Cryptanalyzing the Polynomial Reconstruction based Public-Key System under Optimal Parameter Choice

Aggelos Kiayias - Moti Yung

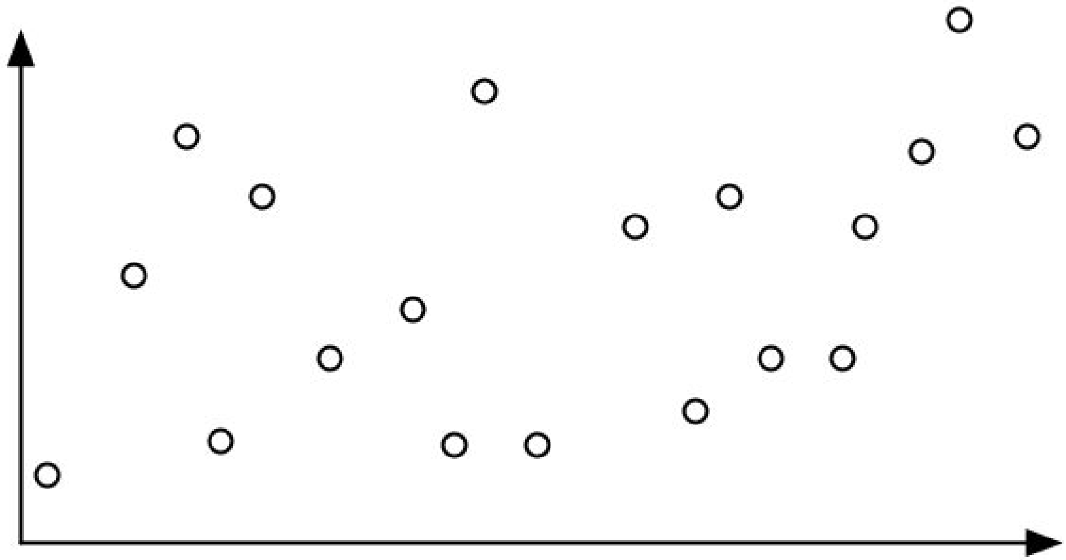
U. of Connecticut - Columbia U.

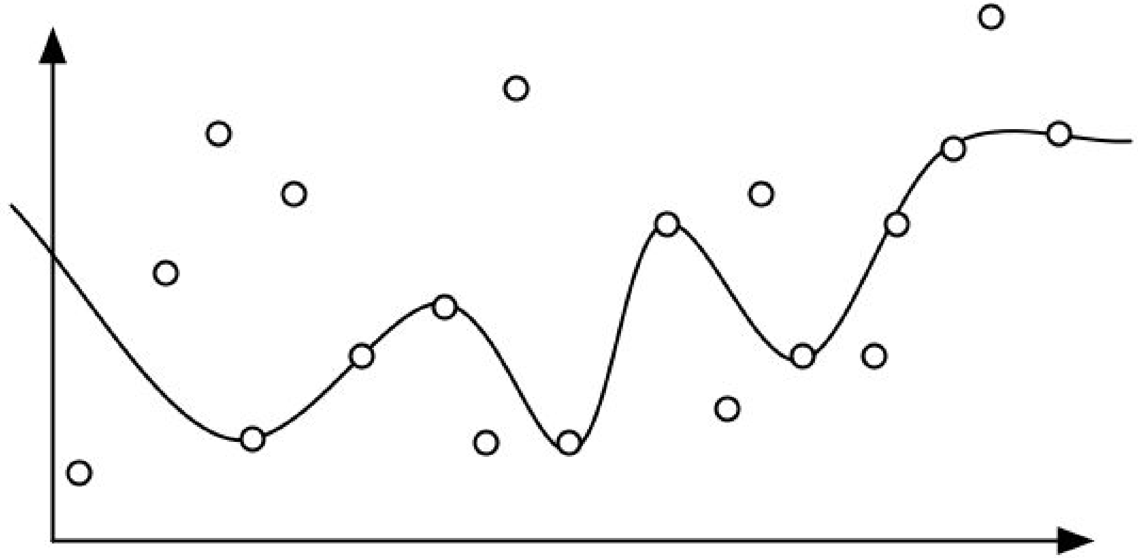
(Public-Key) Cryptography

- intractability assumptions
  - usually:
    - Based on number theoretic problems.
    - factoring, discrete-logarithm.
  - motivation for alternative assumptions:
    - diversity.
    - efficiency.
    - resistance against quantum attacks.
- Hardness of Error-Correcting Code problems?

Polynomial Reconstruction

- Given a set of points over a finite field $\{z_i, y_i\}$ i=1,…, n and parameters n, k, w recover all polynomials p of degree less than k such that $p(z_i)=y_i$ for at least n-w indices of
  {1,…, n}





Polynomial Reconstruction, II

- Reed-Solomon Codes Decoding:

$$w \leq \frac{n-k}{2}$$

- Guruswami-Sudan List-Decoder

$$w < n - \sqrt{n(k-1)}$$

PR-Based Cryptography

- Advocated in [KY02]
- pros:
  - efficiency: matrix arithmetic over binary extension fields.
- cons:
  - difficulty on building primitives.
- KY02: semantically secure <u>symmetric</u> cryptosystems with provable properties, two party secure computation protocols, pseudorandom number generators.

PR-based One-Way Fuction

- cf. Reed Solomon Decoding:
  - message = coefficients of polynomial *p*.
  - encoding = evaluation of polynomial on points $z_1 \ldots z_n$
- hardness:
  - add *W* random errors.
  - must make instance unsolvable by (list-)decoding techniques.

$$W > n - \sqrt{n(k-1)}$$

The AF03 Cryptosystem

- Public-Key: PR-instance $[n, k+1, w]$

$$\{z_i, y_i\} : \exists p,\ \text{monic}\ ,\ \#\{p(z_i) \neq y_i\} \leq w$$

- Secret-key: *p* / *w* error point locations.
- Message-space: F^k
- Encryption: choose *a* random from F

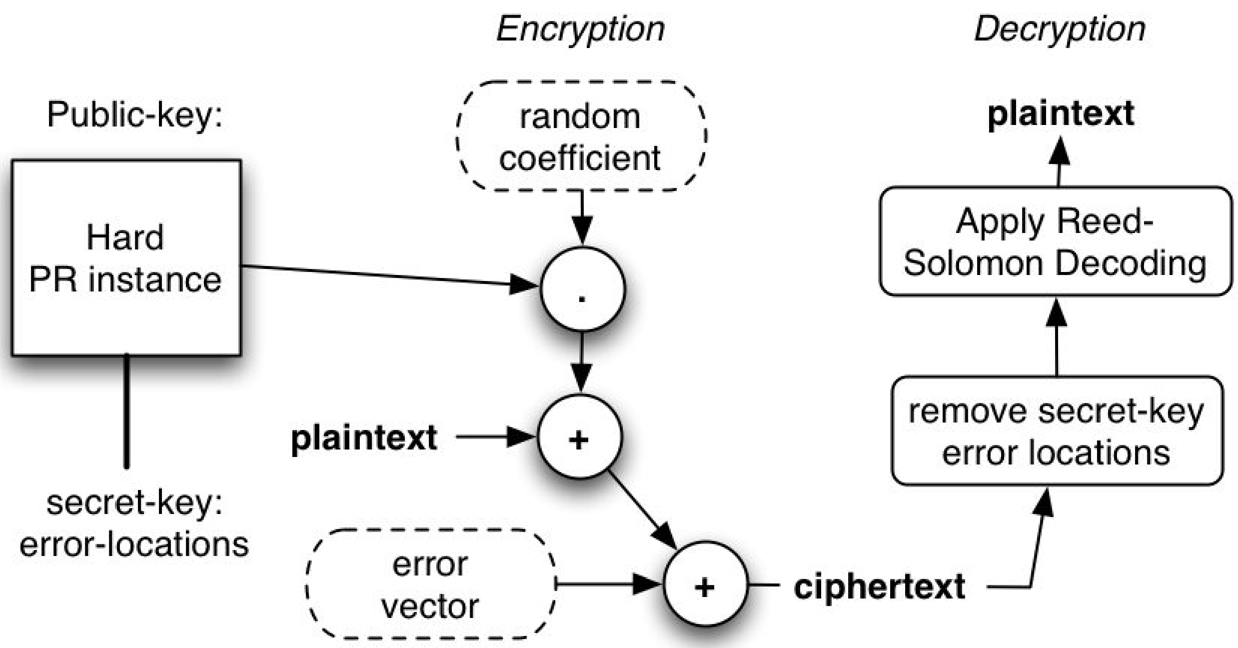
$$y_{i,cipher} = a \cdot y_i + p_{msg}(z_i)$$

- Decryption: remove error points + decode to recover: $a \cdot p(x) + p_{msg}(x)$

AF03 Cryptosystem Break

- Coron PKC04
- ciphertext-only attack against the AF04 cryptosystem for their specific parameter choice.

The AF03 General Approach



Points to optimize:

- Decoding Algorithm employed during Decryption.
- Number of errors introduced during Encryption.

AF03: choose errors applying a worst-case analysis (!) and standard Reed-Solomon Decoding for decryption - as opposed to using the state of the art.

Research Direction.

- Important to understand the power of the general approach.
- Does optimizing the parameters of the AF cryptosystem thwart the attack of Coron?

Recall Decryption Operation:

- Given ciphertext $\{\langle z_i, y_i' \rangle\}_{i=1}^{n}$
- Let *I* be the "good points" of the public-key
  - say *n*-*W*
- Keep only good points: $\{\langle z_i, y_i' \rangle\}_{i \in I}$
- The resulting sequence of points

must be decodable; *w* = errors introduced during encryption

- AF03:

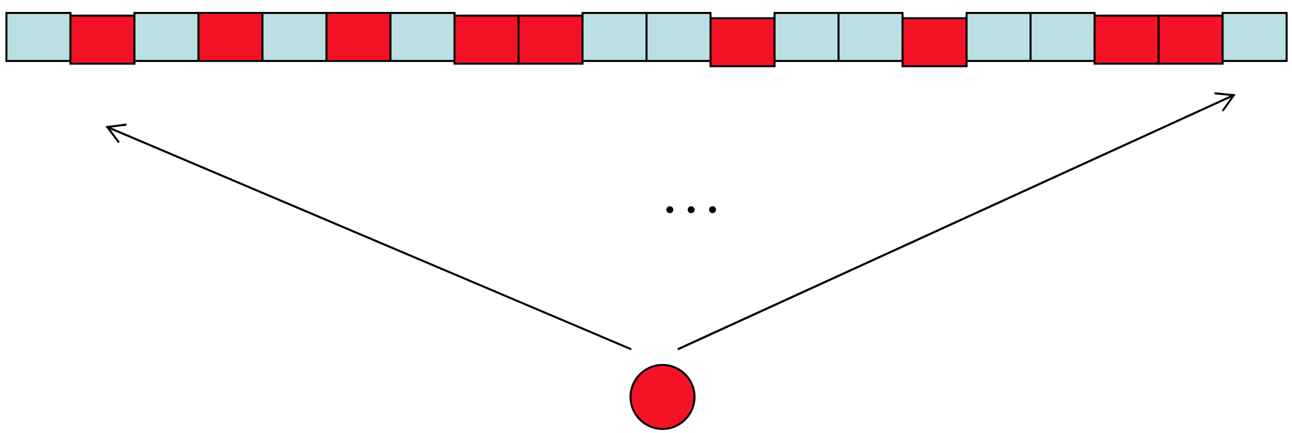
ensures decodability even if all encryption errors corrupt good locations

$$w \leq \frac{n - W - k - 1}{2}$$

overkill!

A simple boost of the Encryption error Parameter

- Recall encryption operation
- Public-key has “good” and “bad” locations.



When sender selects *w* error-locations there is significant probability that he will be corrupting already corrupted locations.

Modeling the probability

- Number of good points in the ciphertext that will be corrupted by the sender follow the
  - Hypergeometric distribution
    - mean = (n-W)/n (ratio of good points)
    - number of trials = w
  - We apply the Chvatal bound for the tails of the hypergeometric (cf. Chernoff).
  - by this simple trick (without changing the original cryptosystem at all) we achieve: **238%** more errors allowed in the encryption function -- in a typical parameter setting
    - n=2000, *k*=100, *W*=1600, *w(AF)=171, w(HERE)=407*
  - Analysis of Coron's attack fails.

Still insecure though

- a probabilistic analysis of Coron’s attack (that we perform) shows that it still manages to break the scheme with high probability.

Allowing even more errors during encryption…

- How?
  - use state of the art in Reed-Solomon *list-decoding*.
  - Is it possible?
    - yes! list-decoding is unambiguous with high probability in the random noise setting.
    - -> decryption will be unambiguous with high probability.

Lemma

- Let $\{\langle z_i, y_i\rangle\}_{i=1}^{n}$ a PR instance.

with parameters $[n, k, e]$

random errors.

- probability of more than one solution is less than

$$\binom{n}{t}^2 / |\mathbf{F}|^{n-e-k}$$

Optimal Variant of AF03

- Optimized parameter *w* (errors during decryption).
  - according to the state of the art of RS decoding.
- Employ list-decoding (unambiguous with high probability) for decryption.
- We achieve improvement **777%** on the selection of the number of encryption errors *w*.
  - n=2500, k=101, W=2063,
  - w[AF03]=167
  - w[HERE] = 1298
  - Beyond bounds for Coron's attack.

The bad news: New Attack!

- Suppose public-key: $\{\langle z_i, y_i \rangle\}_{i=1}^{n}$
- ciphertext: $\{\langle z_i, y_i' \rangle\}_{i=1}^{n}$
- We know that the following is a [n, k, W] PR-instance: $\{\langle z_i, y_i - z_i^k \rangle\}_{i=1}^{n}$
- and:

$$\exists a \in \mathbf{F} : \{\langle z_i, y_i' - a \cdot y_i \rangle\}_{i=1}^{n}$$

is a [n,k,w] PR-instance.

The attack

- Denote: $\hat{y}_i = y'_i - \lambda y_i$ where λ is free
- define the system:

$$\forall i = 1, \ldots, n \quad \sum_{j_1 \geq 0, j_2 \geq 0, j_1 + (k-1) j_2 < l} q_{j_1, j_2} z_1^{j_1} \hat{y}_i^{j_2} = 0$$

- with uknowns q_{j_1, j_2}

- **Lemma.** number of unknowns >= $\dfrac{l(l-1)}{2(k-1)}$

Attack Explanation.

- The matrix of the system is a function of λ
- denote: $A[\lambda]$

- **THEOREM 1.** the matrix $A[a]$ where *a* is the random coefficient selected by the sender is *singular*.

- Singularity implies $\det(A[\lambda])$ is a polynomial whose roots include *a*.
- **unless** $\det(A[\lambda])$ is the zero-polynomial

Attack Explanation, 2

- **THEOREM 2.** The probability **P** that the polynomial is the zero-polynomial satisfies:

$$\mathbf{P} \leq 2s(n-l)/|\mathbf{F}|$$

- where

$$s = \lfloor \frac{l-1}{k-1} \rfloor$$

- bounded away from 1 for all reasonable parameter settings.
- Attack succeeds discovers *a*.

Attack Explanation, 3

- Once *a* is known, recall the condition on the public-key and the ciphertext:

$$\exists a \in \mathbf{F} : \{\langle z_i, y_i' - a \cdot y_i \rangle\}_{i=1}^{n}$$

- is a [n,k,w] PR-instance.
- which are decodable parameters - that decode to the message polynomial!
- with high probability Guruswami-Sudan will give you the plaintext polynomial.

Conclusions, I

- AF03 cryptosystem broken in the optimized parameter setting.
- Ciphertext-only attack.
- On a more positive note though:
  - present work demonstrate the power of employing probabilistic analysis when selecting parameters for Polynomial Reconstruction and Coding theoretic cryptosystems.
  - unnoticed in the previous work.
  - must use all tools available if we are to design a secure cryptosystem.

Conclusions, II

- Open problem:
  - design a public-key cryptosystem based on PR.
- Wrong design strategies in previous work:
  - did not employ probabilistic analysis.
  - did not employ/consider state-of-the-art decoding methods.
  - did not employ / understand the Provable Security framework for “Polynomial Reconstruction Based Cryptography”

    that has been put forth [KY-icalp2002]