제한된 파형을 이용한 향상된 RSA-CRT 부채널 분석

박종연*, 한동국*, 이옥연*, 최두호**
*국민대학교 수학과
**한국 전자 통신 연구원
e-mail : flyfree@kookmin.ac.kr

Improved Side Channel Attack using Restricted Number of Traces on RSA-CRT

Jong-Yeon Park*, Dong-Guk Han*, Okyeon Yi* and Dooho Choi*
*Dept. of Mathematics, Kookmin University
**Electronic and Telecommunications Research Institute(ETRI)

요 약

RSA-CRT 알고리즘은 RSA 의 지수승 연산의 효율성을 향상시키기 위해 널리 사용되고 있으며, CRT 를 적용한 알고리즘은 다양한 방법의 부채널 분석(Side Channel Analysis)으로부터 약점이 노출되어 왔다. 그 중 Boer 등에 의해 발표된 MRED 분석 방법은, 등 간격의 데이터(Equidistant Data)를 이용하여 CRT 의 모듈러 리덕션 연산(Modular Reduction)결과로부터의 약점을 활용하여 일반적인 DPA 분석 법을 적용시킨 방법이다. 우리는 리덕션 결과의 데이터에 의존한 분석에서 벗어나, 리덕션 알고리즘 중간 연산 과정을 공격하는 새로운 공격 방법을 개발하였으며, 새로운 공격은 오직 "256×n 개"의 파형만으로 키 공간을 상당히 줄일 수 있기 때문에, 제한된 평문 수에서 이전에 알려져 있던 일반적인 MRED 분석 방법보다 향상된 분석 성능을 제공한다. 본 논문은 리더션 연산과정을 이용한 새로운 전력 분석 방법을 실제 MCU Chip 을 이용한 분석 결과를 제안한다.

1. 서론

부채널 분석(Side Channel Analysis)이란, 암호 알고리즘의 입력 값과 출력 값을 이용한 암호 분석 방법이 아닌, 암호 장비의 연산에서 발생되는 전력, 또는 전자파의 일반적인 특성 또는 통계적인 분석에 의해 암호 장비의 키를 찾는 방법이다[1]. 그 중 차분 전력 분석(Differential Power Analysis, DPA)은 부채널 분석 중 가장 강력한 분석 방법 중 하나로써, 상관 전력 분석(Correlation Power Analysis, CPA) 등의 방법으로 발전, 연구되어 왔다[2,3]. 이러한 분석은 공격 자가 키를 추측하고 추측한 키와 전력 모델에 의해 계산된 연산 값과 중간 파형 정보의 통계적인 분석을 이용하는 방법이다.

CRT 기반의 RSA 의 전력 분석은 초기 리덕션 단계에서의 비밀 소수 p 를 공격자가 알 수 없기 때문에 일반적인 DPA 분석 방법을 이용하여 분석 하는 것이 불가능 하다. 하지만 CRT 알고리즘의 초기 리덕션 단계와 재조합 단계가 반드시 필요하다는 약점 때문에 RSA-CRT 에 특성화된 여러 가지 SPA(Simple Power Analysis) 또는 DPA 기반의 전력 분석 방법이 소개 되어 왔다. 그 중 Boer 등이 제안한 전력 분석 방법은 등 간격의 평문을 이용하여 초기 리덕션 단계를 공격하는 방법으로써, MRED(Modular Reduction on Equidistant Data)분석 방법으로 알려져 있으며, 등 간격의 중간 값을 이용하여 DPA 분석이 적용된다[4]. CRT 의 재조합 단계를 분석하는 방법으로는 SPA 기반의 Novak 방법과, DPA 기반의 Amiel 방법 등이 있으며, Garner 의 CRT 재조합 연산에서 발생하는 약점을 이용하였다[5,6].

본 논문은 RSA-CRT 분석 방법 중 MRED 분석 방법을 변형한 새로운 MRED 분석 방법인 NMRED(New Modular Reduction on Equidistant Data)방법을 제안한다. NMRED 는 리덕션 알고리즘 중 뺄셈 단계의 borrow 값이 발생하는 구간에서의 알고리즘의 변화를 이용하며, 기존의 분석 방법보다 향상된 결과를 보였다. 본 논문은 다음과 같이 구성된다. 2 장에서는 기존의 MRED 전력 분석에 대하여 소개하며, 3 장에서는 우리의 새로운 공격 방법을 이론적으로 규명하고, 4 장에는 실험 결과와 기존의 분석 결과의 비교를 통해 새로운 공격 방법의 장, 단점을 논한다. 마지막 5 장은 본 논문을 결론짓는다.

2. MRED 전력분석

MRED 전력 분석은 등 간격의 입력 데이터 수열 $\{x, x-1, x-2, \ldots x-N+1\}$을 이용하여 등 간격의 리덕션 출력 수열$\{r, r-1, r-2, \ldots, r-N+1\}$을 만들어 낼 수 있는 다음의 성질 의하여 분석이 가능하다[4].

$$x-i \bmod p = x-i \quad (1)$$

공격자는 p 를 정확히 알 수 없더라도 식 (1)이 성립한다는 성질을 이용하여, 바이트 블록 단위의 등 간격 추측 key 수열은 {j, $j-1$ mod256, $j-2$ mod256,…,$j-N+1$mod256}의 각 원소의 헤밍웨이트를 계산하여 실제 파형상의{r,$r-1$ mod256, $r-2$ mod256,…, $r-N+1$ mod256}과의 최대 상관계수를 계산한다.

알고리즘 1: MRED(k 번째 byte)
입력 값 : k 번째 등 간격 평문에 의한 전력 정보 T=(s_0,s_1,s_2,…,s_{n-1}) 출력 값 : Key candidates
Step 1 For j from 0 to 255 *Setp 1.1* A_j={$H(j)$, $H(j-1\text{mod}256)$, $H(j-2\text{mod}256)$,…, $H(j-n+1 \text{ mod } 256)$} *Step 1.2* $\rho_j=\rho(A_j, T)$ *Step 1.3 If* $j=0$ *then* $key=0$, $\rho_r=\rho_j$ *Otherwise; if* $\rho_r<\rho_j$ *then,* $r_k=j$, $\rho_r=\rho_j$ *Step 2* Return r_k

<**알고리즘** 1> MRED **알고리즘**

<알고리즘 1>은 MRED 분석 알고리즘이다. $H(r)$는 r 의 헤밍웨이트를 의미하며, 추측한 j 값이 r 과 일치한다면, 가장 높은 상관관계를 가질 것이기 때문에, CPA 분석을 통하여 직접 키를 찾아내는 것이 가능하다. 위 공격 방법은 다음의 수식(2)에 의해 최하위 바이트에서 임의의 크기까지 확장 시킬 수 있다.

$$x - i(256)^k \bmod p = r - i(256)^k \qquad (2)$$

이러한 방법으로 계산된 중간 값을 이용하여 r 에 대한 DPA 분석을 시행한다. r 값을 이용하여, GCD($p \times q$, $x-r$)을 계산하면 공격자는 쉽게 RSA 의 비밀 소수인 p 를 찾아 낼 수 있다. 분석에 필요한 파형 수는 각 블록의 분석에 필요한 파형 수의 평균 m 에 블록 수의 곱인 $m \times n$ 이다.

3. Reduction 알고리즘을 이용한 NMRED 분석

기존의 MRED 는 등 간격 평문에 대한 헤밍웨이트를 적용하였다. 하지만 우리의 새로운 방법은 헤밍웨이트를 계산하지 않고 단순 등 간격의 중간 값을 이용하는 NMRED(New MRED)방법을 제안한다.

알고리즘 2: NMRED(k 번째 byte)
입력 값 : k 번째 등 간격 평문에 의한 전력 정보 T=(s_0,s_1,s_2,…,s_{n-1}) 출력 값 : Key candidates
Step 1 For j from 0 to 255 *Step 1.1* A_j={j, $j-1$mod256, $j-2$mod256,…, $j-n+1$ mod 256} *Step 1.2* $\rho_j=\rho(A_j, T)$ *Step 2* Return {$\rho_{jk} \mid j \in Z_{256}$}

<**알고리즘** 2> NMRED **알고리즘**

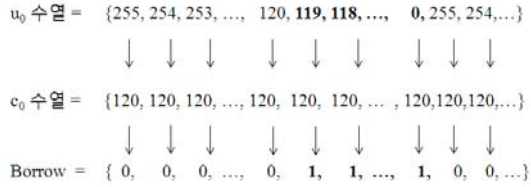
<알고리즘 2>는 NMRED 분석 방법을 나타낸 것이다. k 번째 바이트 분석 결과인 256 개의 상관계수 값을 이용해 공격자는 키의 후보를 결정하게 된다. *Step 1.1* 의 A_j 는 <알고리즘 1>에서와 같이, 파형과의 상관계수를 계산하기 위한 중간 연산 값이다. A_j 는 헤밍웨이트에 의해 계산되지 않고 단순히 등 간격의 평문을 활용하여 계산된다.

리덕션 알고리즘은 일반적으로 곱셈과 덧셈, 뺄셈 등의 연산으로 구성되어 있으며[7], 연산의 마지막 단계는 상수 또는 '몫에 의존한 상수'에 의한 뺄셈 단계를 가지고 있다. NMRED 방법은 상수에 의한 뺄셈 알고리즘의 블록 단위 연산의 borrow 가 존재한다는 성질을 이용한 것이다.

알고리즘 3: 큰 수 뺄셈
입력 값 : u, c (n+1 byte 크기) $u \geq c$ 출력 값 : $u - c$
Step 1 Borrow = 0 *Step 1.1* For i from 0 to n do the following *Step 1.2* $r_i = (u_i - c_i + \text{borrow}) \bmod 256$ *Step 1.3* If$(u_i - c_i + \text{borrow}) \geq 0$ then borrow=0; otherwise borrow= –1 *step 2* Return((r_n,r_{n-1}…,r_1,r_0))

<**알고리즘** 3> **큰 수 뺄셈 알고리즘**

<알고리즘 2>는 큰 수 뺄셈 알고리즘을 나타낸 것이다[7]. *Step* 2.2 에서 뺄셈 연산 시 상위 자릿수로부터 빌려오는 값인 borrow 값을 수정한다. 이 때에 조건문에 의해서 0 과 –1 의 선택에 따라 두 가지 파형의 형태로 분류가 된다. 뺄셈의 최하위 바이트인 $u_0 - c_0$ 연산만을 주목하여 보자. Borrow 값의 발생은 $u_0 - c_0$ 연산 시 $u_0 < c_0$ 인 경우에 발생하게 된다.



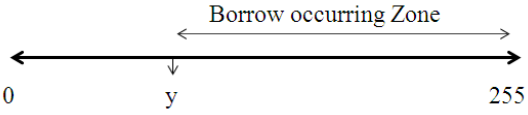
(그림 1) $u_0 - i - c_0 = r_0 - i$ 연산의 borrow 발생

(그림 1)은 뺄셈의 최하위 바이트에서 발생하는 $u_0 - c_0$ 연산에서의 'borrow 발생'을 $u_0 = 255$, $c_0 = 120$ 으로 설정하였을 때 계산 값을 나타낸 것이다. 등 간격 평문 공격에 의해서 발생되는 u_0 수열은 {u_0 mod 256, u_0-1 mod 256,u_0-2 mod256,…u_0-N+1 mod 256}이며, c_0 는 몫에 의해 고정된다. Borrow 발생이 u_0 수열의 값에 의해 변함에 따라 파형은 borrow '0' 과 borrow '–1' 로 구분될 것이다. 따라서 공격자가 borrow "0"과 borrow "–1" 파형으로 분류 할 수 있다면, DPA 또는 CPA 분석이 가능하다. 하지만 공격자는 나누는 값인 p 를 알 수 없으므로, borrow 의 발생을 정확하게 예측할 수 없고, 따라서 c_0 값을 알 수 없다. NMRED 방법에서는 c_0 값이 고정 상수임을 이용하여, u_0 수열의 값에 따라 확률적으로 borrow 수열을 구성한다. Z_{256} 에 속하는 임의의 값 y 와 임의의 상수 c 의 최하위 블록 c_0 에 대하여 $y - c_0$ 를 계산할 때 borrow 가 발생할 확률은 다음과

같이 계산 된다.

P(borrow occurring)
$= P(y < c_0) = (255 - y) / 256$ (3)

즉, u_0 수열에 속한 고정된 y 에 대하여 borrow 가 발생할 확률은 $(255-y)/256$ 이다. (그림 4)는 균등 확률 분포를 갖는 c_0 와 고정된 x 에 의한 Borrow 발생 경우를 그림으로 나타낸 것이다.



(그림 2) 균등 확률 분포를 갖는 c_0 와 고정된 x 에 의한 Borrow 발생의 경우

단, $c_0 \in Z_{256}$ 는 균등 확률 분포(Uniform Distribution)를 따른다고 가정하자. c_0 는 고정된 몫과 비밀 소수 p 의 곱에 의해 결정되기 때문에, RSA 프로토콜의 소수 p 의 생성에서의 난수성(Randomness)이 충분히 보장 된다면, c_0 는 균등분포를 갖는다.

따라서 c_0 를 알 수 없는 공격자가 수식 (3)의 확률을 이용하여 u_0 를 추측하기 위한 수열인 B_j 를 구성하여 <알고리즘 2> 의 *Step 1.1* 에 해당하는 NMRED 분석을 위한 중간 값을 계산한다. 예를 들어, B_{255} 는 다음과 같다.

B_{255} = {0/256, 1/256, 2/256, … ,254/256,255/256,0/256,…}

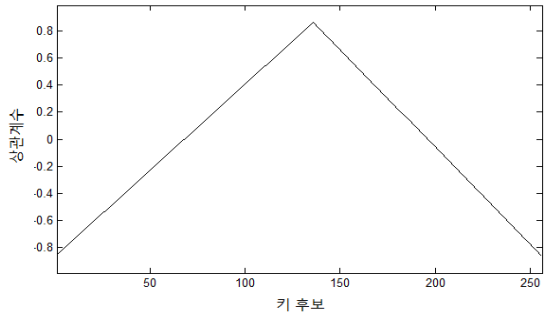
이러한 방법으로 구성된 B_j 는 u_0 를 찾기 위한 수열이라고 볼 수 있다. 하지만 공격자가 찾는 공격 대상 변수는 u_0 가 아닌 r_0 이기 때문에, NMRED 의 최종 중간 값 수열은 u_0 를 찾는 B_j 를 r_0 를 찾는 A_j 로 변형해야 한다. 상수 분모 255 를 제외한 B_j 은 (그림 1)의 u_0 수열의 255 에 대한 보수 관계이므로 Pearson 상관계수 $\rho(u_0$ 수열, $B_j) = -1$ 이다. 한편, $u_0 - c_0 = r_0$ 이므로, $u_0 = r_0 + c_0$ 이다. 그러므로 $u_0 - c_0 = r_0$ 연산에서 borrow 가 발생하는 것은 $u_0 = r_0 + c_0$ 의 연산에서 덧셈 carry 가 발생하는 것과 같다는 성질을 이용하여, B_j 계산과 같은 방법으로 r_0 를 추측하기 위한 중간 값으로 A_j 을 확률적으로 계산할 수 있다.

P(carry 발생 $| z \in Z_{256}$)
$= P(256 \le z + c_0) = P(256 - z \le c_0) = z / 256$ (4)

예를 들어 $u_0 = 255$ 이며 $c_0 = 120$ 이므로 $r_0 = 135$ 이다. 따라서 A_{135} 는 다음과 같이 표현된다.

A_{135} = {135/256, 134/256, 133/256, …, 1/256,…}

A_j 는 키 r_0 를 찾는 새로운 중간 값 집합이며, borrow 가 발생되는 파형과의 최대 상관계수를 계산하여 키를 찾아 낼 수 있다.



(그림 5) $A_j (r \in Z_{256})$과 borrow 발생과의 상관계수

(그림 5)는 (그림 3)과 같은 값으로 $u_0 = 255$, $c_0 = 120$, 로 설정하여 borrow 발생과 A_j 과의 상관계수를 계산한 것이다. 결과 가장 높은 상관계수를 갖는 키는 135 이며, 찾는 r_0 값임을 알 수 있다. 반면 가장 낮은 상관계수를 갖는 키는 255(=u_0)이며, $\rho(A_j, B_j) = -1$ 이기 때문에 발생한다. 결과 적으로 $\rho(A_{255}$,carry 발생) = $-\rho(A_{135}$, carry 발생) 이다.

4. 실험 결과와 한계

3 장에서는 borrow 가 발생하는 시간 영역을 이용하여 분석을 할수 있음을 이론적으로 검증 하였다. 본 장에서는 실제 실험 결과를 바탕으로 분석 성능을 확인하고 한계점을 제시한다.

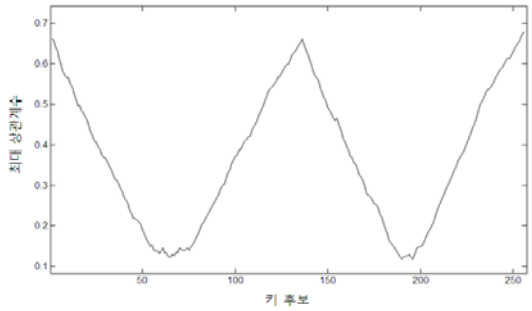
파형 수집	디지털 오실로스코프
프로세서	MCU chip – software board
샘플링 레이트	250MS/s
알고리즘	8 비트 연산의 모듈러 리덕션 알고리즘 $x - i(256)^k \bmod p = r - i(256)^k$
변수 크기	32byte 등 간격 입력 평문 16byte 소수 'p'

<표 1> **실험 환경**

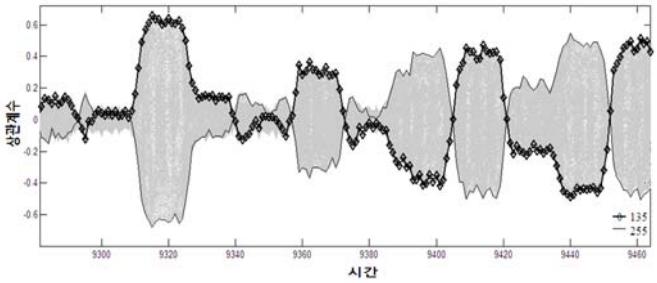
<표 2>는 실험 환경을 나타낸 것이다. 본 실험은 MCU chip 소프트웨어 보드에서 구동되는 모듈러 리덕션 알고리즘에서 진행되었다.

(그림 6)는 carry 발생을 확률적으로 계산한 A_j 을 중간 값으로 계산하여 뺄셈 단계를 분석한 결과 256 개의 r_0 후보에 대한 결과 값이다. 분석 결과, (그림 5)의 시뮬레이션 결과에 대한 절대값으로 분석되었다. 그림에 의하면 두 개의 높은 peak 주변으로 높은 상관계수가 분포하는 것을 알 수 있으며, 두 개의 높은 peak 는 u_0 와, r_0 의 후보라고 볼 수 있다.

(그림 7)은 전체 키에 대한 공격 구간에서의 상관계수를 나타낸 것이다. 가장 편차가 큰 두 개의 선은 A_{255} 상관계수와 A_{135} 의 상관계수이다. 대칭 적으로 가장 큰 상관도를 보이는 두 개의 키 사이에 다른 키가 분포되어 있음을 알 수 있다. 따라서 이론적으로 가장 낮은 상관도를 갖는 키 u_0 와 가장 높은 상관계수를 갖는 r_0 가 전체 시간 영역에서 최대 상관계수가 가장 높았으며, 두 개의 키 후보로 압축 가능하다.



(그림 6) borrow 발생을 이용한 최하위 바이트 공격 결과



(그림 7) 전체 키에 대한 상관계수 형태

	1st 바이트	2nd 바이트	3rd 바이트
최소 파형 개수(New)	256 개	256 개	256 개
최소 파형 개수(일반 MRED)	2800 개 이상	3000 개 이상	1800 개 이상

<표 2> NMRED 와 MRED 와의 분석 최소 파형 수 비교

<표 3>에서 알 수 있듯, 일반적인 MRED 분석에 수천 개의 파형이 필요했었지만, NMRED 분석 에는 오직 256 개의 파형이면 한 블록을 분석하는 파형 수로 충분하며, 256 개는 분석 중간 값을 구분 할 수 있는 최소 파형 개수 이므로, 최소한의 파형 수 만으로 분석이 되는 놀라운 성능을 보여 주었다. 전체 블록의 키를 찾는 데에는 $256\times n$(블록 수)의 파형이 필요하므로, 기존의 분석 방법보다 훨씬 적은 수의 평문 정보만을 이용한 분석이 가능하다. 하지만, NMRED 분석 결과는 (그림 6)에서와 같이 r_i 와 u_i 를 구분 할 수 없기 때문에 2 개 또는 여러 개의 키 후보를 주는 역할을 할 뿐, 전력 파형 수를 높임으로써, 하나의 유일한 키를 찾아내는 방법은 아니다. 하지만 아주 적은 량의 파형 수로 키 후보를 좁히는 효과가 있으므로, 제한된 파형 수의 분석에 효과가 있다. NMRED 의 또 다른 약점은 상수가 c_k=0 인 경우에 borrow 가 발생하지 않기 때문에 분석 방법이 적용되지 않으며, c_k=1, c_k=255 와 같이 $c_k \in Z_{256}$ 값이 0(=256) 에 가까울 경우에는 분석 성능이 저하될 가능성이 있다.

5. 결론

본 논문은 기존의 등 간격 평문을 이용한 CRT 분석 위치를 새로 정의하고 키를 찾아내기 위한 새로운 중간 값 수열을 생성한 NMRED 분석 방법을 제안 하였다. 우리의 방법에 의하면 $256\times n$ 개의 파형 수 만으로 $2''$ 개의 키 후보로 압축 할 수 있음을 보였다. 또한 NMRED 는 비교적 넓은 영역에서 피크가 나타나므로, 단순 데이터에 의존한 MRED 분석에 비해 분석 결과를 얻기가 쉽다. 하지만 NMRED 는 키 후보를 줄 뿐이며 단 하나의 키를 확정 할 수 없으며, C_k=0 일때에 분석이 되지 않는다는 단점이 있다. 따라서 향후에는 이를 보완하여 향상된 NMRED 분석 방법을 연구 할 것이다.

Acknowledgement

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(20100024870)
This work was supported by SCARF project which is the R&D program of KCA.
[Development of the Technology of Side Channel Attack Countermeasure Primitives and Security Validation]

참고문헌

[1] P,Kocher, J.Jaffe, and B. Jun, “Introduction to differential power analysis and related attacks”, 1998, White paper, Cryptography Research, http://www.cryptography.com/dpa/technical, 1998.

[2] P. Kocher, J. Jaffe, and B. Jun, “Differential power analysis,” Advances In Cryptology - CRYPTO' 99, LNCS 1666 Springer-Verlag, pp. 388-397, Santa Barbara, USA, August 1999.

[3] E. Brier, C. Clavier, and F. Olivier, “Correlation power analysis with a leakage model,” Cryptographic Hardware and Embedded Systems 2004. LNCS 3156 Springer-Verlag, pp. 16-29, 2004.

[4] B. D. Boer, K. Lemke, and G. Wicke, “A DPA attack against the modular reduction within a crt implementation of RSA”, Cryptographic Hardware and Embedded Systems 2002, LNCS 2523 Springer-Verlag, pp. 228-243, 2002.

[5] Roman Novak “SPA-Based Adaptive Chosen -Ciphertext Attack on RSA Implementation”, Public Key Crptography 2002, LNCS 2274 Springer-Verlag, pp. 252-262, 2002.

[6] Frederic Amiel, Benoit Feix, and Karine Villegas, “Power Analysis for Secret Recovering and Reverse Engineering of Public Key Algorithms”, International conference on Selected area in cryptography 2007, LNCS 4876 Springer-Verlag, pp 110-125, 2007.

[7] A.J Menezes, Paul C.van Oorschot and S.A Vanstone, “Handbook Applied Cryptography”, CRC press ISBN: 0-8493-8523-7, 1996.