

Fall 2014 SEI Research Review

October 28–30, 2014

Summaries of Projects, Presentations, Workshops, and Posters

Copyright 2014 Carnegie Mellon University

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Department of Defense. References herein to any specific commercial product, process, or service by trade name, trade mark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by Carnegie Mellon University or its Software Engineering Institute.

NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

This material has been approved for public release and unlimited distribution except as restricted below.

Internal use:* Permission to reproduce this material and to prepare derivative works from this material for internal use is granted, provided the copyright and "No Warranty" statements are included with all reproductions and derivative works.

External use:* This material may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other external and/or commercial use. Requests for permission should be directed to the Software Engineering Institute at permission@sei.cmu.edu.

* These restrictions do not apply to U.S. government entities.

Carnegie Mellon® and CERT® are registered marks of Carnegie Mellon University.

DM-0001786 Copyright 2014 Carnegie Mellon University

CONTENTS

Welcome to the SEI's Annual Software and Cybersecurity Research Review . . .	3	Edge-Enabled Tactical Systems (EETS)	48
Projects and Presentations		Cybersecurity Expert Performance and Measurement	50
Elicitation of Unstated Requirements at Scale (EURS)	6	Automated Cyber-Readiness Evaluator (ACE)	51
Investment Model for Software Sustainment	8	Profiling, Tracking, and Monetizing: Analysis of Internet & Online Social Network Concerns	52
Value-Driven Incremental Development (VDID)	10	Aligning Software Architectures and Acquisition Strategies	53
Agile Adoption in the Department of Defense (DoD)	12	Workshops and Posters	
Acquisition Dynamics	14	Insider Threat Workshop	56
Quantifying Uncertainty in Early Lifecycle Cost Estimation (QUELCE)	16	Developing and Maintaining a Skilled Cyber Workforce Workshop	57
Software Model Checking for Verifying Distributed Algorithms.	18	Agile in Acquisition Workshop	58
Verifying Evolving Software	20	Incremental Lifecycle Assurance Through Architecture-Centric Virtual System Integration Workshop	59
Contract-Based Virtual Integration and CPS Analyses	21	Cyber Intelligence Research Consortium	60
High-Confidence Cyber-Physical Systems (HCCPS)	22	Data-Driven Decision Making for Software-Reliant Systems: The SEI Empirical Research Office	62
Software Assurance Engineering—Integrating Assurance into System and Software Engineering	24	References	64
Secure Coding	26		
Vulnerability Discovery	28		
Simulating Malicious Insiders in Real Host-Monitored Background Data	30		
Insider Threat Mitigation	32		
Deep Focus: Increasing User Depth of Field to Improve Threat Detection	34		
Malware Analysis	36		
Malware Distribution Networks	38		
Behavior-Based Analysis and Detection of Mobile Devices	39		
Data-Intensive Systems	40		
Graph Algorithms on Future Architectures	42		
Real-Time Mobile Applications in Intermittently Connected Networks	44		
Probabilistic Analysis of Time-Sensitive Systems	46		



Carnegie Mellon University

Software Engineering Institute

Welcome to the SEI's Annual Software and Cybersecurity Research Review



Kevin Fall, PhD
Deputy Director,
Research, and CTO
Carnegie Mellon University
Software Engineering Institute
kfall@cmu.edu

Welcome to Carnegie Mellon University and the Software Engineering Institute (SEI). Our Research Review is intended to bring together the government, academic, and industrial communities with whom we work and interact to highlight our research activities.

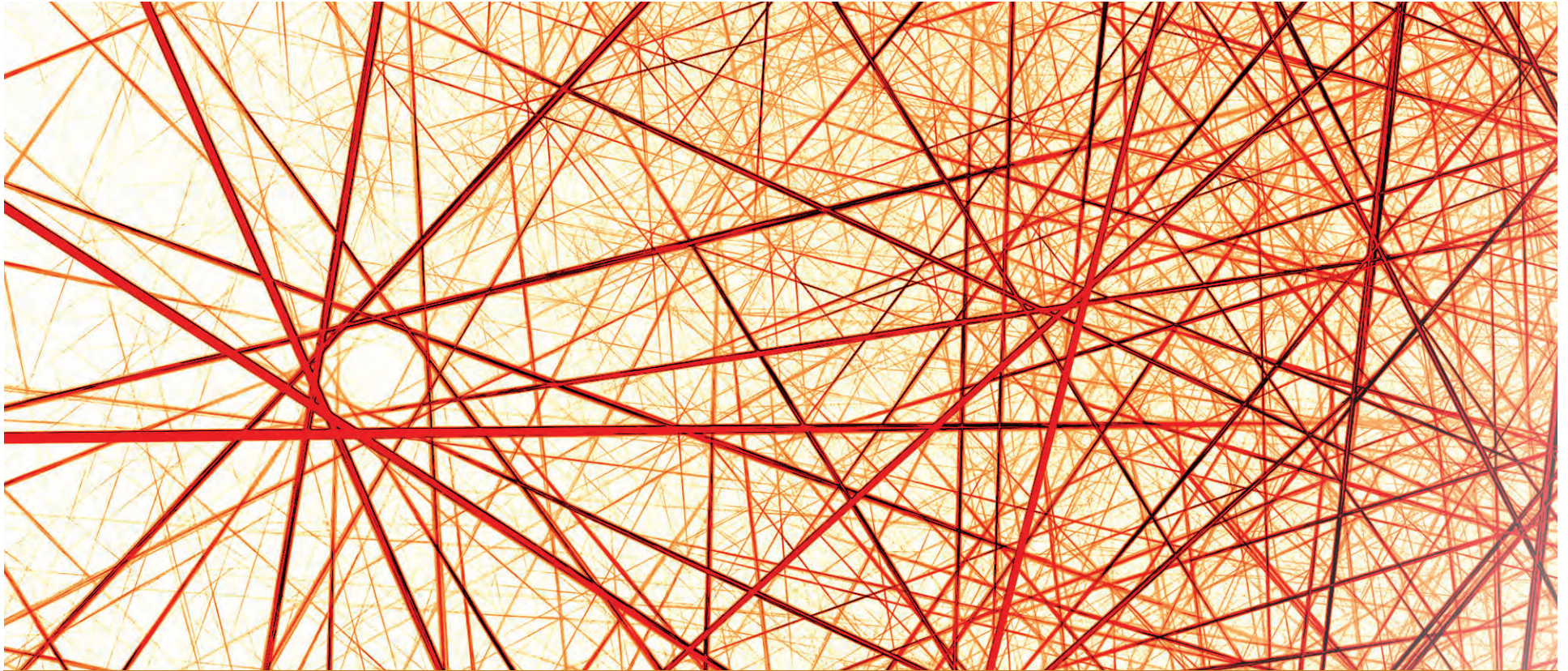
The SEI is a federally funded research and development center (FFRDC) sponsored by the U.S. Department of Defense. We endeavor to apply the best combination of thinking, technology, and methods to the most deserving government software-related problem sets, free from conflict of interest.

While other FFRDCs and research centers are also attentive to the government's problems, the SEI brings its unique combined capabilities in cybersecurity and software together with its university affiliation and industry access to bear on important and challenging software-related problems—in acquisition, development, testing, security, safety, operations, and sustainment.

To provide the capabilities it offers, the SEI's workforce maintains expertise in the following technical areas: software and systems engineering, cybersecurity and software assurance, computer science, applied mathematics, measurement and analysis, and acquisition of software-reliant systems. The SEI's work products include research reports, methods, software prototypes, and educational courses.

This booklet contains summaries of the research projects comprising the SEI's research portfolio, in addition to interactive workshops and other activities at the SEI. We encourage you to reach out to the authors, presenters, and other members of the SEI's staff for additional information, discussion, and future collaboration opportunities.

Thank you.



Projects and Presentations

Elicitation of Unstated Requirements at Scale (EURS)

Principal Investigators

Michael Konrad, PhD
Software Solutions Division
mdk@sei.cmu.edu
(412) 268-5813

Robert Stoddard
Software Solutions Division
rws@sei.cmu.edu
(412) 268-1121

Internal Investigator

Nancy Mead, PhD
CERT Division

Traditionally, requirements are communicated through specifications. Yet stakeholders often have requirements that they are not aware of, so they do not specify them. Uncovering these unstated requirements can be challenging; yet they can be a source for innovative system features and key architectural drivers. Thus, they can fundamentally affect the design, implementation, performance, and evolution of a complex software system.

Due to the increased interconnectedness brought about by continuing technological improvement and its application, today's larger software systems and the stakeholders that rely on them are better described as sociotechnical ecosystems (STEs) composed of diverse sets of organizations competing and collaborating around technology platforms.

Few requirements elicitation methods undertake a systematic approach to determining the unstated needs of stakeholders, and those that do so engage a collocated team of requirements analysts in time-boxed fashion. Requirements elicitation methods reported in the literature either elicit unstated needs from a small number of stakeholders on a small set of features (e.g., through prototyping and simulation) or perform a more broad requirements elicitation from a large number of stakeholders but without a systematic approach to discovering

unstated needs (e.g., StakeRare [Lim 2012]).

In addition, this input is typically analyzed in a collocated team setting, which further limits the comprehensiveness of the requirements analysis.

These sacrifices in scope and attention often result in an incomplete identification of key system requirements and priorities. For example, this incompleteness often manifests as a lack of attention to critical

security requirements because software developers often overlook trends in markets or the emergence of new technologies that may negatively affect the system and its stakeholders.

This project will continue an effort started in FY13 to extend a requirements elicitation method (called "KJ") for determining unstated needs. We aim to develop and validate a scalable method for determining the unstated needs of the multiple stakeholders typical of today's STEs. This method, tentatively called "KJ+," will be scalable to address the needs of multiple categories of stakeholders; be usable by a diverse, non-collocated team of requirements analysts; and result in a more complete set of requirements as the basis for subsequent system design, implementation, and continued sustainment.

Our approach includes the following tasks:


1. Re-Assess Requirements Engineering methods, technologies, and collaboration opportunities to establish a multi-year research roadmap for implementing KJ+.
2. Select and train an external collaborator in KJ+.
3. Establish measures for coverage of unstated needs.
4. Conduct KJ+ experiment with external collaborator.
5. Toward achieving a greater level of virtual collaboration and scale, conduct an internal-SEI experiment using an Ideation platform (e.g., IdeaScale) in conjunction with textual and network analysis support.
6. Leverage the lessons learned from the SEI and external collaborator experiments to conduct KJ+ at much larger scale with an appropriate public community.
7. Evaluate KJ+ using the measures defined in Task 3 to determine technical soundness of the tool-assisted KJ+.

Artifacts Developed and Research Outcomes

The KJ+ Method
KJ+ Method Training (for use in pilots and in transition)
Pilot results, evaluation, and report

Resources

EURS webpage: <http://www.sei.cmu.edu/measurement/research/eliciting-requirements/>



We aim to develop and validate a scalable method for determining the unstated needs of the multiple stakeholders typical of today's STEs. This method, tentatively called "KJ+," will be scalable to address the needs of multiple categories of stakeholders; be usable by a diverse, non-collocated team of requirements analysts; and result in a more complete set of requirements as the basis for subsequent system design, implementation, and continued sustainment.

Eliciting Unstated Requirements at Scale

KJ+ Training can itself be delivered virtually

Session 1

Introduce the KJ Method and describe the SEI's approach for using this method in a virtual (non face-to-face), distributed setting.

Session 2

Explain and practice KJ interviewing techniques, emphasizing the critical importance of capturing context information regarding good and bad extremes of experience. Provide examples of KJ report statements.

Session 3

Explain and practice KJ affinity grouping technique, emphasizing grouping by non-obvious themes of experience. Explain and provide examples of innovative solutions and unstated needs.

Session 4

Explain and practice Kano analysis.

Overview of SEI Approach



Example of Traditional Affinity Grouping (Hotel)

ID#	Traditional Interviewing Statement	Checkin/Checkout Affinity	Room Quality Affinity	Room Service
1	Clean Room		X	
2	Reliable Room Service Delivery			X
3	No-Hassle Check-In/Out	X		
4	Friendly Staff	X		X
5	Room Service Food Fresh & Hot			X
7	Room Service Available	X		X
8	Nice Towels		X	
9	New Bathroom		X	
10	Good Room Service Selection			X
11	Mini-Refrigerator in Room		X	
12	Attractive Furnishings		X	
13	Big TV		X	
14	Express Checkout	X		
15	Quiet Heater/Air Conditioning		X	
16	Non-Smoking Room Available	X		

Traditional Responses with Added KJ Contextual Data

ID#	Traditional Interviewing Statement with Added KJ Context from Probing
1	Prefer a Clean Room with a fresh smell to give my hotel stay a pleasant start
2	Expect Reliable Room Service Delivery so I don't have to keep calling on status
3	No-Hassle Check-In/Out helps me avoid tracking a lot of detail during a business trip
4	Friendly Staff pick up my spirits when I am tired on a business trip
5	If my Room Service Food is not fresh and hot, I have to spend time finding a local restaurant
6	Don't Lose Reservation is a message I don't want to hear because I do not have access to my travel agent
7	I like it when Room Service is Available because I can avoid worrying about logistics
8	Nice Towels put me in a good mood when I have to get up early in the morning
9	New Bathroom gives me a clean feeling and adds energy to my day
10	Good Room Service Selection keeps my stress level down and reduces anxiety about my diet
11	Mini-Refrigerator in Room gives me choices as I decide about food and snacks while working in my room
12	Attractive Furnishings put me in an energetic mood, enabling me to get more work done in my room
13	Big TV helps me see hotel area traffic, whether I am in bed or on the hotel room balcony
14	Express Checkout helps me a lot as I am forgetful about the time and logistics to check out
15	Quiet Heater/Air Conditioning enables me to think creatively on hard problems without distraction
16	Non-Smoking Room Available is a must or I will have a headache while trying to work in my room

Cell Phone Use Exercise

I'm a physician and need to receive phone calls at all times

I go to the symphony pretty often, and must keep my cell phone on vibrate.

I must wear my phone to feel it vibrate, but my dressy clothes have no pockets.

I want to keep my phone in my purse.

I hate wearing a belt to hold my phone with my dressy clothes.

Theme of needing to receive incoming cell phone calls when quiet is required.

Consider offering a ring sound of someone sneezing or gently coughing.

KJ Affinitization Resulting from Added KJ Contextual Data Hotel Example 1

ID#	Traditional Interviewing Statement	One theme of experience could be:
2	Expect Reliable Room Service Delivery so I don't have to keep calling on status	As a very busy traveler, I need help in looking up information, contacting remote agencies and tracking a lot of detail, without human assistance or delay.
3	No-Hassle Check-In/Out helps me avoid tracking a lot of detail during a business trip	
5	If my Room Service Food is not Fresh & Hot, I have I have to spend time finding a local restaurant	An innovative solution could be: A free application on a smart phone (or hotel issued device), which enables precise SIRI-like queries, and which also communicates with my TV and interactive displays throughout my room, balcony and other areas of the hotel, taking advantage of sensing my location.
6	Don't Lose Reservation is a message I don't want to hear because I do not have access to my travel agent	
7	I like it when Room Service is Available because I can avoid worrying about logistics	
13	Big TV helps me see hotel area traffic, whether I am in bed or on the hotel room balcony	
14	Express Checkout helps me a lot as I am forgetful about the time and logistics to check out	

Hotel Example 2

ID#	Traditional Interviewing Statement	One theme of experience could be:
1	Prefer a Clean Room with a fresh smell to give my hotel stay a pleasant start	I need to recover from a busy, stressful day and re-generate my entire being during my stay in the hotel.
4	Friendly Staff pick up my spirits when I am tired on a business trip	
7	I like it when Room Service is Available because I can avoid worrying about logistics	
8	Nice Towels put me in a good mood when I have to get up early in the morning	
9	New Bathroom gives me a clean feeling and adds energy to my day	
10	Good Room Service Selection keeps my stress level down and reduces anxiety about my diet	
11	Mini-Refrigerator in Room gives me choices as I decide about food and snacks while working in my room	An innovative solution could be: I need a complete, relaxing and rejuvenating experience during my presence in the hotel based on a strategic treatment of my five senses including sensors in my vicinity that can read and provide feedback when things are amiss.
12	Attractive Furnishings put me in an energetic mood, enabling me to get more work done in my room	
14	Express Checkout helps me a lot as I am forgetful about the time and logistics to check out	
15	Quiet Heater/Air Conditioning enables me to think creatively on hard problems without distraction	
16	Non-Smoking Room Available is a must or I will have a headache while trying to work in my room	

Investment Model for Software Sustainment

Principal Investigator
Robert Ferguson
Software Solutions Division
rwf@sei.cmu.edu
(412) 268-9750

The allocation of sustainment funds for the Armed Services is targeted, by law, to be no more than 50 percent to acquisition contractors and at least 50 percent organic (to service members, civilian employees and contract employees). In order for this organic workforce to be effective and efficient, the Service invests in personnel, tools, processes, and facilities to perform the work. Organic sustainment organizations have ready access to funds for most product maintenance and enhancement, but it is difficult for them to obtain funding for these internal improvements. How can the sustaining organization make the business case for these investment funds?

The work arises out of a specific modernization contract that had been underfunded so that no operational test kit was made available to the sustainers. Consequently, the efficiency and effectiveness of the sustainers was adversely affected, and platform availability was significantly reduced.

This research intends to show that small differences in the timing and amount of infrastructure investment funds can result in dramatic changes in the efficiency and effectiveness of the sustainment organization. Further, we hypothesize that a reduction in sustainment performance has long-lasting effects on the utility of the system and, hence, on warfighter readiness and capability.

There were two primary goals for the FY14 work and a secondary objective:

- Calibrate the current system dynamics (SD) model¹ with the support of an actual sustainment organization by instrumenting the operations of the sustainer, monitoring the external demand for sustainment work, and correlating sustainment output to mission performance.
- Extend the SD model to include sustaining organizations with responsibility for a portfolio of products. Also, calibrate this model.

¹ We developed the current SD model during a prior research project. That model reasonably represented the behavior of a sustainment organization with responsibility for a single product.



This research intends to show that small differences in the timing and amount of infrastructure investment funds can result in dramatic changes in the efficiency and effectiveness of the sustainment organization. Further, we hypothesize that a reduction in sustainment performance has long-lasting effects on the utility of the system and, hence, on warfighter readiness and capability.

- The secondary objective is to investigate the potential for using a catastrophe theory model² that would show how economic forces are out of balance, using fewer parameters and eliminating the need for a time-based simulation to show proximity to the tipping point.

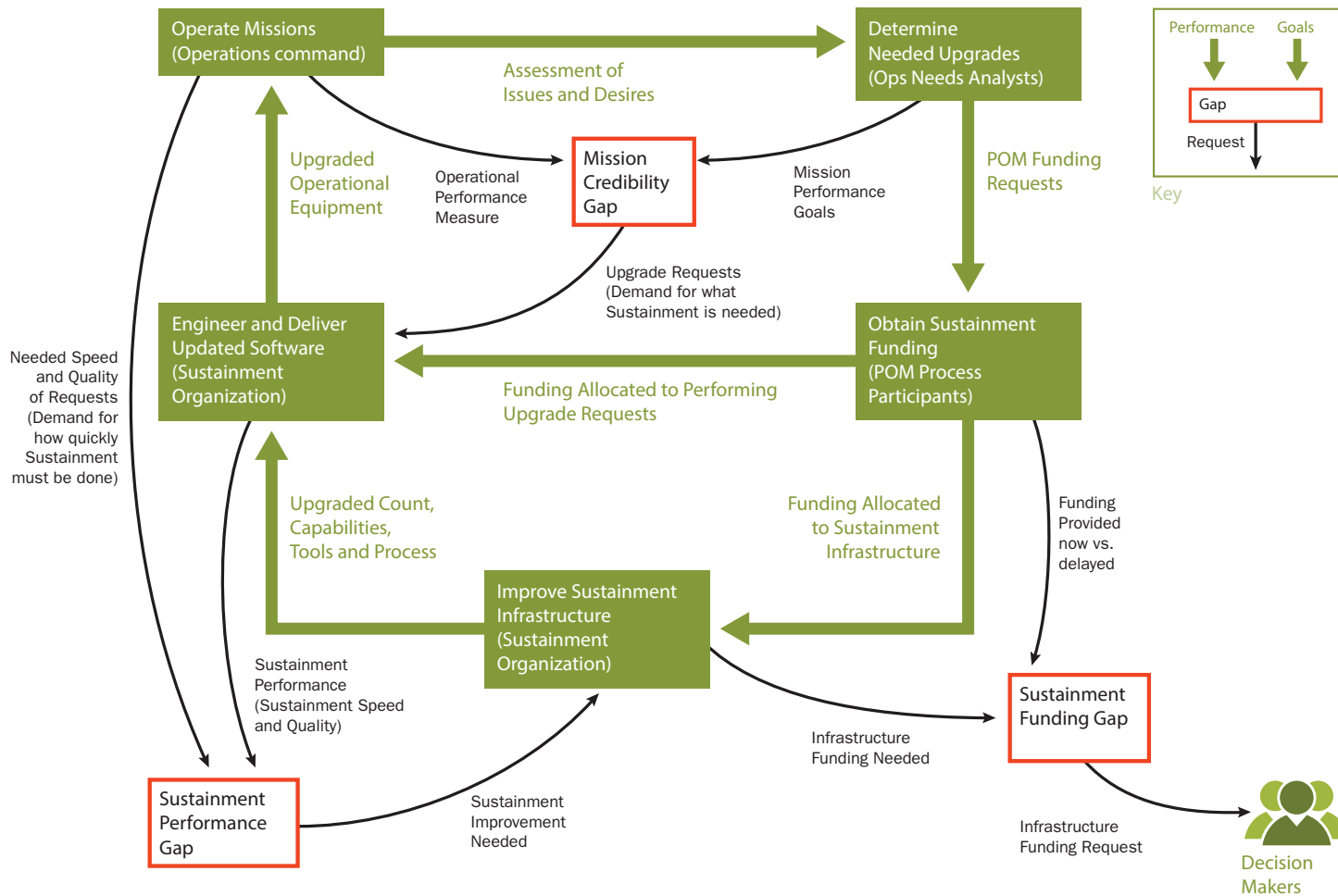
Research Outcomes

- An SD model using the Vensim³ systems dynamics modeling tool. This model has sliders that are used to adjust inputs and other parameters and produces graphs that show the resulting performance of the system.
- Model was successfully calibrated and validated with a collaborating sustainment organization. Write a report describing the actual calibrated parameters and the activity required to acquire the measurement data.
- Develop a catastrophe theory model using the economic forces described by the SD model as performance gaps and warfighter demand.

² http://en.wikipedia.org/wiki/Catastrophe_theory

³ <http://vensim.com/>

Comparing Organizational Performance to Goals



Tracking gaps between performance and goals

Value-Driven Incremental Development (VDID)

Principal Investigator

Ipek Ozkaya, PhD
Software Solutions Division
ozkaya@sei.cmu.edu
(412) 268-3551

Focus Area Leads

Ipek Ozkaya, PhD
Software Solutions Division

Chuck Weinstock, PhD
Software Solutions Division

Robert Nord, PhD
Software Solutions Division
(co-PI)

Lengthy and up-front requirements, design, integration, test, and assurance cycles delay delivery, resulting in late discovery of mismatched assumptions that results in system-level rework [Feiler 2010]. Industry data show that 70 percent of errors are introduced during requirements and architecture design, while 80 percent of errors are discovered during system integration test or later, with a rework cost that is 300 or more times the cost of discovering and correcting the errors earlier [NIST 2002].

Failure to integrate architecture analysis with development efforts early and continuously leads to costly increases in rework in maintaining systems. In DoD and elsewhere, architecture analysis and assurance activities are conducted neither frequently nor early enough to give ongoing insights into the quality of the system being developed, because such activities are not well connected with other software artifacts (code, requirements, design documents, etc.) and require additional resources to create. This results in unanticipated rework that lengthens testing and integration cycles, costly re-assurance activities when changes occur, and the inability to plan for achieving stringent quality attribute concerns relating to performance, modifiability, safety, and so on. These key operational challenges are worsened by technical challenges that include assurance efforts focused on the entire system rather than what changed and an invisible set of architectural information.

To ameliorate this, the Value-Driven Incremental Development (VDID) project investigates how quality attribute requirement allocation and dependency analysis informs the incremental development and assurance of the architecture through managing rework.

Research Outcomes

Focus areas and the outcomes for FY14 included the following:

- **Architectural dependency management** focus area investigated augmenting dependency structure modeling techniques with architectural information. The motivation to understand structural dependencies in software systems is rooted in controlling the cost of change and

evaluating modification impacts [Nord 2013]. Analysis of an implementation view of the system can miss important architectural dependencies that leads to costly rework. The goal was to make key architectural information (e.g., fault-propagation dependencies) available to developers both during architecture modeling and development through tool support. We developed an approach that includes dependencies associated with multiple perspectives or views of the architecture [Nord 2014]. We created a dependency guide and a unified model to guide the identification of dependencies. Our pilot studies applied the approach to a typical scenario in industry for managing cost of change and safety-critical testing costs. We observed that the model-driven engineering tools focusing on state transitions allow the engineers to focus on data-flow and events but cause them to miss data entity relationship, virtual resource behavior, and deployment-related dependencies. Mapping key dependencies, identified using a multi-view dependency analysis, to module view elements allows developers to concretely assess the impact of change and recognize system elements that need to be developed further. We piloted our approach using the Architecture Analysis and Design Language (AADL), demonstrating the extraction of hidden dependencies that impact fault propagation, design time code change, and testing.

- **Incremental assurance** focus area defined eliminative argumentation as a core concept that is a basis for arguing confidence and in establishing the theory of confidence. In addition to providing a basis for evaluation, the approach provides a method for constructing an argument in which one can have confidence. We demonstrated the ability to generate assurance cases and confidence maps from AADL architecture models annotated with requirement specifications and verification activities and piloted with examples from our industry collaborators.
- **Quality attribute requirement allocation** focus area investigated the incremental evolution of quality attribute requirements. Previous multi-project studies analyzed integrated architecture and agile practices that illustrated how architecture supports prototype experimentation and rapid tradeoff analysis [Bellomo 2013]. This motivated the

Value-Driven Incremental Development (VDID)

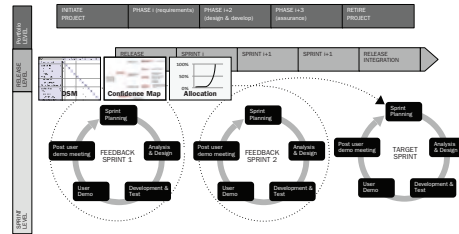
next study where we analyzed four industry projects to extract how they break cross-cutting concerns such as performance, security, and availability into manageable increments [Bellomo 2014]. Developers refined performance requirements by ratcheting response to stimuli in a context to explore feasible increments that could be allocated to releases. This refinement supports ongoing exploration of the requirements and solution, and evolutionary development when new information is acquired. We also conducted organizational surveys with developers to understand how architectural refinement and technical debt are related.

Our findings demonstrated that technical debt occurs regardless of software development processes followed, with a significant portion of debt coming from architectural sources and bad architecture decisions. The ability to measure the impact of such small refinements on an ongoing basis remains a challenge despite many tools for developers. This will be the focus of our work going into FY15.

Value-Driven Incremental Development Integrating Architecture Analysis and Assurance With Development

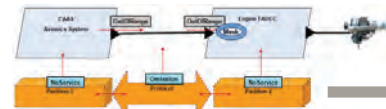
Objective

Investigate how quality attribute requirement allocation and dependency analysis inform incremental development and assurance through managing rework during development.



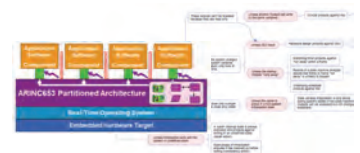
Multi-dimensional Analysis

What is the design implication of a release decision?



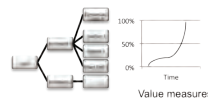
Architecting for Incremental Assurance

What are the assurance implications of a release decision?



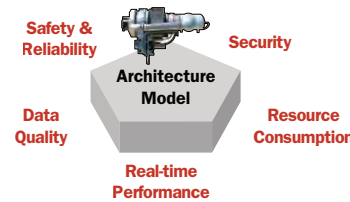
Quality Attribute Allocation

How do we break architectural features into increments; what measures are needed to make good release decisions?



Solution Approach

- Collect quality attribute requirements using architecture-tactics questionnaires
- Create models for deployment view augmented with partitioning and fault-tolerance information
- Generate an experiment environment where models can be seamlessly exchanged
- Apply modifiability and fault-propagation metrics
- Validate that augmenting with fault-tolerance information provides information about propagating rework
- Validate whether incremental-assurance information can be contained within architecture changes

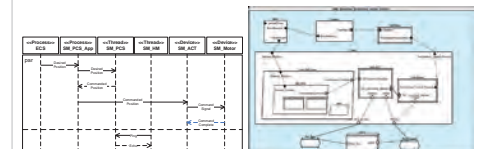


Root	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
NAV	1	LTc						LS							
COMM	2		LTc						LS						
LUM	3			LTc						LS					
SU1	4				LTc										
SU2	5					LTc					LS				
SU3	6						LTc					LS			
CND_M	7							LTc	SR	SR	SR	SR	SR	SR	SR
IMC_D	8	LD							SR	LTa					SR
COMM_D	9	LD									LTc				SR
LUM_D	10	LD										LTc			SR
SU1_D	11			LD					SR				LTc		SR
SU2_D	12				LD					SR				LTc	SR
SU3_D	13					LD			SR						LTc
REND	14						SR	SR	SR	SR	SR	SR	SR	SR	LTc
Comp. Impl. Layer	15	c	c	c	c	c	c	c	c	c	c	c	c	c	Tc

Multi-Dimensional Analysis Drives Increment Value Assessment

Selected FY14 Results

- Improved rework analysis by making **architectural dependency** information (e.g., fault-propagation dependencies) available to developers during architecture modeling and development
- **Eliminative argumentation** defined as a core concept that is a basis for arguing confidence and in establishing the theory of confidence
- Incremental evolution of **quality attribute requirement allocation** using architecture tactics-based data collection occurs through small refinements and ratcheting of response measures. Empirical studies and surveys with organizations revealed architectural rework occurs in such context and can be managed by better quantification of technical debt.



	Requirement	Architecture
SMR-Req-1: desired position	1	1
SMR-Req-2: actual position	1	1
SMR-Req-3: actual, req2 controlled position	1	1
SMR-Req-4: command completion	1	1
SMR-Req-5: command duration	1	1
SMR-Req-6: response delay	1	1
SMR-Arch-1: SMS_SMA_PCS	D	A
SMR-Arch-2: SMS_SMA_PCS CommandPositionData	D	A
SMR-Arch-3: SMS_SMA_PCS CommandPositionData	D	A
SMR-Arch-4: SMS_SMA_ACT	D	A
SMR-Arch-5: SMS_SMA_ACT StatusData	D	A
SMR-Arch-6: SMS_SMA_ACTOR	D	A
SMR-Arch-7: SMS_SMA_VECTOR ActualPositionData	D	A
SMR-Arch-8: SMS_SMA_HM	D	A

Multi-view analysis allows developers to see different types of dependencies that need to be investigated when changes occur

Collaborators

Carnegie Mellon University, Clemson University, University of British Columbia, University of Pennsylvania, DoD and industry partners



Contact: <http://www.sei.cmu.edu/architecture/research/info@sei.cmu.edu>

©2014 Software Engineering Institute

Representation of a poster developed to describe this project

Agile Adoption in the Department of Defense (DoD)

Principal Investigators

Mary Ann Lapham
Software Solutions Division
mlapham@sei.cmu.edu
(412) 268-5498

Suzanne Miller
Software Solutions Division
smg@sei.cmu.edu
(412) 268-9143

Though they are starting to gain acceptance in the U.S. Department of Defense (DoD) community, Agile methods are still relatively new to the DoD acquisition community. As recently as four years ago, most programs did not willingly or openly admit to using Agile methods even if they were. These methods can be disruptive for the DoD acquisition community since they are somewhat antithetical to the standard modes of operation within the DoD acquisition and development communities.

Today, more government programs are openly professing to employing Agile methods. However, knowledge of the fact and character of diffusion (how many adopters have been reached) and the infusion (how routine has use of the new practices become) progress would be helpful to multiple stakeholders for three reasons.

- Potential adopters want to know how “safe” it is to adopt the new practices— “Is Agile here to stay?”
- Policy makers want to know if the practices are sufficiently useful and feasible to implement before changing policy to accommodate them.
- Acquisition professionals dealing with vendors/contractors who have adopted the practices want to know if this is a set of practices they need to be prepared to oversee on an ongoing basis.

Our project objective is to identify ways to enable practitioners to develop validated tools, techniques, and practices that correlate with Agile and lean concepts for use within DoD. The overwhelming focus of our research project is to understand and overcome the perceived and actual technical and cultural barriers to these methods.

On the one hand, acquisition professionals are familiar with and understand how to navigate, even if slowly, the acquisition lifecycle and its attendant constraints using traditional waterfall-based lifecycles. That path minimizes their personal risk—if something negative happens, at least they were following the DoD-accepted practices for program management. On the other hand, if they embrace Agile’s iterative and incremental methods that have had continually growing success in the commercial industry, they will have to produce their own guidance and training, at a minimum. In addition, a worst case could include accusations of breaking acquisition regulations, with attendant personal and program consequences.

Research Outcomes

Two main tasks are part of this work:

- **Task 1** is to produce guideline documents on particular topics that our customer engagements show are problematic and on topics that our discussions with the SEI Agile Collaboration Group indicate are needed by the acquisition practitioner community.
- **Task 2** is to produce a State of Agile Adoption report that summarizes progress in Agile adoption in the DoD in the last several years. This report will include both survey data and mini-case summaries of Agile use in government settings.

Adoption of Agile/Lean Methods in Government Settings

State of Adoption

People can be threatened by a new method... **RETRAINING RESISTANCE OR FEAR of FIRING**

EARLY ADOPTERS CHASMS EARLY MAJORITY

Study Results

According to Contractors:

AGILE

QUALITY

WAY BETTER!

SCHEDULE

AGILE

"Covert agile"

You've been awfully flexible lately...

Uh-oh, they're on to me!

Barriers

WE WANT YOU TO BE A RISK-TAKER, SOMEONE WHO'S TAKING CUTTING EDGE, AND PROACTIVE GO AHEAD...

OHMYGAWD! LOOK OUT! SLOW DOWN! WHAT!? I'VE NEVER DONE IT THAT WAY! WE'RE ALL GOING TO DIE!

AGILE is a seed requiring nurture... and protection.

THE LIMITS OF WHITE-BOARDING

Guidelines

BALANCE

SPEED STABILITY

AGILE

WE WANT AGILE BEING DONE THE RIGHT WAY

TEST COVERAGE

BURN DOWN

FORMALITY

AGILE AT SCALE

How do we get it right and not be too late!

Issues

Acquisition Process

BUYER CONTRACTOR

INSIGHT VS OVERSIGHT

Enablers

Training less stultified

GIVING PEOPLE A CLEAR IDEA ABOUT HOW THE ROLE CHANGE WILL (UNWELL) REDUCE FEAR OF THE UNKNOWN

WE ARE AGILE. WE WILL ADAPT.

WHEN WE SAY AGILE, WE MEAN PROCESSES USING THE "A" PRINCIPLES IN INCREMENTAL, RECURSIVE AND REFINED

4 levels

LOW CEREMONY

SEMI-FORMAL

HIGH CEREMONY

INCREMENTAL REVIEW FORMALITY

FROM PROCESS POLICE TO ADAPTIVE LEADERSHIP

Business Teams

Technical Teams

To succeed, all levels of an organization (if the customer) need to support the agile process

Acquisition Dynamics

Principal Investigators

William Novak
Software Solutions Division
wen@sei.cmu.edu
(412) 268-5519

Andrew P. Moore
CERT Division
apm@cert.org
(412) 268-5465

The failures of software-intensive joint programs cost the U.S. Department of Defense (DoD) many millions of dollars per year in cancellations and cost overruns—not including the impacts of late delivery, reduced deployment, and inadequate performance and functionality. Since the problems facing joint programs are structural and inherent, they will continue to recur until they are properly addressed.

Joint acquisition programs experience exacerbated cost, schedule, and quality failures—and joint acquisition management personnel often have limited breadth of experience to recognize that such failures are ubiquitous, that programs fail repeatedly for the same reasons, and that there are known corrective and preventative techniques.

This work is creating a virtual laboratory for simulating joint program behaviors—allowing us to develop and test mitigation and solution approaches to evaluate their efficacy in resolving the problems. In addition, gaining a deeper understanding of the dynamics at work in joint programs will help Joint Program Office (JPO) staff anticipate issues. Enabling better decision making on the part of acquisition leaders and staff will produce better program outcomes for the DoD.

This effort seeks to improve the operation of joint acquisition programs by understanding the forces that produce poor outcomes and by testing methods for overcoming those forces that include policy changes (for JPOs, services, and the DoD) and improved educational methods (for Defense Acquisition University, et al).



This work is creating a virtual laboratory for simulating joint program behaviors—allowing us to develop and test mitigation and solution approaches to evaluate their efficacy in resolving the problems.

The key technical ideas involved in this work include extending prior “Acquisition Archetypes” work using systems thinking to model acquisition program dynamics¹

¹ For more information on Acquisition Archetypes, listen to William Novak’s podcast in the SEI Podcast Series (<http://www.sei.cmu.edu/podcasts/>).

- applying social dilemma analysis to joint acquisition programs to develop and model solution approaches
- using system dynamics modeling to characterize joint acquisition program problems and test candidate solutions
- using an interactive online scenario-based survey to understand decision-making behaviors in joint program contexts and using that data to tune the behavior of the system dynamics model

This work consists of two tasks:

Task 1: Modeling Joint Acquisition Dynamics and Candidate Solutions: Develop and analyze the behavior of the Joint Program model and the effectiveness of modeled mitigation/solution approaches in reducing the adverse consequences

Task 2: Understanding Joint Acquisition Decision-Making via Experimental Scenarios: Gather data on the actual decision-making behaviors of experienced acquisition staff in the context of hypothetical joint program scenarios

Research Outcomes

- Demonstrate the ability to improve decision-making for joint acquisition program participants through an understanding of the incentives driving joint acquisition program behaviors
- Assess the efficacy of mitigation and solution approaches to the joint program social trap, with conclusions on their ability to be applied to joint programs
- Develop an approach for modeling the dynamics of acquisition program behaviors that applies to a wide range of specific acquisition program contexts

The project will produce the following artifacts:

- A validated system dynamics model of the joint program problem, as well as candidate mitigation/solution models
- Historic performance data on two DoD joint programs to enable model validation
- Analysis of the effectiveness of candidate mitigation/solution models at improving joint program performance

Acquisition Dynamics

- A web-hosted, scenario-based survey that collects realistic joint program stakeholder decision-making data
- Analysis and summary of experimental results of joint stakeholder decision-making scenarios

Applications

The system dynamics model can be the basis of a management “flight simulator” that provides hands-on simulations of program behaviors. This can help DoD acquisition staff gain a deeper understanding of program dynamics to help them make better decisions.

The system dynamics model can be used to help component acquisition executives and other DoD policymakers better understand the potential implications of both current and proposed acquisition policy, and thus help shape more effective policy that can produce better outcomes for both individual programs and the broader acquisition system.

Acquisition Dynamics System Dynamics Model

Carnegie Mellon Software Engineering Institute

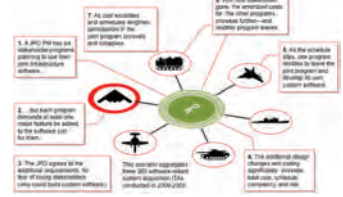
William E. Novak Jay D. Marchetti
 Andrew P Moore Matthew L. Collins
 Julie B. Cohen
 Dr. Cleotilde Gonzalez, Carnegie Mellon

Candidate Solutions

Use the system dynamics acquisition model to simulate and compare the effectiveness of different candidate solutions from the academic literature and from acquisition staff.

The Problem: A Social Trap

We see the following joint program story play out all too often:

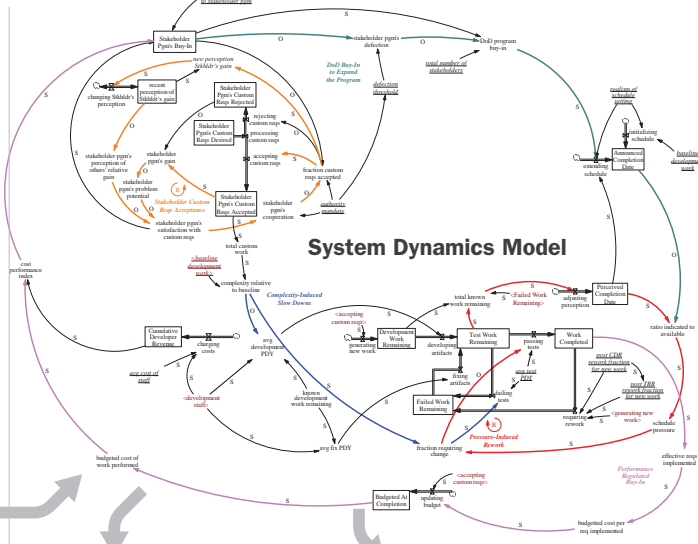


The Rationale

- Declining defense budget for acquisition
- Need for cost savings through joint programs
- Higher rate of problems in joint programs
- Joint program ideas apply to common infrastructure development
- Growing need for interoperability
- Software increasingly key for interoperability

System Dynamics Model

The system dynamics model of a joint acquisition program has two major segments: **Stakeholder Program Interactions** segment (upper left): Models the complex dynamics among participating stakeholder programs, the JPO, and the Services involving willingness to participate, schedule pressure, confidence in the JPO, fairness, pressure to cooperate, and many other factors. **System Development** segment (lower right): The dynamics of software development as affected by developer experience, schedule pressure, system complexity, requirements changes and volatility, and many other factors.

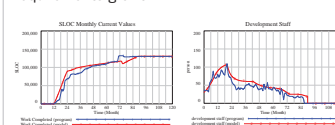


Model Simulation Results— Behavior Over Time

The development portion of the model generates:

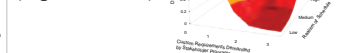
- monthly SLOC values that correspond closely to actual program development data
- monthly staffing levels that also track closely to those of an actual program

Both datasets are based on actual joint program requirements growth



Model Simulation Results— Program Behavior

Visualizing developer productivity as a function of schedule realism and late-addition requirements shows a “tipping point” is crossed when these factors combine. A JPO accepting custom requirements late in development to appease stakeholder programs can lead to the program’s collapse when schedules then slip unacceptably. This no-win situation for the JPO is the basis for the joint program “social trap”



Authority

- Regulates the good, prevents overuse
- Unpopular to enforce a mandate

Altruistic Punishment

- Pay to penalize uncooperative partners
- May escalate and cause retaliation

Shared Destiny

- Incentivize based on program outcome
- Incentive is in the future

Potential Future Applications

Acquisition Management Decision Support: Use the system dynamics model to run “what if?” scenarios on programs in progress, analyzing potential decision outcomes.



Interactive Acquisition Learning: Help acquisition staff learn to make decisions in complex situations by using model-based training simulations.



Analyze Policy: Help reshape how acquisition programs are conducted by analyzing proposed policies and outcomes at the OUSD(AT&L) and Service Component Acquisition Executive (CAE) levels.

Software Engineering Institute | Carnegie Mellon University

Contact: William Novak wen@sei.cmu.edu
 ©2014 Software Engineering Institute

Representation of a poster developed to describe this project

Quantifying Uncertainty in Early Lifecycle Cost Estimation (QUELCE)

Principal Investigator


Robert Stoddard
Software Solutions Division
rws@sei.cmu.edu
(412) 268-1121

U.S. Department of Defense (DoD) officials and the Government Accountability Office (GAO) have frequently cited poor cost estimation as a significant contributor to cost overruns, and the problems are increasing. The growth in major defense acquisition programs' (MDAPs) Research, Development, Test, and Evaluation (RDT&E) costs from the initial estimate rose from 27 percent in 2000 to 54 percent in 2011. Furthermore schedule slip in delivering initial capabilities rose from 16 to 23 months.

Presenter

Robert Ferguson
Software Solutions Division

Existing methods of early estimation usually assume a set of fixed input parameters and apply a risk factor at the end of the estimation process. The method we are developing, called QUELCE, explicitly represents the uncertainty inherent in various inputs and decisions, leading to probability distributions for the inputs to the estimate. QUELCE allows for the use of previously calibrated DoD cost estimation relationships (CERs) to calculate the resulting risk (cost and schedule impacts). DoD decision makers reviewing MDAPs can then make informed choices and fund programs at levels consistent with the magnitude of risk to achieving success, leading to fewer and less severe program cost overruns. QUELCE also focuses attention on the specific risks that may cause the largest problems and minimizes the time spent analyzing risks of lesser impact.



The method we are developing, called QUELCE, explicitly represents the uncertainty inherent in various inputs and decisions. QUELCE allows for the use of previously calibrated DoD cost estimation relationships (CERs) to calculate the resulting risk (cost and schedule impacts). DoD decision makers reviewing MDAPs can then make informed choices and fund programs at levels consistent with the magnitude of risk to achieving success, leading to fewer and less severe program cost overruns.

QUELCE synthesizes the use of several well-known tools and methods including scenario modeling, Bayesian Belief Network (BBN) modeling, Dependency Structure Matrices (DSM), and Monte Carlo simulation. QUELCE allows us to quantify uncertainty, use subjective inputs, depict influential relationships, control the scale of the problem, and document the assumptions underlying the estimate. The use of existing CERs means that estimation relationships do not have to be recalibrated.

Our efforts are now focused on the research and technologies needed to move the method from a laboratory setting into application within active MDAPs.

The remaining technical challenges to the QUELCE work for FY15 include

- Additional data mining of sensitive cost-variance data from DoD MDAP and MAIS programs
- Characterizing the uncertainty of group expert judgment for use in the BBN
- Developing more detailed change drivers for sustainment and modernization programs
- Developing a prototype, supervised machine learning mechanism towards the hopes of a semi-automated approach to maintain a “living” repository of program change-driver experiences

Research Outcomes

In FY14, we mapped the BBN output nodes to an additional cost-estimation vendor tool in support of a multi-year retrospective, conducted the on-site portion of the QUELCE workshop with a live MDAP, and performed repeatability experiments focused on rater agreement of highlighted change driver excerpts from DoD program textual artifacts.

Research into technologies allowing the implementation of many of the QUELCE steps in virtual ways will enable broad, cost-effective participation by subject matter experts who are not co-located. DoD domain-specific reference points will be semi-automatically data-mined with a machine learning mechanism. In turn, this will provide a continually updated and sophisticated repository query capability, to support expert judgment in future QUELCE workshops.

Quantifying Uncertainty in Early Lifecycle Cost Estimation (QUELCE)

QUELCE Workshop

The Quantifying Uncertainty in Early Lifecycle Cost Estimation (QUELCE) workshop enables a client to convene a set of domain experts to formulate early life-cycle cost estimates expressed as cost distributions rather than single points. The QUELCE method involves a five-step process that begins with identifying potential future changes to nominal program execution that will influence program cost. This is followed by probabilistic modeling of the interrelationships of the program change drivers and Monte Carlo simulation of cost model inputs to create program cost estimate distributions. Because many of the inputs are based on subject-matter expert judgment, this workshop also involves a novel approach to calibrating expert judgment through a series of training exercises.

Data Requirements

- Pre-workshop access to existing planning artifacts, such as AoA and ICD/CDD
- Access to domain experts who can anticipate different reasons for cost changes during program execution

Time Frame

- SEI preparation of 1–2 weeks to review available documentation with two SEI staff members
- Two SEI staff members on site for 5–7 days to facilitate five 3-hour workshops with both technical and financial program office staff
- 5–7 days to prepare baseline estimate and suggested scenario-based estimates
- Typically, 3–5 days to assist program office staff with explaining estimates as needed

Expected Results

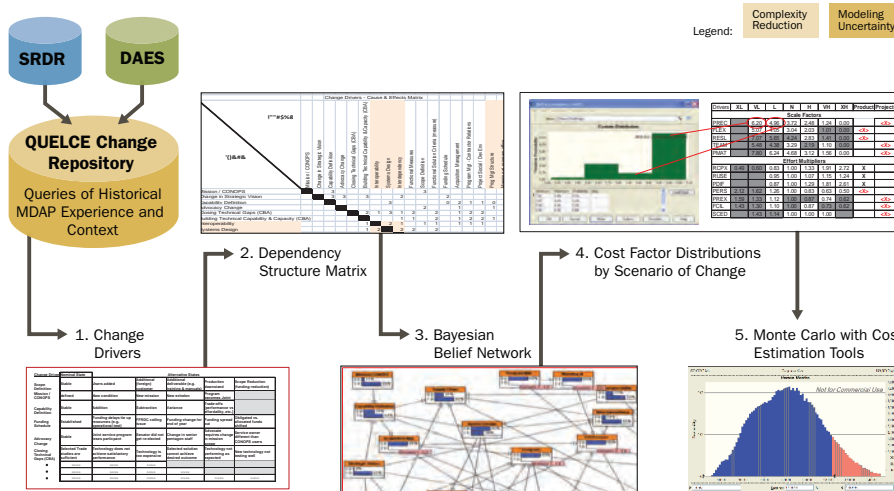
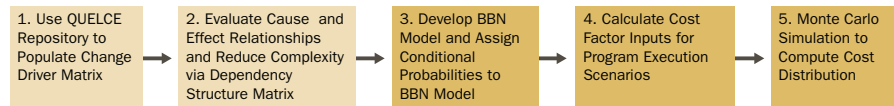
QUELCE produces a cost estimate that is represented as a distribution from which a decision maker can understand the level of risk associated with a particular cost value. It also produces an executable model that can be used to run alternative scenarios and that can be updated in the future for reestimation purposes. The model and information developed also provide good documentation of the basis of the estimate.

Publications

Quantifying Uncertainty in Early Lifecycle Cost Estimation (QUELCE)
www.sei.cmu.edu/library/abstracts/reports/11tr026.cfm

Quantifying Uncertainty in Expert Judgment: Initial Results
www.sei.cmu.edu/library/abstracts/reports/13tr001.cfm

Improving the Reliability of Expert Opinion within Early Lifecycle Cost Estimation
blog.sei.cmu.edu/post.cfm/improving-the-reliability-of-expert-opinion-within-early-lifecycle-cost-estimation



QUELCE Research

- Objective**
 Quantify expert judgment of anticipated program execution uncertainties and enable more accurate inputs to existing cost models.
- Description**
 Continuing research into the QUELCE method includes
1. Calibrating group judgments of the probabilities of change driver occurrence and co-occurrence
 2. Expanding the QUELCE change-driver taxonomy to include detailed sustainment change drivers
 3. Prototyping of supervised machine learning to enable the automatic processing of a future stream of DoD program artifacts. This will help create a "living" domain reference point repository benefiting ongoing DoD cost estimation.

Collaboration Opportunities

- Calibrating expert judgment in a group setting – Hubbard-style calibration to create more stability in elicited parameters
- Designing and mapping QUELCE BBN output nodes to cost model inputs
- Defining and classifying program change drivers
- Expanding the use of QUELCE in MDAP and PMO risk management programs to enhance the identification of future risks

Software Model Checking for Verifying Distributed Algorithms

Principal Investigator

James Edmondson, PhD
Software Solutions Division
jedmondson@sei.cmu.edu
(412) 268-8905

Presenter

Sagar Chaki, PhD
Software Solutions Division

Distributed software plays a key role in controlling many safety-critical and mission-critical systems, including cyber-physical systems (e.g., autonomous coordinated multi-robot missions) [NSF]. Therefore, verifying safe operation of distributed cyber-physical system (DCPS) is an important challenge. However, this problem is also notoriously complex and not addressed adequately by existing verification & validation regimes, which are largely based on testing.

For example, the combination of concurrency (e.g., message ordering) and sensitivity to timing (e.g., thread scheduling and message transmission delay) renders simulation and testing inadequate for verifying safety of DCPS with high confidence. Moreover, using the correct middleware semantics is also critical for sound analysis. In essence, the DCPS has a very large state space (which increases exponentially with the number of nodes), only a miniscule fraction of which is tested.

In this project, we have developed a new approach to producing high-assurance distributed software. Our approach, a form of verifying compilation, consists of two steps:

1. Verification: We have developed a new domain-specific language (called DASL) for writing distributed algorithms. A DASL program is first verified for correctness using two steps:

a. Sequentialization: In this step, the distributed algorithm (which is inherently concurrent) written in DASL is translated to an equivalent single-threaded C program in a provably correct way [CE14a]. For this step, we extend prior work on sequentialization [Lal 2009] to the “synchronous” model of computation.

b. Model Checking: The C program is then verified using an off-the-shelf software model checker [Jhala 2009]. Recent years have seen a lot of progress in applying model checking to verify software, in terms of algorithms as well as tools (e.g., CBMC, UFO.) For our experiments, we used the model checker CBMC [Clarke 2004].

2. Code-Generation: Once the DASL program is verified, it is translated to C++ code that uses the MADARA¹ middleware for communication. To guarantee the synchronous model of computation (used for verification) we have developed a new synchronizer protocol based on barriers and proved its correctness [CE14a]. This synchronizer is incorporated into the generated C++ code along with the algorithm. To demonstrate the effectiveness of our approach, we also generate code that interacts with V-REP² for visual simulations with realistic robot models.



We have developed a new approach to producing high-assurance distributed software. Our approach, a form of verifying compilation, consists of verification using a new domain-specific language for writing distributed algorithms and code-generation.

Research Outcomes

All our tools and examples are publicly available³ as open-source. Documents and tutorials are included. In addition, we have two publications [CE14a, CE14b] in peer-reviewed venues reporting on our research.

¹ <http://madara.googlecode.com>

² <http://www.coppeliarobotics.com>

³ <http://mcda.googlecode.com>

Verifying Synchronous Distributed Applications

Motivation

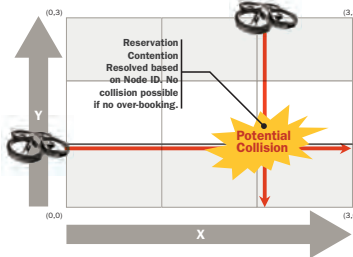
Distributed algorithms have always been important

- File Systems, Resource Allocation, Internet, ...
- Increasingly becoming safety-critical
- Robotic, transportation, energy, medical
- Prove correctness of distributed algorithm implementations
- Pseudo-code is verified manually (semantic gap)
- Implementations are heavily tested (low coverage)



Model-Driven Verifying Compilation of Synchronous Distributed Applications, Sagar Chaki, James Edmondson, Proceedings of MODELS 2014

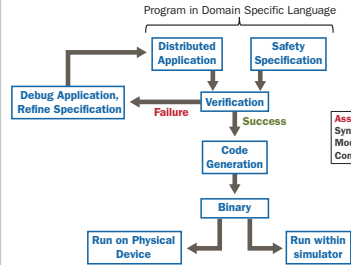
Example: Synchronous Collision Avoidance



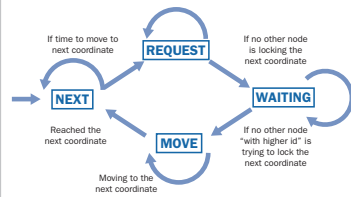
Tool Usage

- Project webpage (<http://mcda.googlecode.com>)
- Tutorial (<https://code.google.com/p/mcda/wiki/tutorial>)
- Verification
 - `dasic-nodes 3 -seq-rounds 3 -seq-dbl-out tutorial-02.c tutorial-02.dasil`
 - `cbmctutorial-02.c` (takes about 10s to verify)
- Code generation & simulation
 - `dasic -nodes 3 -madara -vrep -out tutorial-02.cpp tutorial-02.dasil`
 - `g++ ...`
 - `mcda-vrep.sh 3 outdir/tutorial-02 ...`

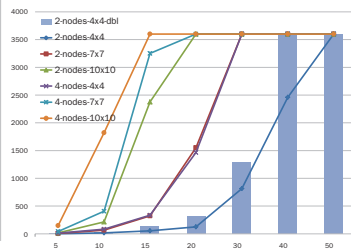
Approach : Verification + Code Generation



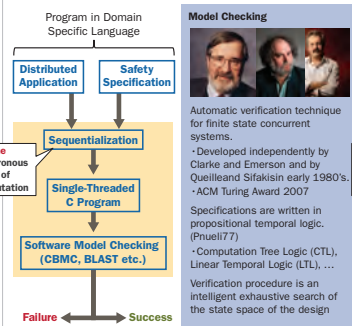
Collision Avoidance Protocol



Results: Collision Avoidance



Verification



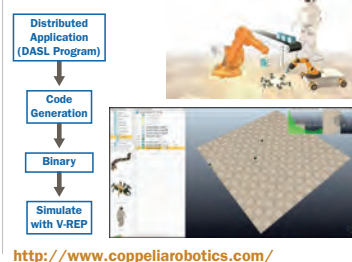
Synchronous Collision Avoidance Code

```

// ...
void move() {
    // ...
}
// ...

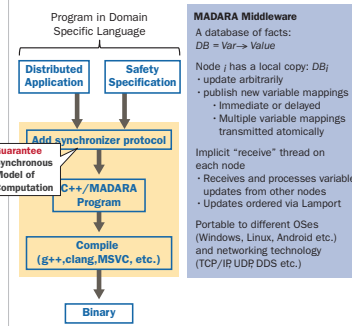
```

Simulation with V-REP



<http://www.coppeliarobotics.com/>

Code Generation



Synchronous Collision Avoidance Code

```

// ...
// ...
// ...

```

Future Work

- Improving scalability and verifying with unbounded number of rounds
- Verifying for unbounded number of nodes (parameterized verification)
- Paper to appear at SPIN'2014 Symposium
- Asynchronous and partially synchronous network semantics
- Scalable model checking
- Abstraction, compositionality, symmetry reduction, partial order reduction
- Fault-tolerance, uncertainty, ...
- Combine V&V of safety-critical and mission-critical properties

Verifying Evolving Software

Principal Investigator

Arie Gurfinkel, PhD
Software Solutions Division
arie@sei.cmu.edu
(412) 268-7788

Dealing gracefully with evolution by efficiently re-verifying a program after a change is a major challenge to program verification techniques such as Software Model Checking. In general, even a small change has a profound impact on program behavior, triggering an expensive re-verification of the whole program.

Internal Investigator

Sagar Chaki, PhD
Software Solutions Division

In this project, we are addressing this challenge by developing and evaluating algorithms for propagating verification results through human- and machine-generated evolution. We build on the recent advancements in proof-based verification, regression verification, and upgrade checking [Albarghouthi 2013, Godlin 2013, Fedyukovich 2013, Sery 2012]. Our key insight is that current proof-based analysis techniques generate explicit proofs, also known as, verification certificates, (as inductive invariants in First Order Logic) that can be migrated across evolution boundaries.

For the purpose of this research project, we have restricted our attention to machine-generated evolution. In particular, we focused on propagating verification results through compiler optimizations used in LLVM compiler and UFO verification engine.



We are developing and evaluating algorithms for propagating verification results through human- and machine-generated evolution. In particular, we have developed and evaluated key building blocks of a new compiler architecture that combines an optimizing compiler and a verifier under a common front-end.

As the first step in this ambitious research direction, we have explored applications of our techniques to mitigating the semantic gap hazard of formal verification. The semantic gap—the difference in semantics (i.e., interpretation of the program) between the verifier and the compiler—is one of the major threats to validity of verification results [Gurfinkel 2013]. For example, it is one of the main reasons why compiler optimizations are prohibited in most safety-critical domains.

In particular, we have developed and evaluated key building blocks of a new compiler architecture that combines an optimizing compiler and a verifier under a common front-end. In this architecture, verification certificates are propagated to the lowest level of intermediate representation of the executable. In the long term, our work will improve usability of the verifier by presenting a user with more understandable verification messages; allow for validation of correctness of compiler optimization; and allow producing self-certifiable executable code.

Research Outcomes

We have developed, within the UFO framework, two components that propagate verification results across code transformations.

Our first component, called FrankenBit, propagates verification results between idealizing semantics of arithmetic used by the verifier, and the machine semantics of arithmetic used by the compiler [Gurfinkel 2014a]. FrankenBit has participated in the 3rd Software Verification Competition (SV-COMP) and has received Silver and Bronze medals [Gurfinkel 2014b].

Our second component, called Niagara, propagates verification results across challenging LLVM optimizations that preserve program's loop structure [Fedyukovich 2014]. We are currently working on extending Niagara to apply to more complex optimizations.

Contract-Based Virtual Integration and CPS Analyses

Principal Investigator

Dionisio de Niz, PhD
Software Solutions Division
dionisio@sei.cmu.edu
(412) 268-9002

Internal investigators

Sagar Chaki, PhD
Software Solutions Division

Julien Delange, PhD
Software Solutions Division

Peter Feiler, PhD
Software Solutions Division

External Investigators

Prof. David Garlan, PhD
School of Computer Science,
CMU

Ivan Ruchkin
Graduate Student, ISRI, CMU

Bradley Schmerl, PhD
School of Computer Science,
CMU

Models and analyses are heavily and routinely used during the development of today's cyber-physical systems (CPS). This trend is expected to strengthen. Models enable the creation of a design-time description of a CPS before any of the parts are physically built. This also allows different groups to create parts of the model independently to be integrated into the full model at a later time. Analyses enable verifying (and modifying) the models at design time to guarantee important quality attributes, such as control stability, schedulability, power consumption, safety and security [Astrom 2011, Buttazzo 2011, Klein 1994, Dagle 2012, Moreno 2012, Tang 2008, de Niz 2011, Nam 2011]

Unfortunately, analyses are developed independently and often focus on different abstractions of the CPS model. This is due to the fact that they originate mostly from different scientific communities, such as real-time, control theory, security, formal verification, etc. As a result, these analyses make different assumptions about, and work on different abstractions of, the system that are not always compatible with one another. Thus, there is a need to rigorously discover dependencies and consistencies between analyses so that the results of applying them can be trusted.



Our aim is to solve the problem of untrustworthy analysis results by specifying contracts for each analysis and analyzing the contracts to discover dependencies and inconsistencies among them.

This leads to two problems that render analysis results untrustworthy:

1. analyses make inconsistent assumptions
2. one analysis alters the model, thereby violating assumptions made by another

Even worse, these problems manifest themselves very late at physical integration time, leading to costly fixes. This issue is particularly prevalent in multi-tier industries, such as avionics and automotive, where systems are integrated from independently developed parts whose designs have been analyzed with a mish-mash of tools.

Our aim is to solve this problem by: (1) specifying contracts for each analysis and (2) analyzing the contracts to discover dependencies and inconsistencies among them.

Research Outcomes

- An analysis contract language implementation in an Architecture Analysis and Design Language (AADL) annex
- An algorithm to discover conflicts and inconsistencies between a set of analyses based on their contracts
- Implementation of our approach on top of the Open Source AADL Tool Environment (OSATE 2) and demonstration on an industrial example
- Peer-reviewed publications and presentations in relevant venues (AFRL S5 2014, EMSOFT 2014, AADL Standards Meeting 2014)

High-Confidence Cyber-Physical Systems (HCCPS)

Principal Investigators

Sagar Chaki, PhD
Software Solutions Division
chaki@sei.cmu.edu
(412) 268-1436

Dionisio de Niz, PhD
Software Solutions Division
dionisio@sei.cmu.edu
(412) 268-9002

The High-Confidence Cyber-Physical Systems (HCCPS) project is predicated on the dire need for techniques to certify software-reliant systems that interact with the physical world—a.k.a. cyber-physical systems (CPS). Examples include avionics systems—e.g., the need to certify unmanned aerial vehicles before their insertion into civilian airspace was highlighted by a recent report [DSB 2012]—missile systems, and autonomous vehicles.

Effective certification of CPS requires objective evidence, such as through rigorous verification (e.g., as mandated by DO-178C). Moreover, the correct behavior of CPS cannot be verified by looking only at the software. For example, the navigator of an autonomous vehicle must perform obstacle avoidance maneuvers correctly, and on time. The software must not only be synchronized with the physical environment, but indeed can only operate correctly when such synchronization exists.

Therefore, effective analysis for CPS must take into account assumptions about, and interactions with, the physical environment. These include time, dependent physical parameters, functionality, and coordination between distributed entities.



The goal of the HCCPS project is to enable the development of CPS whose behavior we trust. In FY14, we developed theories and practical techniques for verifying CPS correctness in three areas: timing, functionality, and coordination (all foundational for CPS) and relevant to the DoD [BAH 2010, DSB 2012, CTSB 2010].

The goal of the HCCPS project is to enable the development of CPS whose behavior we trust. In FY14, we developed theories and practical techniques for verifying CPS correctness in three areas: timing, functionality, and coordination (all foundational for CPS) and relevant to the U.S. Department of Defense (DoD) [BAH 2010, DSB 2012, CTSB 2010]. Specifically, we developed theories and

analyses that help assure these conditions: timing correctness via real-time schedulability analysis, functional correctness via model checking, and quantitatively assured coordination via probabilistic model checking.

Our techniques are founded on mathematical principles, produce quantitative results, and are validated through continuous interaction with DoD-relevant stakeholders.

Research Outcomes

Sub-project 1

In FY14, we pursued three sub-projects. The first sub-project addressed the scheduling of parallel real-time tasks in multicore processors while accounting for potential interference due to shared memory. In order to do this, we first developed a memory partitioning mechanism and timing analysis that allowed different degrees of memory partition sharing between programs. This work won the best paper award in RTAS 2014 [KIM 2014].

The memory partitioning mechanism was implemented within the Linux/RK¹ kernel. We then created a memory partition allocation and timing analysis algorithm based on Mixed-Integer Linear Programming for tasks running under the Global Earliest-Deadline First scheduler. This work is in process of been published.

This work included the development of a global earliest deadline first (EDF) scheduler and a region partitioning mechanism to allocate different regions of memory from a program to different memory partitions under Linux/RK, and a memory-profiling tool based on the Valgrind profiling framework.²

Sub-project 2

The second sub-project addressed the scalability of functional verification of real-time software by combining two complementary techniques—counter-example guided abstraction refinement (CEGAR) and proof-based abstraction (PBA). CEGAR works with abstractions (over-approximations) and uses counterexamples to obtain finer abstractions iteratively. In contrast, PBA works with under-approximations and weakens them iteratively until an appropriate one is found.

¹ <https://rtml.ece.cmu.edu/redmine/projects/rk>

² <http://valgrind.org/>

High-Confidence Cyber-Physical Systems (HCCPS)



We developed theories and analyses that help assure these conditions —timing correctness via real-time schedulability analysis, functional correctness via model checking, and quantitatively assured coordination via probabilistic model checking.

We created a new model checker that combined over-and-under-approximations, and used it as the backend engine for our sequentialization-based tool (REK) for verifying real-time software. This research is reported in publications [KGC14, CGS14]. The model checker (and examples we used for validation) are available publicly.³

Sub-project 3

The third sub-project developed a new algorithm to predict the probability of success of a coordination algorithm when used by multiple heterogeneous agents operating under uncertain environment. The technique also provides a confidence interval about the prediction so that we know how accurate it is. To this end, we used a new “fuzzy sampling” technique that we developed as part of this project. We validated the technique on an example involving a coordinated mine-detection mission by a group of Kilobots.

Using our approach, we were able to make predictions about which group of Kilobots would have the highest likelihood of detecting the mine and reporting it back to the base station. Our research is reported in a publication [CGKL14], and our tools and examples are available publicly.⁴



We endeavor to apply the best combination of thinking, technology, and methods to the most deserving government software-related problem sets, free from conflict of interest.

³ <http://spacer.bitbucket.org>, <http://www.andrew.cmu.edu/user/arieg/Rek>

⁴ <https://db.tt/Wc9tBsNd>

Software Assurance Engineering—Integrating Assurance into System and Software Engineering

Principal Investigator

Carol Woody, PhD
CERT Division
cwoody@cert.org
(412) 268-9137

Internal Investigator

Chris Alberts,
CERT Division

Software assurance (SwA)—implementing software with a “level of confidence that software functions as intended and is free of vulnerabilities, either intentionally or unintentionally designed or inserted as part of the software, throughout the lifecycle”¹—has been legislatively mandated for the U.S. Department of Defense (DoD). This worthy goal needs to be translated into practical actions that designers and developers can execute and the DoD can validate. We have identified two critical research tasks to further DoD capabilities in addressing SwA across the software lifecycle.

Task 1: Analyzing Cybersecurity Risk Early in the Software Lifecycle

Task Lead: Chris Alberts

During the acquisition and development of software-reliant systems, DoD program personnel normally focus on meeting functional requirements, often deferring security to later life-cycle activities. In fact, security features are usually addressed during system operation and sustainment rather than being engineered into a system.

Operational security vulnerabilities generally have three main causes: (1) design problems, (2) implementation/coding problems, and (3) system configuration problems. This proposal is focused primarily on analyzing design vulnerabilities that cannot be corrected easily during operations. Early detection and remediation of design vulnerabilities would help reduce residual security risk when a system is deployed. Scenarios relevant to target operational missions will be developed and analyzed to identify security risks, needed mitigations, and confirm requirements to address mitigations.

By applying this approach, acquisition and development organizations should be able to identify a more complete set of security requirements by moving beyond compliance to consider cybersecurity risks from a mission/operational perspective.

¹ From Section 933 of the NDAA 2013

Research Outcomes

In this effort, we have focused on improving capabilities for characterizing early lifecycle security risk. We formalized the Security Engineering Risk Analysis (SERA) method and published a paper describing it in *CrossTalk: The Journal of Defense Engineering*, September/October 2014, titled “Evaluating Security Risk Using Mission Threads.” In addition, we developed a training course that will be presented at the Annual Computer Security Applications Conference (ACSAC) in December 2014. A technical note describing the method in detail has been drafted for publication later in the fall. This research project includes collaboration with Travis Breau, PhD, Carnegie Mellon University (CMU) Institute for Software Research. The CMU research team explored ways to express the scenarios, risks and mitigations that support effective analysis and published the following papers:

- Hibshi, H.; Breau, T. D.; Riaz, M.; & Williams, L. “Towards a Framework to Measure Security Expertise in Requirements Analysis,” 13-18. In *Proc. IEEE 1st International Workshop on Evolving Security and Privacy Requirements Engineering (ESPRE)*. Karlskrona, Sweden, Aug. 25, 2014. IEEE, 2014.
- Hibshi, H.; Breau, T. D.; Riaz, M.; & Williams, L. “Discovering Decision-Making Patterns for Security Novices and Experts.” In Submission: *International Journal of Secure Software Engineering*.

Task 2: Using Software Quality Models to Support Software Assurance

Task Lead: Carol Woody

Quality models exist that have implemented systems with order of magnitude improvement in removing quality defects; specialization of these models has been explored for safety, which has similarities to security, and results have been good. This research project analyzed data from successful high software quality projects to determine what can be gained by applying specialized quality models to address safety and security defects. The resulting approach will allow projects to use defect data during the development process in order to predict appropriate progress toward effective operational security and safety. Building on known and validated modeling

Software Assurance Engineering—Integrating Assurance into System and Software Engineering

capabilities for quality provides feasibility and credibility for the use of modeling to address operational safety and security.

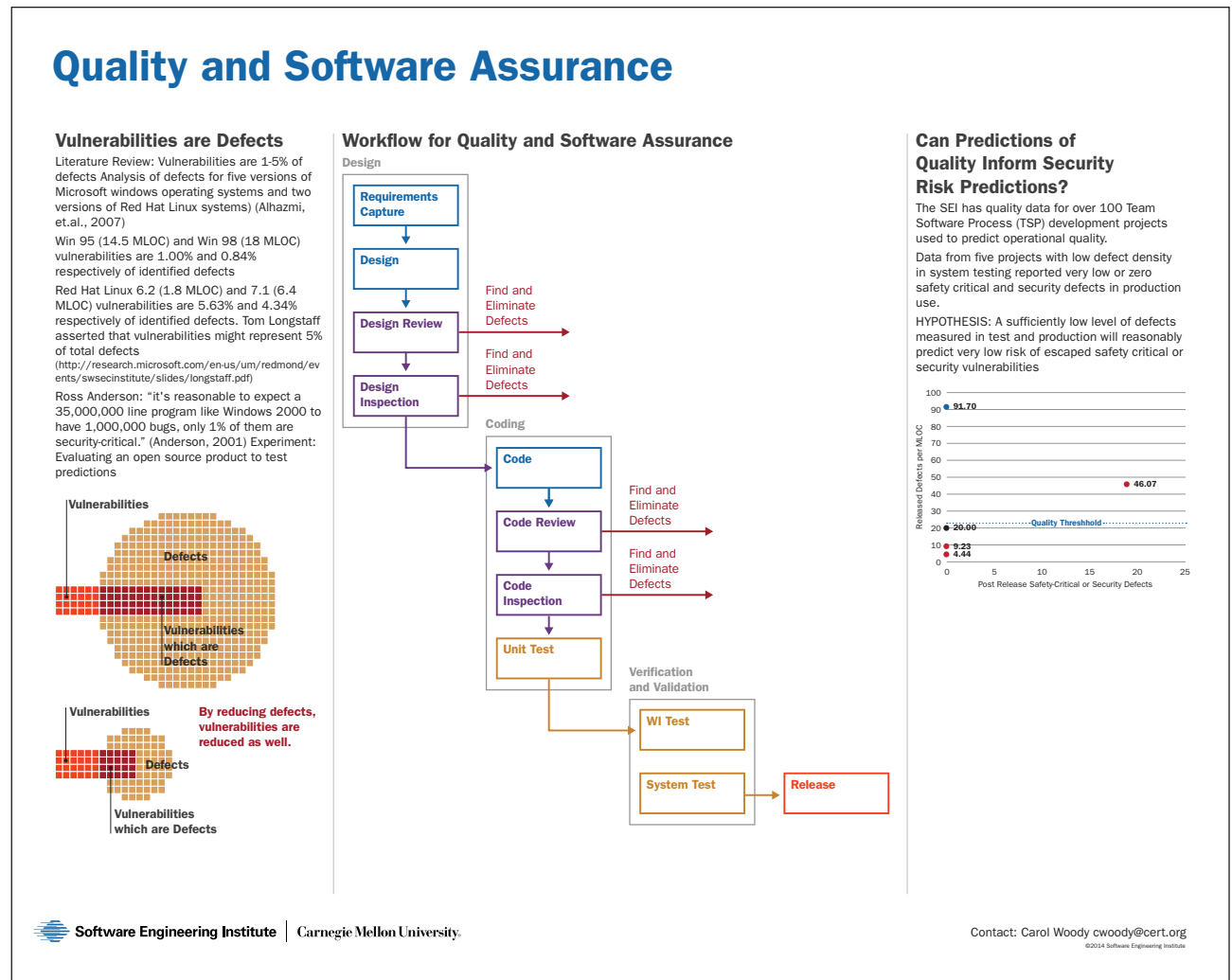
In this task, we considered the following key operational and technical challenges:

- Can software quality models be specialized to appropriately increase confidence that the software in development will be sufficiently secure and function as intended?
- Can available quality and vulnerability data (public or shared through collaborators) be used to effectively calibrate a specialized quality model to track and project security defects?

Although high-quality code is not necessarily secure, poor quality defective software cannot be secure. Therefore, some minimum level of quality software may be considered necessary for secure code. There is general agreement that good quality is an essential condition for software with security requirements. The level of necessary quality is an open question.

Research Outcomes

This research evaluated projects with high quality that also had excellent security and safety results to determine how these results were achieved and their potential applicability to DoD assurance needs. We conducted a workshop with representatives from the projects in our evaluation and participants from government, industry, and academia to review the results of our analysis and consider challenges of broader applicability. A technical note expected to be published later this fall has been drafted describing the projects evaluated, the challenges of security prediction and key findings from the workshop.



Representation of a poster developed to describe this project

Secure Coding

Principal Investigator

Robert C. Seacord
CERT Division
rcs@cert.org
(412) 268-7608

Because of limited resources, computer security incident response teams (CSIRTs) are typically unable to respond to the large number of vulnerabilities reported each year.

The goal of the CERT Secure Coding Initiative is to reduce the number of vulnerabilities to a level that can be mitigated fully in DoD operational environments. This will be accomplished by preventing coding errors or discovering and eliminating security flaws during implementation and testing.

CERT has been extremely successful in the development of secure coding standards that have been adopted at corporate levels by companies such as Cisco and Oracle and the development of the Source Code Analysis Laboratory (SCALE) that supports conformance testing of systems against these coding standards. The success of the secure coding standards and SCALE contributed to the impetus for the inclusion of software assurance requirements in the National Defense Authorization Act (NDAA) for Fiscal Year 2013.

Research Outcomes

Compiler-Enforced Buffer Overflow Elimination

Buffer overflow is the leading cause of software security vulnerabilities. It is responsible for 14 percent of all vulnerabilities and 35 percent of critical vulnerabilities (Common Vulnerability Scoring System score of 10) over the past 25 years, as reported by Sourcefire [Younan 2013].

CERT has completed a multiyear effort to modify the LLVM compiler to enable hoisting bounds checks from loops and functions [Keaton 2014]. This proof-of-concept prototype has been used to demonstrate how these optimizations can be performed reliably on bounds checks to improve their performance. However, the performance of bounds propagation is the dominant cost, and the overall runtime cost for bounds checking for C remains expensive, even after these optimizations are applied. Nevertheless, optimized bounds checks are adequate for non-performance-critical applications, and improvements in processor technology may allow optimized bounds checking to be used with performance-critical applications.

A valuable follow-on study would be to explore performance gains on future Intel processors that include MPX hardware assistance. If performance proves adequate, then buffer overflow checking could be left in place in deployed software systems.

C and C++ Thread Safety Analysis

With the rise of multi-core processors, concurrency has become increasingly common. The broader use of concurrency, however, has been accompanied by new challenges for programmers, who struggle to avoid race conditions and other concurrent memory access hazards when writing multi-threaded programs. CERT's approach is to allow developers to define a thread usage policy, which specifies which threads are permitted to execute particular code segments or to access particular data fields. CERT and Google developed thread safety analysis for Clang, which uses annotations to declare and enforce thread safety policies in C and C++ programs [Hutchins 2014]. Static analysis tools can help developers by allowing threading policies to be formally specified and mechanically checked to detect potential race conditions and deadlocks.

Clang is a production-quality C++ compiler that is available on most platforms, and the analysis can be enabled for any build with a simple warning flag: `-Wthread-safety`.

The analysis is deployed on a large scale at Google, where it has provided sufficient value in practice to drive widespread voluntary adoption. The need for annotations has not been a liability; it even confers some benefits with respect to software evolution and maintenance [Hutchins 2014].

Secure Coding

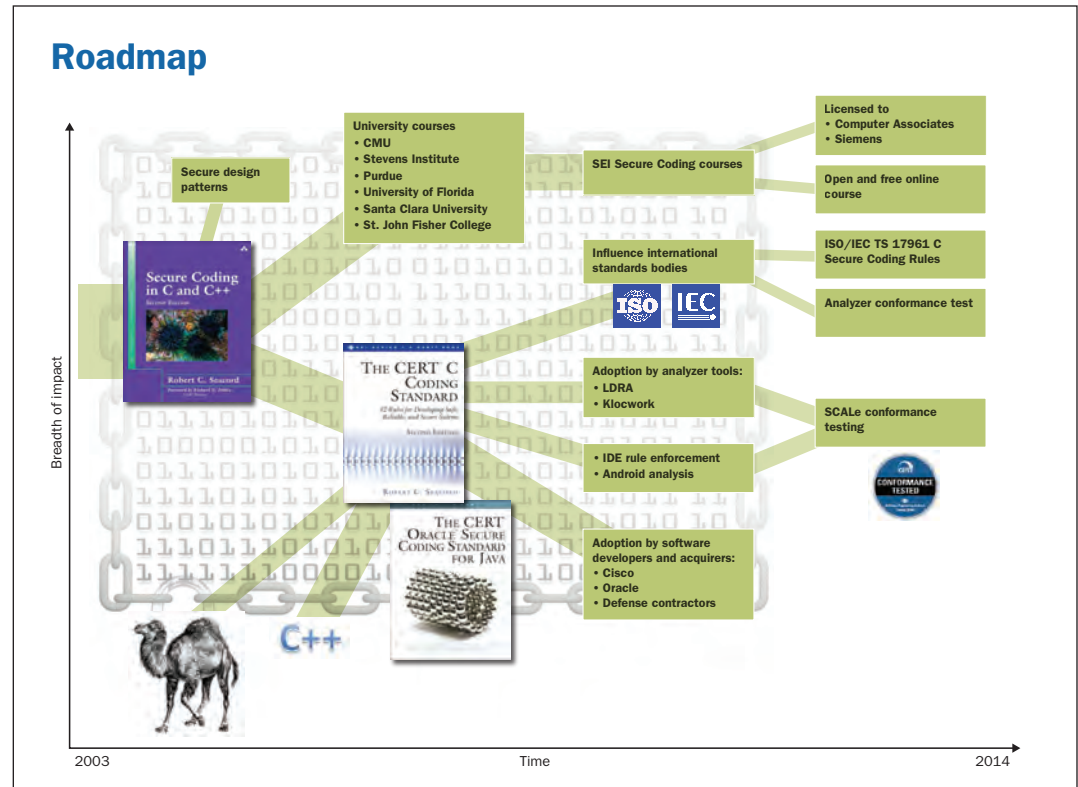
Android Taint Flow Analysis for App Sets

Mobile apps can potentially access a variety of sensitive information, such as a user's location, contacts, and the unique identifier of the phone (IMEI). Applications such as social networking and banking apps can additionally collect and store a large amount of sensitive data. A significant concern in this setting is exfiltration of sensitive data, which may violate users' privacy and allow undesired tracking of users' behavior. Popular Android apps have been shown to leak sensitive information including location, IMEI number, phone number, and the SIM card ICC-ID [Enck 2010]. Static analysis prior to this work simply analyzed tainted data flow across a single component. Consequently, malicious app developers evaded detection by ensuring that sensitive data flowed across multiple components before arriving at a restricted sink such as a website or an outgoing text message.

CERT developed the DidFail static taint analysis¹ for Android that combines and augments the FlowDroid and Epicc analyses to precisely track both inter-component and intra-component data flow in a set of Android applications to detect potential information leaks [Klieber 2014]. The analysis occurs in two phases: given a set of applications, we first determine the data flows enabled individually by each application, and the conditions under which these are possible; we then build on these results to enumerate the potentially dangerous data flows enabled by the set of applications as a whole. This two-phase approach allows fast analysis when installing a new app (phase 2) using pre-computed results from phase 1.

Potential future work includes enhancing the inter-component part of the taint flow analysis to include additional component types and data channels such as static shared fields and file system data flows. Adding context sensitivity to the analysis should improve the precision of the tool, making it practical for commercial software development projects and deployment in app stores.

¹ Available for download from <http://www.cert.org/secure-coding/tools/didfail.cfm>



SEI work in secure coding since 2003

Vulnerability Discovery

Principal Investigator


David Warren
CERT Division
dwarren@cert.org
(412) 268-9569

Vulnerabilities are pervasive in software-based systems and protocols, both in traditional IT networks and networks that support critical U.S. infrastructure such as weapons systems, supply chain, physical plant operations, and robotics. Some of these systems and protocols have been and remain isolated; however, the increasing connectivity of U.S. Department of Defense (DoD) networks continues to broaden the attack surfaces of many critical systems.

In addition, functionality is the driving force behind the DoD software acquisition process. Security is often a secondary concern, mainly due to the costs and technical shortcomings associated with security analysis of software and suppliers.

The goal of the FY14 Vulnerability Discovery project was to reduce the number of vulnerabilities in critical DoD and U.S. government (USG) systems by advancing and transitioning novel research in vulnerability discovery to high-impact DoD and U.S. government stakeholders.

This project is focused on advancing the state of the art in research and the state of practice of stakeholder operations in two categories of DoD-critical system security: (1) sound vulnerability discovery in traditional computing platforms and (2) vulnerability discovery in low-power, low bandwidth networked systems.



This project is focused on advancing the state of the art in research and the state of practice of stakeholder operations in two categories of DoD-critical system security: (1) sound vulnerability discovery in traditional computing platforms and (2) vulnerability discovery in low-power, low bandwidth networked systems.

the proposed techniques into the DoD software acquisition and support processes will result in software applications that are hardened—and more secure—before and after they are deployed into the DoD infrastructure.

Research Outcomes

This project seeks to mitigate these weaknesses by (1) advancing and facilitating the adoption of an automated process for sound vulnerability discovery and prioritization for traditional computing platforms (2) developing and transitioning vulnerability discovery techniques for networked control systems. Adoption of

Automatic and sound vulnerability discovery

- *Formalizing exploit types*

Formally defining new types of exploits so that we can soundly discover related vulnerabilities is an ongoing and challenging problem. However, our team has a track record of making steady progress in this space.

- *Advancing binary analysis*

Engineering a working automatic and sound exploit generation has involved innovations in modular symbolic execution, binary analysis, black-box fuzzing, corpus distillation, application command line interference, and several other research areas. Continuing systems research in binary analysis supports solutions to other key technical challenges, streamlines resource utilization, and supports deploying the system in operational contexts.

- *Executing on real code*

Operational relevance and real-world applicability are cornerstones of this research. While we measure the performance of our system against benchmark and research corpuses, we ensure that our system works against applications currently running in the DoD and other U.S. government organizations.

Low-power, low-bandwidth networked system vulnerability discovery

- *Developing generalizable vulnerability discovery techniques*

Developing novel techniques for vulnerability discovery in low-power, low-bandwidth, networked systems is an emerging area in security research. In order to impact a significant portion of DoD networks, our work will focus on developing techniques that can be generalized to multiple stakeholder-critical protocols and systems.

Vulnerability Discovery

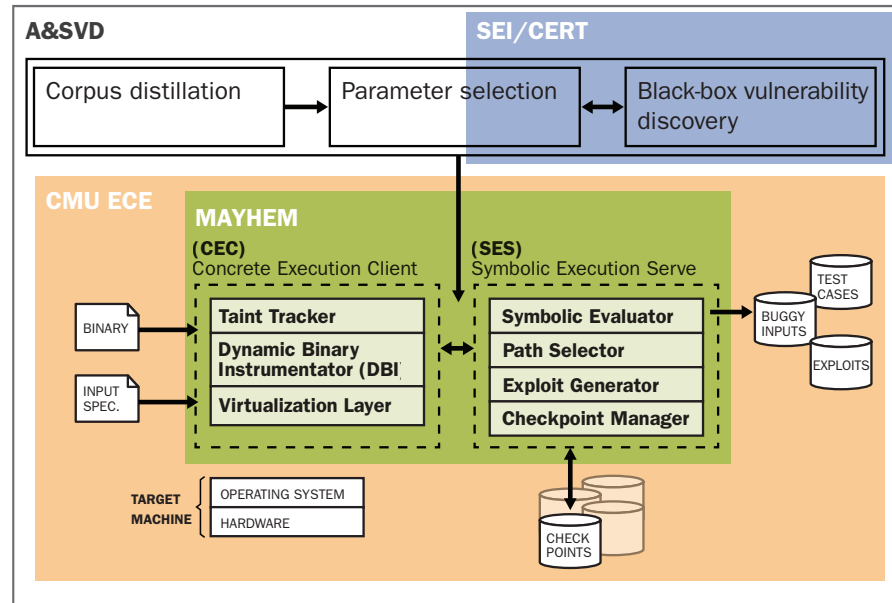
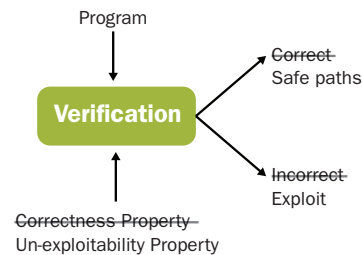
Task 1: Automated and Sound Vulnerability Discovery

Vision: Automatically check DoD software systems for exploitable bugs

Discover vulnerabilities automatically in compiled x86 applications

- “Zero false positives”
- Automatically generate an exploit for each vulnerability; no source code necessary

This project combines expertise from both SEI/CERT and CMU



Task 2: Low-Power Low-Bandwidth System Vulnerabilities

Focuses on vulnerability discovery in Things that have enough compute power to pose a threat to a network they are attached to, but not enough for somebody to think of them as a computer.

1. How do vulnerability discovery techniques for LPLB systems differ from those for traditional computing systems? Are there techniques that do not transfer well? Are there techniques that have shown efficacy in traditional computing but have not showed up on the LPLB side? Why?
2. What processes, tools, and techniques are most effective at improving the security of LPLB systems? For developers and creators of such systems For acquirers, deployers, and operators of such systems
3. What metrics can be applied to assess the efficacy and/or efficiency of those processes?

Simulating Malicious Insiders in Real Host-Monitored Background Data

Principal Investigators

Kurt Wallnau, PhD
CERT Division
kcw@cert.org
(412) 268-3265

Brian Lindauer
CERT Division
lindauer@cert.org
(512) 666-5438

Our task is to provide insider threat test data for a research program that is developing a new generation of (anomaly-based) insider threat detectors. The program has at its disposal a unique research resource: a secure data facility operated by an industry partner on behalf of the program that contains real (background) data gathered from approximately 5,000 employees from host-monitored computers deployed in the workplace. The data currently contains more than 2.5×10^8 user events and counting, with approximately 3 million new events gathered each day.

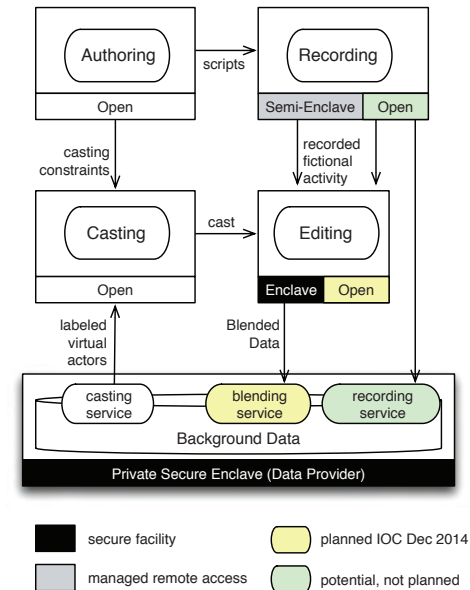
As others have done, we create threat data from a combination of synthetic threat data overlaid on real background data. The novelty of our approach is in the way in which we make constructive use of background data to create threat data with plausible threats that exhibit richly textured social realism. We regard background as an existing, real and deep fabula (the myriad details of fictional story that are implied but not described in fictional stories) as well as a source of actors whose real behavior can be augmented to simulate the activity of fictional characters in an insider threat drama.

There is solid grounding in the idea that fictional narratives play a fundamental role in the way we humans abstract, compress, simulate and share meanings, and a rich, largely untapped and surprisingly formal reservoir of dramatic theories upon which effective “cyber-social experiences” can be constructed.

Threats are specified as insider threat *cyberplays* with plausible plots and well-developed characters. Screenplays are compiled into user-level programs that simulate fictional characters on host-monitored computers in a virtual recording studio. In parallel, a central casting service is used to select from background real users to play

fictional roles that best match their job roles, social networks and patterns of activity. Fictional activity is then blended into the activity of real users in the cast. The cast of augmented real users and unmodified background users constitute simulated dramatic performances in test windows, with each performance constituting a single test case. Performances by different casts of users, or by the same cast of users in different test windows, constitute distinct test cases.

Indirect measures of the merits of the approach are encouraging, both in the number of peer-reviewed research articles published from the research (over 70 at last count) and in the low number of data artifacts (no confirmed reports in the past 20 months). Although dramatic and narrative theory is not often encountered in cybersecurity test and evaluation literature, we think this is likely to change. There is solid grounding in the idea that fictional narratives play a fundamental role in the way we humans abstract, compress, simulate and share meanings, and a rich, largely untapped and surprisingly formal reservoir of dramatic theories upon which effective “cyber-social experiences” can be constructed.



Simulating Malicious Insiders in Real Host-Monitored Background Data

		←- threat dramas →-																																					Feature
2	3	4	5	6	7	9	10	11	13	14	15	16	17	21	23	24	27	28	29	30	31	32	33	34	38	40													
			x																										AFFILIATED										
														x					x	x	X							CEMAILU											
x		x		x	x	x					x					x		x	x		x						COLOCATED												
	x										x					x											COTIMEZONES												
x						x								x				x	x			x					COWORKERS												
																											CONSYSADMIN												
						x																					CONTRACTOR												
x																		x									DEVELOPER												
													X		x	x										x	EMAILCOMS												
			x																x								EXTERNU												
																				x							FOREIGNCOM												
	x			x	x						x	x															FRIENDS												
														x													HTTPU												
											x							x							x		x	IMCOMS											
	x			x	x							x		x								x	x				IMU												
			x																								INTERNU												
																										x	LOCATION												
																										x	MANYHOSTU												
			x						x																		NETDRIVEU												
																							x				OFFLINE												
																x											POPULAR												
																										x	PVTEMAILCOMS												
																	x										RUNSPGM												
			x																								SUBCOLLABS												
		x		x		x									x			x		x			x			x	SUPERVISES												
																						x					SUPERVISOR												
							x		x														x	x	x		SYSADMIN												
x							x			x						x			x		x	x					TECH												
	x													x													WEBMAILU												
																											WEBSTORAGEU												

ID	Threat Drama	Threat Construct	#Perf
1	Anomalous Encryption	Theft of IP	3
2	Blinded Me with Science	Benign	1
3	Bollywood Breakdown	Espionage	1
4	Bona Fides	Espionage	3
5	Breaking the Stovepipe	Unauth Discl; IP Theft	3
6	Byte Me!	Fraud	2
7	Byte Me! Middleman	Fraud	1
8	Circumventing Sureview	IT Sabotage	2
9	Conspiracy Theory	IT Sabotage	2
10	Credit Czech	Fraud	1
11	Czech Mate	Fraud; Espionage	1
12	Exfil With Steganography	Theft of IP	1
13	Exfil Before Layoff	Theft of IP	2
14	Exfil With Screenshots	Theft of IP	2
15	From Belarus With Love	Theft of IP; Espionage	1
16	Gift Card Bonanza	Fraud	1
17	Hiding Undue Affluence	Espionage	3
18	Indecent RFP	Fraud	1
19	Indecent RFP 2	Fraud	2
20	Insider Startup	Theft of IP	2
21	Job Hunter	Theft of IP, Espionage	1
22	Layoff Logic Bomb	IT Sabotage	2
23	Manning Up I	IT Sabotage	2
24	Manning Up II	Espionage	1
25	Masquerading I	Misc	1
26	Masquerading II	Misc	1
27	Naughty by Proxy	Misc	4
28	Outsourcer's Apprentice	Fraud	3
29	Panic Attack	Espionage	2
30	Parting Shot	Theft of IP, IT Sabotage	1
31	Parting Shot: Deadly Aim	Theft of IP, IT Sabotage	1
32	Passed Over	IT Sabotage	3
33	Selling Login Credentials	Fraud	1
34	Snowed In I	Espionage	1
35	Snowed In II	Espionage	3
36	Stealing Login Credentials	Fraud	1
37	Strategic Tee Time	Benign	1
38	Survivor's Burden	Theft of IP	3
39	The Big Goodbye!	Benign	1
40	What's the Big Deal?	Theft of IP	1

Above: Cross-reference of a sample of threat dramas and the "central casting" features used to select users to play dramatic roles. Not shown: bespoke casting features developed for specific threat dramas (those that involve non-automatable corpus research) or threat dramas developed early in the program that emphasized technical exploits rather than social complexity.

At right: Threat dramas, the threat construct sampled by the drama, and the number of times the drama was performed (each performance in a different 30-day test window)

Insider Threat Mitigation

Principal Investigators

Andrew P. Moore
CERT Division
apm@cert.org
(412) 268-5465

William R. Claycomb, PhD
CERT Division
claycomb@cert.org
(412) 268-8931

Despite the high impact of insider attacks, organizations struggle to implement effective insider threat programs. The stove-piped nature of many organizations combined with the tendency to view insider threat as an information technology problem has made it difficult for organizations to deal effectively with the socio-technical nature of the insider threat problem. The challenges to secure systems against the malicious insider are significant:

- Insider attacks are low frequency but high impact events, making scientific validity hard to establish.
- The behaviors of malicious actors and good employees can be extremely difficult to distinguish, making insider threat signal detection challenging.
- Insider attacks can use unprecedented tactics, often with new and emerging technologies.

Our goal in FY14 was two-fold:

1. Evaluate and improve insider threat mitigation pattern effectiveness through piloting in partner organizations (Task 1)
2. Develop and evaluate technical solutions for emerging insider threats, with a focus on insider-facilitated espionage (Task 2)

By moving from the ongoing threat analysis into trends that help characterize the emerging threat, we are able to understand the requirements for next-generation enterprise system architectures needed to defend effectively against insider attacks.

Task 1 of our effort pilots individual insider threat mitigation controls in partnering organizations. Key partners are sharing data and acting as pilot sites for insider threat controls developed. Lessons learned during the pilots will be compared with data analysis results previously conducted and used to refine the controls tested, described as insider threat mitigation patterns. Key pilot performance metrics include the total number of alerts per indicator, false/true positive rates, average analyst time per event, and the number of events referred for

secondary analysis. In this phase of our work, establishing field significance during pilot testing will be challenging due to the low base rate for testing control effectiveness. We have already started to review the existing architecture of partners' insider threat programs to determine possible pilot testing opportunities.

Task 2 of our effort leverages CMU's experience in social network analysis to better understand the potential for using insiders' social networks as a basis for insider threat risk indicators and early warning. Insiders' expanded use of online social networks (OSNs) makes this a potentially rich source



By moving from the ongoing threat analysis into trends that help characterize the emerging threat, we are able to understand the requirements for next-generation enterprise system architectures needed to defend effectively against insider attacks.

of improved risk information. We hypothesize that, over time, insider social networks exhibit weakening of internal ties (i.e., social connections), promoting the strengthening of external ties to adversaries. Although the strength of internal ties may decrease, the variety of those ties may increase as the insider accumulates information of value to the external ties. The provision of information to external ties strengthens those ties. External ties to suspicious individuals increase associated risk measures.

Some potential ways of measuring tie strength include the reciprocity, frequency, duration, emotional intensity, and honesty of communication between individuals. Social network analysis can play an important role in quantifying the risk due to malicious as well as inadvertent insider threats.

Insider Threat Mitigation

Research Outcomes

The objective of this continuing work is to develop specific insider threat mitigations that form an architectural foundation for next-generation U.S. Department of Defense (DoD) enterprise systems and technologies. Mitigation patterns and pattern languages developed through this research will enable coherent reasoning about how to design and implement DoD enterprise systems to protect against insider threat, benefiting federally mandated insider threat programs being formed across the U.S. government.

Instead of being faced with vague security requirements and inadequate security technologies, DoD enterprise system designers, evaluators, and incident investigators will be armed with a validated set of mitigation patterns that will enable them to develop and implement effective strategies against the insider threat in a timelier manner and with greater confidence.

Insider Threat Mitigation Project

A Dynamic Network Approach

Emergence of Threat – Ego centered analysis of specific cases

Approach:

- Semi-automated coding with fine-tuning to add dates
- Extract meta-networks one per year
- Comparison at “role” level
- Apply network analytics and visualization

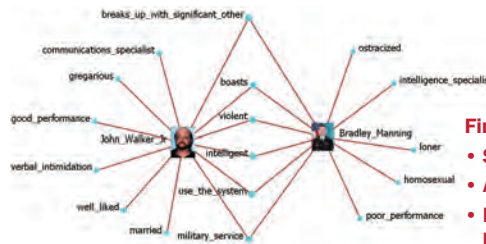
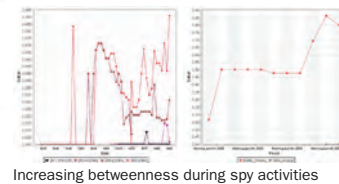
Walker – Gang example

Case records/searches (open-source)



Manning – Lone Wolf example

open-source



- #### Findings on Insiders:
- **Special characteristics**
 - **Access**
 - **Increasing betweenness**
 - **Disrupted family network**

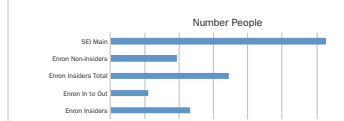
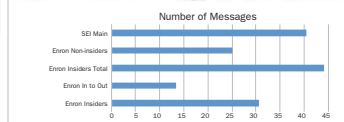
Software Engineering Institute | Carnegie Mellon University

Emergence of Threat – Email centered analysis of possible anomalies

Approach:

- Networks formed from meta-data
- One network per year
- Segment internal from internal-to-external communication
- Remove suspected distribution lists
- Identify “normal behavior” using Enron
- Develop pattern for “insiders” in contrast to “normal” using Enron
- Apply to anonymized SEI email

- CMU-CS (and CASOS):**
- Dr. Kathleen Carley
 - Neal Altman
 - Geoff Morgan
 - Matt Benigni
- SEI:**
- Matthew Collins
 - Andrew Moore
 - Dr. William Claycomb



- #### Findings on SEI -v- Enron:
- **SEI—more email, proportions similar**
 - **Both—dominant dense core with numerous stars**

Enron core for 2001—Newman group coloring



SEI core for 2013—Newman group coloring



- #### Findings on “Insiders”—those accused:
- **Are not “top” network actors**
 - **Form a densely connected sub-group**
 - **High level of in-group communication**
 - **Low out-group communication**

Center for Computational Analysis of Social and Organizational Systems

Contact: Andrew P. Moore apm@cert.org
Kathleen Carley kathleen.carley@cs.cmu.edu

©2014 Software Engineering Institute

Representation of a poster developed to describe this project

Deep Focus: Increasing User Depth of Field to Improve Threat Detection

Principal Investigator

William R. Claycomb, PhD
CERT Division
claycomb@cert.org
(412) 268-8931

Internal Investigators

Jason Clark, PhD
CERT Division

Brian Lindauer
CERT Division

Bronwyn Woods, PhD
CERT Division

The need to detect malicious behavior and unauthorized information disclosure on sensitive systems is of paramount importance to the U.S. Government. This was recently reinforced by high-profile classified information leaks by Bradley Manning and Edward Snowden.

As a recognized leader in insider threat research, CERT is leading the way in finding answers to improve detection capabilities and prevent future leaks. We believe the next step in the insider threat research roadmap is developing a fundamental understanding of individual users. While current detection efforts rely on matching signatures of suspicious behaviors across all users, ongoing research like DARPA's ADAMS project is significantly improving anomaly detection at the user level.¹

However, it is unclear whether the significantly high false positive rates still encountered by researchers can be overcome with existing analysis techniques and available data. Simply put, real user behavior in actual operational environments likely has a high frequency of anomalous events, overwhelmingly benign, that cannot be distinguished from malicious events without additional context.

The usefulness of incremental approaches in solving this problem appears to be limited; we believe an innovative solution is necessary. We propose development of new analysis techniques that focus on data representing ordinary user behaviors that users are unlikely to realize are being monitored.

Examples include, but are not limited to: patterns of application launch, use of specific shortcuts (Ctrl+C vs. Edit -> Copy), web browsing behavior, typing speed and error rate, mouse use [Shen 2012], system calls by user applications [Song 2013], network connections on behalf of the user [Kent 2013], or sentiment/linguistic analysis of email or instant messages.

We believe that combinations of these behaviors can at best uniquely identify users among their peers with high confidence, and at worst provide a high-confidence measure that the individual interacting with the IT system is not the authorized owner of the account being used. By understanding the unique way each user interacts with IT systems, we can detect account misuse (masquerading) as well as significant deviations from normal behavior that when combined with signature or anomaly based threat detection data strongly indicate malicious behavior versus a benign anomaly.


Goals

We have two specific goals:

1. Develop a measure of confidence that the person currently interacting with the IT system is or is not the authorized user. Visualize this metric for analyst use.
2. Provide context by which insider threat and anomaly detection engines can determine with higher confidence that suspicious behavior is malicious.

Research Outcomes

- Algorithms to analyze statistical properties of user behavior that distinguish users from one another
- Methods for incorporating algorithms into operational threat detection programs
- A tool to visualize individual user behavioral attributes, including perceived risk to the organization. This task will be conducted jointly with the Cybersecurity Center at Oxford University.
- Large data sets of features describing user behavior, for future research (owned by and hosted at partner sites, though anonymized data sets may be retained by SEI)



As a recognized leader in insider threat research, CERT is leading the way in finding answers to improve detection capabilities and prevent future leaks. We believe the next step in the insider threat research roadmap is developing a fundamental understanding of individual users.

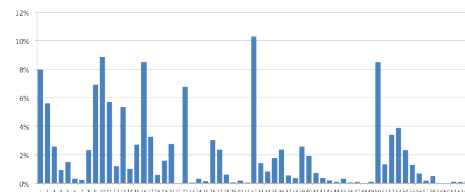
¹ The DARPA Anomaly Detection at Multiple Scales (ADAMS) project creates, adapts, and applies technology to anomaly characterization and detection in massive data sets. For more information, visit [http://www.darpa.mil/Our_Work/120/Programs/Anomaly_Detection_at_Multiple_Scales_\(ADAMS\).aspx](http://www.darpa.mil/Our_Work/120/Programs/Anomaly_Detection_at_Multiple_Scales_(ADAMS).aspx)

Deep Focus—Increasing User “Depth of Field” to Improve Threat Detection



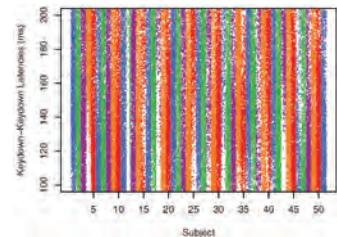
Linguistic Patterns

Characteristics of a user’s speech or writing can be measured both structurally and linguistically. Using these metrics, researchers have shown the feasibility of identifying anonymous authors [Narayanan 2012]. Others have observed measurable changes in linguistic patterns of known insiders [Taylor 2013].



Keystroke/Mouse Biometrics

Using metrics like keystroke latency or mouse dynamics, researchers have shown how individual users can be identified [Shen 2013]. Furthermore, evidence suggests changes in a user’s personal state, such as increased stress, is also detectable.



Our Approach

Current insider threat detection techniques often focus on specific indicators of malicious activity, such as unusual file copy or printing activity. While current work shows promise in reducing high false-positive rates for this type of detection, we find that many insiders operate “under the radar,” carrying out malicious activity within the scope of authorized activity not detected by standard signature or anomaly-based detection mechanisms.

Our approach considers various metrics of user behavior and interaction with IT systems that are also “under-the-radar”, or cannot be easily manipulated or spoofed by a determined attacker. Examples include network authentication, keystroke & mouse biometrics, linguistic patterns, and host-based interaction. We intend to create profiles of user behavior that can identify malicious behavior and/or provide critical contextual information to existing insider threat detection mechanisms.

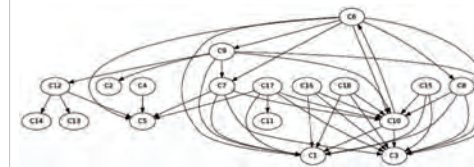
We have three specific goals: develop a measure of confidence that the person currently interacting with the IT system is or is not the authorized user, increase the efficiency with which anomalous benign behavior is distinguished from malicious behavior, and visualize our approach for analyst use.

Visualization

Issues of concern must be visible and apparent to analysts. In cooperation with the Cyber Security Centre at Oxford, we will extend an existing insider threat visualization toolkit to represent data and anomalous activity in a clear and actionable manner.

Network Authentication Graphs

These directed graphs represent a user’s authentication activity between networked computers over a predefined period. Research shows that administrative users generally have larger more complex graphs than normal users [Kent 2013]. Furthermore, it is possible to profile each user’s authentication activity, resulting in the ability to detect abnormal, potentially malicious, activity. Empirical research on malicious insiders shows that many insiders engage in reconnaissance and gathering activities, accessing numerous network locations that often differ from the insider’s normal work activity.



The network authentication graph from a typical user with administrative access. This user accessed 18 computers with 41 authentication arcs. This is a more complex authentication graph than those of general users. [Kent 2013]

References

[Taylor 2013] Taylor, P.; Dando, C.; Ormerod, T.; Ball, L.; Jenkins, M.; Sandham, A.; Menacere, T. “Detecting insider threats through language change.” Law and Human Behavior, Vol 37(4), Aug 2013, 267-275

[Kent 2013] Kent, A.; Liebrock, L.M., “Differentiating User Authentication Graphs,” IEEE Security and Privacy Workshops, May 2013

[Killourhy 2008] Killourhy, K. and Maxion, R. “The Effect of Clock Resolution on Keystroke Dynamics.” In 11th International Symposium on Recent Advances in Intrusion Detection (RAID-08), 2008.

[Narayanan 2012] Narayanan, A.; Paskov, H.; Gong, N.Z.; Bethencourt, J.; Stefanov, E.; Shin, E.C.R.; Song, D., “On the Feasibility of Internet-Scale Author Identification,” IEEE Symposium on Security and Privacy, May 2012

[Shen 2013] Shen, C.; Cai, Z.; Guan, X.; Du, Y.; Maxion, R., “User Authentication Through Mouse Dynamics,” IEEE Trans. on Info. Forensics and Security, Jan. 2013

Team CMU/SEI:
William R. Claycomb, Ph.D.
Bromwyn Woods, Ph.D.
Brian Lindauer

Jason Clark
Roy Maxion, Ph.D.

U. Mass at Amherst:
David Jensen, Ph.D.

University of Oxford:
Sadie Creese, Ph.D.
Jason Nurse, Ph.D.
Phil Legg, Ph.D.

Los Alamos National Labs:
Joshua Neil, Ph.D.
Alex Kent

Partners



Software Engineering Institute | Carnegie Mellon University

Contact: Bill Claycomb <claycomb@cert.org>
©2014 Software Engineering Institute

Representation of a poster developed to describe this project

Malware Analysis

Principal Investigator

Ed Stoner
CERT Division
ers@cert.org
(412) 268-6187

Internal Investigators

William Casey, PhD
CERT Division

Sagar Chaki, PhD
Software Solutions Division

Cory Cohen
CERT Division

Jeffrey Gennari
CERT Division

Arie Gurfinkel, PhD
Software Solutions Division

Jeff Havrilla
CERT Division

Charles Hines
CERT Division

Leigh Metcalf, PhD
CERT Division

Soumyo Moitra, PhD
CERT Division

Rhiannon Weaver, PhD
CERT Division

Presenter

Jonathan Spring
CERT Division

Task 1: Automatic static analysis of malware binaries

Automation of static analysis of malicious binaries amplifies the effort of a limited pool of malware analysts and accelerates insight generation captured by higher-level abstractions accessible to more network defenders within the U.S. Department of Defense (DoD).

Analyzing large numbers of malware attacking the DoD worldwide infrastructure is a time-consuming process. Malware analysis requires specialized skills, and when confronted with novel malware binaries, malware analysts can spend days (or even weeks) reverse-engineering a single sample.

This bottleneck in the process of deriving actionable insights by understanding the threat presented by malware can be mitigated by both automating repetitive tasks and providing more semantically rich abstractions used by a malware analyst and others who use his or her results.

Program analysis techniques can be useful in providing automated understanding of software behavior. Similar techniques have been used to characterize both non-malicious and malicious binary executables. We propose using a compiler transformation framework called ROSE [Quinlan 2000] to leverage well-established program analysis techniques to analyze malware binaries at a larger scale than is currently done by human analysts.

Task 2: Suffix trees¹

In order to advance DoD capabilities to reason about uncharacterized data, it is fundamentally important to optimize the performance characteristics of basic operations supporting data mining and novel similarity studies for large-scale data. Specifically we are focusing on the goal of optimizing suffix-tree data structures for the identification of longest common substring (LCS) in the area of malcode analysis (code-clones) and zero suppressed binary decision diagrams (ZDDs) for compact representations of set families.

¹ The design of optimized suffix-data and ZDD data structures with application to malware threat discovery and code clone detection



Program analysis techniques can be useful in providing automated understanding of software behavior. Similar techniques have been used to characterize both non-malicious and malicious binary executables. We propose using a compiler transformation framework called ROSE [Quinlan 2000] to leverage well-established program analysis techniques to analyze malware binaries at a larger scale than is currently done by human analysts.

Improvements made to address larger scale data impact DoD problems of malware threat discovery and product chain validation. For malware threat discovery, the attribution problem is an instance of determining the existence of a shared provenance given a set of known malicious binary files. The concept of additive cloning could resolve questions of code derivation and allow the DoD to answer the question of which malware components came first in a set of malware in related intrusion sets. For product chain validation, provenance inference can assert that discrepancies from an expected code-base (the gold standard including version history) are either acceptable variations or anomalies for additional testing.

Malware Analysis

Task 3: Malicious behavior and model checking

We aim to formally describe the common or invariant behaviors which malicious software express and further develop an effective methodology to identify and classify malicious behaviors by resolving separable features of software behavior in general. The methodology to be developed aims to enhance data-discovery, rapid characterization, and determination of malicious behavioral signatures—thereby enhancing current research addressing the outstanding practical problems to identify and classify malware by behavior.

The challenge is to describe formally software behavior and be able to determine if the behavior is malicious. In order to achieve this goal we shall conduct the following tasks:

1. Construct an accurate binary instrument for trace capture called a system trace monitor; test its capability on control software and known software.
2. Apply the trace monitor to a set of publicly available software and a set of publicly available malware data. Upon completion of this task our set of traces held in ASCII alpha-numeric form are encoded to numerical sequences to facilitate the analysis and pattern matching tasks.
3. Apply data analysis techniques to the trace data. Analyses of trace data include model checking, machine learning, clustering analysis, and string/pattern matching with the intent to determine features that link software by behavior.

By drawing the links between patterns within software's trace and its behavior, we more formally develop methods to classify software traces as malicious or benign within the formal language of hyperproperties. To interpret the meaning of ordered behavior patterns, we utilize model checking and machine learning techniques and explore its feasibility.



While other FFRDCs and research centers are also attentive to the government's problems, the SEI brings its unique combined capabilities in cybersecurity and software together with its university affiliation and industry access to bear on important and challenging software-related problems—in acquisition, development, testing, security, safety, operations, and sustainment.

Malware Distribution Networks

Principal investigator

José Andre Morales, PhD
CERT Division
jamorales@cert.org
(412) 268-9392

William Casey, PhD
CERT Division
wcasey@cert.org
(412) 268-3002

Presenter

Aaron Volkman
CERT Division

An operational challenge in malware detection is to efficiently identify persistent URLs and potential cyber-attacks early enough to execute proper action.

In this research, we attempt to graph URL-based malware distribution networks (MDNs) by leveraging the Google and Bing search engines and the Google Safe Browsing (GSB) data set as our primary data source. Graphing MDNs by leveraging search engines and their public services is novel in the field. Other approaches focus on identifying malicious URLs but do not deep-dive into constructing their respective MDN graphs.

The potential impact of this work is the identification of URLs and subnetworks that should be shut down or blocked and the prediction of potential upcoming cyber-attacks.

Task 1: Build and analyze MDN graphs

Our technical challenge in this task is to create and analyze MDN graphs correctly using data from GSB. Studying GSB reports can provide detailed information on why GSB flags a URL as suspicious. These explanations may include observed traffic of the URL, leading to the assessment of one or more of its roles.

Graphing MDNs multiple times a day will give a window into the intra-day and daily changes of various MDNs. This facilitates (1) the identification of persistent subnetworks and URLs that may be critical to the MDN's operations and (2) trend analysis to understand dynamic behaviors of an MDN, its ability to resist and recuperate from attacks, and its potential future cyber-attacks.

Identifying persistent URLs and subnetworks facilitates the discovery of potentially critical infrastructure of an MDN that can greatly disrupt operation if blocked or shutdown. Identifying persistence in this manner enhances current state-of-the-

art, which identifies malicious URLs but does not consistently pursue the associated MDN and potential persistent sub-components.

Trend analysis of MDNs can reveal certain topological changes required when preparing for, carrying out, or winding down a malicious cyber-attack. Knowing these changes can assist in detecting when a future malicious cyber-event is being prepared, allowing for appropriate action to prevent or mitigate it. Trend analysis can also provide fundamental understanding of MDN resilience to publicly known attempts to disrupt or take down its infrastructure. This form of resilience analysis is novel in the field and can help in choosing techniques for future takedown/disruption attempts. Identifying the malware traversing an MDN can provide attribution of its use in large-scale malicious cyber campaigns—providing deeper understanding of its role, administrators, and use within malware distribution.

Task 2: Create an interactive MDN mapping system

By graphing an MDN with a predetermined suspicious URL set from GSB, we will create a mapping system that will identify the role of a URL in the MDN and its connectivity to other URLs. The technical challenge is to provide efficient rendering of MDNs, allowing requests for trend analysis, MDN structural changes, and persistent URLs and subnetworks. To meet this challenge we will

1. design and implement the necessary framework to display MDNs as they are created throughout the day
2. give access to archives for various tasks such as trend analysis and persistent URL and subnetwork identification

Research Outcomes

The expected outcome is a real-time MDN graphing system that can alert to potential cyber-attacks and identify persistent URLs and subnetworks via trend analysis. Our research fundamentally enhances the field by providing a way to graph MDNs accurately using already discovered malicious URLs. This allows for rapid graphing and analysis of MDNs over long periods of time.



Our research fundamentally enhances the field by providing a way to graph MDNs accurately using already discovered malicious URLs. This allows for rapid graphing and analysis of MDNs over long periods of time.

Behavior-Based Analysis and Detection of Mobile Devices

Principal Investigator

José Andre Morales, PhD
CERT Division
jamorales@cert.org
(412)268-9392

Presenter

Joseph Yankel
CERT Division

Both the number of Android OS-enabled mobile device users and the malware targeting Android OS are growing exponentially. This poses a clear and present danger to users of these devices that are currently poorly protected and highly vulnerable to malware infection.

The number of new Android apps appearing on authorized app markets such as Google Play is growing daily and there is not a practical, effective approach to analyze them for suspicion assessment, which heightens the probability of a malicious app getting past any security filters and ending up being installed on multiple devices. The purpose of our research is to (1) disallow malicious apps from ever being available for download via app markets and (2) detect and triage apps exhibiting malicious behaviors on an actual mobile device.

This research provides impact to U.S. Department of Defense (DoD) and other related agencies in two significant ways:

1. facilitate the identification of potential malware early enough to avoid damage to a mobile device
2. provide fast and accurate suspicion assessment within a controlled environment of an Android app to an analyst

In general, in order to assess suspicion in an object, which is an Android app for the purposes of this research, we take the approach of determining “who you are” and “ what you do” with the following meanings:

- Who you are = source of download
- What you do = the sequence of interactions with the operating system



This research seeks to develop an effective and efficient behavior-based analysis approach capable of accurate suspicion assessment of software for mobile devices.

The technical challenge in answering these two questions is assessing how much and what kind of data is required to give a high confidence answer. Our approach to addressing this technical challenge is as follows:

1. Assess the market from which the Android app under analysis was downloaded.
2. Build activity graphs based on data collection at the Linux kernel level. An activity graph represents the key processes and threads associated with a currently executing Android app.
3. Track and document instances of execution events for each node on the activity graph. The execution events we track deal with the file system, other non-related processes, memory, network activity, and use of mobile device specific features such as Global Positioning System (GPS) and Short Message Service (SMS) messaging.
4. Analyze each node’s activities to identify implementations of suspicious execution events and then map those events to abstract suspicious behaviors.

Research Outcomes

This research seeks to develop an effective and efficient behavior-based analysis approach capable of accurate suspicion assessment of software for mobile devices.

Data-Intensive Systems

Principal Investigators

John Klein
Software Solutions Division
jklein@sei.cmu.edu
(412) 268-5981

Ian Gorton, PhD
Software Solutions Division
igorton@sei.cmu.edu
(412) 268-6872

The need to capture, query, and manage data in petascale repositories has become pervasive in U.S. Department of Defense (DoD) mission areas such as C4ISR, healthcare, flight data management, and logistics. These data-intensive systems [Gorton 2012a] present significant software engineering challenges. They are typically systems of systems, comprising separately developed and evolved data repositories and applications deployed on independently operated data center and network infrastructures.

The fundamental design principles for data-intensive systems are poorly understood and the wide spectrum of design and implementation options exacerbates the design complexity—leading to suboptimal technical solutions prone to instability, lack of scalability, and cost/schedule overruns.



Our unique focus is on the tight coupling of software, data, and deployment architectures that exists in these big data systems [Gorton 2014], an area of investigation that has not been previously been explored in the research community.

In this project, we are addressing these issues through codifying enduring architecture design principles for data-intensive systems. Our unique focus is on the tight coupling of software, data, and deployment architectures that exists in these big data systems [Gorton 2014], an area of investigation that has not been previously been explored in the research community. Specifically, we are pursuing the following activities.

Task 1:

To support software architects' navigation through the data-intensive systems design space, this project will capture, organize, and codify design knowledge so that the broad community of DoD software architects can leverage it.

Task 2:

We will extend this rich collection of quality attribute scenarios and related tactics by associating them with concrete realizations and code examples for a single data management technology of interest to our collaborators (e.g., MongoDB, Cassandra¹).

Task 3:

To codify this collection of design knowledge into a semantic wiki, we will leverage and extend successful work in architecture knowledge management description and organization [Kruchten 2004, Ali-Babar 2007, Buschmann 2007] and our experience in building scientific knowledge management systems [Gorton 2012b].

Task 4:

To validate this initial architecture knowledge repository, we will design experiments based on presenting experienced architects with a design scenario and enabling them to navigate the design space using the wiki.

Research Outcomes

The project outcomes will directly support the U.S. Army's Telemedicine and Advanced Technology Research Center (TATRC), for which we are evaluating alternative software and data architectures for an integrated electronic healthcare record (iEHR) system.²

This project has produced the following artifacts:

- a Web-facing knowledge base that encodes software and data architecture design principles and relates these to specific big data technology realizations in terms of software patterns
- one journal article (published) [Gorton 2014] and one conference paper (in review)
- a half-day day tutorial that was presented at several conferences

¹ For more information, visit <http://www.mongodb.org/> and <http://cassandra.apache.org/>.

² Mentioned with permission

Scalable Data-Intensive Systems

Framing

State of the Practice: Limited building scalable systems

- Netflix, Google, Facebook, et al—huge investment, still more art than science

Horizontal scaling and NoSQL provide the pieces

- Some assembly required
- “Just download Hadoop” doesn’t solve most real problems

Architecture principles have changed— “convergence of concerns”

- Can’t abstract away underlying topology
- Application architecture, data model, and deployment topology are tightly coupled

Need to enable the “average architect” to design these scalable systems

Today’s Warfighter has access to an ever-increasing number of sensors, imagers, internet artifacts, open source and other sophisticated collection devices, to the point that a major challenge has become how to sift through this massive amount of information to find the most critical and actionable items of intelligence. ‘Big Data’ tools, techniques, and technologies seek to provide the means to analyze, exploit and share conclusions drawn from this seemingly overwhelming information load.

Knowledge Capture and Dissemination in Software Engineering



Johannes Gutenberg, circa 1450



in Science (e.g. biology - <http://www.ncbi.nlm.nih.gov>)



Project Method and Approach

Prototype knowledge repository

- Build using Semantic MediaWiki
- Custom semantic model to support two use cases:
 - Top-down architecture design
 - Bottom-up technology selection
- Populate using experience from MHS prototyping experiments and public domain
- Build visualizations to support typical workflows

Validate by observing use of the repository

- Test subjects solving model problems
- Real architects developing real systems

QuABase – Quality At Scale

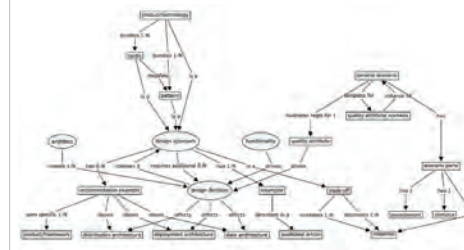


- Semantic-based Information Model
- General model of software architecture knowledge
 - Populated with specific big data architecture knowledge

Dynamic, generated, and queryable content

Knowledge Visualization

QuABase Semantic Model



Graph Algorithms on Future Architectures

Principal Investigator

Eric Blaine Werner
Emerging Technology Center
ebwerner@sei.cmu.edu
(412) 268-1984

Deputy Principal Investigator

Scott McMillan, PhD
Emerging Technology Center

Investigators

Matt Gaston, PhD
Emerging Technology Center

Jonathan Chu
Researcher, Emerging
Technology Center

Ed Morris
Advisor, Software Solutions
Division

Ed Stoner
Advisor, CERT Division

Andrew Lumsdaine, PhD
Collaborator, Indiana
University

There are many U.S. Department of Defense (DoD) and intelligence community (IC) applications that utilize graph algorithms at the core including ISR¹, knowledge representation, and route planning. Advances in graph algorithm implementations will have positive impact on the efficiency and cost of these systems.

In FY14, our research focused on Graphic Processing Units (GPUs) and multi-GPUs with a vision toward eventual support for heterogeneous processors, FPGAs², and other emerging hardware platforms. There are many research questions to be answered in designing and developing a graph library that supports multiple hardware architectures and several challenges.

- Is it possible to support various architectures and abstract the specific implementation complexities for those architectures?
- How do we maximize performance and utilization for applications that use the general-purpose graph library?
- What are the patterns, bindings, and implementation strategies that lead to a usable and widely used graph library?

Primary Hypothesis

It is possible to develop (incrementally) a general-purpose graph library that supports multiple emerging hardware platforms, exploits the performance gains of these architectures, and separates the concerns of expressing algorithms and developing for specific architectures.

Experiment design and analysis

Over the course of the project, various experiments will be designed to test incrementally the hypothesis (and sub-hypotheses); these will include

- measuring supportability of graph data structures and algorithms by hardware type
- measuring and assessing the amount of hardware specific concepts that must be exposed to developers using the library
- measuring performance and utilization via the library compared to “direct” implementations

Research Outcomes

We will draw on existing research implementations of graph algorithms for “newer” architectures and existing libraries (BGL, PBGL, MTGL, thrust, cusp, etc.)³ to find common/preferred implementations approaches for data structures and algorithm.

We will analyze possible approaches and design the appropriate APIs and binding. We will design and develop activities that will be incremental and iterative with built-in validation and experimentation processes.

We will provide accessible and proven libraries that government and Defense Industrial Base organizations can use to efficiently implement applications with graph analysis requirements and effectively exploit emerging computing platforms to do so.

Our goal is to produce and publish at least one peer-reviewed conference paper on design and implementation approaches and results of experiments; the library is intended to be available as an open source project.

¹ Intelligence, Surveillance, and Reconnaissance
² Field Programmable Gate Arrays

³ Boost Graph Library, Parallel Boost Graph Library, MultiThreaded Graph Library

Graph Algorithms on Future Architectures

Motivation

- Fast, efficient graph analysis is important and pervasive
- Heterogeneous hardware is coming here
- We are building a library that helps developers use both

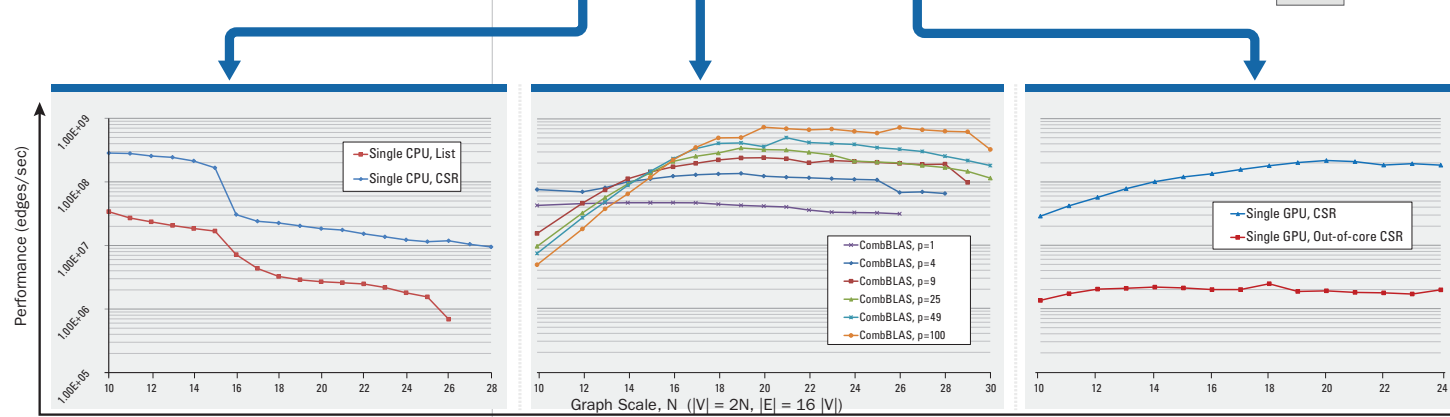
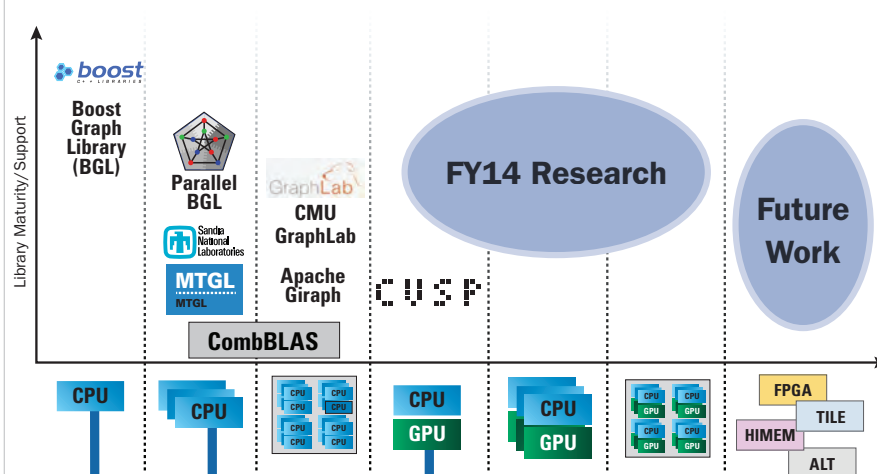
Current Results

- Survey of current graph algorithm libraries
- Baseline implementations on CPU architectures
- CombBLAS library shows promise of scalability
- Initial implementation of GPU algorithms

Current Work

- Designing library for GPU and Multi-GPU graph library (in collaboration with Indiana University)
- Borrowing concepts from BLAS community
- Mapping to unique GPU architecture for performance

Maturity/Support of graph libraries across the different types of computer architectures



Representation of a poster developed to describe this project

Real-Time Mobile Applications in Intermittently Connected Networks

Principal Investigator

Jeffery Hansen, PhD
Software Solutions Division
jhansen@sei.cmu.edu
(412) 268-9565

Internal Investigators

Jeffrey Boleng, PhD
Software Solutions Division

Scott Hissam
Software Solutions Division

Marc Novakouski
Software Solutions Division

Lutz Wrage
Software Solutions Division

First responders and soldiers are often asked to operate in environments that lack access to reliable, ubiquitous, and high-bandwidth network infrastructure. This lack can be due to the destruction of existing infrastructure by natural disasters (e.g., earthquake in Haiti or the tsunami in Japan), operating in rural locations with little or no infrastructure (e.g., Midwestern United States), or operating in military environments (e.g., host nation's infrastructure may be unavailable or denied).

Lack of such access, or sufficient capacity, exacerbates the responders' or soldiers' ability to provide aid or execute their mission to the degree possible, as the information needed to convey progress and status, receive operational changes, and coordinate amongst peers becomes increasingly difficult to acquire.

These environments are sometimes called disconnected, intermittent, and low-bandwidth (DIL) [Scott 2011]. Operations in DIL environments by first responders or soldiers experience a range of connectivity issues—from intermittent loss of connectivity to long-term disconnections and no end-to-end connectivity. One approach that has been proposed for operating in such environments is Delay Tolerant Networking (DTN) in which information is cached and forwarded as connectivity events occur [Fall 2003, 2008].

Dealing with DIL environments is a key challenge for the U.S. Department of Defense (DoD). Situational awareness (SA) applications such as Blue Force Tracking (BFT) or those in support of fire control demand real-time or near real-time connectivity and are typically not disruption tolerant [Parikh 2005]. In a connected state, the application will behave as expected, but once the application enters a zone where it is no longer connected it will stall, fail, or reduce in functionality—if not become useless [CSTB 2010]. These disruptions can result in decisions being made based on old or unsynchronized information, leading to mission failure.

The *U.S. Air Force Cyber Vision 2025* highlighted the need for “ensured operations in congested, competitive, contested, and denied environments” and the “ability to avoid, survive, and



This research project will have two primary types of outputs: architectural guidance and software enhancements to the SEI's Information Superiority at the Edge (ISE) research platform.

recover from disruption” whether the disruption be sudden or sustained, natural or manmade [Maybury 2012]. These disruptions introduce mission risk in any network-centric operation such as ground-based tactical environments.

Research Outcomes

This research project will have two primary types of outputs: architectural guidance and software enhancements to the SEI's Information Superiority at the Edge (ISE) research platform. We anticipate that integrating DTN into the ISE platform and testing it in real DIL environments will illuminate potential enhancements to the DTN architecture or, alternatively, lead to changes in the mobile application reference architecture for which ISE is a reference implementation. We will publish and update to that reference architecture as needed.

We will add the following capabilities to the ISE platform:

- DTN communications layer
- Prediction models for time/location/connectivity information for the user in disconnected mode
- Connectivity assistant for attempting synchronization for the user in disconnected mode
- Bandwidth optimization techniques for synchronization
- Smart Caching for disconnection preparation
- Evaluation framework for testing each capability

These modifications will be available to transition partners of the SEI interested in the Edge-Enabled Tactical Systems suite of applications and may potentially one day be made open source with the rest of the ISE platform.

Real-Time Mobile Applications in Intermittently Connected Networks

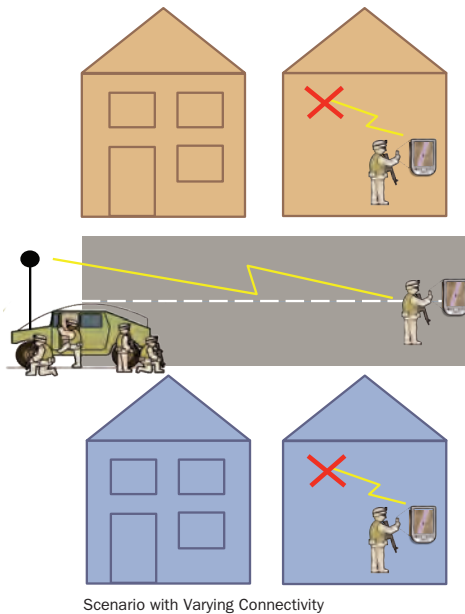
Problem Statement

Problem

- Real-time distributed applications depend on reliable communications
- Tactical environments are often characterized by disconnected, intermittent and low-bandwidth (DIL) communications

To address this problem, we seek to develop methods that

- Enable real-time shared group context in a DIL environment
- Keep information synchronized in real time despite communication outages
- Apply group context to make these more effective

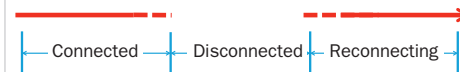


Scenario with Varying Connectivity

Approach

Keep network users productive

We consider three communication states



Connected State

Goal

Maintain shared group context

Make best use of available bandwidth

Techniques

Pre-cache data likely to be relevant later in the mission
Delay transmission of noncritical data

Disconnected State

Applications continue to function

Predict state where possible

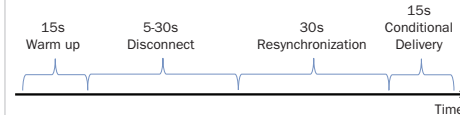
Predict team location based on mission plan
Provide connectivity map to help the user reconnect

Reconnecting State

Re-establish shared group context as quickly and accurately as possible

Prioritize synchronization of critical messages
Eliminate redundant messages

Experiment Design



Independent Variables:

- Disconnect time [10, 20, 30, 40, 50]
- Metadata extensions [disabled, enabled]
- Messages per second [5, 10, 20]

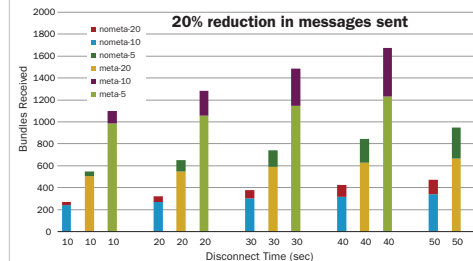
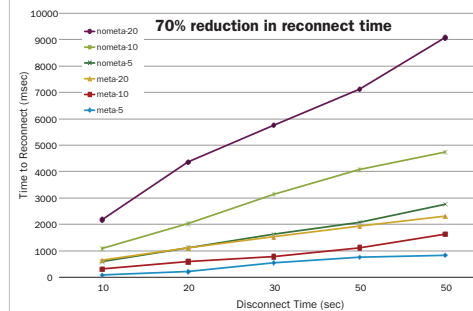
Results

Implementation: ISE+

- DTN (Delay Tolerant Networking) protocol for message delivery
- DTN Metadata Extension Blocks
 - expiration, replacement, redundancy elimination, conditional delivery
- Position prediction for disconnected operation
- Dynamic connectivity map construction
- Pre-caching of mission-relevant data



Metadata Extension Measured Improvements



Probabilistic Analysis of Time-Sensitive Systems

Principal Investigator

Jeffery Hansen, PhD
Software Solutions Division
jhansen@sei.cmu.edu
(412) 268-9565

Internal Investigators

Sagar Chaki, PhD
Software Solutions Division

Dionisio de Niz, PhD
Software Solutions Division

Ian Gorton, PhD
Software Solutions Division

Lutz Wrage
Software Solutions Division

Time-sensitive systems operating in uncertain environments have complex stochastic behaviors that are difficult to analyze and predict in part due to uncertainty. Autonomous systems are not analytically validated—either prototypes are fielded or they are tested until budget is exhausted. Data-intensive systems achieve best-effort latencies validated through exhaustive, time-consuming testing of the final software system on its target execution platform. This problem is fundamental to computer science and has recently become critical due to the emergence of big data and distributed autonomous systems such as Unmanned Aircraft Systems (UASs).

We will explore predictive techniques to verify that these systems satisfy their requirements. Our verification techniques will provide greater confidence in the achievement of mission goals by complex distributed stochastic systems fielded by the U.S. Department of Defense (DoD). For large-scale data analysis applications, our techniques will make it possible to cost-effectively provision systems on “right-sized” computational platforms, providing predictable performance at optimum cost. We will pursue two tasks: (A): for real-time safety-critical systems, we will use statistical model checking and uncertainty quantification to predict their behavior; (B): for data-intensive applications, throughput-oriented techniques will be used.

Task A

We will apply statistical model checking (SMC) to collections of UASs, each running a mixed-criticality Zero-Slack Rate Monotonic (ZSRM) scheduler. This is a unique application of a verification technique (SMC) that models uncertainty to an adaptation mechanism (ZSRM) that adapts to uncertainty while protecting critical tasks. We will compare ZSRM with Rate Monotonic Scheduling (RMS) in this setting. We expect SMC to show that deadlines of critical tasks are better protected with ZSRM compared to RMS leading to improved performance against mission goals.



For large-scale data analysis applications, our techniques will make it possible to cost-effectively provision systems on “right-sized” computational platforms, providing predictable performance at optimum cost. We will use statistical model checking and uncertainty quantification to predict the behavior of real-time safety-critical systems. In addition, we will use throughput-oriented techniques for data-intensive applications.

Task B

We will create calibrated predictive performance models for Hadoop-based applications based on measurements from testbeds. This will allow us to validate the accuracy of the models on realistic data-intensive applications.

Research Outcomes

- Task A: prototypes; simulation results; empirical observations on suitability of statistical model checking for real-time safety-critical task in uncertain environments; publication; demonstration
- Task B: generic performance measurement testbed for Hadoop; methodology for calibrating a Hadoop infrastructure; initial validation results; publication; demonstration

Probabilistic Analysis of Time Sensitive Systems

Problem Statement

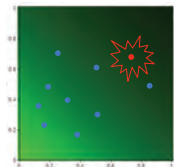
Time-sensitive systems in uncertain environments have complex behaviors. How do we assure correctness of such systems?

- Exact probabilistic verification is infeasible due to model size
- Black box testing does not yield bounded predictions
- Need formal approach for dealing with uncertainty
- Accurate, bounded, probabilistic results
- In reasonable time even for rarely occurring errors

Stochastic Model Checking (SMC)

SMC is a rigorous simulation-based approach for estimating that a property holds in a system.

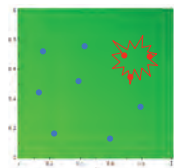
- System properties described in formal language (BLTL, etc.)
- Property is tested on "sample trajectories" (sequence of states)
- Each outcome treated as a Bernoulli trial (i.e., coin flip)



SMC Basics

- Indicator function $I(\vec{x}) = 1$ if property holds for input \vec{x} .
- Relative Error $RE(\hat{p}) = \frac{\sqrt{\text{Var}(\hat{p})}}{E[\hat{p}]}$ is measure of accuracy.
- Draw random samples from input distribution $f(\vec{x})$ until target Relative Error is met.
- Estimated probability that property holds is:

$$\hat{p} = \frac{1}{N} \sum_{i=1}^N I(\vec{x}_i) = \frac{1}{10} = 0.1 \quad RE(\hat{p}) = \frac{0.32}{0.1} = 3.2$$



Importance Sampling

- Modify input distribution to make rare properties more visible.
- Weighting function $W(\vec{x})$ maps solution back to original problem.
- Reduced relative error with same number of samples.

$$\hat{p} = \frac{1}{N} \sum_{i=1}^N I(\vec{x}_i) W(\vec{x}_i) = \frac{0.2 + 0.5 + 0.3}{10} = 0.1$$

$$RE = \frac{0.18}{0.1} = 1.8$$

Semantic Importance Sampling A New Approach to Importance Sampling

Input Specification in C

```
#include "osmosis_client.h"
//@dist a=uniform(min=0,max=5)
//@dist b=normal(mean=3,std=1,min=0,max=5)
void simple()
{
    double a = INPUT_D("a");
    double b = INPUT_D("b");
    double c = a + b;
    double d = (a - b)/2.0;
    ASSERT(sin(c)*cos(d) < 0.995);
}
```

Translate C model to SMT2 for Analysis.

SMT2 Model

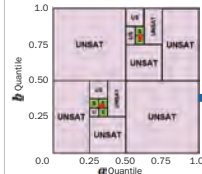
```
(set-logic QF_NRA)
(declare-fun a () Real)
(declare-fun b () Real)
(declare-fun a_1 () Real)
(declare-fun b_1 () Real)
(declare-fun c_1 () Real)
(declare-fun d_1 () Real)
(assert (>= a 0))
(assert (<= a 5))
(assert (>= b 0))
(assert (<= b 5))
(assert (= a_1 a))
(assert (= b_1 b))
(assert (= c_1 (+ a_1 b_1)))
(assert (= d_1 (/ (- a_1 b_1) 2.0)))
(assert (not (< (* (sin c_1) (cos d_1)) 0.995)))
(check-sat)
(exit)
```

Input Cube

ASSERT()

Recursively invoke dReal SMT checker to build abstract model of specification.

Abstract Indicator Function $I^*(\vec{x})$



Weight function $W(\vec{x}_i)$ is probability p^* that \vec{x} is in $I^*(\vec{x})$.

Abstract Probability

$$p^* = \frac{5}{2^8}$$

Number of cubes in $I^*(x)$ (5) and Level of cubes (2^8) are indicated.

Final Probability Estimate

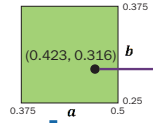
$$\hat{p}_{raw} = 0.024$$

$$RE(\hat{p}_{raw}) = 0.01$$

Use $I^*(x)$ to generate random input vectors:

- Randomly pick SAT cube
- Randomly pick point in cube

Input Generation



Apply inverse CDF on each input variable

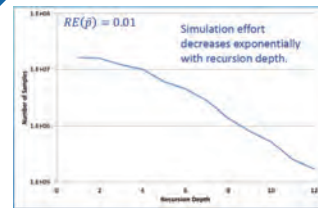
$(a, b) = (2.115, 2.503)$

Apply generated inputs to original C model to calculate bounded failure probability estimate.

Raw Probability Estimate

$$\hat{p} = p^* \hat{p}_{raw} = 0.00047$$

$$RE(\hat{p}) = 0.01$$



Edge-Enabled Tactical Systems (EETS)

Principal Investigator

Grace A. Lewis
Software Solutions Division
glewis@sei.cmu.edu
(412) 268-5851

Focus Area Leads

Group Autonomy for
Mobile Systems: James R.
Edmondson, PhD, Software
Solutions Division

Information Superiority to the
Edge: Jeffrey L. Boleng, PhD,
Software Solutions Division

Edge Analytics: Soumya
Simanta, Software Solutions
Division

Cyber-Foraging in Resource
Constrained Environments
(Tactical Cloudlets): Grace
A. Lewis, Software Solutions
Division

The Edge-Enabled Tactical Systems (EETS) project investigates efficient and easily deployable mobile solutions for teams of soldiers and first responders operating in “edge” environments characterized by dynamic context, limited computing resources, high levels of stress, and poor network connectivity.

EETS develops architectures, advanced prototypes, and other strategies for allocating resources and managing information flow through efficient use of cyber-foraging exploitation of context information; incorporation of sophisticated, user-directed sensors; and rapid analysis of data. Together, these techniques address the challenges of limited and uncertain computational resources, and the need for real-time situational awareness and rapid analysis of data by personnel operating in tactical situations.

Some of the technical challenges addressed by this work include

- identification, characterization, and development of solutions that address critical quality attributes (e.g., survivability, resiliency) that are largely ignored by commercial mobile solutions
- development of situational awareness and decision support solutions that incorporate increasing numbers/types of sensors and data feeds without causing information overload
- development of solutions that can work in environments where computing resources and power are limited and/or intermittent
- development of solutions that require minimal operator guidance to control groups of semi-autonomous devices operating in direct support of warfighters/first responders performing tactical missions
- development of solutions that are adaptable to intelligence, combat, peacekeeping, disaster response, and other unanticipated missions

Research Outcomes

Focus areas and outcomes for FY14 included the following:

- **Group Autonomy for Mobile Systems (GAMS)** develops middleware and algorithms to enable a single human operator to control a heterogeneous swarm of sensors, tailored to

mission contexts. The goal for FY14 was to create prototype portable middleware, distributed algorithms, and tools to support warfighter-directed groups of autonomous sensors and systems. Developed capabilities include area coverage techniques that specialize in prioritized zones and mission-focused swarm formations such as shielding of important areas, people, or moving objects.

- **Information Superiority to the Edge** develops prototypes, architectures, and algorithms that apply advanced information processing and sharing capabilities, filter data to reduce cognitive load, and integrate advanced activity-recognition techniques to automatically determine the user’s situation. Capabilities developed in FY14 include integration with external software for better awareness at the edge, gathering and use of individual sensor data to perform activity recognition, demonstration of group context-awareness that results in reduced cognitive load and improved task completion, and demonstration of multiple link layer DTN routing/forwarding in DIL network environments.
- **Edge Analytics** develops architectures, algorithms, and prototypes to enhance the situational awareness of warfighters and first responders in near real-time (seconds to minutes) by analyzing data streams (e.g., social media streams from Twitter). The goals for FY14 were to identify, develop, and apply inference algorithms to streaming data; to adapt the algorithms and architecture to maximize resource utilization and elasticity by using non-blocking and asynchronous architectures, taking clues from end users to perform on-demand processing to balance resource utilization; and to increase the usability to the system by supporting end user control for visualizations and real-time data exploration.
- **Cyber-Foraging in Resource-Constrained Environments** develops architectures and prototypes for Tactical Cloudlets — forward-deployed, discoverable, virtual-machine-based servers that can be hosted on vehicles or other platforms and provide (1) infrastructure to offload computation, (2) forward data-staging for a mission, (3) data filtering to remove unnecessary data from streams intended for dismounted warfighters, and (4) collection points for data heading for enterprise repositories. Our research in tactical cloudlets

focuses on providing cloud computing capabilities at the edge for computation offload and data staging that consider a wide range of critical quality attributes not considered by the commercial mobile ecosystem, such as survivability, resiliency, and security. The goal for FY14 was to deliver a spectrum of cyber-foraging solutions and prototypes optimized/adapted for use in resource-constrained environments, with an initial focus on survivability. Developed capabilities include optimal cloudlet selection, capability migration between cloudlets, a light-weight management layer, and support for disconnected operations.

Edge-Enabled Tactical Systems

Carnegie Mellon Software Engineering Institute: Advanced Mobile Systems Team

Current Capabilities

Group Autonomy for Mobile Systems (GAMS)

Portable middleware, distributed algorithms and tools to support warfighter-directed groups of autonomous sensors and robotic systems. The focus for FY14 was on area coverage techniques that specialized in prioritized zones and mission-focused swarm formations such as shielding of important areas, people, or moving objects.

Information Superiority to the Edge (ISE)

Mobile application prototype that supports small edge units of soldiers or first responders by (1) sharing individual context information derived from sensors and manual input about events and activities; (2) improving accuracy and timeliness of task completion and reducing cognitive load by providing targeted information, group-coordination capabilities, and task guidance; and (3) capturing individual and group context for leveraging of resource (sensing, battery, processing, etc.) optimization models and activity-recognition algorithms. Recently added capabilities allow fine grained network and data optimization in Disconnected, Intermittent, Low-Bandwidth (DIL) environments by leveraging Delay Tolerant Networking (DTN) protocols and meta-data extensions for store-carry-forward data transmission and policy definition to shape context routing and forwarding in the network.

Edge Analytics

System that provides real-time situational awareness to warfighter and first responders units based on open source and social media data streams by (1) performing timeliness-accuracy tradeoffs to provide faster results in analyzing high velocity data streams; (2) providing macro trend analysis (sentiments, topics, named entities, location) on stream slices; (3) analyzing stream slices to incrementally identify network structure and metrics; and (4) supporting interactive visualizations to allow operators to understand and digest high volumes of fast-moving data.

Cloudlet-Based Cyber-Foraging

Forward-deployed, virtual machine (VM) cloudlets that can be hosted on vehicles or other platforms and provide (1) infrastructure to offload computation, (2) forward data-staging for a mission (3) data filtering to remove unnecessary data from streams intended for dismounted warfighters; and (4) collection points for data heading for enterprise repositories.

FY14 Research Focus

Group Autonomy for Mobile Systems (GAMS): Develop middleware and algorithms to enable a single human operator to control a heterogeneous swarm of sensors, tailored to mission contexts

- Create algorithms for distributed prioritized and pheromone-based area coverage
- Create algorithms for swarm formation flying and target protection/tracking/swarming
- Support new VREP simulation platforms and Drone-RK quadcopter and Platypus boat real-world robotics platforms
- Create middleware for networked periodic applications with extensible platforms and distributed algorithms for C++, Java, Android, ARM, Intel and other architectures

Information Superiority to the Edge (ISE): Develop prototypes, architectures, and algorithms that apply advanced information processing and sharing capabilities; filter data to reduce cognitive load; and integrate advanced activity-recognition techniques to automatically determine the user's situation

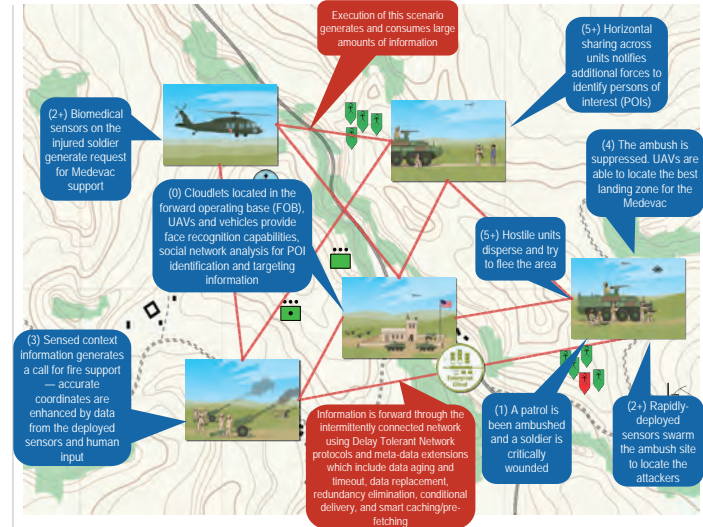
- Integrate with external software for better awareness at the edge
- Gather and use individual sensor data to perform activity recognition
- Demonstrate group context-awareness results in reduced cognitive load and improved task completion
- Demonstrate multiple link layer DTN routing/forwarding in DIL network environments

Edge Analytics: Enhance the situational awareness of edge users in near real-time (seconds to minutes) by analyzing social media streams and sensor streams to provide actionable intelligence, trends, and summaries

- Improve inference algorithms (semantics, multi-stream sensor)
- Develop adaptive algorithms and architecture for use in edge environments
- Increase usability by supporting end-user control

Cloudlet-Based Cyber-Foraging: Demonstrate that tactical cloudlets can increase the survivability of mobile software systems in the field

- Extend tactical cloudlet implementation by adding capabilities targeted at increasing mobile systems survivability, such as optimal cloudlet selection, cloudlet handoff (live migration), ease of management and deployment, and support for disconnected operations
- Validate new capabilities against a set of survivability metrics for mobile systems in edge environments



FY15 Research Focus

Establishing Trusted Identities in Disconnected Tactical Environments

We will develop trusted identity solutions that work within the constraints of DIL environments in which there is no consistent access to third-party online trusted authorities that validate the credentials of the requester or a certificate repository. Developed solutions will be validated and integrated in the tactical cloudlet implementation.

Assigning Credibility Scores to Social Media Streams in Real-Time

Trust in the credibility of information provided by social media channels is a key challenge. We will develop an algorithm that can assign a credibility score quickly (seconds) and provide a human-understandable chain of reasoning in the end user's vocabulary to evaluate the veracity of data.

Fusion of Social and Physical Sensor Data

This effort will improve trust in situational awareness by developing techniques to fuse data from social media with non-textual data and data from physical opportunistic sensors. We will develop algorithms that extract sensor metadata, other contextual data about the particular sensor, and where possible, analyze non-textual and sensor data to infer context to generate and assign new metadata.

Group Autonomy for Mobile Systems

GAMS is transitioning to the FY15 LENS ELASTIC and FY15 LINE DART projects. ELASTIC will focus on middleware and algorithms for distributed autonomous systems that dynamically respond to user needs, resources, and mission contexts. It will also develop complementary algorithms, called accents, which amplify a core algorithm to allow the autonomous agent to work on multiple missions simultaneously. DART focuses on the verification of distributed cyber-physical systems and will be using the GAMS middleware as a target platform.

Cybersecurity Expert Performance and Measurement

Principal Investigator

Jennifer Cowley, PhD
CERT Division
jcowley@cert.org
(412) 268-4461

Internal Investigator

Bronwyn Woods, PhD
CERT Division

Training the cybersecurity operator to some criterion of excellence has been at the forefront for the U.S. Department of Defense (DoD). However, it is unclear what experts are and what valid metrics exist to distinguish experts from other levels of expertise. Beyond these concerns with identifying experts, training programs are interested in understanding what individual attributes (e.g., performance monitoring abilities, prior work experiences, motivation and conscientiousness, risk attitudes, etc.) predict expertise and expert performance.

We argue that individual performance is difficult to evaluate separately from team performance in cybersecurity operations, because of required teamwork skills necessary for team success. We aim to define and identify experts and assess individual attributes that impact the development of individual expertise nested within teams.

The implications of this work are three-fold:

1. This work will help the DoD make informed decisions about selecting a future cohort of cybersecurity operators.
2. Since metric development will include the identification of expert/novice differences in overt behavioral task work (and respective knowledge structures), training programs may be able to teach these differences to novices and possibly expedite the development of expertise.
3. Many automated systems are built from expert judgment, yet it is not clear whether valid experts were used to model their judgments. With metrics that identify experts vs. lower levels of expertise, we hope to improve automated decision aids and other expert systems with ensuring the appropriate individuals are modeled.

The technical challenge is generating valid metrics: metrics that identify experts (e.g., human and team performance) and metrics of attributes that impact expertise development. We recognize that cybersecurity operations is a dynamic domain

where the skillset of cybersecurity operators may evolve with the changing adversarial capability. Thus, individual attributes may be shifting, further complicating the ability to establish a relationship between individual attributes and performance over time.

The operational challenge is not only fielding generated metrics for the assessment of metric reliability and validity but also to convey the importance of our results in a way that improves the selection and training of cybersecurity personnel. This research is best suited for the SEI because of the security training simulation that real cybersecurity operators use in training and evaluation, such as Cyber Guard and Cyber Flag.¹ These cyber operators come from a variety of organizations in the government, military, and commercial sectors.


This research fits into sustainment and operations with potential feedback to systems development. Sustainment includes ensuring that systems are patched and properly configured to defend these systems against cyber-attacks. Cyber operators who have been properly trained can reduce the effectiveness of attacks and improve our ability to sustain and operate the system.

Because this research will produce new knowledge about the factors that affect operator performance, we envision the results of this research affecting system design. For example, understanding how operators use information to assess and perceive risks and how they experience heavy workloads could lead to new insights into how to design better user interfaces.

Research Outcomes

The products of this research are (1) new metrics to assess attitudes, beliefs, and performance and (2) research findings that identify the factors that affect or predict human performance in a training simulation.

¹ For more information on Cyber Guard and Cyber Flag exercises, see "Cyber Guard Exercise Tests People, Partnerships" at <http://www.defense.gov/news/newsarticle.aspx> and "Cyber Flag exercises sharpen DOD cyber operations and defense" at <http://gcn.com/articles/2013/12/09/cyber-flag.aspx>.



Because this research will produce new knowledge about the factors that affect operator performance, we envision the results of this research affecting system design.

Automated Cyber-Readiness Evaluator (ACE)


Principal investigator

Rotem Guttman
CERT Division
rdguttman@cert.org
(412)268-9216

The U.S. Department of Defense (DoD) has a need for assessing the capability and combat readiness of its cyber workforce. However, because cyber is a relatively new mission area for the DoD, it does not yet have a robust, objective assessment platform that it can use to validate the hands-on, technical knowledge and skills of its cyber workforce. Our research aims to address this challenge problem by developing a system that can interpret the hands-on activities a user is performing on a computer screen and translate those actions into quantifiable measures of cyber-based knowledge and skill.

Workforce improvement is a priority for the DoD as outlined in DoD Directive 8570.1¹, "Information Assurance Training, Certification, and Workforce Management," and an integral part of that is the ability for the DoD to assess and verify the technical capability of its cyber workforce. In fact, two of the six workforce management objectives (C1.3.2 and C1.3.5) defined in The DoD Information Assurance Workforce

Improvement manual focus on assessment activities. Specifically, these objectives aim to (1) establish a baseline of skills among personnel performing Information Assurance (IA) functions across the DoD and (2) verify IA workforce knowledge and skills through standard certification testing.²



We are developing an automated cyber-readiness evaluation capability that can interpret the actions a user is performing on a screen within a defined desktop environment and based on those actions, objectively measure an individual's competence within a defined knowledge and skill set.

Despite this need to baseline and verify cyber knowledge and skills, the DoD does not yet have an adequate platform or methodology for assessing the knowledge and skills of its cyber workforce.

To address this challenge problem, we are developing an automated cyber-readiness evaluation capability that can

- interpret the actions a user is performing on a screen within a defined desktop environment
- based on those actions, objectively measure an individual's competence within a defined knowledge and skill set

Research Outcomes

- Computer vision system that can convert activities occurring on a computer screen for a known, defined environment into a human readable format of actions
- Intelligent tutor system that can parse output from the vision system and translate it to objective measures of knowledge and skill for a defined skill set
- Published articles for both the vision system and the intelligent tutoring system

¹ DoD Directive 8570.1, "Information Assurance Training, Certification, and Workforce Management," August 15, 2004.

² DoD Manual 8570.01-M, "Information Assurance Workforce Improvement Program," December 19, 2005 (updated January 24, 2012).

Profiling, Tracking, and Monetizing: Analysis of Internet & Online Social Network Concerns

Principal Investigator

Jason Clark, PhD
CERT Division
jwclark@cert.org
(703) 247-1362

This research explores concerns facing Internet, specifically Online Social Network, users. The attacks we discuss can lead to identity theft, biased and tailored website content delivery, geolocation threats, monetization, and an overall lack of privacy. We introduce a profiling and tracking attack that correlates a user's online persona that is captured from seemingly innocuous website traffic (e.g., operating system, search engine, browser, time spent on website, etc.) with that of the same user's real Facebook profile through analytics captured from a custom Facebook Fan Page. We show how an adversary might identify the personally identifiable information of the user given only their online persona.

The protection of one's identity is paramount especially for users working in the intelligence community. As a result, these organizations are currently employing privacy-preserving technologies as part of their standard network defenses to anonymize their outbound traffic.

Our results show that while network-level anonymity systems are better at protecting end-user privacy than having no privacy preserving technology, they are unable to thwart de-anonymization attacks aimed at applications and private data of end-users. We demonstrate and substantiate our claims using a targeted experiment using actual scenarios of real-world users who are relying on a privacy preserving technology.

Our Approach

To this end, we execute multiple attacks associated with network monitoring, phishing, and Online Social Networks. We also discuss how a user can be monetized through an attack vector such as spam. Spam is a profit-fueled enterprise, and cybercriminals are focusing more of their efforts at growing Online Social Networks. One of the common methods of monetizing Online Social Network spam is to entice users to click on links promising free gift cards and iPads. However, these links actually lead to ad networks that bombard users with surveys in an attempt to collect personal and contact information that is then sold to other marketers.

To date, we lack a solid understanding of this enterprise's full structure. We examined the survey scam process to determine the affiliates that are behind this lucrative scam by performing an analysis of five months of Facebook spam data. We provide the first empirical study and analysis of survey scams and demonstrate how to determine which ad networks are sponsoring the spam.

Next, we focus on why people act in an insecure way when specifically handling their passwords and personal images. We believe this is a major problem as seen in sextortion-related cases. By using a combination of well-known human-computer interaction methods such as surveys and exit interviews, combined with custom software, we show that study participants act differently if they visually see the threat associated with their security behavior. We analyze responses from 30 Craigslist participants via a set of three surveys and an exit interview.

Furthermore, we analyze the results of Cloudsweeper which is designed to scan Google Mail accounts and report any cleartext passwords, their associated monetary value, and provides the option to allow for such passwords to be encrypted and redacted.

Additionally, we introduce for the first time the Google Image Extractor, which is designed to extract selected images from the participants Google Mail account and provide the opportunity for users to delete their images seamlessly. Our contributions will help determine if there is a need for applications such as Cloudsweeper and the Google Image Extractor or if an overhaul of the traditional password management strategy is necessary.

All of this research highlights the importance of education on prevalent attack vectors for compromising client systems and violating user privacy. We show the extent to which information made freely available on the Internet, can negatively impact the organization and users.

Research Outcomes

Upon completion of the experiments, we compiled the results and presented it as security awareness briefings.

Aligning Software Architectures and Acquisition Strategies

FY14 research not presented at this review

Principal Investigator

Lisa Brownsword
Software Solutions Division
llb@sei.cmu.edu
(703) 908-8203

Internal Investigators

Cecilia Albert
Software Solutions Division

David Carney
Software Solutions Division

Patrick Place
Software Solutions Division

While software is increasingly important to the success of government programs, there is little consideration of its impact on the software architecture and on the acquisition strategy when key program decisions are being made early in the program (pre- Milestone B). We have observed specific instances where misalignment between the software architecture and the acquisition strategy has resulted in program delays and cost overruns—and, in some cases, program cancelation. Further, we have observed that alignment between software architecture and acquisition strategy does not occur naturally. Our research focuses on answering the question, *Can we improve the probability of a program's success through a method, to be used by government program offices, that produces mutually constrained and aligned program acquisition strategy and software architecture?*

We view our research in four phases.

1. Work in FY12 was intended to answer the question of whether misalignment of acquisition strategy and software architecture was an issue and, if so, why misalignment occurs. We found that misalignment between acquisition strategy and software architecture was a culprit in several expensive programmatic failures. We also discovered seven patterns of failing behavior, or “antipatterns,” that are major contributors to this misalignment. Finally, we developed a model that links key concepts and their relationships and maps weakness in the relationships to the anti-patterns.
2. Work in FY13 was intended to answer the question of whether *acquisition quality attributes* (as a key entity postulated in phase 1) could be of use. We found that business goals imply a set of acquisition quality attributes that can be used to judge the effectiveness of the acquisition strategy—analogueous to mission goals expressed in program-specific software quality attribute scenarios that can be used to judge the effectiveness of the software architecture. In addition, we found that acquisition quality attribute scenarios could be expressed in program-specific scenarios parallel to those defined for software quality attributes. Finally, we

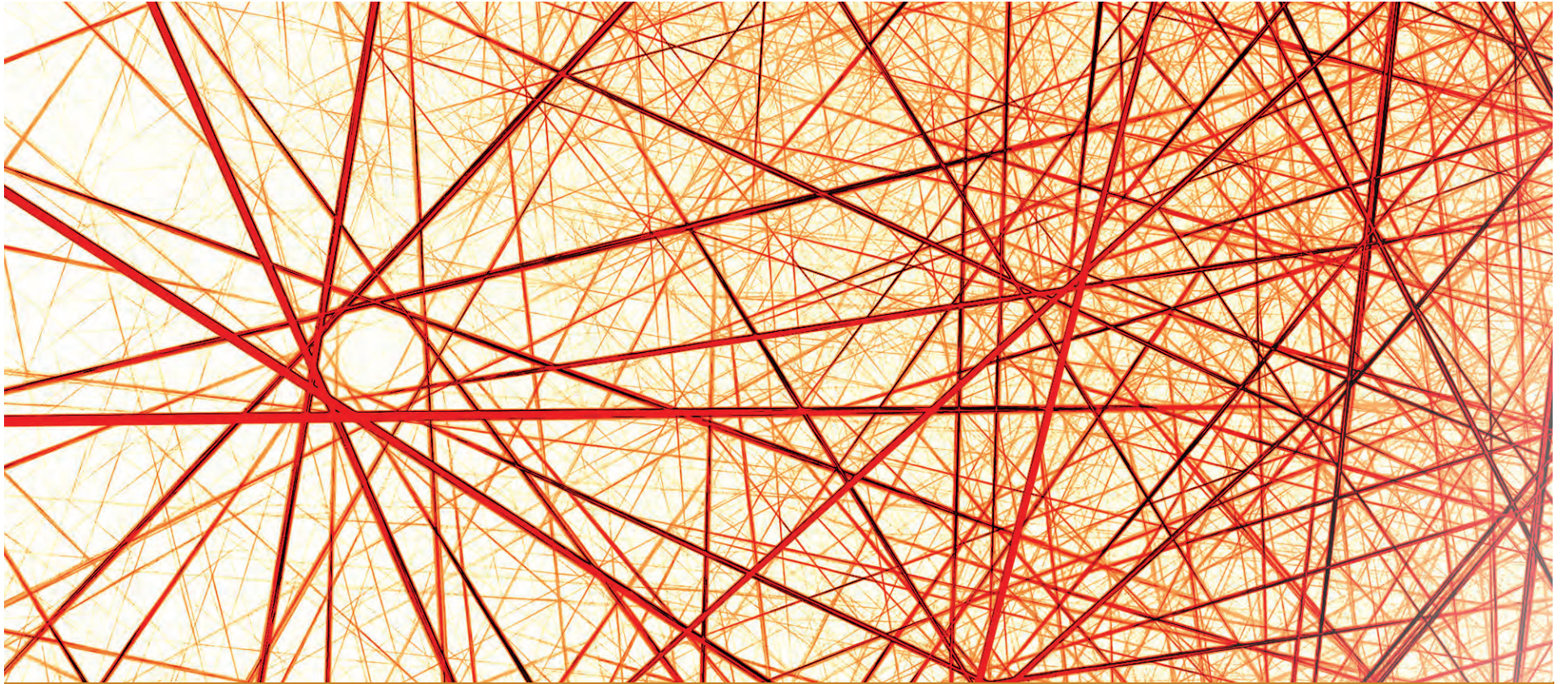
demonstrated that the Quality Attribute Workshop (QAW), suitability extended, is a reasonable vehicle for collecting a useful set of program-specific acquisition quality attribute scenarios.

3. Work for FY14 was to develop a facilitated method that can be used by a Program Office prior to Milestone A or B. We envision that this initial alignment method will:
 - Elicit, document, and prioritize the acquisition-related mission and business goal scenarios associated with the program
 - Elicit the acquisition and software quality attribute scenarios associated with the set of acquisition-related mission and business and goals
 - Identify conflicts among the acquisition strategy and software-architecture-relevant quality attribute scenarios for resolution by the program office
4. Our future work includes developing a technique that objectively measures the quality of the relationship between software architecture and acquisition strategy, and extending the alignment method to include a set of acquisition tactics that can be used to successfully address a set of significant and pervasive acquisition quality attributes.

Research Outcomes

Our FY12 technical report *Isolating Patterns of Failure in Department of Defense Acquisition* described the seven anti-patterns we discovered along with the model of key entities and their relationships. Our FY13 report, *Results in Relating Quality Attributes to Acquisition Strategies*, captured our investigations into the use of acquisition quality attributes. The principal products for FY14 were as follows:

- Initial alignment method built by adapting and extending SEI architecture methods (e.g., QAW, Architecture Tradeoff Analysis Method, Pedigreed and Pedigreed Attribute eLicitation Method)
- Proof-of-concept using the method in a program engagement.



Workshops and Posters

Insider Threat Workshop

Facilitator

Randall Trzeciak is the technical manager of the CERT Division's Enterprise Threat and Vulnerability Management initiative. His team focuses on insider threat research, threat analysis and modeling, assessments, and training.

The Insider Threat Center, a part of the CERT Division at Carnegie Mellon University's Software Engineering Institute (CMU SEI), has been researching insider threats since 2001. The SEI at CMU is a Federally Funded Research and Development Center (FFRDC), sponsored by the U. S. Department of Defense.

Uniquely positioned as part of the SEI, the CERT Insider Threat Center serves as a trusted broker to assist the community in the short term and to conduct ongoing research in the long term. We have compiled a database that contains hundreds of actual insider threat incidents.

Our research focuses on both the technical and behavioral aspects of actual compromises. The goals of our work are to raise general awareness of the risks of insider threats and to help identify the factors that influence an insider's decision to act, the indicators and precursors of malicious acts, and the countermeasures that will improve the survivability and resiliency of the organization.



Our goal for the workshop is that participants leave with a clear understanding of actionable steps that can be taken to better manage the risk of insider threat in their organizations.

In the Insider Threat Workshop, we discuss these topics:

- Overview of Insider Threats
- Insider Threat Technical and Behavioral Patterns
 - IT Sabotage
 - Theft of Intellectual Property
 - Fraud
 - Non-Malicious (Unintentional) Insider Threats
- Best Practices for Mitigating Insider Threats
- Considerations for Building an Insider Threat Program



This workshop consists of discussions that encourage and enable participants to understand the importance of including an assessment of insider threats into their enterprise-wide risk assessment process. Our goal for the workshop is that participants leave with a clear understanding of actionable steps that can be taken to better manage the risk of insider threat in their organizations.

For additional information about CERT's work concerning insider threat, please visit <http://www.cert.org/insider-threat/>. On that website, you will learn about our research, publications, tools, best practices, and services. Our offerings include the Insider Threat Vulnerability Assessment and the Insider Threat Program Evaluation.

In addition, we offer insider threat training and certificate programs to educate professionals on how to

- help organizations identify and manage their insider threat risks
- measure organizational preparedness to defend against insider threat risks
- evaluate an organization's insider threat program
- build and operate an insider threat program

Developing and Maintaining a Skilled Cyber Workforce Workshop

Facilitator

Josh Hammerstein oversees the research agenda for the CERT Division cyber workforce development team and focuses on initiatives that help our customers develop skilled, competent workforce personnel. These initiatives include practices that facilitate a better transfer of knowledge and skill to cyber personnel, methodologies for objectively and accurately evaluating cyber personnel, and technologies that enable high-fidelity simulations for training purposes. In addition to his work at the SEI, Josh is adjunct faculty at Carnegie Mellon University teaching a graduate-level course on cyber forensics and incident response.

A *competent* cybersecurity workforce is critical for protecting the economic and national security interests of the United States. Organizations in government and industry face the ongoing challenge of ensuring that their cybersecurity personnel possess the knowledge, skill, and experience to protect their organizations' data and cyber assets.

Our workshop on achieving competency in cybersecurity is provided in two parts.

Part 1: SEI Workforce Development Strategic Initiatives

The first half of the workshop focuses on initiatives at the SEI to improve the state of the practice of cyber workforce development. These initiatives include the following:

- **Workforce Improvement:** facilitating better transfer of knowledge and skill to cybersecurity personnel
- **Assessment/Evaluation:** developing evaluation capabilities that are accurate, objective, and scalable to a large workforce
- **Cyber-Modeling and Simulation:** researching and developing technologies that enable high-fidelity simulations for training and testing purposes

Part 2: Panel Discussion

The second half of the workshop is designed as a panel discussion with individuals from government, industry, and academia. Their perspectives on the challenges they face in developing and maintaining skilled cyber workforce personnel enrich everyone's understanding of key issues. In addition, their thoughts on strategic approaches for addressing those challenges can open the way to greater insight about innovative solutions.



Agile in Acquisition Workshop

Facilitators

Mary Ann Lapham works to improve the acquisition of software-reliant systems through research and applying technologies. For DoD programs this means working with the Program Office to assist and advise on software issues at the system and/or segment level.

Suzanne Miller is developing guidance for DoD adoption of agile methods. Suzanne also performs research and development to build work products related to systems of systems governance and acquisition

This workshop addresses common myths and misconceptions while providing a basic overview of what Agile is and is not. The following topics are addressed.

Why do we need Agile or software methods anyway?

- Can software development be managed the same as hardware? Why? Why not?
- Discussion of hardware “is a part of” versus software “is used by” constructs

Key Components of Agile Development

- We discuss the Agile Manifesto and its associated 12 principles. Myths and misconceptions of these principles are identified, discussed, and debunked.

Traditional and Agile Acquisition Lifecycles

- A short discussion and comparison of traditional methods versus agile methods will describe what is similar and what is different
- Some of the Myths discussed will include
 - Agile is a fad
 - Agile is just spiral
 - Agile is just incremental
 - Agile is “cowboy programming”
 - Agile is only for small projects



Common Agile Methods

- A quick overview of the most popular Agile methods or those methods termed “Agile”

Scrum—the Most-Adopted Agile Method

- Discussion of the sweet spot for Scrum, key elements of Scrum, and the Scrum framework

Challenges to Agile Adoption in DoD

- “Agile can be adopted without changing your behavior.” We discuss five cultural differences between Agile and traditional methods in highly regulated environments.

Suggestions for Successful Use of Agile Methods in DoD Acquisition

- Attributes of Agile Success
- Enabling Agile software development success
- A short look at agile in the certification and accreditation process

Incremental Lifecycle Assurance Through Architecture-Centric Virtual System Integration Workshop

Facilitator

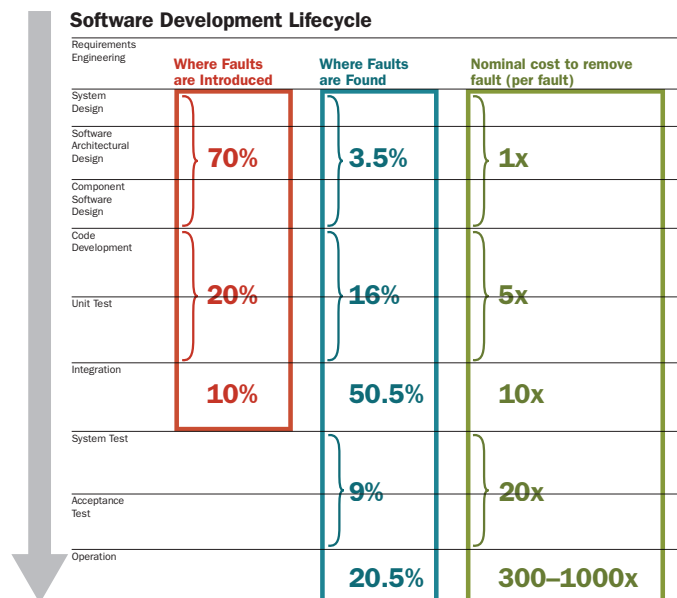
Peter Feiler, PhD, is a Principal Researcher of the Architecture Practice (AP) initiative at the Software Engineering Institute (SEI). His current research interests to improve the quality of safety-critical software-reliant systems and reduce rework and qualification costs through architecture-centric virtual system integration and incremental lifecycle assurance. Peter has been the technical lead and main author of the SAE Architecture Analysis & Design Language (AADL) standard.

Mission and safety critical software-reliant systems—also known as Cyber-Physical Systems—face the increasing challenge of an exponential increase in verification-related software rework cost. Industry studies show that 70 percent of defects are introduced in requirements and architecture design, while 80 percent are discovered post-unit test.¹

The international industry standard Architecture Analysis & Design Language (AADL) was targeted to address these issues through virtual system integration. In virtual system integration, it is possible to discover—analytically, early in the lifecycle—system-level issues concerning operational system properties.

In this workshop, we discuss a four-part improvement strategy for incremental lifecycle assurance of software-reliant systems

1. architecture-led requirements specification to improve the quality of requirements
2. architecture refinement and incremental virtual system integration to discover issues early
3. compositional verification through static analysis as assurance evidence
4. incremental verification and testing throughout the lifecycle to continuously improve confidence



¹ Studies by Boehm, NIST, Dabney, and Galin quantified phase-based percentages of defect introduction and detection and provide rework cost factors [Boehm 1981, RTI 2002, Dabney 2003, Galin 2004]

Cyber Intelligence Research Consortium

Technical Lead

Jay McAllister
jjmcallister@sei.cmu.edu
412-268-9193

In June 2014, the SEI Emerging Technology Center (ETC) established the Cyber Intelligence Research Consortium—a member-funded initiative that researches and develops technical solutions and analytical practices to help people make better judgments and quicker decisions with cyber intelligence. We define cyber intelligence as the acquisition and analysis of information to identify, track, and predict cyber capabilities, intentions, and activities to offer courses of action that enhance decision making.

One of the first results of our work—a conceptual framework—puts this definition into practice. To develop it, we combined aspects of conceptual models in traditional intelligence analysis, risk management, and cybersecurity with our three years of experience studying how organizations perform cyber intelligence. The resulting framework balances the rigor, agility, and creativity needed to conduct comprehensive analysis in the complex and ever-changing cyber domain. This approach engages the art and science of analytical tradecraft to analyze a cyber issue in context, from the way it functions, to its strategic impact on a target, and everything in between.



We developed our cyber intelligence conceptual framework by combining aspects of conceptual models in traditional intelligence analysis, risk management, and cybersecurity with our three years of experience studying how organizations perform cyber intelligence. The resulting framework balances the rigor, agility, and creativity needed to conduct comprehensive analysis in the complex and ever-changing cyber domain.

Our work on the conceptual framework is informed by the experiences and perspectives of our members. Since June 2014, six industry, government, and academic organizations have joined the Consortium. Their practitioners and decision makers are a part of this endeavor because they want access to cost effective resources for cyber intelligence workforce development and technology scouting, awareness of analytical practices across the domain, and insight into the skills and capabilities of the SEI.

As our membership evolves, so too does our approach. We engage with members in a variety of ways. Our in-person events include tradecraft labs that showcase relevant technologies and a crisis simulation that enables members to apply analytical techniques and technologies to a simulated cyber attack. Virtually, we interact with members through webinars, biweekly emails, and newsletters. These interactions allow for continuous dialogue on our technical and analytical research and development efforts—work based on trends we identify by studying the methodologies, processes, tools, and training that our members employ in their daily operations. Over the next seven months, the Consortium will host two tradecraft labs, four webinars, and one crisis simulation, as well as provide members with multiple emails and newsletters focused on improving their cyber intelligence analysis capabilities.

Cyber Intelligence Research Consortium

Advancing the art and science of cyber intelligence

The Consortium is a member-funded initiative that researches and develops technical solutions and analytical practices to help people make better judgments and quicker decisions with cyber intelligence.

Consortium Offerings to Members

-  Steering Committee: Guide Consortium activities and plan for future success
-  Cyber Threat Baseline: Anonymized research of members' cyber threat environments to identify common challenges and associated best practices
-  Tradecraft Labs: Workshops to advance cyber intelligence capabilities and showcase relevant technologies
-  Implementation Frameworks: How-to guides for navigating key analytical practices and technologies
-  Crisis Simulation: Capture-the-flag exercise to apply analytical techniques and technologies to a simulated cyber attack
-  Intelligence Insights: Biweekly emails and bimonthly newsletters on topics relevant to the practice of cyber intelligence

Environmental Context

Provides scope for the analytical effort

- Highlights the importance of context - technical and nontechnical, internal and external to an organization

Reporting and Feedback

Offers courses of action to enhance decision making

- Represents the communication of and subsequent responses to cyber intelligence
- Takes into account audience background and technical expertise

Macroanalysis


Assesses the strategic implications of a cyber issue

- Adds perspective, context, and depth
- Enables proactive and predictive intelligence
- Provides appropriate insight for technical and nontechnical audiences

Analytical Acumen

Facilitates timely, actionable, and accurate intelligence on a cyber issue

- Conceptualizes an analyst's interactions with the other components
- Represents analytical tradecraft—the 'art' and 'science' of doing cyber intelligence analysis



Consortium Conceptual Framework

Cyber Intelligence: the acquisition and analysis of information to identify, track, and predict cyber capabilities, intentions, and activities to offer courses of action that enhance decision making.

This framework puts this definition into practice. It emphasizes the rigor, agility, and creativity needed to analyze threats in the complex and ever-changing cyber domain.

Data Gathering

Acquires and aligns data for analysis

- Asks the right questions to get the right data, and the right amount
- Relies on domain expertise (self and others) and technology

Microanalysis

Assesses the functional implications of the cyber issue

- Examines the issue's nature, ability, and quality
- Enables reactive intelligence

Representation of a poster developed to describe this project

The SEI Empirical Research Office: Data-Driven Decision Making for Software-Reliant Systems

In Washington, DC

Forrest Shull, PhD
Assistant Director for
Empirical Research
Software Solutions Division
(703) 247-1372

In Pittsburgh, PA

Dave Zubrow, PhD
Associate Director for
Empirical Research
Software Solutions Division
(412) 268-5243

Anita Carleton
Deputy Director
Software Solutions Division
(412) 268-7718

Overview

Combining a solid understanding of the compelling challenges facing DoD programs with expertise in empirical research methods and data analysis, the Empirical Research Office (ERO) will apply and advance research in data science, data mining, and related fields to provide credible and practical advice to DoD decision makers, along with measures and heuristics that Program Managers can use to assess program health and evaluate emerging technologies.

Our motivation

Software is essential to the DoD and its programs, and it forms the very backbone of our critical national infrastructure. Yet software's contributions to mission failures and the well-publicized difficulty contractors have in meeting cost and schedule commitments show that we need new ways to acquire and build software. Many voices, not least that of the Under Secretary of Defense for Acquisition, Technology, and Logistics, have called for a data-driven approach based on the measured performance of development processes, technologies, and acquisition policies.

The SEI, as the nation's only federally funded research and development center (FFRDC) focused on software, has created an Empirical Research Office to serve as the DoD's "go to" place for empirically grounded information regarding software engineering, acquisition, and sustainment.

Activities

The ERO engages in four broad classes of activities:

1. **Analysis**—providing decision makers with credible and practical advice for policy, as well as measures and heuristics for assessing program health and evaluating emerging technologies. All work will proceed from a solid understanding of the compelling challenges facing DoD programs and expertise in empirical research methods and data analysis. We will use a fact-based approach to help determine which technologies are working and which are not and how improvements can be made in the development, acquisition, and sustainment of software-intensive systems.



The ERO aims to improve the capability delivered for every dollar of DoD investment in software systems by grounding policy and program decision-making on high quality data and analysis.

2. **Data collection**—maintaining data that is meaningful and descriptive. To get the data needed to make analytic work relevant and actionable, we will engage directly with DoD programs and other appropriate sources and consult with programs on contractor data requirements.

3. **Tool and infrastructure development**—creating appropriate infrastructure and tool support for managing and organizing the data and supporting low cost analyses. This includes making appropriate subsets of the data available for collaborative research projects with other FFRDCs and the larger research community.

4. **Community engagement**—convening researchers and subject matter experts to bring the best analytical methods and expertise to bear on DoD problems and elevate the quality of research and overall capability. We will also leverage the broad research community to identify emerging trends and potentially disruptive new technologies.

For More Information

ssd-empirical-research@sei.cmu.edu

The SEI Empirical Research Office

Data-driven decision making for software-reliant systems

Our Research

The ERO will apply and advance research in data science, data mining, and related fields to:

- Baseline the current state of the practice
- Conduct technology evaluations
- Provide measures for assessing program health and risk
- Quantify costs and benefits of emerging technologies
- Formulate guidance and policy
- Create tools to support analyses

Example analyses include:

1. What types of software development are most costly or uncertain?
2. What is the difference between best-in-class and worst-in-class performance?
3. Which software technologies are proving problematic for software sustainment?
4. How can return on investment of new software development paradigms, like model-based engineering, be calculated?
5. What measures of a program's technical debt can be extracted automatically from a software code base?



Engagement Opportunities

The ERO will obtain and curate a comprehensive collection of data, findings, and benchmarks related to DoD software engineering, sustainment, and acquisition concerns.

This allows us to engage with stakeholders to:

- Help design measurement programs and data collection approaches for your environment and goals.
- Access existing data and bring it to bear on your problems.
- Improve data quality.
- Provide benchmarks and baselines.
- Host new datasets for the good of the community.
- Conduct data-driven analyses.
- Produce context-specific factbooks and state-of-the-practice reports.

Common process for obtaining data and ensuring quality.

Common tools and infrastructure for trusted analysis.

Segmented data and repositories with different levels of openness as appropriate.

References

[Albarghouthi 2013]

Albarghouthi, Aws; Gurfinkel Arie; Li, Yi; Chaki, Sagar; & Chechik, Marsha. "UFO: Verification with Interpolants and Abstract Interpretation" (Competition Contribution), 637-640. *19th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS)*. Rome, Italy, 2013.

[Ali-Babar 2007]

Ali-Babar, M. & Gorton, I. "A tool for managing software architecture knowledge," 63-69. *Proceedings of the 2nd Workshop on Sharing and Reusing architectural Knowledge – Architecture, Rationale, and Design Intent (SHARK/ADI)*. Minneapolis, MN (USA), May 2007. IEEE, 2007.

[Astrom 2011]

Astrom, K. J. & Wittenmark, B. *Computer-Controlled Systems: Theory and Design*. Dover Books on Electrical Engineering Series. DOVER PUBN Incorporated, 2011.

[BAH 2010]

Booz Allen Hamilton. *Systems2020–Final Report*. Office of Scientific Research, August 16, 2010. <http://www.acq.osd.mil/se/docs/BAH-Systems-2020-Report-Final.pdf>.

[Bellomo 2013]

Bellomo, Stephany; Nord, Robert L.; & Ozkaya, Ipek. "A study of enabling factors for rapid fielding: combined practices to balance speed and stability," 982-991. *International Conference on Software Engineering (ICSE 2013)*. San Francisco, CA (USA), May 2013. ICSE, 2013.

[Bellomo 2014]

Bellomo, Stephany; Ernst, Neil; Nord, Robert L.; & Ozkaya, Ipek. "Evolutionary Improvements of Cross-Cutting Concerns: Performance in Practice." *30th International Conference on Software Maintenance and Evolution (ICSME 2014)*. Victoria, British Columbia, Canada. Sept. 28-Oct. 3, 2014.

[Boehm 1981]

Boehm, B. W. *Software Engineering Economics*. Prentice Hall, 1981.

[Buschmann 2007]

Buschmann, F.; Henney, K.; & Schmidt, D. C. *Pattern-Oriented Software Architecture. Vol. 4: A Pattern Language for Distributed Computing*. Wiley, May 2007.

[Buttazzo 2011]

Buttazzo, Giorgio C. *Hard Real-Time Computing Systems*. Springer, 2011. ISBN: 978-1-4614-0676-1.

[CE14a]

Chaki, Sagar & Edmondson, James R. "Model-Driven Verifying Compilation of Synchronous Distributed Applications," 201-217. *ACM/IEEE 17th International Conference on Model Driven Engineering Languages and Systems (MoDELS 2014)*. Valencia, Spain, Sept. 28-Oct. 3, 2014.

[CE14b]

Edmondson, James R. & Chaki, Sagar. "Toward parameterized verification of synchronous distributed applications." 109-112. *International SPIN Symposium on Model Checking of Software (SPIN 2014)*. San Jose, CA (USA), July 2014.

[CGKL14]

Sagar, Chaki; Giampapa, Joseph Andrew; Kyle, David; & Lehoczky, John P. "Optimizing Robotic Team Performance with Probabilistic Model Checking," 134-145. *International Conference on Simulation, Modeling, and Programming for Autonomous Robots (SIMPAN 2014)* in *LNAI Lecture Notes in Artificial Intelligence*. Bergamo, Italy, October 2014. Springer, 2014.

[CGS14]

Chaki, Sagar; Gurfinkel, Arie; & Sinha, Nishant. "Efficient Verification of Periodic Programs using Sequential Consistency and Snapshots." *14th Formal Methods in Computer-Aided Design (FMCAD 2014)*. Lausanne, Switzerland, October 2014.

[Clarke 2004]

Clarke, Edmund M.; Kroening Daniel; & Lerda, Flavio. "A Tool for Checking ANSI-C Programs," 168-176. *10th International Conference on Tools and Algorithms for the Construction and Analysis of Systems*. Barcelona, Spain, Mar. 27-Apr. 2, 2014. ETAPS, 2004.

[CTSB 2010]

Computer Science and Telecommunications Board, Committee for Advancing Software-Intensive Systems Producibility, National Research Council. *Critical Code: Software Producibility for Defense*. National Academies Press, 2010.

[Dabney 2003]

Dabney, J. B. *Return on Investment of Independent Verification and Validation Study Preliminary Phase 2B Report*. NASA, 2003.

[Dagle 2012]

Dagle, J. E., "Cyber-physical system security of smart grids," *Innovative Smart Grid Technologies (ISGT) 1, 2* (Jan. 2012): 16-20.

[de Niz 2011]

de Niz, Dionisio & Wrage, Lutz. "A Criticality Decomposition Architecture to Integrate Encrypted Sensor Data in the Smart Grid." *2nd Workshop on Analytical Virtual Integration of Cyber-Physical Systems (AVICPS 2011)*. Vienna, Austria, November 2011.

[DSB 2012]

Defense Science Board, Department of Defense. *The Role of Autonomy in DoD Systems*. Office of the Undersecretary of Defense for Acquisition, Technology, and Logistics, July 2012. <http://www.fas.org/irp/agency/dod/dsb/autonomy.pdf>.

[Enck 2010]

Enck, W.; Gilbert, P.; Chun, B.-G.; Cox, L. P.; Jung J.; McDaniel, P.; & Sheth, A. "TaintDroid: An Information-Flow Tracking System for Realtime Privacy Monitoring on Smartphones." *Communications of the ACM*, 57 3: 99-106.

[Fall 2003]

Fall, Kevin. "A delay-tolerant network architecture for challenged internets," 27-34. *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*. Karlsruhe, Germany, August 2003. ACM, 2003.

[Fall 2008]

Fall, Kevin & Farrell, Stephen. "DTN: An Architectural Retrospective." *IEEE Journal on Selected Areas in Communications*, 26 (5), June 2008.

[Fedyukovich 2013]

Fedyukovich, Grigory; Sery Ondrej; & Sharygina, Natasha. "eVolCheck: Incremental Upgrade Checker for C," 292-307. *19th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS)*. Rome, Italy, Mar. 2013. ETAPS, 2013.

[Fedyukovich 2014]

Fedyukovich, Grigory; Gurfinkel, Arie; & Sharygina, Natasha. "Incremental Verification of Compiler Optimizations," 300-306. *Sixth NASA Formal Methods Symposium (NFM 2014)*. Apr. 29-May1, 2014.

[Feiler 2010]

Feiler, P. "Challenges in Validating Safety-Critical Embedded Systems." *SAE Int. J. Aeronautics* 3, 1(2010): 109-116. doi:10.4271/2009-01-3284

[Galín 2004]

Galín, D. *Software Quality Assurance: From Theory to Implementation*. Pearson/Addison-Wesley, 2004.

[Godlin 2013]

Godlin, Benny & Strichman, Ofer. "Regression verification: proving the equivalence of similar programs." *Softw. Test., Verif. Reliab.* 23, 3(2013): 241-258.

[Gorton 2012a]

Gorton, I. & Gracio, D., eds. *Data-Intensive Computing: Architectures, Algorithms and Applications*. Cambridge University Press, 2012 (ISBN-13: 978-0521191951).

[Gorton 2012b]

Gorton, I; Sivaramakrishnan, C.; Black, G. D.; White, S. K.; Purohit, S; Lansing, C. S.; Madison, M. C.; Schuchardt, K. L.; & Liu. Y. "Velo: A Knowledge Management Framework for Modeling and Simulation." *Computing in Science & Engineering* 14, 2 (2012):12-23.

[Gorton 2014]

Gorton, Ian & Klein, John. "Distribution, Data, Deployment: Software Architecture Convergence in Big Data Systems." *IEEE Software*. <http://doi.ieeecomputersociety.org/10.1109/MS.2014.51>

[Gurfinkel 2013]

Gurfinkel, Arie. "Trust in Formal Methods Toolchains." *VeriSure: Verification and Assurance*. St. Petersburg, Russia, July, 2013.

[Gurfinkel 2014a]

Arie Gurfinkel, Anton Belov, and João Marques-Silva. "Synthesizing Safe Bit-Precise Invariants," 93-108. *20th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS)*. Grenoble, France, 2014.

[Gurfinkel 2014b]

Arie Gurfinkel and Anton Belov. "FrankenBit: Bit-Precise Verification with Many Bits - (Competition Contribution), 408-411. *20th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS)*. Grenoble, France, 2014.

[Hutchins 2014]

Hutchins, DeLesley; Ballman, Aaron; & Sutherland, Dean. "C/C++ Thread Safety Analysis." *IEEE 14th International Working Conference on Source Code Analysis and Manipulation*. IEEE, 2014. DOI 10.1109/SCAM.2014.34.

[Jhala 2009]

Jhala, Ranjit & Majumdar, Rupak. "Software model checking." *ACM Comput. Surv.* 41, 4 (2009). <http://doi.acm.org/10.1145/1592434.1592438>

[Keaton 2014]

Keaton, David & Seacord, Robert C. *Performance of Compiler-Assisted Memory Safety Checking*. Software Engineering Institute: Carnegie Mellon University (CMU/SEI-2014-TN-014), August 2014.

[Kent 2013]

Kent, Alexander & Liebrock, Lorie. "Differentiating User Authentication Graphs." *IEEE Symposium on Security and Privacy Workshop on Research for Insider Threat (WRIT)*, May 2013.

[KGC14]

Komuravelli, Anvesh; Gurfinkel, Arie; & Chaki, Sagar. "SMT-Based Model Checking for Recursive Programs," 17-34. *26th International Conference on Computer Verification (CAV 2014)* in *Lecture Notes in Computer Science, Vol. 8559*. Vienna, Austria, July 2014. Springer, 2014.

[KIM 2014]

Hyoseung, Kim; de Niz, Dionisio; Andersson, Bjorn; Klein, Mark; Mutlu, Onur; & Rajkumar, Ragunathan (Raj). "Bounding Memory Interference Delay in COTS-Based Multicore Systems." *20th IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS 2014)*. Berlin, Germany, April 2014. IEEE, 2014.

[Klein 1994]

Klein, Mark; Ralya, Thomas; Pollak, Bill; Obenza, Ray; Harbour, Michael González. *A Practitioner's Handbook for Real-Time Analysis: Guide to Rate Monotonic Analysis for Real-Time Systems*. Kluwer Academic Publishers. 1994. ISBN-13: 978-0792393610.

[Klieber 2014]

Klieber, William; Flynn, Lori; Bhosale, Amar; Jia, Limin; & Bauer, Lujo. "Android taint flow analysis for app sets," 1-6. In *Proceedings of the 3rd ACM SIGPLAN International Workshop on the State of the Art in Java Program Analysis (SOAP '14)*. Edinburgh, UK, June 2013. ACM, 2014. DOI=10.1145/2614628.2614633 <http://doi.acm.org/10.1145/2614628.2614633>

[Kruchten 2004]

Kruchten, P. "An ontology of architectural design decisions in software intensive systems," 54-61. *Proceedings of the 2nd Groningen Workshop on Software Variability Management (SVMG 2004)*. December 2004.

[Lal 2009]

Lal, Akash & Reps, Thomas W. "Reducing concurrent analysis under a context bound to sequential analysis." *Formal Methods in System Design* 35, 1(2009): 73-97. <http://dx.doi.org/10.1007/s10703-009-0078-9>

[Lim 2012]

Lim, S. L. & Finkelstein, A. "StakeRare: Using Social Networks and Collaborative Filtering for Large-Scale Requirements Elicitation." *IEEE Transactions on Software Engineering* 38, 3(2012): 707-735.

[Mar 2008]

Mar, R. A. & Oatley, K. "{The function of fiction is the abstraction and simulation of social experience." *Perspectives on psychological science* 3, 3(2008): 173-192.

[Maybury 2012]

Maybury, M. "Air Force Cyber Vision 2025," Presentation, Air Force Association's Monthly Breakfast, Arlington, Va., 17 July 2012. <http://www.afa.org/events/Breakfasts/MayburyPPT.pdf>

[Moreno 2012]

Moreno, Gabriel & de Niz, Dionisio. "An Optimal Real-Time Voltage and Frequency Scaling for Uniform Multiprocessors." *IEEE International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA 2012)*. Seoul, South Korea, August 2012. IEEE, 2012.

[Nam 2011]

Nam, Min-Young; de Niz, Dionisio; Wrage Lutz; & Sha, Lui. "Resource Allocation Contracts for Open Analytic Runtime Models." *Ninth ACM International Conference on Embedded Software (EMSOFT 2011)*. Taipei, Taiwan, October 2011. <http://toc.proceedings.com/13052webtoc.pdf>

[NIST 2002]

National Institute for Standards and Technology (NIST). *The Economic Impacts of Inadequate Infrastructure for Software Testing* (NIST Planning report O2-3, 2002). NIST, Washington, DC. <http://www.nist.gov/director/planning/upload/reportO2-3.pdf>

[Nord 2013]

Nord, Robert L.; Ozkaya, Ipek; Sangwan, Raghvinder S.; Delange, Julien; Gonzalez, Marco A.; & Kruchten, Philippe. "Variations on Using Propagation Cost to Measure Architecture Modifiability Properties," 400-403. *29th IEEE Conference on Software Maintenance (ICSM 2013)*. Eindhoven, The Netherlands. September 2013, IEEE, 2013.

[Nord 2014]

Nord, Robert L.; Ozkaya, Ipek; Sangwan, Raghvinder S.; & Koontz, Ronald J. "Architectural dependency analysis to understand rework costs for safety-critical systems." *ICSE Companion 2014*: 185-194.

[NSF]

National Science Foundation. *Cyber Physical Systems*. http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503286

[Parikh 2005]

Parikh, Salil, & Durst, Robert C. "Disruption tolerant networking for marine corps CONDOR." *Military Communications Conference, 2005. MILCOM 2005*. IEEE. IEEE, 2005.

[Quinlan 2000]

Quinlan, D. "ROSE: A Preprocessor Generation Tool for Leveraging the Semantics of Parallel Object-Oriented Frameworks to Drive Optimizations via Source Code Transformations," 383-397. *Proc. Eighth Int'l Workshop on Compilers for Parallel Computers (CPC '00)*. Aussois, France, Jan. 2000, École Normale Supérieure, 2000.

[RTI 2002]

RTI International. *The Economic Impacts of Inadequate Infrastructure for Software Testing* (Planning Report O2-3). NIST, 2002.

[Scott 2011]

Scott, K.; Refaei, T.; Trivedi, N.; Trinh, J.; & Macker, J.P. "Robust communications for disconnected, intermittent, low-bandwidth (DIL) environments," 1009-1014. *2011 Military Communications Conference*. Baltimore, MD (USA), November 2011. IEEE, 2011.

[Sery 2012]

Sery, Ondrej; Fedyukovich Grigory; & Sharygina, Natasha. "Incremental Upgrade Checking by Means of Interpolation-based Function Summaries," 114-121. *International Conference on Formal Methods in Computer-Aided Design (FMCAD)*. Cambridge, UK, Oct. 2012.

[Shen 2012]

Shen, Chao; Cai, Zhongmin; Maxion, Roy; Xiang, Guang; & Xiaohong, Guan. "Comparing Classification Algorithms for Mouse Dynamics based User Identification." 61-66, *IEEE 5th International Conference on Biometrics: Theory, Applications and Systems (BTAS 2012)*, Washington, DC, September 2012, IEEE Press, 2012.

[Sommer 2010]

Sommer, R. & Paxson, V. "Outside the closed world: On using machine learning for network intrusion detection," 305-316. *2010 IEEE Symposium on Security and Privacy (SP)*. Oakland, CA (USA), May 2010. IEEE, 2010.

[Song 2013]

Song, Yingbo; Ben Salem; Malek; Hershkop, Shlomo; & Stolfo, Salvatore. "System Level User Behavior Biometrics Using Fisher Features and Gaussian Mixture Models." *IEEE Symposium on Security and Privacy Workshop on Research for Insider Threat (WRIT)*, May 2013.

[Tang 2008]

Tang, Qinghui; Gupta, Sandeep K. S.; & Varsamopoulos, Georgios. "Energy-Efficient Thermal-Aware Task Scheduling for Homogeneous High-Performance Computing Data Centers: A Cyber-Physical Approach," *IEEE Transactions on Parallel and Distributed Systems* 19, 11(Nov. 2008): 1458-1472.

[Wallnau 2014]

Wallnau, Kurt; Lindauer, Brian; Theis, Michael; Durst, Robert; Champion, Terrance; Renouf, Eric; & Petersen, Christian. "Simulating Malicious Insiders in Real Host-Monitored User Data." *7th Workshop on Cyber Security Experimentation and Test (CSET 14)*, San Diego, CA (USA), Aug. 2014. USENIX Association, 2014.

[Whiting 2008]

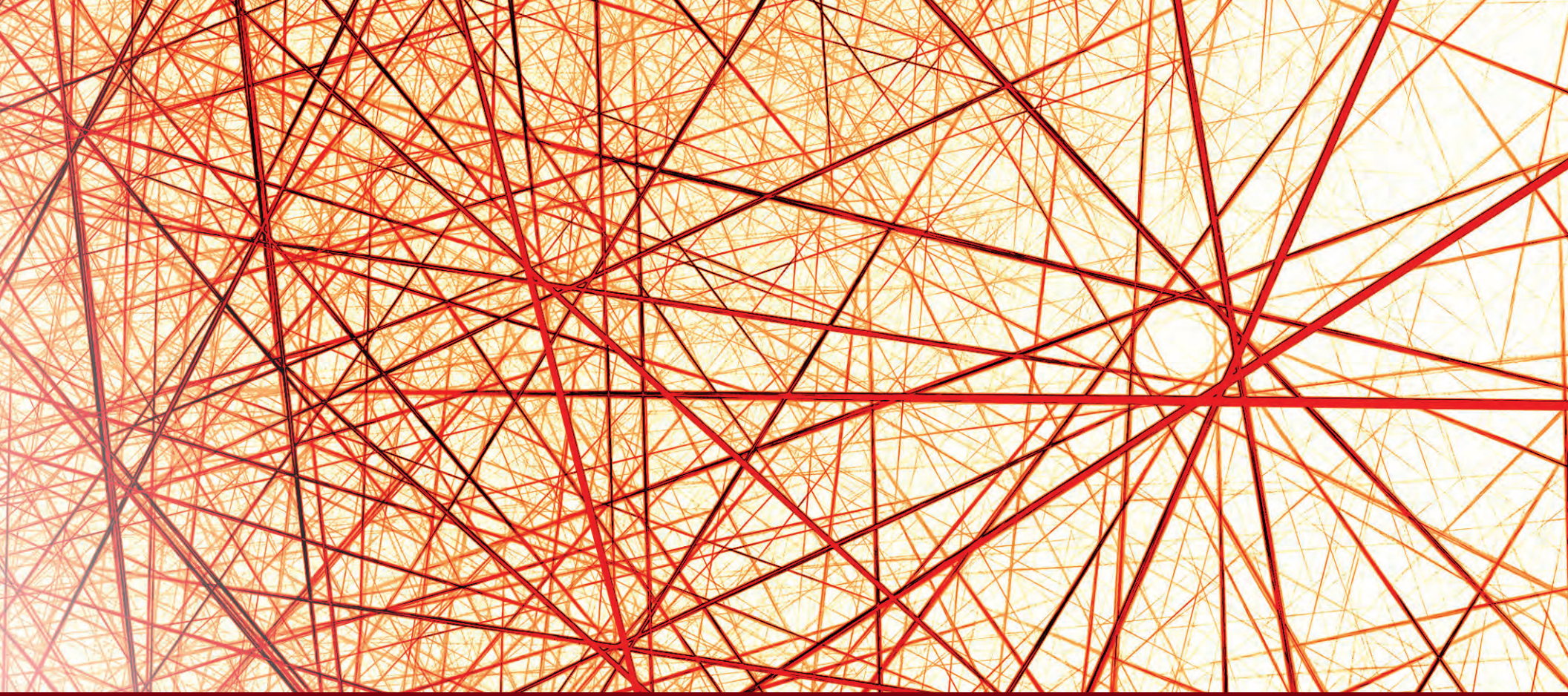
Whiting, M. A., Haack, J., and Varley, C. "Creating realistic, scenario-based synthetic data for test and evaluation of information analytics software," 8. In *Proceedings of the 2008 Workshop on Beyond Time and Errors: Novel Evaluation Methods for Information Visualization* (2008). Florence, Italy, April 2008. ACM, 2008.

[Younan 2013]

Younan, Yves. "25 Years of Vulnerabilities: 1988–2012." Sourcefire Vulnerability Research Team, 2013. <http://vrt-blog.snort.org/2013/03/25-years-of-vulnerabilities-1988-2012.html>.

The Software Engineering Institute (SEI) is a federally funded research and development center (FFRDC) sponsored by the U.S. Department of Defense and operated by Carnegie Mellon University.

The SEI mission is to advance the technologies and practices needed to acquire, develop, operate, and sustain software systems that are innovative, affordable, trustworthy, and enduring.



Software Engineering Institute
Carnegie Mellon University
4500 Fifth Avenue
Pittsburgh, PA 15213-2612

SEI Washington, DC
Suite 200
4301 Wilson Boulevard
Arlington, VA 22203

SEI Los Angeles, CA
2401 East El Segundo Boulevard
El Segundo, CA 90245