File 20110204.0348: Weekly activity report 0174:

weekly activity report 174 (loughry)
Joe Loughry
Sent: 04 February 2011 03:48
To: Niki Trigoni; Andrew Martin; Joanna Ashbourn
Cc: otaschner@aol.com; anniecruz13@gmail.com; andrea@hpwtdogmom.org;
chip.auten@comcast.net; edloughry@aol.com; diane@dldrncs.com;
Joe Loughry; mmcauliffesl@comcast.net; tom.a.marso@lmco.com

Weekly activity report no. 20110203.1205 (GMT-7) sequence no. 0174, week 3 HT

The R'' developer has been experimenting with virtual machine hosted CDS for some time. Recently, technical and certification guidance from NSA and other C&A authorities have shifted from a prior stance of hostility towards virtual machine monitors---for example, the NetTop discussion in my notes of the AEHF MCS Ground Segment review during the R_0 time frame---to embracing the same technology of late for space, weight and power reasons. (It is likely as well that the state of hardware support for virtualisation on X86 processors was poor in 1999 compared with today.) CDS certifiers still consider the cascade problem to be paramount [NSA, 2011b], but the latest design guidance has been to recommend use of a minimised hypervisor for the next TOE of R'' (which will not be RM 5.0.1) and the developer will probably follow suit. The new information is significant because this design guidance comes from back-channel sources and references the favourable impression of [redacted], another top-level certification authority. I need to figure out how to represent this narrative event in ATLAS.ti.

Use of the Scientific Software tool for qualitative analysis of the R'' certification effort narrative material is under way. Metrics of progress are poorly defined at this time, but there are 345 records, 417 links to R'' in the hermeneutic unit, and 22 names with more than one link associated. I began with all of the records in one file, but it has proved necessary to split the file into pieces for ease of importing into ATLAS.ti. Every time I turn around I find another opportunity to use another paper artefact from the dead file, although some of those are very large and mainly consist of minor diffs from a previous version, so I am trying various ways of coding them and will go back later to fix it. I am getting a sense from the reading of the way a theory is supposed to emerge from the ground truth, but I confess I do not completely understand it yet. New references added to my list of things to read included Denzin and Lincoln (2005), Miles and Huberman (1994), Backhouse and Dhillon (1996), Adams and Sasse (2001) and Martin and Turner (1986). Some day I will have to do a grounded theory analysis of these reports together with my lab notebook and my daily log. It might yield a plan for finishing a DPhil more quickly.

The Air Force funding sponsor was pleased enough with the new report and my proposed metrics and testing framework to request abstracts for follow-on work. However, after the sponsor's Technical Council review, the sponsor came back today asking for even more information beyond. I will respond to that request tomorrow morning in support of the sponsor's briefing next Tuesday. Further to obtaining more funding for my research, I noticed a Broad Agency Announcement (BAA) from a government agency requesting proposals in fourteen Technical Topic Areas (TTA). The 'insider threat' (TTA #4) and 'digital provenance' (TTA #10) are potentially applicable to one or the other of my research topics. I have asked for bid and proposal help from Lockheed to write at least one funding proposal this month.

Travel arrangements for March have been set. I will arrive Friday, 11th March and meet with Dr Ashbourn that day, then go to RAF Fairford for some planned work and one other location over the weekend. Monday through Wednesday are available to meet with my supervisor a few times and to do library research before flying out Thursday morning. Getting started on writing Chapters 1--3, for the past week I have been reading papers and books with stylistic observation in mind, trying to improve my academic writing style. I have one writing style for this report, another that I use for formal reports, and yet another writing style for email.

Part of the reason for the side trip in March is related to the requirements of DoD 8570.01-M; in an engineering staff meeting today the R'' developer related that they have not seen any contractual requirements for 8570 compliance yet, but another programme in the building has instituted a policy, probably at the accreditor's request, that no installation personnel are allowed to go out any more without at least an IAT Level I certification. For the R'' programme, because it is a network enclave CDS, at minimum an IAT Level II---and for some installations, IAT Level III---are required. The corporate parent still is not paying for much training; I predict that prioritisation will become apparent soon, affecting CDS programmes first; this will be forced by the contract language in DFARS 252.239-7001 (Jan 2008).

The Oxford Security Reading Group did not meet this week. Dr George Danezis from Microsoft Research in Cambridge spoke to the Information Security and Privacy Programme seminar on 'Privacy preserving smart-metering'. Dr Danezis' research speciality is traffic analysis; there are some good papers on his web site [Danezis and Clayton, 2007; Danezis 2011]; he was previously in Ross Anderson's research group with Markus Kuhn. Cornelius provided a phone link into the seminar room that enabled me to listen through Skype; the presenter began talking about Norwich Union's use of GPS for monitoring driving patterns to inform insurance billing. There are privacy differences between that sort of use of information and similarly architected schemes used for congestion charging in the City or for over-the-road freight in Europe. The Norwich Union trial failed; people resisted fine-grained reporting of vehicle position data and the offered protections of the information gatherer. DRM systems like iTunes have similar privacy difficulties.

Traditionally, electricity billing records are aggregated over about a three month period. Smart meters reduce that aggregation time to thirty minutes or less; fine grained measurements then become available to third parties. Using traffic analysis, it is possible to infer the occurrence of events inside an individual house: how many people are living there, what their schedules are, what times the house is unoccupied, how often they bathe, even a guess at what types of food they eat (from the detectable difference between the cycle time of a microwave oven and a hob, the power draw profiles of other types of cookers, or thermodynamic effects of refrigerator door opening events). The thing that is special about electricity billing, unlike car insurance, is that smart metering is now required by law in the U.S. In the context of insurance, the location of their car can be privacy sensitive information to some people. Similarly, DRM records are sensitive information, even affecting careers (example: a recent MP in the expenses scandal). The leakage of information in half-hour blocks from electricity meters is very small, but over long sample times with statistical methods it could be used to build up a much finer-grained picture when data are correlated with other time-varying events available in the environment. An example of other time-varying events would be TV show schedules. Additionally, smart meters work in two directions; control features can be used to remotely turn off a residence for non-payment, or to switch the

household to a pre-payment plan. Dr Danezis is interested in designing security architectures that leak a controlled amount of information in two directions. Interestingly, although traffic analysis was mentioned repeatedly throughout the seminar, covert channel analysis (CCA) was not.

Fundamentally, one party at least is---by nature---entitled to access to all of the usage information: the consumer. Even that simple statement is not a given, however. Example: within a family, the 'consumer' extends over more than one person. Which of the members should have access to the information? There could be risk of privacy violations even in such a small, close-knit group of people. Student houses are another, even more complicated situation. Making the equation more difficult, electricity meter readings are not only used for billing purposes, but for maintenance monitoring, infrastructure planning, and troubleshooting. Consequently, any architecture that does not take into account those needs will never be rolled out by the utility providers.

Returning to the subject of satnav, for an application as simple as congestion pricing, an example of a privacy preserving architecture is to compute the bill inside the box. No information is leaked or transmitted to the billing authority except for the amount of time spent inside the congestion pricing area. In principle, the box should be simple enough that the user can audit the box. In 2007, exactly this was proposed. The insurance companies refused to buy into the scheme.

[Temporarily lost connection to the mobile telephone in the seminar room for a few minutes. Reconnected at 17.34 GMT.]

What the Microsoft researchers in Danezis' group have proposed can be shown to apply their design principles throughout. The smart meter (electricity or car insurance is irrelevant; the principles are the same in any case) must be tamper-resistant and must be trusted by the user not to leak information. (This is not quite the same as not leaking any information: see below.) Every fifteen minutes, the meter provides its readings to a user-controlled device such as a PDA. There is no direct channel to the provider; all communication takes place in the clear and in a form intelligible to the user. By law, meter readings must be preserved in the meter for six months; this provides redress for the user in cases of disputed billing disagreements. The meter provides to the user a 'certified' set of meter readings, and cryptographically ensures that readings are amenable to further processing. Readings are digitally signed so that none can fake the evidence that a particular set of readings came from a certain meter during a specified time. The researchers use a zero-knowledge proof to demonstrate that the final bill is a correct representation of the electricity usage.

[Question from the audience: 'what is the lifetime of the meter? It affects the security design. The end-user has incentive to lower the electricity bill, and has physical possession of the equipment; he can try to take it apart. Over time, parts of the security design or assumptions will break.'] This is an interesting question because it has an unexpected answer. The expected part of the answer is that lid switches and tamper alarms are well-understood technology to solve the detection-of-reverse-engineering problem. The reverse engineering problem is solvable in this instance, unlike the situation with DVD players, because both legally and practically there is an ongoing transactional and contractual relationship between electricity provider and consumer, with real goods [electricity] going in one direction and monetary value [bill payments] going in the other. By that justification, there is no privacy issue here. Nevertheless, the implementation of the meters is provided with two levels of security; there is a channel for tamper

evidence and the privacy policy has an exception for communication over that channel.

Secondly, what about the availability of billing information to the provider? What if a user never comes back on-line? What if he hides, refusing to answer the door and never paying his bill? The solution is that the meter continues to store information about usage and if the utility company can ever manage to break the door down and retrieve the readings, then there is no privacy at issue in this instance either. The purpose of the stored information is to enable retroactive billing; there is no privacy related difference between that and the aforementioned disputation of billing for the purposes of storing the information.

[Comment from the audience: 'end users cannot effectively audit a black box. The meter could be lying to them.'] Answer from the presenter: we do not have a really good answer to this objection. Certainly, some users might have the knowledge and ability to audit the box, but not my grandmother; my grandmother does not do cryptography. But under this scheme, the meter can be made very simple. The rest of the protocol happens in plain sight, in the clear on the user's device. Like open-source software, if you do not have the time or ability to audit the code yourself, at least you have the option of choosing to trust someone else who does have the capability to audit the code.

[Another question from the audience: 'who is the certification authority?'] Answer: usually, certification in this industry has to do with accuracy of readings, not privacy. As a user, how should you decide whether to trust one provider or another? The Dutch population, notably, did not accept the technology. The courts, if this information exists, would tend to use it in criminal and civil prosecutions. An entire country decided not to implement smart metering within their borders for this reason among others.

The presenter gave details of the cryptographic protocol. There are two protocols, he said, the first one a very general protocol that can compute arbitrary functions. [There was no reference to Gentry's fully homomorphic encryption, but it seems applicable, especially since the two operations that a billing system needs are addition and multiplication for accumulation of usage data and the application of tariffs, respectively.] The second protocol is less general but very simple; its definition fits on a single slide. The researchers provide commitments to meter readings, cryptographically signed; they can do additions and multiplications on the values before sending the readings on to the providers. The proof code is very simple, but how do we know it is trustworthy? Three implementations are provided, one in C, one in C# that runs in a web browser, and a provable C# implementation that has been verified not to leak information to the provider.

In conclusion, other than electricity billing and congestion pricing there are many other applications of this technology. Forecasting, demand response, verification of demand response to preclude cheating, and measuring the efficiency of distribution networks are all immediate applications in the electric power industry. The project provides functions to compute all of these and to leak them without actually revealing the usage data. 'We are not crazy about privacy,' said the presenter. If there is a need to leak information, they do. They are flexible about the tariff policies they can apply: non-linear functions, taxation, even very weird ones in the case of usage-based pricing. But the meters are always trivially simple; this is future-proofing. A well-known principle of web security is that the programmer should never trust anything that comes from the user. The present situation

```
is a different paradigm.  If you can trust the certification offered
by the meter, then you can do arbitrarily complex calculations on the
client device owned by the user and still be able to trust the results.
[Personally, I would like to see a more detailed justification
of that last statement.]  The UK is unusual in that there are four
different entities in the chain responsible for electricity generation,
distribution, supply and delivery.  [I lost the audio connection again
at this point, but the seminar was almost over.]

Meeting with Dr Martin scheduled for early next week.

My current tasks, in priority order, are:

1. Coding raw narrative in ATLAS.ti and making links.  See above for
progress metrics.

2. Reading more papers on grounded theory.

3. Continue building the RDP hermeneutic unit in the tool; begin a new
one for RZ [RP is waiting].

4. Re-do the way that event traces are coded.  The way it is now failed
to work (no way to associate an event with previous event).

5. [Waiting] Figure out if I can use the chronological record in my lab
notebook as a source for the 'memo writing' activity that occurs
later [not done yet]

6. Plot tasks on a new Google Calendar as blocks in a 168-hour week.
Establish limits on non-thesis work times.  [This is not done yet
due to more requests for reports and clarifications from the funding
sponsor.]

7. Outline the survey journal article that the assessors asked for.
I still need to find some good examples of how to write a survey
article [not started yet]

8. Outline new Chapter 1, revised Chapter 3, and completely new Chapter 4
[I want to have these done for my trip to Oxford in March.]

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

End of WAR 0174.

References