TCHo: a hardware-oriented trapdoor cipher

Jean-Philippe Aumasson
Matthieu Finiasz, Willi Meier, Serge Vaudenay

University of Applied Sciences Northwestern Switzerland
School of Engineering

ENSTA

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

ASYMMETRIC ENCRYPTION

There are "security proofs" for public-key encryption: reductions to integer factorization, discrete log, lattice problems, etc.

But. . .

1) on quantum computers, RSA, ECC, ElGamal, etc. are broken
2) on hardware, slow and difficult to implement

On the other hand, LFSR-based stream ciphers fit well lightweight environments.

TCHo

- encrypts with only a LFSR and pseudorandom bits
- decrypts with simple linear algebra over GF(2)
- is semantically secure
- is not known to be harmed by quantum computers

TCHo AND RSA

Public key:

- **TCHo**: irreducible polynomial P
- RSA: composite integer $n = pq$

Private key:

- **TCHo**: a sparse multiple of P
- RSA: the prime factors of n

Hard problem:

- **TCHo**: finding a sparse multiple (polynomial)
- RSA: finding a prime factor (integer)

Encryption:

- **TCHo**: probabilistic
- RSA: deterministic

DESCRIPTION OF **TCHo**

ENCRYPTION

10101001 . . . 10101001	repetition of $m\|\|m\|\| \ldots \|\|m$
$\oplus$	
01110110 . . . 01101110	output of a LFSR with random state
$\oplus$	
00100100 . . . 00100010	random bits with bias $\gamma = \Pr(0) - \Pr(1)$

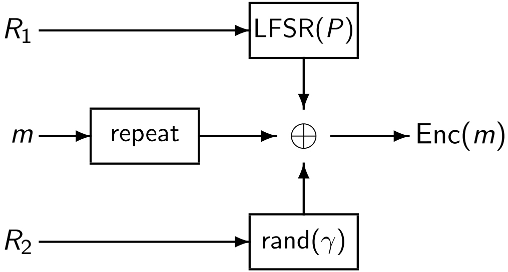
such that

- LFSR feedback polynomial is the public key P
- $\gamma > 0$ (more zeros than ones)
- the ciphertext is a ℓ-bit string, with $\ell \geq \deg(K)$

$$\mathrm{Enc}(m) = m|| \ldots ||m \oplus \mathrm{LFSR}(P) \oplus \mathrm{rand}(\gamma)$$

ENCRYPTION

Implementation is built on three independent components, fed with two random (unbiased) samples R_1 and R_2
$\Rightarrow$ parallelizable



DECRYPTION

$$\begin{array}{ll} K & \text{private key, sparse multiple of } P \\ \otimes & \\ 10011011\ldots10101011 & c = m||\ldots||m \oplus \mathrm{LFSR}(P) \oplus \mathrm{rand}(\gamma) \\ = 0100\ldots1101 & m'||\ldots||m' \oplus \mathrm{rand}(\gamma^{\mathrm{w}(K)}) \end{array}$$

$\Rightarrow$ can compute m' (count majority), and recover m:

$$m \leftarrow \psi(m')$$

ψ is a linear mapping defined by K

PRODUCT POLYNOMIAL$\otimes$BITSTRING

Let $K = \sum k_i x^i$, and a bitstring $u = (u_0, \ldots, u_{\ell-1})$, then $K \otimes u = v$, with v of $\ell - \deg(K)$ bits, and

$$v_i = u_i k_0 + \cdots + u_{i+\deg(K)} k_{\deg(K)}$$

$\approx$ sequence of dot products

Properties exploited in decryption (recall $K = P \times P'$)

- $K \otimes \Big(\text{output of LFSR with feedback } P\Big) = 0 \ldots 0$
- $K \otimes \Big(\text{output} \ldots \oplus \text{rand}(\gamma)\Big) \approx \text{rand}(\gamma^{\text{w}(K)})$

DECRYPTION

K	private key, sparse multiple of P
$\otimes$	
$10011011\ldots10101011$	$c = m\|\|\ldots\|\|m \oplus \mathrm{LFSR}(P) \oplus \mathrm{rand}(\gamma)$
$= 0100\ldots1101$	$m'\|\|\ldots\|\|m' \oplus \mathrm{rand}(\gamma^{\mathrm{w}(K)})$

$\Rightarrow$ can compute m' (count majority), and recover m:

$$m \leftarrow \psi(m')$$

ψ is a linear mapping defined by K

DECRYPTION RELIABILITY

$\psi(m)$ repeated

$$N = \frac{\ell - \deg(K)}{|m|} \text{ times}$$

Decrypt incorrectly $\Leftrightarrow$ majority logic fails $\Leftrightarrow$ at least one bit of $\psi(m)$ is noised more than half the times.

$$\Pr[\text{bad decryption}] \approx |m| \cdot \varphi\Big(- \sqrt{\frac{N\gamma^{2w}}{1-\gamma^{2w}}}\Big)$$

with φ the cumulative distribution of $\mathcal{N}(0,1)$.

KEY GENERATION

Problem: find a pair (K, P), with K a sparse multiple of P, of given degree and weight, and P of degree in $[d_{\mathrm{min}}, d_{\mathrm{max}}]$.

Until a suitable P is found, repeat

- pick a random K of given degree and weight
- factorize it
- look for an irreducible P of suitable degree in K's factors

(in practice large degrees: $\deg(K) > 15\,000, \deg(P) > 5\,000$)

EXAMPLE OF PARAMETERS

For 80-bit security,

- plaintext of $|m| = 128$ bits
- ciphertext of $\ell = 56\,000$ bits
- public-key is polynomial P of degree $\in [7\,150, 8\,000]$
- private-key is polynomial K of degree $24\,500$ and weight 51
- noise has bias 0.98
- decryption fails with probability 2^{-23}

SECURITY OF **TCHo**

PRIVATE KEY RECOVERY

We can decrypt

- if we recover K, sparse multiple of the polynomial P, OR
- if we find another sparse multiple of degree $\leq \deg(K)$

Computational problem **LWPM**

- Parameters: w, d, d_P, $0 < d_P < d$ and $w \ll d$.
- Instance: P of degree d_P
- Question: find a multiple of P of degree $\leq d$ and weight $\leq w$.

Strategies: exhaustive search, generalized birthday paradox, syndrome decoding.

In **TCHo**, the existence of a solution is guaranteed !

PRIVATE KEY RECOVERY

Computational problem **LWPM**

- Parameters: w, d, d_P, $0 < d_P < d$ and $w \ll d$.
- Instance: P of degree d_P
- Question: find a multiple K of P, s.t. $\deg(K) \leq d$ AND $\mathrm{w}(K) \leq w$.

Strategies: exhaustive search, generalized birthday paradox, syndrome decoding.

In **TCHo**, the existence of a solution is guaranteed !

LWPM requires $\Omega(2^{\lambda})$ operations if

$$\binom{d}{w-1} \leq 2^{d_P} \quad \text{and} \quad w \log \frac{d}{d_P} \geq \lambda$$

BASIC SECURITY PROPERTIES

TCHo is trivially malleable,

$$\mathrm{Enc}(m) \oplus \Delta = \mathrm{Enc}(m \oplus \Delta)$$

TCHo can be inverted by a CCA adversary: given challenge ciphertext c, just query for $m \leftarrow \mathrm{Dec}(c \oplus \Delta)$, and recover original message $m \oplus \Delta$.

TCHo can be used as a KEM in hybrid encryption scheme, to provide IND-CCA security.

SEMANTIC SECURITY

Consider the problem of distinguishing

$$c = \mathrm{LFSR}(P) \oplus \mathrm{rand}(\gamma) \oplus (m || \ldots || m)$$

for a chosen m, from

$$\mathrm{rand}(0)$$

(real-or-random game)

challenge XORed with m gives either

$$\mathrm{LFSR}(P) \oplus \mathrm{rand}(\gamma) \text{ OR } \mathrm{rand}(0)$$

Reduction to **Noisy LFSR**: distinguish (ℓ-bit strings)

- $\mathrm{LFSR}(P) \oplus \mathrm{rand}(\beta)$ from
- $\mathrm{rand}(0)$

SEMANTIC SECURITY

Noisy LFSR: distinguish

- LFSR(P) $\oplus$ rand(β) from
- rand(0)

$P \otimes$challenge$=$ either rand($\gamma^{\mathrm{w}(P)}$) or rand(0).

$\Rightarrow$ **Noisy LFSR** solvable if can distinguish rand($\gamma^{\mathrm{w}(P)}$) from rand(0)

If we know P' such that $\mathrm{w}(PP') < \mathrm{w}(P)$,
$(PP')\otimes$challenge$=$ either rand($\gamma^{\mathrm{w}(PP')}$) or rand(0).

$\Rightarrow$ **Noisy LFSR** solvable if can distinguish rand($\gamma^{\mathrm{w}(PP')}$) from rand(0)

but less bits than with P!

SEMANTIC SECURITY

With the previous method, we get a ratio $\frac{\text{advantage}}{\text{complexity}}$

$$\max_{w \in [0, d_P], N \geq 1} \sqrt{\frac{N}{2\pi}} \frac{\gamma^w}{wN + 2^{\deg(P)} \left(\frac{\ell}{d_P}\right)^{w-1} \binom{\ell}{w}^{-1}}$$

with N the number of bits with bias $\gamma^{\mathrm{w}(PP')}$ used,

Theorem
Assuming the hardness of **LWMP** and **Noisy LFSR**,
TCHo is semantically secure.

PERFORMANCES OF **TCHo**

PERFORMANCES

Recall parameters: $|m| = 128$, $|\mathrm{Enc}(m)| = 56\,000$,
$\deg(P) \in [7\,150, 8\,000]$, $\deg(K) = 24\,500$, $\mathrm{w}(K) = 51$, $\gamma = 0.98$.

Average timings with C++ & NTL, gcc 3, over Intel P4 1.5GHz.
NTL used for matrix inversion and polynomial factorization (Cantor-Zassenhaus).
Biased random bits generated in 2 steps: 1) pick weight k w.r.t. γ, 2) pick word of weight k.

Timings:

- Encryption: 90ms (bottleneck = LFSR output computation)
- Decryption: 65ms (bot. = product ciphertext$\otimes K$)
- Key generation: 30min (bot. = factorization)

(timings include precomputation of ψ)

PERFORMANCES

Flexible parameters (trading-off ciphertext length, key gen. time, enc/dec. time, etc.). For example with parameters $|m| = 128$, $|\mathrm{Enc}(m)| = 150\,000$, $\deg(P) \in [6\,000, 8\,795]$, $\deg(K) = 17\,600$, $\mathrm{w}(K) = 81$, $\gamma = 0.9766$.

- Encryption: 228ms
- Decryption: 424ms
- Key generation: 2min20s

These are software timings, **TCHo** is for hardware!

PERFORMANCES

"Why do you give software timings for a hardware cipher??"
$\rightarrow$ did not have the opportunity to implement HW.

Expected much faster on hardware devices, because of

- efficient LFSR
- only GF(2) linear algebra
- parallelization

but key generation. . .

CONCLUSION

SUMMARY

TCHo is. . .

- based on the hardness of recovering a sparse polynomial multiple
- semantically secure
- post-quantum
- flexible
- fast in hardware (except key gen.)

FURTHER WORK

more experiments. . .

- benchmarks on FPGA, ASIC, etc.
- suitable for passive RFID tags ?

more analysis. . .

- speed-up key generation
- replace huge LFSR by. . . ?
- weak instances ?
- solve **LWPM** efficiently ?
- solve **Noisy LFSR** efficiently ?

TCHo: a hardware-oriented trapdoor cipher

Jean-Philippe Aumasson
Matthieu Finiasz, Willi Meier, Serge Vaudenay

University of Applied Sciences Northwestern Switzerland
School of Engineering

ENSTA

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE