Fee:
The registration fee is $550 (Early Bird Registartion by February 29 is $496 US).
Registration pays for:
- Conference rooms • Opening dinner
- Lunch, Break food and drinks • CD & conference materials
- New one-year membership card with quote • Fun

The Data Farming CD/DVD, if provided, will be attached here. For additional copies of the CD or of the Scythe please contact Ted Meyer (tedmeyer@mac.com)



Scythe - Proceedings and Bulletin of the International Data Farming Community

Seed Center for Data Farming
273 Glasgow Hall ~ Naval Postgraduate School ~ Monterey, CA 93950

Issue 3 - Workshop 15