THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

Software Security Expert Judgment Elicitation Workshop

Attending this workshop will benefit all persons performing or estimating software reverse engineering tasks

- Do you perform or lead software reverse engineering (RE) or vulnerability assessment tasks?
- Have you ever been asked to estimate for a software reverse engineering or vulnerability assessment task?
- If so, was it challenging determining the answer?

If you answered yes to any of the above questions, this problem area is important to you!

All software reverse engineering and vulnerability assessment personnel, including project managers, to participate in a research study titled, "Towards Understanding the Effects of Analysis Environment upon Security Vulnerability Discovery in Large Software".

JHU-APL's Reuben Johnston (AOD/QWD) is leading this George Washington University (GWU) academic dissertation research study under the direction of Dr. Thomas Mazzuchi of the Department of Engineering Management and Systems Engineering at GWU. The study will explore the effects of various software characteristics and analysis environments[1] upon the post-release discovery of security vulnerabilities[2]. Ancillary purposes are to introduce structured expert judgment data gathering techniques and to explore the application of Bayesian computed proportional intensity models (PIM[3]) towards vulnerability discovery prediction. The former, elicitation of structured expert judgment, aids the data scarcity problem. The latter is a step towards addressing the primary deficiency of "black-box"[4] reliability modeling, which is their inability of normalizing discovery predictions to other post-release operational environments.

Participants will be asked to attend **one** 4-hour elicitation workshop at the Johns Hopkins University Applied Physics Laboratory. Through this collaborative learning experience volunteer participants will help advance the science of software vulnerability prediction modeling. Participants will benefit from:

- REuben's crash course introduction to approximately 50 metrics which possibly influence software analysis and reverse engineering (plus, reflection time to ponder their levels of significance)
- RE practice exercises on decompiled JAVA code
- RE and vulnerability assessment knowledge exercises (plus, time to think about your personal software analysis and reverse engineering process)
- Practice estimating discovery tasks, given certain different product release scenarios and analysis environments
- Concluding group discussions on
  - Metric influences upon discovery
  - Software analysis and reverse engineering processes used by the different participants

[1] Analysis environment-set of unique variables defining the operational environment that influences the discovery of security faults for a particular software release (Johnston, 2014)

[2] Vulnerability-an instance of a mistake in the specification, development, or configuration of software such that its execution can violate the explicit or implicit security policy (Ozment, 2007; Krsul, 1998)

[3] The Cox proportional intensity model (PIM), also known as the modulated Poisson process, is based upon the assumption that the mean value function for the $\boldsymbol{X}_i$ environment, $\Lambda_i(t;\boldsymbol{\theta},\boldsymbol{X}_i)$, is exponentially related to the baseline $\boldsymbol{X}_0$ environment's mean value function $\Lambda_0(t;\boldsymbol{\theta},\boldsymbol{X}_0)$, via $\Lambda_i(t;\boldsymbol{\theta},\boldsymbol{X}_i) = \Lambda_0(t;\boldsymbol{\theta},\boldsymbol{X}_0)e^{\boldsymbol{\beta}^T\boldsymbol{X}_i}$ (Cox, 1972)

[4] "Black box" reliability models only utilize historical fault/failure data for similar software and their predictions do not take into account internal characteristics of the software (e.g., size, complexity, or structural dependencies) or the levels of resources applied to analysis (Johnston, 2014)

- What makes the estimation process difficult

Volunteering will help towards advancing the science of software vulnerability prediction modeling. The benefits to science and humankind that might result from this study are:

- Whether or not certain analysis environment covariates influence post-release vulnerability discovery, as well as their levels of influence
- A solution towards meeting the need for model normalization to the analysis environment of interest (allowing for vulnerability discovery rate estimation and comparison across differing product release scenarios and analysis environments)
- Application of structured expert judgment elicitation and Bayesian analysis techniques towards addressing the vulnerability discovery data sparseness problem
- A methodology, to assist software company decision makers in predicting the number of post-release discoveries for different environments

Participation is not limited to only JHU-APL employees and readers are encouraged to invite qualified colleagues to participate in the workshop. It is also allowable to attend the workshops and not participate (if space is available).

During the elicitation session, questions will be based upon **crafted** release scenarios for **fictional** software products. Main elicitation sessions will ask the participants what the expected discovery counts would be within specified time intervals for various release scenarios and analysis environments. In addition, there will be a separate set of RE and security analysis practice exercises (listed above).

Table 1 details the elicitation tasks involved in this research, should you choose to participate. No questions will solicit personally identifiable information. Reported results of this research study will not name or identify attendees. Workshops will be no more than 4 hours (not including the two 15-minute breaks) and this event is strictly voluntary (i.e., not billable). Participants may refuse to answer any of the questions and may also stop at any time. Possible risks or discomforts which could be experienced during this study include: loss of confidentiality and minimal psychological stress (estimated to be comparable to experiences from undergraduate third or fourth year level computer laboratory examinations). The impacts from loss of confidentiality will be controlled as is described in the paragraph following Table 1 below.

Table 1-Overview of Tasks

Information questionnaire (completed prior to start)	• Participants will receive a brief questionnaire with general questions on formal education (degrees only, numbers of classes which covered software security related topics), relevant work experience statistics (only the total months of software security assessment experience), information on professional training completed (only the total days of relevant software security sessions attended), and knowledge about various technology (such as processor architecture knowledge and software languages known)
	• Participants will complete the questionnaire
Session I, RE task estimation practice exercise (30 minutes)	• Attendees will receive a brief instructional overview of the RE task estimation practice exercise • Participants will be provided release scenario and analysis environment descriptions, including product descriptions and available personnel to perform the work • Participants will be provided a short list of task descriptions to estimate
	• Participants will be asked to provide estimates for the time to complete the tasks given the described release scenario and analysis environment
Session II, expert elicitation	• Attendees will receive a crash course on metrics

session A (60 minutes)	• Attendees will receive a brief instructional overview of the NHPP parameter elicitation session • Participants will receive definitions for five separate variants of a baseline release scenario and analysis environment combination
	• Participants will complete the elicitation session for estimating the PIM parameters $\boldsymbol{\theta}$ in the baseline environment's NHPP mean value function $\Lambda_0(t;\boldsymbol{\theta},\boldsymbol{X_0})$, given each of the release scenario and analysis environment combinations ○ For each of the five variants, distributions[5] for the expected discoveries per interval will be elicited for five consecutive 10-week intervals (see Figure 1)
Break (15 minutes)	• Attendees will have a 15-minute break
Session III, expert elicitation session B (60 minutes)	• Attendees will receive a brief instructional overview of the PIM parameter elicitation session • Participants will receive definitions for 50 separate pairwise sets of baseline release scenario and analysis environment combinations
	• Participants will complete the elicitation session for estimating the PIM parameter vector $\boldsymbol{\beta}$, which will involve pairwise comparisons of approximately multiple pairs of environments (each pair of environments differs slightly between each other) ○ For the j-th pair ($j = 1 \mathinner{..} 50$), one environment $\boldsymbol{X}_j^2$ will be assumed to have, on average, a specified number expected discoveries $E[N(t)]$ after 50 weeks of analysis ○ Experts will be simply asked to estimate the expected discoveries $E[N(t)]$ for the other environment, $\boldsymbol{X}_j^1$, after 50 weeks of analysis (see Figure 2)
Session IV, RE and vulnerability analysis knowledge exercise (30 minutes)	• Attendees will receive a brief instructional overview of the RE and vulnerability analysis knowledge exercise • Participants will receive a knowledge exercise on software reverse engineering and vulnerability analysis ○ Exercise will include general questions in the domain of software vulnerability assessment and reverse engineering
	• Participants will complete the knowledge exercise on software reverse engineering and vulnerability analysis
Break (15 minutes)	• Attendees will have a 15-minute break
Group discussions and wrap-up (30 minutes)	• Attendees will discuss metrics • Attendees will discuss RE process and vulnerability analysis • Attendees will discuss RE task estimation
Session V, RE practical exercise (30 minutes, participants may leave when finished)	• Attendees will receive a brief instructional overview of the RE practical exercise • Participants will receive instructions and a computer for the practical RE exercise including one reverse engineering and one change task, on decompiled obfuscated JAVA source code • Participants will receive a brief survey rating experiences from the exercise
	• Participants will complete the timed practical RE exercise • Participants will complete the brief survey rating experiences from the exercise

[5] For each interval, experts will be asked to provide distributions via point estimates defining the 5%, 50%, and 95% likely values for the expected discoveries $E[N(t)]$ for several possible baseline environments $\boldsymbol{X_0}$

Exit questionnaire (completed before leaving)	• Participants will be provided a brief questionnaire querying their experiences in the study
	• Participants will complete the questionnaire

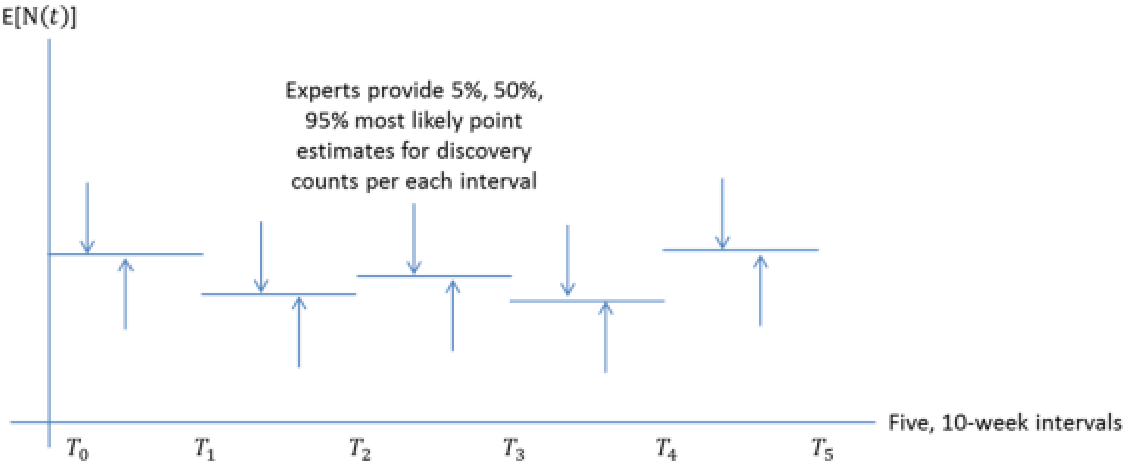
E[N(t)]
Experts provide 5%, 50%, 95% most likely point estimates for discovery counts per each interval
Five, 10-week intervals
T_0
T_1
T_2
T_3
T_4
T_5

Figure 1- Illustrative technique for eliciting the expected number of discoveries per time interval

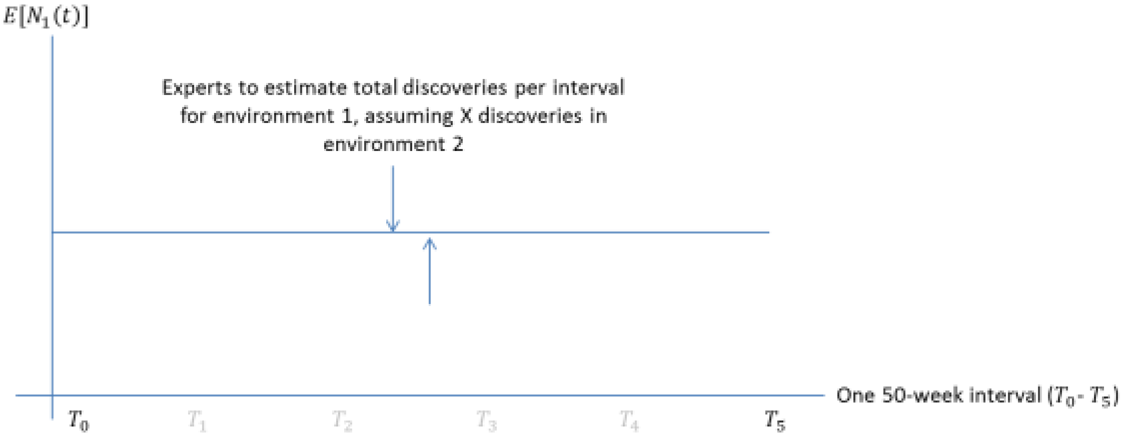
$E[N_1(t)]$
Experts to estimate total discoveries per interval for environment 1, assuming X discoveries in environment 2
One 50-week interval (T_0- T_5)
T_0
T_1
T_2
T_3
T_4
T_5

Figure 2-Illustrative technique for eliciting data to estimate β

Taking part in this research is entirely voluntary and there will be no monetary compensation for participating. The status of your employment will not, in any way, be affected should you choose not to participate, or if you decide to withdraw from the study at any time. Every effort will be made to keep your information confidential, however, this cannot be guaranteed. To mitigate this risk, each participant will be randomly assigned a number and all data gathered will be identified solely by these identification numbers. Additional precautions have been taken to limit the impact of loss of confidentiality, with the creation of questions based upon constructed scenarios[6], and not upon anything which would be considered business sensitive. If results of this research study are reported in journals or at scientific meetings, the people who participated in this study will not be named or identified.

The Office of Human Research of George Washington University, at telephone number (202) 994-2715, can provide further information about your rights as a research participant. Further information regarding this study may be obtained from the contacts listed in Table 2, and at http://www.reubenjohnston.com/Research.php.

Table 2-Contact Information

Name	Role	Telephone	Email
Reuben Johnston	Researcher	240-228-5869, 443-858-2416 cell	reuben.johnston@jhuapl.edu
Dr. Thomas Mazzuchi	Principle Investigator	202-994-7541	mazzu@gwu.edu

Your willingness to participate in this research study is implied if you proceed with completing the survey/interview. It is also allowable to attend the workshops without the requirement of elicitation form completion (i.e., for informational purposes only). We ask that you not discuss the research outside of the workshop, until notified by Reuben Johnston that the research has been completed (to prevent biasing of participants attending alternate workshop dates).

*Please keep a copy of this document in case you want to read it again.

[6] Constructed scenarios will describe the analysis environment for realistic, but fictional, software releases. The main reason for this approach is that there are no possible repercussions for having security discussions on the post-release vulnerability discoveries from these scenarios