File 20100319.0358: Weekly activity report 0128:

```
weekly activity report 128 (loughry)
Joe Loughry
Sent: 19 March 2010 03:58
To: Niki.Trigoni@comlab.ox.ac.uk; Andrew Martin; Joanna Ashbourn
Cc: otaschner@aol.com; andrea@hpwtdogmom.org; chip.w.auten@lmco.com; Joe Loughry;
mmcauliffesl@comcast.net
Attachments:
Weekly activity report no. 20100318.2001 (GMT-6) sequence no. 0128, week 8+1 HT

I presented at Reading Group this week the paper 'Computing Arbitrary
Functions of Encrypted Data' by Craig Gentry (Comm. ACM 53(3),
pp. 97--105, 2010).  I talked for an hour about crypto maths, history,
and what a neat result this is.  I think it will be as important
in future as RSA.  Remember that when Public Key Encryption (PKE)
was first proposed by Diffie and Merkle around 1974, the knapsack
problem was a toy example---it was completely infeasible to implement,
a mathematical curiosity; no one could see a way to implement it as
a practical cryptosystem.  RSA was a speed hack that came along later
(1978) and made PKE practical to implement for the first time.

Back then, people looked aghast at RSA because it was so computationally
inefficient compared to alternatives like DES.  Even today, limited
performance processors like smart card CPUs and the TPM (which uses
similar technology) are restricted to relatively small RSA keys of
2048 bits (this information due to John Lyle).  I think that Craig
Gentry's Fully Homomorphic Encryption (FHE), in either the ideal lattice
implementation or the (just barely practical) integer implementation
given in van Dijk et al. (2009) will be the Diffie--Merkle ancestors of
practical FHE some day.

I began by describing the homomorphicity of RSA under multiplication and
my experiments with OpenSSL's implementation of RSA encryption.  I never
did get it to work under OpenSSL, I think because of padding, for which
there are several options in the openssl rsautl library.  I can make the
homomorphicity under multiplication appear by doing the operations by
hand in Scheme, but it doesn't work in OpenSSL with a 1024-bit RSA key.
I think it's either a byte ordering mismatch, a bit ordering mismatch,
or a padding problem.  I don't have enough time to play with it right
now to figure it out, but the experiments were fun and I was able to
report them to Reading Group.

Several practical encryption functions are known to be homomorphic under
either multiplication or addition (but not both); this doesn't cause
problems in practice because they deliberately aren't used in modes
of operation that would leak information that way.  RSA is homomorphic
under multiplication---this is obvious from the equation---but none are
homomorphic under both multiplication and addition at the same time.
If it were, you could do Boolean algebra with it, and that's the key
to computing arbitrary functions on encrypted data.  Given addition and
multiplication, that gives you OR and AND on binary numbers, and that lets
you construct any Turing machine just like wiring together a combinatoric
circuit using electronic gates.  The insight that it's possible to compute
arbitrary functions on encrypted data given homomorphicity under both
addition and multiplication is not new---Rivest, Adelman, and Dertouzos
published about it in 1978---but Gentry's contribution is an encryption
scheme that is homomorphic enough to implement itself using functions
that it can compute before the encrypted information is overwhelmed by
the accumulation of noise bits $r$ at each step.  The decryption key for
step $n+1$ is embedded inside a doubly-encrypted message at step $n$,
```

bootstrapping itself along. The encryption function is probabilistic and so lossy that it allows only a handful of operations to be done at each step. The breakthrough is a series of iterative refinements that keep improving the process to a point where it just barely works.

A digression followed about known weaknesses in CBC and ECB modes of operation and how in practice large prime numbers are chosen for use in PKE. FHE doesn't use prime numbers, only 'large' integers, and the size of them is not really disclosed yet.

The author develops a series of 'somewhat homomorphic' crypto functions ε, $\varepsilon^\dagger$, $\varepsilon^\star$, showing at each step that there is a slight flaw preventing going on to the next step (either because it's too inefficient or because the noise bits r pile up too fast). Then he fixes that flaw and carries on. In the penultimate incarnation he simplifies the integer arithmetic to a single XOR of the least significant bits of the cyphertext and key, putting him within reaching distance of implementing ε within the set of functions it can itself execute. The last little bit he needs is to add a tiny amount of 'grease' by deliberately leaking a small amount of keying material to the following stage of the bootstrap sequence. That is not unprecedented in the crypto literature: 'server assisted cryptography' is well known and useful in situations where the sender is severely resource-constrained and needs to offload some of the work to the receiving end. It's reminiscent of Shannon's 1948 paper on the mathematical theory of communication and the idea of forward error correction. It doesn't necessarily make the cryptosystem less secure, although that is certainly something to watch out for.

There is a neat trick in the paper on page 102 where the algorithm decrypts the inner payload of a doubly encrypted package using key pk_1 without ever knowing pk_2. I described how this is related to the Encrypt-Decrypt-Encrypt (EDE) operation in 3DES, which exists for two reasons: firstly, because it reduces to single DES if the same key is used for all three operations, but secondly because many crypto systems are known to be weakened by multiple applications. Jun mentioned that RSA is one of the algorithms known to be vulnerable to the multiple encryption effect, which I did not know. It's believable, though, from looking at the way the RSA algorithm works.

Jun asked about the key size parameter λ and the rate at which the number of noise bits r grows at each encryption step. I replied that the answer doesn't appear in this paper (Gentry 2010) but a proof is given in van Dijk et al. (2009) that r remains less than the threshold of probability p so that the encryption scheme ε is guaranteed to be information preserving (van Dijk et al.). It really is necessary to read all three papers (Gentry (2010), van Dijk et al. (2009), and Gentry (2009)) to get the whole concept. Watch for the author's next paper. His pattern seems to be that each new paper is more understandable than the last. Gentry (2010) explains the core concept well but there are still a few points left unexplained or difficult to understand. I think his next paper will explain the circular security property and semantic security concepts better.

There must be a fourth paper, because it's clear from Gentry (2010) and van Dijk et al. (2009) that there is an implementation of this algorithm running now, which means that Gentry knows the size of λ. The value of the key size integer that works in practice is not revealed in any of the three papers, and I suspect that some other researcher will soon explore the parameter space and figure out what the range of allowable values is. The author clearly knows what the value is,

but he hasn't told us yet.

The analogy that Gentry uses of glove boxes that can be stuffed inside other glove boxes but can only be used for one minute before the gloves go stiff is both good and useful. This analogy, which is missing from the other papers, is used skilfully throughout to tie the concept together. I thought the analogy was extraneous on first reading, but later I realised that the physical analogy is important. I didn't understand the details of Gentry's 2009 paper on FHE using ideal lattices when it first came out---that's the reason I didn't propose it for reading group at the time although I was excited by the possibilities---but I do understand the new integer scheme in van Dijk et al. (2009) by reference to the explanation of (a slightly different form of) the algorithm in Gentry (2010).

The paper does not claim that this cryptosystem is especially secure, only that it works in concept. I think it will be developed into a secure system by other people before long. Cryptographers will have a lot of fun attacking this one, because there are several apparent chinks in the armour that a good archer with a longbow could put an arrow through: the semantic security, the 'grease', the circular security property, and the approximate gcd problem all come to mind.

Finally, the author discusses efficiency in the last section of the paper. Before I read this paper, I believed other reports out there that FHE was hopelessly inefficient, but now I think it's quite implementable. The author says that the efficiency is proportional to λ^6, which certainly looks bad until he shows that RSA is also proportional to λ^6 when you take into account all the operations required. So the efficiency is about the same as RSA. In other words, not very good, but acceptable in practice.

I look forward to his next paper when we can learn what λ is.

Afterwards, I sent an email to Dr Martin notifying him that the VALID 2010 organisers have extended their deadline, which is good because I am not finished with my paper. I spent too much time playing with homomorphism in RSA and reading four papers on FHE when I should have been reading just one. It was a challenging paper; I shouldn't have spent so much time on it because it put me behind schedule. But it was so interesting I got caught up in it.

Back to work on the paper for VALID 2010, now due 30th March. I will send it to Dr Martin to look over as soon as possible. Tomorrow, I hope.

I am preparing a talk to be given at work about my research and preliminary results. In accordance with Dr Martin's advice I will make this talk different from previous talks and also correspond to a chapter or section of a chapter out of my dissertation outline. Which reminds me, I need to look over the dissertation outline again and begin filling it in with the next level of detail after all the changes I have made in the past months.

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

1. VALID 2010 paper (based on preliminary results from first case study) due 30th March 2. Update dissertation outline 3. Methodology chapter to be finished by 10th April. 4. Crosstalk journal paper (based on methodology chapter and my talk in Oxford) to be submitted by 15th April. 5. Begin writing progress report for confirmation of status. 6. CT&E practitioner survey (I need to have data and preliminary interpretation

```
by end of summer).  7. Update the schedule.  8. Apply for confirmation
of status---I want to send in the application in June.

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

End of WAR 0128.

References