File 20100212.0602: Weekly activity report 0123:

weekly activity report 123 (loughry)
Joe Loughry
Sent: 12 February 2010 06:02
To: Niki.Trigoni@comlab.ox.ac.uk; Andrew Martin; Joanna Ashbourn
Cc: andrea@hpwtdogmom.org; Joe Loughry; mmcauliffesl@comcast.net
Attachments:
Weekly activity report no. 20100211.1813 (GMT-7) sequence no. 0123, week 4 HT

I met with Dr Martin on Wednesday after reading group by video teleconference. I reported that I had tried the thesis-testing exercise in Rugg and Petre (2004) and found that it helped. The latest iteration of my thesis problem is:

> Is there a difference in the post-CT&E software defect rate (as measured by the number of findings of Category I, II, II, and IV) between the same or newer versions of a given system in subsequent rounds of CT&E by different DAAs?

This, according to the test mentioned earlier, is a solid thesis statement for the following reasons:

1. It has a limited number of clearly distinguishable possible outcomes.

2. The occurrence of any of the outcomes would tell us something interesting.

3. Before collecting and analysing the data, it is not obvious which outcome is the most likely.

In reality, this is the same thesis statement as before, just expressed in a form more amenable to statistical analysis. I honestly can't say which outcome I think is more likely. I can think of convincing arguments why any of them might be true. The first outcome---more findings over time---is made more probable by the continual improvement of testing tools, development of new attacks, adoption of new (presumably better) certification criteria, and by having more eyes looking at the subject. I have an example of this occurring in my first case study: a system which had been evaluated many times by one agency, but when a different agency looked at it, they found a weakness that had been overlooked for years. The second possible outcome---fewer findings over time---is arguably what ought to occur if secure software development practices are followed and are successful. This is perhaps the way the smart money would bet. A third possible outcome would be no change in the number of findings over time; I think this may in fact be a subset of the second possibility, but I need to think about the implications of considering it that way some more first. One thing I need to decide is what is meant by ‘no change’, and what are the implications of that decision. Does it refer to the absolute number of findings (which is a multidimensional quantity), or the rate of occurrence, or a trend in the number, or rate?

Cross domain systems are a good example of systems that get repeatedly CT&E'd because these systems go into a new situation with new data owners nearly every time. But what about safety-critical systems? Are they repeatedly re-tested under new criteria every time they are installed in a new type of aircraft, for example? What about updates to the criteria such as the imminent changeover from DO-178B to C? All of these may provide insight into my problem.

Cost (in pounds) is another metric, but one that is going to be difficult to gather data for. I have not seen it clearly broken out in budgets and project managers are loathe to disclose.

Dr Martin, playing the Devil's Advocate, asked: how do you know you're not just measuring noise? The counter-argument is that in this field, people don't get to conduct enough experiments. I think I can answer this objection by using standard statistical methods. I can calculate the number of experiments necessary to get a 95 percent confidence interval, even though my gut feeling is that a CI of 95 percent is not achievable. Conversely, I can show in one of the appendices the CI that is supported by experiments, perhaps thereby anticipating one of the questions an external examiner is likely to ask.

Dr Martin says I am making progress, but I need to watch out for too much introspection. Thoroughness and looking at the problem from all angles is all well and good, but there needs be results as well. I promised to have a set of slides for my talk ready tomorrow (Friday) or so.

Other activity this week:

Following some of Dr Martin's advice, I have been reading the thesis statements in other people's dissertations. When I am in Oxford next week I will look up some more in the library. I have been putting off Lockheed in regard to one particular project, but that will only work for a while. On another note, I have been answering C&A questions more and more often for Lockheed. Other departments are sending people to this department for answers.

Next meeting will be in Dr Martin's office on Friday morning, 19th February. Next time Reading Group meets, I will be on my way to the airport.

Current list of tasks in order of priority, highest priority first:

1. Software engineering talk and slides for 19th Feb. 2. Methodology chapter 3. Renew student visa, including biometrics now required 4. CT&E practitioner survey (waiting until after 19th) 5. New (short) paper (interim???between these two) 6. Crosstalk article (concurrent with above) 7. Update schedule for Dr Martin 8. Apply for confirmation of status this term 9. Must have achieved confirmation of status by end of Trinity term

Joe Loughry
Doctoral student in the Computing Laboratory
St Cross College, Oxford

End of WAR 0123.

References