Local Inversion of maps: Black Box Cryptanalysis

Virendra Sule,
Department of Electrical Engineering
Indian Institute of Technology Bombay
Mumbai 400076, India
vrs@ee.iitb.ac.in

viren.sule@gmail.com

Abstract

This paper is a short summary of the theory of local inversion announced in [6] and results announced in a previous paper on a new universal method for cryptanalysis which uses a Black Box linear algebra approach to computation of local inversion of nonlinear maps in finite fields. It is shown that one local inverse x of the map equation $y = F(x)$ can be computed by using the minimal polynomial of the sequence $y(k)$ defined by iterates (or recursion) $y(k+1) = F(y(k))$ with $y(0) = y$ when the sequence is periodic. This is the only solution in the periodic orbit of the map F. Furthermore, when the degree of the minimal polynomial is of polynomial order in the number of bits of the input of F (called low complexity case), the solution can be computed in polynomial time. The method of computation only uses the forward computations $F(y)$ for given y which is why this is called a Black Box approach.

An application of this approach is then shown for cryptanalysis of several maps arising in cryptographic primitives. It is shown how in the low-complexity cases maps defined by block and stream ciphers can be inverted to find the symmetric key under known-plaintext attack. Then it is shown how the RSA map can be inverted to find the plaintext as well as an equivalent private key to break the RSA algorithm without factoring the modulus. Finally, it is shown that the discrete log computation in finite field and elliptic curves can be formulated as a local inversion problem and the low-complexity cases can be solved in polynomial time.

Introduction

If $F : \mathbb{F}^n \to \mathbb{F}^n$ is a map acting in a cartesian space $\mathbb{F}^n$ over a finite field $\mathbb{F}$, then the *Local Inversion Problem* of F at a given y in $\mathbb{F}^n$ is the problem of computing all x in $\mathbb{F}^n$ such that

$$y = F(x) \tag{1}$$

Such a problem arises in most situations of cryptanalysis of symmetric and public key primitives such as the problem of key recovery (under a known-plaintext attack) in symmetric-key algorithms, plaintext recovery in RSA without factoring the modulus, private-key recovery in RSA without factoring the modulus, or finding discrete logarithms in finite fields and elliptic curves. However, in many such cases, no algebraic or Boolean (in case of $\mathbb{F}$ being $\mathbb{F}_2$) model of Eq. (1) may be available or, if available, is unsuitable for algorithmic solution due to a large number of unknowns in the equations. The map F in such cases can only be practically used as a black box: an algorithm or a hardware machine or a computer code for efficient forward computation of $F(x)$ for a given x.

It is hence suggested that the approach to the solution of cryptanalysis problems using only forward computations by F be called a *Black Box approach* analogous to the well-known Black Box Linear Algebra for solving linear systems of equations. In this paper, we develop such an approach and determine the required conditions for the local inversion to be computable in polynomial time in order $O(n^k)$. Although no assumption such as F being invertible (or a permutation) is made, practically and most often, computing one solution x of a local inversion is a significant achievement even if there may be more solutions. This approach is therefore a universal method to formulate diverse problems of cryptanalysis such as the key recovery of symmetric encryption algorithms, the breaking of RSA without factoring the modulus, and the computation of the discrete logarithm in finite fields and elliptic curves over finite fields as local inversion problems.

Local inversion using the Black Box approach under low complexity data

It is known in the literature that the algebraic approach to solving Eq. (1) poses a very challenging computational problem. Either the algebraic equations expressing the relation (1) involve too many (latent) variables and equations for which most approaches fail to scale up, or building the algebraic (or Boolean model) of (1) with the minimum number of variables by symbolic elimination is yet another challenging computation. Hence, it is worthwhile considering the Black Box approach in which F is

not specified algebraically but the output $y = F(x)$ can be computed for any input x efficiently. It is shown in [6] that the existence of a linear recurrence relation in the periodic sequence $y, F(y), F^{(2)}(y), \ldots$ with *Linear Complexity*[1] (LC) of polynomial size $O(n^k)$ is sufficient to solve one solution x (or a rational root) in polynomial time[2]. This sequence is generated recursively by the map F and the output y at an input x, and is different from the output sequence of a cryptosystem such as a stream cipher. Hence the LC of this recursive sequence is different from that of the LC of the output sequence (often called the LC of the cryptosystem). Cryptosystems are designed with the purpose of having large LC of output sequences. However, even if the output sequences have large LC, the recursive sequence above can have low LC since it also depends on the input (plaintext or IV) and the cipher design does not necessarily consider the constraint of a high LC of this recursive sequence. This result has an important consequence to applications in cryptanalysis.

It turns out that the cryptanalysis of symmetric-key algorithms, RSA cryptanalysis of inverting ciphertext to plaintext, as well as breaking RSA encryption by the same public key from chosen ciphertext attack (CCA) and solutions of Discrete Logarithms in finite fields and Elliptic Curves (ECDLP) are all local inversion problems. Hence, the low LC of respective sequences is a sufficient condition for solving them in polynomial time. Local inversion is also a new method and can be used as a universal attack for solving all of these cryptanalysis problems.

This article is a very short summary of the local inversion method announced in [6] which the reader is expected to refer to for details and proofs of results. The inversion of maps for cryptanalysis has been studied in previous literature by the Time–Memory Tradeoff (TMTO) attack [1, 2, 7]. However, TMTO does not consider finite fields' structure on the domain of the maps. TMTO is also referred to as Rainbow attack [8], which is governed by the square-root bound on the number of points in the image of F, which is of exponential size in the number of unknowns. Hence, this attack is not feasible for realistic sizes of domains (or number of unknowns) of maps F (or Eq. (1)).

Theory of local inversion of maps

It is stated above that local inversion of a map F at y is equivalent to solving the roots of the system of (polynomial) Eq. (1) if such a system is available. Hence, this appears to be a problem solvable by computer algebra, by building the algebraic system of equations corresponding to the map F. Alternatively, we observe that the solutions of (1) generate trajectories of the dynamical system

$$y(k+1) = F(y(k)), y(k) \in \mathbb{F}^n, k = 0, 1, 2, \ldots \quad (2)$$

which, in turn, are in one-to-one correspondence with the recurring sequences

$$S(F, y) = \{y, F(y), F^{(2)}(y), \ldots\}. \quad (3)$$

Thus, the theory behind the local inversion of maps relates three different mathematical domains, maps F defined by algebraic equations in finite fields, dynamical systems defined by maps F, and the ensemble of sequences generated by maps F with respect to points y in their range. The complexity of solving the problem of local inversion depends on the distribution of linear complexities of sequences $S(F, y)$ where y is taken over the range of F. The shift of focus from solving an algebraic system model of $y = F(x)$ to a Black Box model of F generating sequence $S(F, y)$, gives a practically feasible solution to the local inversion problem under the condition of low LC of the sequence $S(F, y)$.

Role of Linear Complexity (LC)

In the literature, LC was proposed as a complexity measure for sequences in finite fields. Such a complexity measure was used for qualifying the complexity of output sequences of maps defined by block ciphers for counter inputs (in counter mode of operation), or for outputs of stream ciphers [3]. The LC of several specially constructed sequences is also studied in Number Theory [4]. Low-LC m of such sequences enables the Berlekamp–Massey attack which allows for prediction of the complete sequence over the whole period from $2m$ terms. In the local inversion approach, however, the sequence $S(F, y)$ defined recursively by F and y is completely different from a sequence of an output stream of the cipher for counter input or output of stream cipher for a fixed IV. The LC of output sequences is used for modeling the sequence as an output sequence of LFSRs. But such modeling does not solve the symmetric key recovery problem. Recursive sequences such as $S(F, y)$ have not been studied previously for known cipher algorithms or general maps from the objective of inversion of maps. The application of the concepts of LC and minimal recurrence for local inversion is a fresh idea proposed in [6] and presents a Black Box approach for the inversion of non-linear maps.

A practical constraint in solving the problem

Due to the finiteness of the field $\mathbb{F}$, the recurring sequences $S(F, y)$ are quasi-periodic for any y. However, even when the sequence is periodic, the full sequence $S(F, y)$ is never practically available for computation since the period may be of exponential order in number of variables n or the field size $n = \log |\mathbb{F}|$. Hence, only a small number of terms in the sequence $S(F, y)$ of order $O(n^3)$ are available for solving x given y. Therefore, the data of a limited number of terms of $S(F, y)$ may be insufficient to compute any solution x or may

[1] i.e. the length of the linear recurrence relation

[2] Polynomial time of a computation: the computation is achievable in $O(n^k)$ steps, equivalently in $O((\log q)^k)$ steps in the field $\mathbb{F}_q$

even compute a false-positive solution which needs to be eliminated by verifying whether $y = F(x)$.

Linear Complexity of a sequence

A given sequence $\widehat{s} = \{s_0, s_1, s_2, \ldots\}$ over a finite field $\mathbb{F}$ is always ultimately periodic due to the finiteness of $\mathbb{F}$, i.e. there exist numbers $r \geq 0$ and $N > 0$ such that

$$s_{(k)} = s_{(k+N)} \text{ for } k \geq r. \tag{4}$$

The smallest r is called the pre-period of $\widehat{s}$ and the smallest N is called period. The sequence is called periodic of period N when $r = 0$. A polynomial can be associated with the periodic sequences which satisfy the relation (4) by defining a one-step right-shift operator on the sequence,

$$X(\widehat{s}) = \{s_{(k+1)}, k = 0, 1, 2, \ldots\}$$

and scalar multiplication of $\widehat{s}$ by constants. Then the relation (4) for periodic sequences can be expressed as

$$s_k = X^N(s_k) \text{ for } k \geq 0.$$

which is equivalent to

$$(X^N - 1)(\widehat{s}) = 0.$$

In general, if

$$p(X) = X^d - \sum_{i=0}^{d-1} \alpha_i X^i$$

is any polynomial in $\mathbb{F}[X]$ such that with the definition of X being a shift operator as defined above,

$$p(X)(\widehat{s}) = 0 \tag{5}$$

then $p(X)$ is called a characteristic polynomial of $\widehat{s}$. Thus, $(X^N - 1)$ is a characteristic polynomial. A characteristic polynomial of smallest degree is unique and is called the *minimal polynomial* of $\widehat{s}$. The degree of the minimal polynomial is called the *Linear Complexity* (LC) of the sequence. Now, getting back to the sequence $S(F, y)$ defined in Eq. (3), let $p(X)$ be a characteristic polynomial of $S(F, y)$. Then the relation (5) signifies a *linear recurrence relation* of degree d satisfied by $S(F, y)$,

$$F^{(d+j)}(y) - \sum_{i=0}^{d-1} \alpha_i F^{(i+j)}(y) = 0 \tag{6}$$

for $j = 0, 1, 2, \ldots$ If $S(F, y)$ is periodic of period N, $(X^N - 1)$ is always a characteristic polynomial. But there is a possibility of a lower-degree characteristic polynomial satisfied by $S(F, y)$. Hence, there exists a minimal polynomial for $S(F, y)$ which is a characteristic polynomial of least degree and is unique because the polynomial is monic.

Solution of the local inversion problem

First consider the theoretical situation in which there are no limitations on the number of terms within the sequence $S(F, y)$ being available for inversion. In the first theorem below, we show that computation of the minimal polynomial solves the local inversion problem when the sequence $S(F, y)$ is periodic. Let the minimal polynomial of a periodic $S(F, y)$ be denoted as

$$f(x) = X^m - \sum_{i=0}^{m-1} \alpha_i X^i. \tag{7}$$

Theorem 1 Let the sequence $S(F, y)$ be periodic of period N and $f(x)$ be its minimal polynomial. Then

1. $f(0) \neq 0$.
2. N is the order of $f(x)$ over $\mathbb{F}$.
3. The solution of the local inverse of $y = F(x)$ in the periodic $S(F, y)$ is

$$x = \frac{1}{\alpha_0} \left[F^{(m-1)}(y) - \sum_{i=1}^{m-1} \alpha_i F^{(i-1)}(y) \right]. \tag{8}$$

Proof 1 Let $f(x)$ as denoted in Eq. (7) be the minimal polynomial of $S(F, y)$. Since the period of the sequence is N, $f(x)|(X^N - 1)$ which implies $\alpha_0 = f(0) \neq 0$ and since N is the smallest number such that the recurrence

$$F^{(N)}(y) = y$$

holds, N is the order of $f(x)$. As $f(x)$ is also a characteristic polynomial of $S(F, y)$,

$$f(x)(S(F, y)) = 0$$

which is equivalent to

$$F^{(m)}(y) - \sum_{i=0}^{m-1} \alpha_i F^i(y) = 0$$

Let x be the solution of $y = F(x)$ in the same periodic orbit of the dynamical system (2). Then $S(F, x)$ is the sequence $S(F, y)$ shifted right by one index due to which the sequence satisfies the same recurrence relation (6) and the relation

$$f(x)(S(F, x)) = 0$$

which is equivalent to

$$F^{(m)}(x) - \sum_{i=0}^{m-1} \alpha_i F^i(x) = 0$$

w.r.t. the minimal polynomial. Substituting for $y = F(x)$ and solving for x from the above equation as $\alpha_0 \neq 0$ gives the solution as stated.

It is important to observe that the solution x of the local inverse is obtained in terms of a linear combination of the sequence $\{y, F(y), F^{(2)}(y), \ldots, F^{(m-1)}(y)\}$.

An incomplete algorithm

We now come to the practically most relevant problem of local inversion of a map F at y. The practical constraint to be faced is that the sequence $S(F, y)$ is not available for the full period because the period is exponential in n (or close to the size of the field when $n = 1$). However, the partial sequence is available up to M terms where M is of polynomial size $O(n^k)$ for some k as the sequence,

$$\{y, F(y), \ldots, F^{(M-1)}(y)\}. \tag{9}$$

In this case, the recurrence relation (6) is satisfied only up to the degree $m = \lfloor M/2 \rfloor$ since there is no further data of the sequence available beyond M terms. Hence, the largest degree of the minimal polynomial is m. As shown in Section 2.6 of [6], the minimal polynomial of the limited sequence up to M terms is obtained from the unique solution $\widehat{\alpha}$ of the linear system

$$H(k)\widehat{\alpha} = h(k+1) \tag{10}$$

where $H(k)$ is the Hankel matrix defined as

$$H(k) = \begin{bmatrix} y & F(y) & \cdots & F^{(k-1)}(y) \\ F(y) & F^{(2)}(y) & \cdots & F^{(k)}(y) \\ \vdots & \vdots & \ddots & \vdots \\ F^{(k-1)}(y) & F^{(k)}(y) & \cdots & F^{(2k-1)}(y) \end{bmatrix}$$

of largest rank for $k \leq \lfloor M/2 \rfloor$. The vector of coefficients of the minimal polynomial of degree k is denoted $\widehat{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_{(k-1)})^T$, and the vector on the r.h.s. of Eq. (10), $h(k+1)$, is the last column of $H(k+1)$ after dropping the bottom entry.

The Incomplete Algorithm 1 is useful to decide whether the local inverse can be computed from the given data. Since the given data is the partial sequence of M terms, which is of polynomial size, the algorithm computes the inverse in polynomial time when the data is sufficient to compute the inverse. The algorithm also checks for false-positive cases of solutions and, if no solution is found, returns a string stating that the data is insufficient.

We can thus conclude:

Theorem 2 Let M be of polynomial order $O(n^k)$. If the sequence (9) has a minimal polynomial satisfying the recurrence relation (6) of degree $m \leq \lfloor M/2 \rfloor$ and the solution x computed in (8) satisfies $y = F(x)$, then the local inverse is computable in polynomial time.

In practice, the polynomial bound $O(n^k)$ is limited to $k = 1, 2, 3$ for realistic cases.

Algorithm 1 Incomplete algorithm to find the unique solution of $y = F(x)$ given y

```
procedure LocalInversion(given M terms of S(F, y))
    Input M of O(n^k) and the sequence (9).
    repeat
        m ← ⌊M/2⌋
        if m = rank H(m) = rank H(m + 1) then
            // A unique minimal polynomial of degree m exists
            Compute min. polynomial and solution x as in (8).
            if x satisfies y = F(x) then
                return x
            end if
        else if rank H(m + 1) > m then
            return "Insufficient data to compute the
              local inverse."
        else   // rank H(m) < m
            M ← M − 1
        end if
    until M = 2
    return "Insufficient data to compute the local inverse."
end procedure
```

Embedding maps

The aforementioned algorithm for local inversion is applicable to maps $F : \mathbb{F}^n \to \mathbb{F}^n$ where the number of bits n of the input and the output are the same. In many practical situations of cryptanalysis, however, the maps arise as $F : \mathbb{F}^n \to \mathbb{F}^m$ where $m > n$. Such a map is called an embedding map. The solution of local inverse $y = F(x)$ for such a map can be found using the theory developed above for regular maps with $m = n$ as explained in the following.

Solution of local inverse for embedding maps

Consider a projection Π from $\mathbb{F}^m \to \mathbb{F}^n$. For a vector y in $\mathbb{F}^m$, $\Pi(y)$ denotes a vector of some of the n components of y. If x is a local inverse of $y = F(x)$ for any such projection, $\Pi(y) = (\Pi \circ F)(x)$. Therefore, x is also a local inverse of $y_1 = \Pi(y)$ under the map $F_1 = \Pi \circ F$. Since F_1 is a standard map in $\mathbb{F}^n$, we can utilize the recurrence sequence generation using F_1 and the minimal polynomial to find a possible local inverse of $y_1 = F_1(x)$. Then such a solution is verified with all other maps to verify whether x satisfies $\Pi_k(y) = F_k(x)$ for all other projections Π_k of $\mathbb{F}^m$ to $\mathbb{F}^n$. A solution is rejected if it fails to satisfy the equation for any k. Hence, we can define the LC of an embedding map to be the smallest LC of the recurrences defined by $y_k = F_k(x)$ over all projections $\Pi_k : \mathbb{F}^m \to \mathbb{F}^n$. In fact, it can be observed that a restricted set of projections on the following subsets of coordinates of y are sufficient to compute and verify the local inverse x,

$$\begin{aligned} \Pi_1(y) &= \{y^1, \ldots, y^n\} \\ \Pi_2(y) &= \{y^2 \ldots, y^{(n+1)}\} \\ &\vdots \\ \Pi_{(m-n+1)} &= \{y^{(m-n+1)}, \ldots, y^m\} \end{aligned} \tag{11}$$

The computational effort depends on the minimal polynomial of recurrences $y_k = F_k(x)$ at the first choice

of k which chooses a projection. Hence, an embedding map can still be locally inverted in polynomial time if there is an index k such that the recurrence defined by $y_k = F_k(x)$ has a LC of polynomial order in number of the variables, and the solution is verified by all other projections.

Typical maps in cryptanalysis

In this final section, we will outline some of the typical maps arising in cryptography which can be considered for local inversion to show that the local inversion approach is a uniform methodology for cryptanalysis. It must be understood, however, that the local inversion is a theoretical methodology – its practical feasibility depends on how small the LC of the local inversion problem as shown in Theorem 2 actually is.

Symmetric encryption algorithms

The two types of algorithms are block ciphers and stream ciphers.

Block ciphers: An algorithm is described by $C = E(K, P)$ where P, C are the blocks of plaintext and ciphertext, respectively, while K is the symmetric key block. In the cases of Known-Plaintext Attack (KPA), the pair (P, C) of blocks is known to the attacker. Thus, the local inversion problem is $y = F(x)$ where $y = C$, $x = K$ and $F(x) = E(x, P)$. The map F is from l bits of K to block length in bits of C. In the chosen plaintext attack, P is chosen such that the inversion problem is easier. For local inversion, P can be chosen such that the LC of $F(x)$ is smallest or polynomially bounded. Such an input may not be easy to compute, though.

Stream ciphers: The algorithm is described by a dynamic system with a state update map $x(k+1) = F(x(k))$ and an output map $y(k) = f(x(k))$. The initial condition $x(0) = (K, IV)$ consists of the symmetric key K (with bit length l) and an initializing vector IV. Hence, the local inversion problem is defined by $y = \widehat{F}(x)$ where y is the vector of output stream $(y(k_0), y(k_0+1), \ldots, y(k_0+l-1))$ while

$$\widehat{F}(x) = ((F^*)^{(k_0+i)} f(x, IV)),$$
$$\text{for } i = 0, 1, 2, \ldots, (k_0 + l - 1)$$

The map $\widehat{F}$ can be turned into an embedding by collecting more samples of the output stream.

For both types the maps for key recovery are available in black-box form with l bits of input x (key length) and output y on which the inversion algorithm can be applied after choosing M bounded by a polynomial order $O(l^k)$.

RSA cryptanalysis

RSA has public keys $n = pq$ where p, q are odd primes and an exponent e such that $\gcd(e, \phi(n)) = 1$. The private keys are p, q, d where $ed = 1 \bmod \phi(n)$. Two local inversion problems can be defined as follows:

Decryption of ciphertext: For the message $m \in [0, n-1]$ the ciphertext is $c = m^e \bmod n$. Let l be the bit-length of n. For an l-bit number x let (x) denote the bit-string in the binary expansion of x and $[x]$ denote the operation of recovering the number x from the binary string (x). Then the map $F : \mathbb{F}_2^l \to \mathbb{F}_2^l$ is defined by $y = (c)$ and $F(x) = ([x]^e \bmod n)$. The sequence $S(F, y)$ in $[0, n-1]$ is

$$\{c, c^e \bmod n, c^{e^2} \bmod n, c^{e^3} \bmod n \ldots\}$$

while the sequence $S(F, y)$ in $\mathbb{F}_2^l$ is

$$\{(c), ([c]^e \bmod n), ([c]^{e^2} \bmod n), ([c]^{e^3} \bmod n), \ldots\}.$$

It is well-known that the sequence $c^{e^k} \bmod n$ is periodic since e is co-prime to $\phi(n)$. The LC is obtained from the sequence of binary vectors. The message m is obtained by solving the local inverse $(c) = F(x)$, $m = [x]$. This way of solving for m shows how RSA ciphertext can be decrypted by local inversion without factoring n. In cases where c causes a low LC of the above sequence, the encryption can be broken in polynomial time as shown by Theorem 2.

Chosen Ciphertext Attack: In this attack, a ciphertext c is chosen by the attacker and the decrypted plaintext m is sought as a verification. The relation is $m = c^d \bmod n$. Hence, the local inverse problem is defined by $y = m$ and $F(x) = c^x \bmod n$. Since d belongs to the ring $\mathbb{Z}_{\phi(n)}$, the number of bits of x is equal to $\phi(n)$. Although $\phi(n)$ is not available to the attacker, we can choose the number of bits to be n. The sequence $S(F, y)$ is

$$\{m, c^m \bmod n, c^{c^m \bmod n} \bmod n, \ldots\}$$

while the map $F : \mathbb{F}_2^l \to \mathbb{F}_2^l$ is defined by $(m) = (c^{[x]} \bmod n)$ and the sequence of binary vectors is

$$\{(m), (c^{[m]} \bmod n), (c^{[c^{[m]} \bmod n]} \bmod n), \ldots\}.$$

Note that the computation modulo n automatically restricts the exponent's modulo $\phi(n)$, hence the sequence generated is correct for the local inversion of the map F. The periodicity of the sequences above is explained in [6]. The sequence is converted to a sequence of binary expansions to compute the LC and the local inverse x. The inverse x is converted back to a number in $[0, n-1]$. Now, an important observation to be noted in this inversion is that the inverse of $m = F(x)$ is x which is not necessarily d but satisfies $ex = 1 \bmod \phi(n)$. Therefore, although the private key d is not exactly recovered, the inverse allows for the decryption of any other ciphertext $\widetilde{c}$ corresponding to the plaintext $\widetilde{m}$ since we have

$$(\widetilde{c}^x \bmod n)^e \bmod n = \widetilde{c}^{ex} \bmod n = \widetilde{c},$$

hence $\widetilde{m} = \widetilde{c}^x \bmod n$. This shows that local inversion by CCA on RSA breaks RSA equivalent to computing the private d key without factoring the modulus n.

Discrete logarithm in finite fields and elliptic curves

The Discrete Logarithm Problem (DLP) has been the first major tool for development of Public Key Cryptography, same as the Diffie–Hellman key-exchange scheme (DH). Later on, the DH scheme was upgraded to elliptic curves, leading to Elliptic Curve Cryptography (ECC). The DLP on Elliptic Curves (ECDLP) has not been found to have an algorithm for solution better than exponential complexity. What we show in this section is that the solution of both of these problems can be addressed by local inversion. Thus, it turns out that at least in cases where the map F and given data y meet the conditions so that the sequence $S(F, y)$ has low LC, the DLPs can be solved efficiently. The situation in case of ECDLP is very complicated because the map F turns out to be an embedding.

DLP on finite fields

In a prime field $\mathbb{F}_p$ for a large prime p, the exponent function with a base a in $\mathbb{F}_p^*$, $\phi : [1, p-1] \to \mathbb{F}_p, x \mapsto a^x \bmod p$ is actually defined as a map from $\mathbb{F}_p^*$ to $\mathbb{F}_p^*$. Taking the binary expansion of numbers in $[1, p-1]$, where l is the bit-length of p, this function expresses a map operation

$$\begin{aligned} F: \quad & \mathbb{F}_2^l \to \mathbb{F}_2^l \\ & (x) \mapsto (a^{[x]} \bmod p). \end{aligned}$$

Here (x) denotes the binary string corresponding to a number $x \in [1, p-1]$ and $[x]$ the reconstruction of the number x from the binary string. Let b be a given element in $\mathbb{F}_p^*$ for a primitive element a. Then the equation for local inversion is $F((x)) = (b)$. The map is available for black box computation. The sequence of binary vectors $S(F, (b))$ is

$$\{(b), (a^{[b]} \bmod p), (a^{a^{[b]} \bmod p} \bmod p), \ldots\}. \qquad (12)$$

As defined by the map iterations $y(0) = (b)$, $y(k+1) = F(y(k))$ for F defined as above on $\mathbb{F}_2^l$. The reader is referred to [6] to see details justifying the periodicity of the sequence.

The solution of the DLP can thus be found as local inversion of this map F at the value (b). If the LC of the above sequence given up to polynomial number of terms $O(l^k)$ is small, then, as described in Theorem 2, the DLP can be solved in polynomial time using Algorithm 1. For the solution of DLP over general field $\mathbb{F}_q$ using local inversion, the reader is referred to [6].

DLP on elliptic curves

In the discrete log problem on an elliptic curve E over $\mathbb{F}_q$, there are given points P and $Q = [m]P$ in E where m is the integral multiplier. It is required to solve for the multiplier m. Define the map

$$F_P : m \mapsto [m]P$$

then the local inversion of $Q = F_P(m)$ solves the ECDLP. However, the map F needs to be expressed in the standard form as before. We can do this by defining the map in $\mathbb{F}_2^l$ for an appropriate l.

Formulation as local inversion

The multiplier m is less than the order of the cyclic group $n = \langle P \rangle$ in E. Sometimes the group E itself has prime order $n = \sharp E$, hence the order of $\langle P \rangle$ is n. To fix the number of bits l in m we consider estimates of the order of E. The well-known bound on the order of E is

$$\sharp E(\mathbb{F}_q) \leq q + 1 + 2\sqrt{q}$$

Assuming $q > 4$ we have $\sharp E \leq 2q$. On the other hand, the point Q in E has two co-ordinates in $\mathbb{F}_q$. Thus, the bit length of $\sharp E \leq 1 + \log(2q) = 2 + \log q = l$ while the bit length of two co-ordinates of a point taken together is $1 + \log(2q) = l$. Consider the map F_P defining the scalar multiplication of P in E

$$\begin{aligned} F_P: \quad & \mathbb{Z}_n \to E \\ & m \mapsto [m]P \end{aligned}$$

In the binary co-ordinate expansion on both sides, this mapping is

$$F_P((m)) = ((Q_x), (Q_y)) \qquad (13)$$

where (m) denotes the coefficients in the binary expansion of m and (Q_x), (Q_y) are coefficients in the binary expansions of co-ordinates of $Q = [m]P$. We shall denote the binary expansion of the co-ordinate pair of Q as (Q).

Formulation of F_P as a map over the binary field

In order to utilize the previous theory of inversion on the map (13), it is necessary to express it as a map in the cartesian spaces of $\mathbb{F}_2$ and verify that it is an embedding.

Let $r = 1 + \log n$ where n is the order of $\langle P \rangle$. Then F_P in (13) represents a map $F_P : \mathbb{F}_2^r \to \mathbb{F}_2^l$ where $r < l$. Thus, F_P is an embedding and it is required to apply the theory of local inversion of embeddings to solve the embedding equation for the local inversion of $F_P(m) = Q$ from the binary representation in (13). The application of Algorithm 1 requires that n the order of P is known. Let $t = l - r$, then the projection equations 11 give $(t+1)$ standard equations

$$\Pi_i \circ F_P(x) = \Pi_i((Q)), i = 1, 2, \ldots (t+1)$$

denoted by $F_i = \Pi_i \circ F_P$ and $y(i) = \Pi_i((Q))$ where Π_i are a projection on r components of y. The details of this projection are described in [6]. Following Theorem 2, we now have:

Theorem 3 If for any of the indices i, $1 \leq i \leq (t+1)$ the projection equation $F_i(x) = y(i)$ has a periodic recurrence sequence $S(F_i, y(i))$ with a LC $m \leq \lfloor M/2 \rfloor$ where M is of polynomial order $O(l^k)$ and the local inverse x satisfies all other projection equations, then the ECDLP is solved in polynomial time by Algorithm 1.

Conclusions

This brief paper summarizes the ideas on a new method which can be referred to as a Black Box approach to cryptanalysis by Local Inversion of maps in finite fields. The methodology is universal in the sense that it is applicable to diverse problems of cryptanalysis of both symmetric as well as public-key cryptography. The most important conclusion arising from this method is that, since the recursive sequences associated with the inversion problems identified by the method have not been studied in the past for their LC for most of the maps F in cryptography, urgent work should be carried out to determine density of points y in the image of F for which the sequences $S(F, y)$ have low LC. These densities should be useful for deciding the grading of security of parameters of cryptographic primitives and should be included as a standard for the design of ciphers.

The algorithm proposed for local inversion always works in polynomial time in the number of unknown variables n of the map F, i.e. when it discovers a correct solution the solution is computed in polynomial time. Another practical advantage of the algorithm is that it does not require an algebraic model of the map equation $y = F(x)$ but only uses forward computation at specified arguments x. Hence this approach is called black box cryptanalysis.

References

[1] Martin Hellman: *A cryptanalytic time-memory trade-off. IEEE Trans. on Information Theory, 26(4), pp.401-406*, 1980.

[2] Howard M. Hays: *Distributed Time Memory Trade-off Attacks on Ciphers.* `http://eprint.iacr.org/2018/123`.

[3] Rainer A. Rueppel: *Analysis and Design of Stream Ciphers.* Springer Verlag, Berlin, Heidelberg, 1986. ISBN 9783540168706.

[4] Harald Niederreiter: *Linear complexity and related complexity measures for sequences. Indocrypt 2003, LNCS 2904, pp.1-17*, Springer Verlag, 2003.

[5] Lawrence Washington: *Elliptic curves, Number Theory and Cryptography.* Chapman and Hall/CRC Press, 2003.

[6] Virendra Sule: *Local inversion of maps: A new attack of Symmetric Encryption, RSA and ECDLP.* `https://arxiv.org/abs/2202.06584v2`, March, 2022.

[7] Mark Stamp and Richard M. Low: *Applied Cryptanalysis.* Wiley-Interscience, 2007.

[8] Rainbow table, Wikipedia: `https://en.wikipedia.org/wiki/Rainbow_table`.