Searching for the Optimum Correlation Attack

Ross Anderson

Computer Laboratory, Pembroke Street, Cambridge CB2 3QG
Email: rja14@cl.cam.ac.uk

Abstract. We present some new ideas on attacking stream ciphers based on regularly clocked shift registers. The nonlinear filter functions used in such systems may leak information if they interact with shifted copies of themselves, and this gives us a systematic way to search for correlations between a keystream and the underlying shift register sequence.

1 Introduction

A number of cipher systems use a nonlinear filter generator to expand a short key into a long keystream. This generator is based on a linear feedback shift register, some of whose state bits are filtered through a nonlinear function to provide the keystream (figure 1):

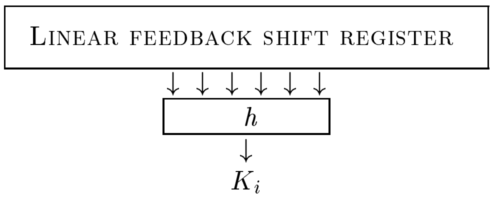

Figure 1 - the nonlinear filter generator

Typical systems have shift registers of between 61 and 127 bits in length, and nonlinear filter functions of varying complexity [MFB] [KBS] [CSh]. Some variants use several functions simultaneously to generate a number of keystream bits in parallel [M1].

The conventional attack on the filter generator [S2] [S3] [MS1] proceeds in two stages. Firstly, we find a function of the keystream which is correlated with the underlying shift register sequence; it can be shown that such a function always exists [XM], even if the combining function possesses memory [G]. The keystream is viewed as a noisy version of the shift register sequence, and is reconstructed by various techniques (figure 2):

$$\boxed{\text{LFSR}} \longrightarrow \boxed{\text{Add noise}} \longrightarrow K_i$$

Figure 2 - the standard model

This 'standard model', which has been the focus of most of the published work on the subject, was first proposed by Siegenthaler [S4]. His original attack involved an exhaustive search through all phases of the shift register to find the highest correlation [S3]; Meier and Staffelbach later showed that iterative reconstruction techniques were much faster, and especially so if the shift register's connection polynomial $f(x)$ is of low weight [MS1], while Mihaljević and Golić proved conditions under which these fast correlation attacks converge [MG].

Where $f(x)$ is not sparse, one can look for a decimation of the sequence whose polynomial is sparse [A1], or more generally a sparse multiple of f (i.e., a low weight parity check) [CS]. Meier and Staffelbach pointed out that low weight checks can be found by meet-in-the-middle techniques [MS1]; and if $f(x)$ has degree n, this will take about $\frac{n}{2}2^{(\frac{n}{2})}$ operations. Recent work by van Oorschot and Wiener has shown that it is feasible to construct special-purpose collision search hardware for n up to 128 or so [VW]. Thus the security of the nonlinear filter generators under consideration boils down to finding good correlations between the keystream and the underlying shift register.

However, the problem of finding an actual correlation tends to have been dismissed with an existence proof. Our principle that robust security depends on explicitness [A2] made us suspicious of this, and inspired us to look for a construction. We would ideally like to have an algorithm which will find the maximum correlation which an attacker can obtain; we can then use this together with the convergence bounds found in [MG] and [M2] to establish whether a given design is vulnerable to a fast correlation attack. Our model is therefore

$$\boxed{\text{Random Noise}} \longrightarrow \boxed{h(x)} \longrightarrow K_i$$

Figure 3 Our model

This is essentially the dual of the problem studied in the standard model; the goal is to find out how much information about an arbitrary signal leaks through $h(x)$ to K_i. Its solution depends on the nonlinear structure of h: if $K_i = S_i + S_{i+1}$, then knowing K_i tells us nothing about S_i, while if $K_i = S_i S_{i+1}$, then whenever $K_i = 1$ we know that $S_i = 1$ too. Note that when attacking a filter generator we can always discount a linear function by moving to a different phase of the underlying shift register, so we are really interested in finding the maximum leakage of $h(l(x))$ over all linear functions $l(x)$.

2 Finding a Correlation

Let us take a concrete example. Suppose that the nonlinear combining function h is given by

$$h(x_1, x_2, x_3, x_4, x_5) = x_1 + x_2 + (x_1 + x_3)(x_2 + x_4 + x_5) + (x_1 + x_4)(x_2 + x_3)x_5$$

This function is used as a primitive in [K] and appears to have been used in other designs too [KBS]; it is distinguished by being as small a function as one can get which is both balanced and correlation immune of degree two.

As already noted, the standard attack on such a cipher would be to look for linear functions of the keystream and of the underlying shift register sequence which are correlated; a variant is the 'best affine approximation' attack of Ding, Xiao and Shan [DXS]. However, both these attacks throw away a lot of information about the nonlinear structure of h, and our goal is to try and identify - and if possible use - all the information which h leaks.

If $K_i = f(S_{i-2}, S_{i-1}, S_i, S_{i+1}, S_{i+2})$, the keystream bits $K_{i-2}, \cdots, K_{i+2}$ all depend on S_i, and if we want to approximate S_i we need a function of the form $S_i = g(K_{i-2}, K_{i-1}, K_i, K_{i+1}, K_{i+2})$. However, the bits $K_{i-2}, \cdots, K_{i+2}$ depend on the nine bits $S_{i-4}, \cdots, S_{i+4}$; and so it is natural to look at the set of input 9-tuples from the shift register sequence which give rise to each possible 5-bit keystream output. We will call the 9 bit to 5 bit function the *augmented function* of h and write it as $\overline{h}$.

When we count the outputs of $\overline{h}$ for each of the 512 possible inputs, we find:

output	# inputs	output	# inputs	output	# inputs	output	# inputs
0	18	8	11	16	16	24	23
1	16	9	17	17	18	25	17
2	14	10	12	18	16	26	10
3	20	11	12	19	18	27	18
4	16	12	23	20	12	28	17
5	14	13	13	21	10	29	15
6	21	14	13	22	15	30	19
7	17	15	19	23	15	31	17

We might have hoped that a good system would have each output generated by sixteen inputs, but the actual table is irregular, and this may give us a way in. For example, there are two outputs (21 and 26) generated by only ten inputs, and if we look at the inputs which generate 26, we find that they are:

```
001010101
001110001
001110010
100110001
100110010
101001011
101110001
101110010
110110001
110110010
```

Casting our eye down these columns, we see that there is only a single zero in the fifth column, and a single one in the sixth and seventh. In other words, if $K_i, ..., K_{i+4} = 11010$, then $S_{i+2} = 1$, $S_{i+3} = 0$ and $S_{i+4} = 0$ with probability 0.9 in each case. The other columns give us correlations of 0.7, 0.8, and so on. Now the fact that the correlation between a nonlinearly filtered sequence and the underlying shift register is uneven was first pointed out by Forré [F], but she did not investigate the matter further. At last we have explained this irregularity — it is simply a matter of counting the input/output stability of the augmented function $\overline{h}$.

Of course, we do not just get information from those inputs which give rise to the least common outputs. For example, when we consider the 17 inputs which generate the output 9, a delightful discovery awaits us: these 17 inputs are all zero in the fourth bit. So whenever we see that $K_i, ..., K_{i+4} = 01001$, we know that $S_{i+1} = 0$.

Now, one might at first think that this yields an optimal correlation attack. After all, if h is an m-bit to 1-bit function, then each shift register bit will contribute to precisely m keystream bits, and we might expect that all the correlation information could be found by examining the augmented function which generates these bits. If we are lucky, we will find correlations of one and be able to solve the cipher outright using linear algebra; otherwise, we should still get lots of correlations of the order of 0.8 or 0.9, with which a probabilistic reconstruction becomes fairly straightforward [M2].

However, we can get more than just correlations between the K_i and the S_i. On looking more closely at the above table, we notice that columns 5 and 6 are inverses of each other. Thus whenever $K_i, ..., K_{i+4} = 11010$, then $S_{i+2} = 1 + S_{i+3}$. Similar relations are to be found in the other output sets; for example, if K_i through K_{i+4} are all equal to zero, then $S_{i+3} = S_{i+4} = S_{i+5}$. Every such equation halves the keyspace which we still have to search.

How well does our technique work against other nonlinear combiners? In recent years, a lot of attention has been paid to bent functions [MS2]. We looked at a typical bent function, $h(x_1, ..., x_6) = x_1 + x_2 + x_3 + x_1x_4 + x_2x_5 + x_3x_6$, and found it to be significantly worse than the correlation-immune function discussed above. In fact, its augmented function never attains nine possible output values (12, 15, 30, 31, 47, 60, 61, 62, 63), while the zero output is attained 100 times. Its information leakage is heavy; for example, the output value 17 is attained 12 times, with all twelve inputs having $S_1 = S_2 = 1$ and $S_4 = S_5 = S_{11} = 0$.

The use of almost bent functions has also been suggested; these are bent functions which have been made balanced by changing a few output values. When we changed the above bent function's output from 0 to 1 for input values of 7, 11, 23, and 27, we found that the columnar behaviour was somewhat less marked, but seven of the nine missing output values were now attained by small numbers of inputs (and the output 31 was only generated by the single input 00010111011). We concluded that attacking a filter generator using a bent or almost bent function would be easy.

The device reported in [M1] and [MFB] uses as its filter what we might call a De Bruijn function — a function on k input bits whose value at the input n is the n-th bit of a 2^k bit De Bruijn sequence. This filter has the interesting property that its augmented function is balanced, with every possible output attained exactly 2^k times; it also has the property that for each output, the last k bits of each of the inputs are identical. For example, with one of the De Bruijn functions implemented in [MFB], the output 21 is attained for 32 inputs of the form ******10110. In other words, the information leakage is total.

One also notices from looking at a few candidates that most randomly chosen balanced filter functions appear to leak rather badly. This suggests that just as it is a bad idea to replace the DES S-boxes with random ones [BS], it is also a bad idea to use randomly chosen filters; and of course a knowledgeable designer can easily place a trapdoor in the filter.

Many further implications remain to be worked out. For example, even if h has too many inputs for exhaustive analysis, it may still have some structure which we can use. It might have a tractable mathematical definition as in [CSh]; if it has a regular tree structure which is key dependent [K] [A3], then these keybits might be deduced by observing which patterns are most common in decimations of the keystream; where an unknown permutation is introduced in the tree structure, the ideas of [MDO] may be useful; and even where none of these tricks can be used, statistical sampling of the augmented function may still give information to the opponent.

Some systems use a number of nonlinear filters in parallel to generate more than one bit of keystream at a time [M1]; but these functions could well interact in a way which facilitates an attack. In fact, in a recent NSA patent on a device for generating simultaneous keystreams, the underlying generator is run at high speed to ensure that the keystreams are linearly independent [S1].

Another implication is that when doing a correlation attack, the 'lumpiness' of the correlation will mean that we have little information about some bits in the shift register sequence, but will know others with high probability. Existing shift register reconstruction algorithms will no longer work, as the problem is now that that of finding S such that given m bits Z and an n by m matrix $\mathbf{A}$, $|\mathbf{A}S - Z|$ is minimised. This problem is tackled in another paper in this volume [M2], which was motivated in part by this work and which extends the techniques pioneered by Meier and Staffelbach.

Clearly, when designing keystream generators, we must write down all the keystream bits K_i to which an arbitrary shift register bit S_i contributes, and then all the equations whereby these bits in turn are generated. We should then consider all the occurrences of S_i in these equations and check for information leakage. On considering a few examples, it appears that using either multiple shift registers or multiple filter functions makes the security harder to evaluate.

3 Conclusions

We have shown a practical method for searching for the best possible local correlations nonlinearly filtered shift register sequences. The key is to look at how the filter function reacts with shifted copies of itself. Many functions react badly — including both bent functions and De Bruijn functions — and if a number of different filter functions are used simultaneously, then their interactions must also be taken into account.

On the theoretical side, we have given students of Boolean functions new problems to investigate, namely what functions have low leakage (defined as the maximum imbalance in any input variable for any given output of the augmented function), and, in general, what properties a set of functions must have in order not to interact in a harmful way.

References

[A1] RJ Anderson, "Derived Sequence Attacks on Stream Ciphers", presented at the rump session of Crypto 93

[A2] RJ Anderson, "Why Cryptosystems Fail", in *Communications of the ACM* v **37** no 11 (November 1994) pp 32–40

[A3] RJ Anderson, "Tree Functions and Cipher Systems", in *Cryptologia* v **XV** no 3 (July 1991) pp 194–202

[BS] E Biham, A Shamir, *'Differential Cryptanalysis of the Data Encryption Standard'*, Springer 1993

[CS] V Chepyzhov, B Smeets, " On a Fast Correlation Attack on Certain Stream Ciphers", in *Advances in Cryptology — Eurocrypt 91*, Springer LNCS v **547** pp 176–185

[CSh] TR Cain, AT Sherman, "How to Break Gifford's Cipher", in *Proceedings of the 2nd ACM Conference on Computer and Communications Security* (ACM, Nov 94) pp 198–209

[DXS] C Ding, G Xiao, W Shan, *'The Stability Theory of Stream Ciphers'*, Springer LNCS v **561** (1991)

[F] R Forré, "A Fast Correlation Attack on Nonlinearly Feedforward Filtered Shift-register Sequences", in *Advances in Cryptology — Eurocrypt 89*, Springer LNCS v **434**, pp 586–562

[G] JD Golić, "Correlation via Linear Sequential Circuit Approximation of Combiners with Memory", in *Advances in Cryptology — Eurocrypt 92*, Springer LNCS v **658**, pp 113–123

[K] GJ Kühn, "Algorithms for Self-Synchronising Ciphers", in *Proc COMSIG 88*

[KBS] G Kühn, F Bruwer, W Smit, " 'n Vinnige Veeldoelige Enkripsievlokkie", supplementary paper to *Proceedings of Infosec 1990*

[M1] G Mayhew, "A Low Cost, High Speed Encryption System and Method", in *Proc 1994 IEEE Computer Society Symposium on Research in Security and Privacy* (IEEE, 1994) pp 147–154

[M2] DJC MacKay, "A Free Energy Minimisation Framework for Inferring the State of a Shift Register given the Noisy Output Sequence" *in this volume*

[MDO] W Millan, EP Dawson, LJ O'Connor, "Fast Attacks on Tree-structured Ciphers", in *Proceedings of Workshop in Selected Areas in Cryptography* (Queen's University, 1994) pp 146–158

[MFB] G Mayhew, R Frazee, M Bianco, "The Kinetic Protection Device", in *Proceedings of the 15th National Computer Security Conference* (NIST, 1992) pp 310–318

[MG] MJ Mihaljević, JD Golić, "Convergence of a Bayesian Iterative Error-correction Procedure on a Noisy Shift Register Sequence", in *Advances in Cryptology — Eurocrypt 92*, Springer LNCS v **658**, pp 124–137

[MS1] W Meier, O Staffelbach, "Fast correlation attacks on certain stream ciphers", in *Journal of Cryptology* v **1** 1989 pp 159–176

[MS2] W Meier, O Staffelbach, "Nonlinearity criteria for cryptographic functions", in *Advances in Cryptology — Eurocrypt 89*, Springer LNCS v **434** pp 549–562

[S1] B Snow, *'Multiple Independent Binary Bit Stream Generator'*, US Patent 5,237,615 (17 August 1993)

[S2] T Siegenthaler, " Correlation Immunity of Nonlinear Combining Functions for Cryptographic Applications", in *IEEE Transactions on Information Theory* **IT-30** no 5 (Sep 1984) pp 776–780

[S3] T Siegenthaler, "Decrypting a Class of Stream Ciphers Using Ciphertext Only", in *IEEE Transactions on Computers* **C-34** no 1 (Jan 1985) pp 81–85

[S4] T Siegenthaler, "Cryptanalysts' Representation of Nonlinearly Filtered m-Sequences", in *Advances in Cryptology — Eurocrypt 85*, Springer LNCS v **219** pp 103–110

[VW] PC van Oorschot, MJ Wiener, "Parallel Collision Search with Application to Hash Functions and Discrete Logarithms", in *Proceedings of the 2nd ACM Conference on Computer and Communications Security* (ACM, Nov 94) pp 210–218

[XM] GZ Xiao, JL Massey, "A spectral characterisation of correlation-immune combining functions", in *IEEE Transactions on Information Theory* **IT-34** (May 1988) pp 569–571

This article was processed using the LaTeX macro package with LLNCS style