File 20100723.0513: Weekly activity report 0146:

weekly activity report 146 (loughry)
Joe Loughry
Sent: 23 July 2010 05:13
To: Niki Trigoni; Andrew Martin; Joanna Ashbourn
Cc: otaschner@aol.com; anniecruz13@gmail.com; andrea@hpwtdogmom.org;
chip.w.auten@lmco.com; edloughry@aol.com; diane@dldrncs.com; Joe Loughry;
mmcauliffesl@comcast.net; tom.a.marso@lmco.com
Attachments:
Weekly activity report no. 20100722.1622 (GMT-7) sequence no. 0146, week 8+5 TT

I met with Dr Martin on Monday of this week. I have been working on a numerical model for the pricing function that is part of my attempt to prove that the accreditor--accreditor interaction model behaves like a classical market. To this end, I have a catalogue of six hundred security controls from the NIST SP 800-53 standard and I am starting to go through the list and rank-order them by the amount of effort, or pain & suffering, that performing each test entails. I do not have a plan yet for assigning a numeric value to each test, but for a start I can at least establish a partial ordering. I want the price of an item to reflect the amount of work that an accreditor would have to do, or conversely the amount of work they would not have to do if another accreditor pre-empted that test by doing it first or offering to do it instead. It might be useful to implement the concept of promising to do such work in future, perhaps even with a discounted 'net present value' type of relation in order to value work that has not been done yet. I should definitely build this feature in to the model, I think. Maybe I will even build in futures and options. Those are useful for risk management.

Dr Martin said he is still trying to decide what to think about it. He wondered if it might evolve into a more complicated piece of game theory later on, i.e., that a simple numerical equilibrium might not be all that is at the core of it, once I really understand the model. I should make sure that the incentives fall in the right places. Actually, I have learnt recently that exactly that is one of the tests that a proposed equilibrium is valid. I have already suggested that accreditor turf wars, something I have observed to exist, might be one of those complicating game theory factors. The type of game is called a Cournot game.

I had an interesting conversation with a retired Air Force accreditor last week. He and I discussed the ways in which a tool such as the one I propose could have made his former job easier. All the accreditors I have talked to so far have said something similar: if I can figure out how to get the other guy to do the work, they are in favour of it. Dr Martin pointed out that fairness is important: it would not do to dump all the work on the little guy. I agreed that there should be a concept of fairness in the tool, and related a lecture by Professor Polak at Yale that I watched the other night, in which he described an example where unfairness was structural. Despite the fact that the equilibrium he showed was more efficient than the case with no equilibrium, it was inherently unfair: the lowest-paid participants ended up getting paid 30 in the more efficient equilibrium, whereas they were paid 32 in the less efficient equilibrium. The richest were not paid any more, but the poor were paid less. The difference went to waste, or friction. It was an interesting lesson. I will try in my tool to get everyone to do a fair share of the work, but efficiently. More efficiently than at present, anyway. My two case studies will serve as a baseline to illustrate what 'at present' means.

We looked at my three-month goals. My first and most important goal is to develop the 'equilibrium' model further to the point where I can begin to implement some of it. Dr Martin pointed out the difference between a model---on which I can perform experiments and what-ifs---and the implementation of a tool for real accreditors to use. Dr Martin cautioned against trying to make it over-general, that I could spend all my time implementing it. I discussed whether to investigate the possible use of systems engineering modelling tools. I know they exist, but not how to use them. I would like to avoid reinventing the wheel if possible. Dr Martin said that looking at modelling tools would be a good idea.

[Update: I talked with Greg Shettlesworth, a very experienced systems engineer at Lockheed, about available simulation tools. He recommended Matlab. He said it has a great interface, and the older modelling and simulation tools I saw him using in 1999 do not exist any more. Mathematica is another possibility, but Matlab is what he recommended.]

My second three-month goal is to finish the analysis of the failed EAL 4+ evaluation. I discussed whether to try to use it to argue the case that if my proposed tool had only existed back then, that the evaluation might not have failed. Dr Martin commented that that argument is not falsifiable. It is all right to hint at conclusions, but primarily I should present my data and conclusions, and let the reader decide the significance. It is still okay to hint, however.

The third goal for the next three months is to write up a similar analysis of the RM 5.0 CT&E effort. It would not be a contribution, however, simply to describe what happened. I need to explain why things happened. I have been taking careful notes of all the meetings I have attended. I think they will probably slip the schedule three weeks, making the certification date the first week in September, but I do not think they will run out of money and I do not think the penetration testers will spring any surprises. I have certainly seen a lot of interesting human interactions, only slightly constrained by the relevant standards documents.

[Update: in Thursday's meeting, they did not slip the schedule.]

The purpose of my three goals is to have preliminary results to show to the assessors at confirmation of status. I may not get to the EAL 4+ analysis before that time. I have all this data to analyse, but it may have to wait until after confirmation.

Dr Martin will be in Australia for the next month, but reachable, at least on a few days notice. I have a good plan and a lot of work to do, so I am going back to work now. I am going to work on the pricing algorithm, the surveys, and the Crosstalk journal article.

Dr Martin said it is coming together pleasingly.

On Wednesday, I presented my paper 'Information Asymmetry in Cross Domain System Accreditation' to the Security Reading Group. Cornelius, John, Mingqiu and Shamal were present. I began by explaining the rationale for attempting to apply what appears to be a completely unrelated field to the problem of determining risk due to security vulnerabilities in a computer system. I am aware that I am taking a risk with this paper, putting forth a not-completely-developed idea at such an early stage of the process of working out the implications, but I must have feedback from other researchers. Specifically, it is so I can validate certain assumptions. Reading Group served the purpose well.

What I saw was that my paper does not contain enough background for non-economists who are unfamiliar with the work of Akerlof, et al. to be able to read it. I was constrained by the four-page limit imposed by this ACM workshop---good for conciseness but bad for a paper such as this that depends on background material that the audience might not have. It is clear to me that I need a larger introduction to bring the justification for the central argument into the forefront instead of leaving it implicit in references to three seminal articles. A rapid introduction to Akerlof's asymmetric information and Spence's market signalling would go a long way towards making the paper easier to understand on its own. I have made notes for extending it.

The paper is under consideration by the 2nd ACM Workshop on Assurable and Usable Security, co-located with ACM CCS in Chicago. Notification is 6th August. I asked if anyone had any ideas where to send it next if SafeConfig 2010 does not accept the paper; the Workshop on Software Engineering Economics (WISE) was mentioned, of course. Shamal suggested NSPW for a longer version of the paper with a more worked-out example. I replied that my model is artificial, but Shamal countered that NSPW dislikes examples from work and would probably enjoy an artificial example more. They want something that workshop participants can hack on and improve. That is why I submitted the paper to the ACM workshop in the first place, hoping for the same kind of interaction. NSPW is more prestigious, but the deadlines worked against me for submitting to NSPW this year.

John suggested looking at who is on the Programme Committee of WISE and to send my paper to other conferences where those same people make up the PCs.

Regarding the central idea of the paper, I saw a lot of glazed eyes. I think people generally agreed that the conclusions I drew in the paper are logical, but they would prefer to see more justification. Mingqiu asked, what is the fundamental assumption: is it the satisfaction of all parties? Cornelius asked, what is the signal that an accreditor sends? Is it that he or she wants to optimise the residual risk to as low a value as possible? More to the point, Cornelius asked, do accreditors cooperate, or do they compete as in a traditional market? I postulated that accreditors cooperate. Shamal instantly pounced on that and asked whether a market solution is the right model. I argued that yes it is, because accreditors still use the market mechanism to negotiate a price. Whether that mechanism is driven by competition or cooperation makes no difference: equal but opposite forces can still drive an equilibrium. Shamal conceded that the market interpretation is valid. He then asked whether a market mechanism was overkill for such a simple result (merely optimising residual risk to as low a value as possible within certain constraints). I replied that the model I presented in this paper is still very simple, lacking some other factors that will make finding an equilibrium level much more interesting. At this point I asked if anyone in the group had any experience with systems engineering modelling and simulation tools. One person suggested Matlab for systems dynamics modelling. I will look for a book on Matlab and see how well suited it seems for my purpose. I have emailed Greg Shettlesworth at Lockheed; I know he uses an expensive commercial tool for systems dynamics modelling; I will find out what he uses and what its capabilities are.

How does an accreditor make another accreditor want to 'buy' the risk he is 'selling'? The answer is implicit in my straw-man pricing function: risk is inversely related to testing effort. Some kinds of tests, for example penetration testing, are very expensive because they require extensively experienced people, sometimes special equipment,

and an open-ended amount of time. (Much like covert channel analysis, penetration tests are never finished; testers simply run out of time or money and are ordered to stop. Left to their own devices, they would generate findings indefinitely at some baseline rate, fuelled by new discoveries, new tools, and the capacity for exhaustive enumeration of a finite state machine.) It was pointed out that an accreditor who is not satisfied can always require more and more assumptions to be tested, more requirements. I tried to shore up my argument by showing that behaviour like that on the part of an accreditor would work to his or her own detriment.

In response to Cornelius's question, I tried to explain better how the limiting case acts to force accreditor behaviour from both directions. If an accreditor were to try to falsely manipulate the 'price' by claiming that the residual risk was lower than the accreditor knew it actually to be, then that accreditor would only be increasing his own personal risk---the risk to his own career of approving an insecure CDS that is likely to fail in the field. Conversely, if an accreditor were to try to manipulate the price in the opposite direction, the result would automatically be to cause more work for himself, because the other accreditors in the market would demand more time and effort in the guise of additional testing. So all of the incentives continue to work in the right directions, something that Dr Martin specifically asked about and also the Economics lecturer at Yale that I watched on video emphasised---that one of the acid tests for the existence of an equilibrium is to verify that the incentives (he called them 'beliefs') work in the right direction. (The other acid test is to assign numerical values and show that the equation balances, which is something I have not done yet. There seems to be an art to picking those numeric values, similar to finding a nice clean integer solution to a system of simultaneous equations.)

There followed some discussion of a possible fallacy. What are the incentives really? In the Common Criteria, the argument goes that Protection Profiles (PPs) are developed by clever and conscientious people who always have the end-user's best interests in mind when they specify requirements in a PP. And yet we know from experience that PPs are far from perfect. Shamal speculated that the fallacy is in assuming a binary outcome, whereas he argued that the real world is a continuum because humans are involved in developing it. I need to think about that some more to come up with a good counterargument.

I emailed Mingqiu afterwards to ask if I had answered her question adequately. I do not think I did so in Reading Group.

In summary, I am now less sure that my paper will be accepted at all; I wish I had had another six pages to expand the introduction, work out the equilibrium numbers formally, and address all the questions that came up in Reading Group. Some of the answers I already knew, just had not had space enough to write them in the paper. Other points were things that I want to put into a new paper, either an extended version of this one (if the present version is not accepted by SafeConfig 2010) or a follow-on article.

In other activities, for the Programme Committee of the comlab DPhil student conference, I found some articles about efficiently running a programme committee and how to review and rank submissions. Other members of the PC have been looking into an inexpensive source of binding for the conference proceedings and a site for the meeting and catering.

I have decided not to go to the UCDMO conference in Boston the second

week of August. I contacted three of the four people I wanted to talk to there, and their recommendation was not to try to catch up with them during the conference. Instead, I am going to make appointments with each individually at their work locations, where I can have guaranteed time. To try to meet all of them at a busy conference full of closed sessions that I do not have clearance for would likely have failed. In addition, the Navy programme office waited so long before granting approval to attend that airline tickets were prohibitively expensive on short notice. I am in the process of arranging appointments in the D.C. area to take place during a stop-over on my next trip in to Oxford. I plan to use the accreditor surveys to prepare for in-person interviews.

I attended the weekly CT&E telecon on the status of my second case study this morning. Present on the call were Kevin Miller, Kori Phillips, myself, Russ Savage, Larry Sampson, Dennis Bowden, NSA I173, NSA I733, Charissa Robinson, Dave Oshman, Emely Martinez, Dan Nichols, and Rob Drake. In return for sitting in on their phone calls, I have lately been giving them a copy of my notes to serve as their meeting minutes.

The first topic discussed was status of government regression testing. USN SSC Charleston reported that testing is still going on; they have found several issues [redacted]. The developer concurred and has been able to reproduce the problem. Severity of new findings is medium. Regression testing is expected to be complete tomorrow.

Dennis Bowden asked the developer to go over the rest of the findings from SSC Charleston. The developer's plan to address the first finding is to change the way the browser is installed, as documented in the installation procedure. This will avoid a new build. The second and third findings are related to each other; the developer had to fix a problem in the source code of one FOSS library. The developer ran 30,000 test messages through the new code this morning and believes that the problem is fixed. The fourth and fifth issues show the system working as designed. The second to last issue is an old one; Microsoft Office files discovered in a directory on the test system are not part of the configuration and should have been removed long ago. They comprise old test data and should not be interpreted as an indication to test functionality that is not in the TOE. The last finding is another case where the system is working as designed. [Editorial note: the regression testing being performed by SSC Charleston essentially duplicates the developer's Factory Acceptance Test procedures; therefore it is not too surprising that repeated findings would show up this week. The testers, as expected, wrote up the issue as a high risk finding. The developer does not believe that Charleston understands the difference between CT&E and ST&E.]

Kevin Miller: we think we [the developer] have good responses to all findings. Either we have a fix in place, or a procedural workaround, or something is working as designed.

Dave Oshman asked whether it is possible to reduce the 15 minute time-out on management workstation sessions. Kevin Miller replied that he would check with the software developer. Currently the session duration matches the screen lock time-out value. Dave Oshman replied that the present value is probably secure enough, but noted that the CDTAB might want to see it reduced, so it would be good to have anticipated their question by having the value be configurable.

Emely next had a question about WinDDS. Are the recommended default

permissions on folders documented in the Trusted Facility Manual (TFM) for sites to use? Russ Savage replied that the permissions are documented in the installer's manual. Kevin Miller allowed as how the folder permissions guidance could be moved to the TFM.

Larry Sampson, the moderator of these telecons, asked when the POA&M would be updated with the latest findings. Dennis Bowden replied that the POA&M will be updated; Kevin Miller stated that the latest findings will be added to the end of the spreadsheet. Larry Sampson asked that the new version be called version 3.

Emely asked for information on how SSC Charleston will receive the new code drop containing fixes for regression testing. The developer replied that they are planning to deliver a total of two files containing critical fixes, with the remaining issues to be handled by a post-5.0 patch.

[Editorial note: NSA won't use the jumpstart procedure for some reason. That might actually be a clue to understanding their pen test methods.]

[Editorial note: Russ Savage will be on-site at STRATCOM for the next four weeks; ST&E will begin three weeks from now. During week 4, the DNI people will arrive. They are still on schedule for a 20th August certification date; that schedule has been maintained for months now. Everyone is striving to keep to the schedule because after 20th August 2010, certain parties will become unavailable. The POA&M states that the developer will fix certain findings from CT&E and ST&E within a certain time after certification; all of the findings listed in the POA&M are minor or medium. The new version after POA&M patch will be called 5.01.]

NSA I173 asked whether the critical fixes will be regression tested, or only tested in ST&E? Kevin Miller replied that the developer will regression-test them.

Rob Drake [accreditor for STRATCOM ST&E] stated that he would regression test them at site. Kevin Miller pointed out that of the fixes, some functionality is not even used at STRATCOM anyway. Dennis Bowden asked that NSA or CT&E test the critical fixes in the two files to be provided by the developer---the library fix and the other one; all of this will become part of the body of evidence.

Larry Sampson asked for a final overview: will all the findings be covered by ST&E? Dennis Bowden replied that the ST&E plan was just posted a couple of days ago.

Rob Drake said we have two weeks; we will definitely test all serious findings. Russ Savage pointed out that the STRATCOM configuration does not use or even include all of the features or message formats covered by CT&E; it would take a special configuration to test some of them, for example VMF. Rob Drake replied that he would test those things at the next site.

Larry Sampson: so we are on schedule? Pressing on? Rob Drake and Dennis Bowden replied in the affirmative.

Larry Sampson then asked about the fact sheet. Dennis Bowden replied that he is working on it. He is also working on another document for Lisa Ackerman, the UCDMO submission document, which incorporates most others by reference.

Finally, Larry Sampson asked if there were any other issues with STRATCOM.

Dennis Bowden: There are no show-stoppers.

Russ Savage [on site]: The build is having no problems. The usual connectivity problems, nothing unusual.

Meeting ended 0827 MDT. The next meeting will be Thursday, 29th July at the usual place and time.

I have not done the defence acquisition process training yet, but the developer is excited about a new contract vehicle that is replacing 'nimso'. Not much experience has been had with Netcent yet, but project managers have expressed happiness to have something other than the former contract vehicle in their critical path. I will try to learn more about what this is, and why it is important.

I have been reading a new book this week: Space Systems Failures, by David M. Harland and Ralph D. Lorenz. It is primarily about flight hardware but contains some software engineering failures as well. It is exhaustive; there are hundreds of failures (and recoveries in some instances) described. Not to the level of detail found in a book like Safeware, only a few paragraphs on each item . All of the sources are cited, though, making it possible to find more details easily. Some relevant quotes that I found in the book:

- On NASA's faster-better-cheaper from a few years ago, Dan Goldin: '...said that if failures do not occur, then NASA was not trying hard enough'.

- 'A more sanguine long-term view might have permitted a few more high-risk attempts, in order to achieve a better determination of the real reliability of lower-cost space systems.'

- 'Other failures can be attributed to overly lean design or operations teams, in which insufficient time, or oversight, or inexperienced (and therefore cheaper) staffing allowed problems to pass unnoticed.'

Source: Harland and Lorenz, 2005, pp. xv--xvi.

My current task list (in priority order, most urgent first):

1. Arrange appointments with Paul Ozura, Frank Sinkular, Dan Nichols, and Dave Wallick in D.C.
2. Develop and test numerical pricing equation. See if both acid tests hold. Test against real-world assumptions for reasonableness.
3. Working on the Crosstalk paper again.
4. Accreditor surveys. I understand the problem much better now. I need to be able to describe these surveys in the Crosstalk paper.
5. Get the other two surveys done for background on the case studies.
6. Finish methodology chapter (waiting on final survey questions).
7. Prepare talk for VALID 2010 in Nice and submit to PIRA for approval.
8. Prepare talk for 13th August at Lockheed on crypto maths.
9. Compare NIST SP 800-53A to ISO 27001/2.

To be done as soon as possible (unchanged from last week):

10. Update dissertation Table of Contents.
11. For Chapter 3 or 4, start writing the interpretation of the first case study results and second case study preliminary results. (This will be needed for both confirmation of status and for answering

```
likely audience questions in France.)
12. Begin writing progress report.
13. Update the schedule.
14. Apply for confirmation of status.

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

End of WAR 0146.

References