ISSN(Print): 2287-5891

http://dx.doi.org/10.3745/KTCCS.2013.2.2.083

An Improved Side Channel Attack Using Event Information of Subtraction

Jong-Yeon Park[†] · Dong-Guk Han[††] · Okyeon Yi[†††] · Jung-Nyeo Kim[††††]

ABSTRACT

RSA-CRT is a widely used algorithm that provides high performance implementation of the RSA-signature algorithm. Many previous studies on each operation step have been published to verify the physical leakages of RSA-CRT when used in smart devices. This paper proposes SAED (subtraction algorithm analysis on equidistant data), which extracts sensitive information using the event information of the subtraction operation in a reduction algorithm. SAED is an attack method that uses algorithm-dependent power signal changes. An adversary can extract a key using differential power analysis (DPA) of the subtraction operation. This paper indicates the theoretical rationality of SAED, and shows that its results are better than those of other methods. According to our experiments, only 256 power traces are sufficient to acquire one block of data. We verify that this method is more efficient than those proposed in previously published studies.

Keywords : RSA-CRT, Side Channel Attack, Equidistant Message Power Analysis, CPA(Correlation Power Analysis)

뺄셈연산의 이벤트 정보를 활용한 향상된 RSA-CRT 부채널분석공격 방법

박 종 연[†] · 한 동 국[††] · 이 옥 연[†††] · 김 정 녀[††††]

요 약

RSA-CRT는 RSA전자서명 알고리즘의 고속화 구현을 위해 가장 많이 사용되고 있는 알고리즘으로, 스마트디바이스에 사용되는 RSA-CRT 알고리즘의 물리적 취약성 검증을 위해 CRT의 각 단계 연산에서 다양한 부채널 분석 이론이 발표되어 왔다. 본 논문에서는 RSA-CRT 구현에 사용되는 뺄셈연산의 이벤트 정보를 활용하여 RSA-CRT의 reduction알고리즘을 분석하는 새로운 SAED(Subtraction algorithm Analysis on Equidistant Data)분석 방법을 제안한다. SAED분석 방법은 알고리즘에 의존한 전력 변화를 이용한 분석 방법이며, 뺄셈 연산을 차분전력분석 방법으로 분석하여 키를 찾아낸다. 본 논문은 SAED분석 방법의 이론적인 합리성을 증명하고, 실험적으로 기존의 분석 방법보다 향상된 결과를 가짐을 보인다. 실험 결과 256개의 파형만으로 하나의 바이트를 분석해 낼 수 있어, 기존 논문보다 효율적인 분석 방법임을 확인 할 수 있었다.

키워드 : RSA-CRT, 부채널 분석, 등간격 평문 전력 분석, 상관관계전력분석방법

1. 서 론

부채널 분석이 최초에 제안된 이후로 다양한 연구가 진행되어 왔다. 연구는 크게 두 가지 측면으로 이루어 지고 있다[1].

그 첫째는, 분석 방법에 대한 대응이다. 하나의 분석 방법이 나오면 그러한 알고리즘이 갖는 취약점을 대응하기 위한 대응책이 나오게 된다. 이러한 연구는 실제 공공기관에서 암호 장비의 안전성을 테스트 하는데 기준으로 삼기도 하며, 대응법에 대한 이론적인 안전성이 충분히 보장 되었는지에 대한 논의도 포함된다. 공격에 대한 대응은 알고리즘 뿐만 아니라 하드웨어 단계에서 공격 가정이 되는 전력 또는 전자파 소비 모델을 무력화 시키는 이른바, 부채널 정보 원천 봉쇄를 목표로 삼기도 하며, 단순히 특정 분석 방법에 대한 하드웨어적인 해결책을 내놓기도 한다.

둘째는 더욱 향상된 분석 방법에 대한 연구이다. 향상된 분석 방법 연구의 일반적인 연구는 분석 구분자(Distinguisher)에 대한 연구이다. 이 연구는 부채널 분석에 대한 연구가 이론

※ 본 연구는 방송통신위원회의 ETRI연구개발지원사업의 연구결과로 수행되었음(KCA-11921-05001).
※ 이 논문은 2012년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임(2012-0007285).
† 준 회 원 : 한국전자통신연구원 사이버융합보안연구단 연구원
†† 정 회 원 : 국민대학교 수학과 조교수
††† 정 회 원 : 국민대학교 수학과 교수
†††† 정 회 원 : 한국전자통신연구원 책임연구원 팀장
논문접수 : 2012년 9월 10일
수 정 일 : 1차 2012년 11월 22일
심사완료 : 2012년 12월 4일
* Corresponding Author : Dong-Guk Han(christa@kookmin.ac.kr)

적인 암호 분석 연구와 가장 큰 차이를 보이는 점이다. 공격자는 키를 찾겠다는 목표를 달성하기 위해 찾아내는 비밀 정보와 공격자가 찾고자 하는 값이 아닌 값들을 구분해 낼 수 있는 기준을 가지고 있어야 하며, 그 기준은 전력 소비 모델(Power consumption model)로 부터 출발한다.

구분자를 무엇으로 하느냐에 따라서 분석 방법은 DPA(Differential Power Analysis)[2], CPA(Correlation Power Analysis)[3], Template Attack[4], Improved DPA[5]과 최근 MIA(Mutual Information Analysis)[6]까지 다양한 기준과 이론이 제시되어 왔다.

향상된 분석 방법의 또 다른 연구는, 새로운 분석 방법 개발에 대한 연구이다. 대응법이 적용된 블록 암호에 대하여 다중라운드 공격을 하는 등의 분석 방법 개발이 있을 수 있으며[7], RSA[8], ECC[9]와 같은 공개키 알고리즘의 경우 수 많은 구현 방법에 특성화된 분석법을 개발 하는 것 역시 분석 방법 개발의 예이다. 예를 들어 RSA-CRT 알고리즘의 경우, 알고리즘의 특성을 이용하여 리덕션(Reduction) 단계를 분석하는 Boer[10]등의 MRED방법, 재조합 단계를 분석하는 Novak[11], Amiel[12]등의 방법이 있다. 특히 MRED의 경우 분석 결과의 해석에 있어서 다양한 유사키 특성이 이론적으로 연구되어 왔다[13,14].

본 논문은 이러한 공개키 알고리즘의 향상된 분석 방법의 연장선에 있으며, 기존에 제안되었던 대표적인 RSA-CRT알고리즘의 부채널 공격법인 MRED로부터 리덕션 알고리즘의 뺄셈 연산에서 발생하는 특성 이벤트가 전력정보로 노출 된다는 것을 이용하였다[13,16]. 본 논문에서 새로 제시한 분석 방법은 SAED(Subtraction algorithm analysis on equidistant data)라 불리며, 헤밍웨이트에 의존해서 나타나는 전력 값이 아닌 산술연산 과정에서 발생되는 이벤트를 확률적으로 분석하여 공격자가 원하는 정보를 얻어내는 방법이다. 또한 새로운 분석 방법에 대한 타당성을 밝히기 위해 우리는 이벤트 발생 기반의 전력모델 가정을 이용한다. 실험결과에 의하면 SAED분석 방법은 기존의 등간격 평문 전력분석보다 훨씬 적은 수의 파형만으로도 키 후보를 상당히 줄일 수 있음을 알 수 있었다.

본 논문의 2장에서는 RSA-CRT분석 방법 중 기존의 방법인 MRED전력 분석에 대해 소개하며, 3장에서는 MRED의 변형된 형태인SAED분석 방법에 대한 이론적 설명을, 4장에서는 소프트웨어 보드에서의 실제 실험을 통해 SAED 분석 결과를 보인다. 5장에서는 본 논문을 결론 짓는다.

2. 배경 지식 - MRED 전력분석

Boer[10]등은 RSA-CRT를 분석하는 방법을 발표하였다. 이 방법은 RSA-CRT에서 $x_p=x \bmod p$를 연산하는 초기 리덕션 단계를 분석하는 방법으로써, Equation (1)이 성립하는 성질을 이용한다.

$$x-i \bmod p = r-i \tag{1}$$

이와 같은 성질을 이용한다면, 공격자는 p값을 모르더라도, r값을 직접 추측하는 것이 가능하다. 공격자는 Equation (1)의 패턴을 전력으로 얻어내기 위하여 등간격 평문을 {x, $x-1$, $x-2$, …}순서로 입력하여, 결과 {r, $r-1$, $r-2$, …}의 전력을 수집할 수 있다. 최하위 바이트를 분석하고자 할 때에, 공격자가 추측하는 r의 최하위 바이트 값은 전체 키 후보의 개수 K, 전체 파형 수 N에 대하여 $\{v_{ij}\}=\{(j-i) \bmod 256 | i=0,\cdots,N-1,\ j=0,\cdots,K-1\}$로 나타난다.

Table 1. Computation of $v_{i,j}$

$v_{i,j}$	x_0	x_1	x_2	x_3	…	x_i
$v_{i,0}$	0	255	254	253	…	$-i \bmod 256$
$v_{i,1}$	1	0	255	254	…	$(1-i) \bmod 256$
$v_{i,2}$	2	1	0	255	…	$(2-i) \bmod 256$
…	…	…	…	…	…	…
$v_{i,255}$	255	254	253	252	…	$(255-i) \bmod 256$

Table 1은 한 바이트의 r을 추측하기 위한 $v_{i,j}$의 구성방법을 구체적으로 보여주며, 만약 r이 맞게 추측이 된다면, $v_{i,j}$의 헤밍웨이트는 r값이 나타나는 구간에서 의 CPA분석을 통해 찾아 낼 수 있을 것이다. Table 2는 $v_{i,j}$의 헤밍웨이트 연산 값인 $h_{i,j}$을 구체적으로 보여준다.

Table 2. $v_{i,j}$-based 8 bit Hamming weight : $h_{i,j}$

$h_{i,j}$	x_0	x_1	x_2	x_3	…	x_i
$h_{i,0}$	0	8	7	7	…	$HW(v_i,0)$
$h_{i,1}$	1	0	8	7	…	$HW(v_i,1)$
$h_{i,2}$	1	1	0	8	…	$HW(v_i,2)$
…	…	…	…	…	…	…
$h_{i,255}$	8	7	7	6	…	$HW(v_i,255)$

두 번째 바이트부터 연산은 동일한 분석 중간 값 $h_{i,j}$을 동일하게 이용한다. 하지만 입력 평문의 간격을 다르게 주어, 각 블록 분석에 특성화된 파형을 이용한 상위 바이트 공격이 가능해진다. Equation (2)는 Equation (1)의 일반화된 형식을 나타내며, 상위 바이트 분석 방법 역시 최하위 바이트 분석과 동일하게 적용될 수 있음을 보여준다. 하지만 수집하는 등차수열의 파형의 공차가 $(2^8)^k$단위로 바뀌어야 하므로, 첫 번째 바이트를 분석할 때와 동일한 CPA분석을 수행하되 간격을 다르게 주어 파형을 수집해야 한다.

$$x-i(256)^k \bmod p = r-i(256)^k \tag{2}$$

(단, $i(256)^d \le p$ 이며, k는 최하위 바이트부터 블록 순서다. 예를 들어 k가 0이면 공격자는 최하위 바이트를 분석하는 것이다.) 즉, 바이트 단위로 저장되는 r은 평문의 간격을 바이트 단위로 줌으로써, 분석이 가능해진다. 8비트 연산에서의 데이터 흐름에 의해 간격을 다르게 주는 선택 평문을 입력 값으로 갖게 되면 등간격의 바이트위치가 결정된다. 한편 Equation (2)는 $r>i(256)^k$가 성립한다는 조건하에서 성립되므로, $r>i(256)^k$가 만족되지 않으면 $r \le t(256)^k$인 t가 존

재하며, Equation (3)에 의해 비밀 소수 p가 연산 가능하다. 여기서 라는 k번째 바이트까지 공격자가 찾는 r값이며, p, q는 RSA알고리즘의 두 개의 비밀소수 값이다.

$$p = GCD(x_0 - F_k - i(256)^k, pq) \quad (3)$$

$$p = GCD(x - r, pq) \quad (4)$$

만약 공격자가 r을 계산해 냈다고 가정하면, Equation (4)의 최대공약수를 계산하여 비밀 소수 p를 찾는 것이 가능하며, 비밀 소수 p를 찾으면 RSA의 비밀 지수를 계산해 낼 수 있다. 즉, CRT의 $r=x \bmod p$값이 노출된다는 것은 RSA알고리즘이 깨지는 것을 의미한다.

3. SAED분석방법

3.1 SAED와 MRED분석 알고리즘의 차이점

알고리즘 1: MRED(v 번째 byte)
INPUT : $s_1,\cdots,s_t$ t개의 등간격 평문에 의한 파형 집합 C_v OUTPUT : r의 v번째 바이트 값
Step 1 For j from 0 to 255 *Step 1.1 set* $h_j=\{HW(i - j) \bmod 256 \mid j=1,2,\cdots,t\}$ *Step 1.2* $r_j=r(C_v, h_j)$ Step 1.3 *If* $j=0$ *then* $key=0$, $r_{key}=r_j$ *Otherwise; if* $r_{key}<r_j$ *then,* $r_v=j, r_{key}=r_j$ *Step 2 Return* r_v

알고리즘 2: SAED(v 번째 byte)
INPUT : $s_1,\cdots,s_t$개의 등간격 평문에 의한 파형 집합 C_v OUTPUT : 나머지 r의 v번째 바이트 단위 블록
Step 1 For j from 0 to 255 *Step 1.1 set* $n_j=\{(i - j) \bmod 256 \mid j=1,2,\cdots,t\}$ *Step 1.2* $r_j=r(C_v, n_j)$ *Step 1.3 If* $j=0$ *then* $key=0$, $r_{key}=r_j$ *Otherwise; if* $r_{key}<r_j$ *then,* $r_v=j, r_{key}=r_j$ *Step 2 Return* r_v

알고리즘 1은 MRED분석 방법을 알고리즘으로 나타낸 것이다. *Step1.1*은 Table 2의 $h_{i,j}$연산을 의미한다. r 의 값을 직접 계산 하는 방법이므로 일반적인 CPA분석 방법과 같이 알고리즘 중간 전력을 예측하여야 하며, 헤밍웨이트 전력 모델을 따른다.

알고리즘 2는 SAED를 이용한 RSA-CRT의 CPA분석 과정을 나타내는 알고리즘이다. 알고리즘 1과는 달리 *step1.1*에서 헤밍웨이트를 계산하지 않는다. SAED분석은 r값의 데이터에 의존하여 나타나는 전력을 분석하는 것이 아니기 때문이며, 중간 데이터에 의존한 전력 모델이 적용되지 않는다. 우리의 공격에서는 파형의 변화는 등간격의 중간 값에 의해 달라지는 알고리즘 변화에 의존한다고 가정한다. 우리는 헤밍웨이트를 계산하지 않고도 분석이 가능하게 된 원리를 설명 할 것이다.

3.2 Event발생 기반의 전력 모형

전력 분석에서 공격자가 알고자 하는 시간적인 위치에 존재하는 연산의 결과는 일반적으로 다음과 같은 전력 모델을 따른다고 알려져 있다[15].

$$P_{total}=P_{op}+P_{data}+P_{el.noise}+P_{const} \quad (5)$$

(P_{op} : 연산의존 전력, P_{data} : 데이터 의존전력, $P_{el.noise}$: 전자 노이즈, P_{const} : 고정 값)

일반적인 전력분석은 중간 데이터를 추측하여 분석하므로, Equation (5)와 같이 P_{data} 의 특성에 의존하며, 일반적으로 P_{data} 는 헤밍웨이트와 비례하여 증가한다고 알려져 있다. 하지만 이러한 분석 가정은 동일한 시간에 동일한 연산이 존재한다고 볼 때에 가능한 분석 방법이며, 많은 단순전력 분석은 동일한 연산이 발생하지 않는 특성을 이용한다[16].

하지만 현실적으로 다른 연산이 발생하는 것을 SPA로 분석해 내기는 상당히 어렵다. 따라서 우리는 Equation (6)과 같은 전력 모델을 기반으로 SPA가 아닌 CPA분석을 적용한다.

$$P_{total}=P_{eo}+P_{total.noise}+P_{const} \quad (6)$$

(P_{eo} : 사건 발생에 의존한 전력, $P_{total.noise}$: 전체 노이즈, P_{const} : 고정 값)

전체 전력은 전력의 모양에 상당한 영향을 주는 이벤트 정보 기반의 전력신호(Event Occurrence based Power signal)와 데이터와 연산에 의존한 전력을 포함한 전체 노이즈 그리고 상수 값으로 결정된다고 가정한다. 파형마다 다른 연산이 발생할 때에는 데이터에 의존한 전력 등은 무시할 정도로 작게 나타나기 때문이다.

3.3 SAED분석이론

알고리즘 3: 다정도 나눗셈 연산
INPUT : positive integers $x = (x_n,\cdots,x_1,x_0)_b$, $p=(p_t,\cdots,p_1,p_0)_b$ with $n \geq r \geq 1$, $p_t \neq 0$. OUTPUT : the quotient $q = (q_{n-t},\cdots,q_1,q_0)_b$ and remainder $r=(r_t,\cdots,r_1,r_0)_b$ such that $x=qp+r$, $0 \leq r < p$
Step 1 For j from 0 to $(n-t)$ do : $q_j \leftarrow 0$ *Step 2* while(x^3 pb^{n-t}) do the following : $q_{n-t} \leftarrow q_{n-t}+1$, $x \leftarrow x - pb^{n-t}$ *Step 3* For i from n down to $(t+1)$ do the following *Step 3.1* If $x_i=p_t$ then set $q_{i-t-1} \leftarrow b-1$; Otherwise set $q_{i-t-1} \leftarrow \lfloor (x_ib+x_{i-1})/p_t \rfloor$ *Step 3.2* While $(q_{i-t-1}(p_tb+p_t-1) > x_tb^2+x_{i-1}b+x_{i-2})$ do : $q_{i-t-1} \leftarrow q_{i-t-1}-1$ *Step 3.3* $x \leftarrow x - q_{i-t-1}pb^{i-t-1}$ *Step 3.4* If $x < 0$ then set $x \leftarrow x+pb^{i-t-1}$ and $q_{i-t-1} \leftarrow q_{i-t-1} - 1$ *Step 4.* $r \leftarrow x$ *Step 5.* Return (q,r)

RSA-CRT의 초기 리덕션 알고리즘은 몫과 나머지를 계산하는 나눗셈 알고리즘과 같다. 알고리즘 3은 가장 흔히 사용되는 다정도 나눗셈 알고리즘으로써, 덧셈 그리고 뺄셈 곱셈등의 연산으로 구성되어 있다[17]. 공격 타겟 알고리즘인 $r=x \bmod p$에 적용하는 알고리즘은 알고리즘 3과 같은 일반적인 나눗셈 연산이라고 가정한다. SAED분석에서 주목하는 단계는 step3.3의 뺄셈 연산이다. $x \leftarrow x - q_{i-t-1}pb^{i-t-1}$는 변수 x에 몫 q와 비밀 키 p의 곱을 빼는 연산을 수행하게 되는데, $q_{i-t-1}pb^{i-t-1}$은 고정된 상수라는 특징이 있다. 그 이유는 비밀소수 p는 RSA프로토콜에서 반드시 고정되어 있으며, 몫은 사실상 고정된 값이기 때문이다.

$$(r-i)+p\times q=x-i \quad (7)$$
$$\Leftrightarrow (r-i)+p+(q-1)\times p=x-i$$

몫 q가 고정되어 있다고 볼 수 있는 이유는 등간격의 평문은 몫에 영향을 주지 않기 때문이다. 몫이 변화할 때의 현상을 보면 다음과 같다. Equation (7)은 Equation (1)을 몫과 나머지로 정리한 수식이며, 최하위 바이트 공격에서의 몫이 q에서 $q-1$로 변화시켰을 때에 결과를 보여준다. $(r-i)+p$는 나머지에 해당하며, 몫과 나머지 정리에서 나머지는 나누는 수 p보다 작아야 하며, $(r-i)+p<p$가 성립해야 한다. 따라서 몫이 바뀌기 위해서는 $r<i$가 만족되야 하며, i는 분석 필요 파형수 만큼 증가하므로, 분석을 위한 등간격의 파형수가 r보다 커야 한다. 일반적인 RSA-CRT알고리즘에서 정수 r의 크기는 512비트, 1024비트등이 주로 사용되므로 최하위 바이트 공격에서의 몫을 변화시키기 위한 파형 수는 대략 2^{512}개라는 것을 알 수 있다. 이정도 수치는 현실적인 분석 파형 크기를 넘어서므로 최하위 바이트 공격에서의 몫 q는 등간격 평문 공격에서 고정된다고 볼 수 있다.

$$(r-i(256)^k)+q\times p=x-i(256)^k \quad (8)$$
$$\Leftrightarrow (r-i(256)^k)+p+(q-1)\times p=x-i(256)^k$$

최하위 바이트인 경우 몫을 변화시키기 위해 필요한 파형수가 크지만, 상위 바이트를 공격할 때는 필요한 파형수가 줄어든다. 그래서 상위 바이트를 분석 할 때에 몫을 변화시키는 파형이 현실적인 수준에서 결정되는지 확인해야 한다. Equation (8)은 Equation (7)을 상위 바이트 공격까지 확장한 것이다. 몫과 나머지 정리에 의해 $r-i(256)^k+p<p$가 성립되어야 하며, $r<i(256)^k$가 만족되는 i까지 q가 상수로 유지된다. 예를 들어, 512비트 r을 찾을 때 몫을 변화시키기 위한 파형 수는 $2^{512}-8^k$개로 계산되며, 가장 적은 파형수로 몫을 변화시킬 수 있는 최상위 바이트를 찾는다고 가정하면 2^8개의 파형만으로도 몫을 변화시키지만, 상위 바이트의 경우 2장에서 언급한 바와 같이 전수조사가 가능하므로 부채널 분석이 적용될 필요가 없다. 따라서 SAED분석에서 몫 q는 상수라고 결론 내릴 수 있다.

몫이 고정된다면 $x \leftarrow x - q_{i-t-1}pb^{i-t-1}$에서 빼는 수 q,p,b 모두가 상수 값이기 때문에, 마지막 단계에서의 $x - q_{i-t-1}pb^{i-t-1}$은 $x-c$로 단순하게 표현 가능하다. 정리하면, $r=x \bmod p$연산인 알고리즘 3의 r을 연산하는 마지막 단계가 *step3.3*의 $x \leftarrow x - q_{i-t-1}pb^{i-t-1}$로 표현가능하고, q와 i,p,t가 고정이므로 $r \leftarrow u-c$ 로 간단하게 표현 가능하다. 따라서 $r=x \bmod p$연산의 마지막은 상수 c에 의한 뺄셈 연산이며, 본 논문에서의 분석 목표가 되는 연산이다.

알고리즘 4: 큰 수 뺄셈
INPUT : positive integers u and c, each having $n+1$ base b digits, with $u \geq c$ OUTPUT : $u - c$ radix 256 representation.
Step 1 $BR=0$ *Step 1.1* For i from 0 to n do the following *Step 1.2* $r_i = (u_i - c_i + BR) \bmod 256$ *Step 1.3* If$(u_i - c_i + BR) \geq 0$ then $BR=0$; otherwise $BR=-1$ *step 2* Return$((r_n, r_{n-1}\cdots, r_1, r_0))$

$r=u-c$를 분석하기 위하여 뺄셈 알고리즘을 살펴보면 알고리즘 4와 같으며[18], SAED는 헤밍웨이트 계산 없이 r값을 예측해야 한다.

알고리즘 4의 *step2.2*에서 $(u_i - c_i + BR) \geq 0$인 경우와, 그렇지 않은 경우에 따라서 파형이 달라짐은 3.2절의 전력모델에 따른다. 따라서 중간 값에 의해 파형의 차이가 발생하기 때문에 공격자는 CPA분석을 할 수 있다. 그러므로 공격자에 의해서 추측된 r 값에 따라서 BR이 발생하는 파형과 발생하지 않는 파형을 구분해 낼 수 있다면, 공격자는 키를 찾아낼 수 있다.

3.4 r 값을 추측하기 위한 확률적인 접근

공격자가 찾는 값의 하나의 블록이 r_i라고 두자, r_i을 계산하기 위한 바이트단위의 뺄셈 연산은 $r_i=u_i-c_i$이다. 다음 정리는 r의 값과 BR의 발생 사이의 관련성을 보여준다.

정리 1]
(Borrow(BR)과 Carry(CR)의 동시 발생적 특성)
a, b, c의 다정도 뺄셈 $a-c=b$의 한 블록 연산에서 Borrow값이 발생하면, 다정도 덧셈 $a=b+c$의 동일한 블록 연산에서 Carry가 발생한다. 그리고 그 역도 성립한다.

증명) $a-c=b$의 한 블록 연산에서 Borrow가 발생했다고 가정하자. 최하위 블록의 연산이라 두어도 일반성을 잃지 않으므로, $a_0-c_0=b_0 \bmod 256$의 연산을 생각해보자. Borrow가 발생 했으므로, $a_0<c_0$이며 $a_0=c_0+b_0-256$라고 볼 수 있다. 그 의미는 $a_0+256=b_0+c_0$이므로 덧셈 연산에서Carry가 발생하게 된다. 역으로 $a_0=b_0+c_0$에서 Carry가 발생했다고 가정하자. 즉, $a_0+256=b_0+c_0$이다. $a_0-b_0+256=c_0$이고, $c_0<256$이므로 $a_0-b_0+256<256$이다. 즉, $a_0-b_0<0$이다. 즉 a_0-b_0연산에서 Borrow가 발생하게 된다.

정리 1은 $r_i = u_i - c_i$에서 발생하는 BR의 발생이 $r_i+c_i=u_i$

연산으로 바꾼 뒤 넘겨주기 값(Carry,*CR*)이 발생하는 것으로 바꾸어도 동일한 파형의 변화를 추측 할 수 있다는 것을 의미한다. 하지만 공격자는 c를 알 수 없기 때문에 언제 이러한 파형의 변화가 발생하는지 정확히 예측하는 것이 불가능하다. 따라서, 우리는 정확하게 *BR*의 발생을 예측할 수 없지만, 높은 확률로 발생될 것이라고 예측되는 지점과, 그에 대한 발생 기대 값을 계산하여 확률적인 기준을 만들 것이다.

$$P(\text{Carry Occurence}) = P(256<r_i+c_i)=(256-r_i<c_i)=r_i/256 \quad (9)$$

Equation (9)는 추측한 변수 r_i 와 고정된 c_i 에 대하여 Carry가 발생할 확률을 계산한 것이다. r_i는 공격자가 예측한 임의의 값이므로 알려진 값(known value)이며, 0부터 255사이의 값을 갖는 모르는 c_i 값에 대한 확률을 계산하면 된다. 예를 들어 공격자의 추측 r_i가 0이라면 P(Carry Occurrence)=0 이다. 추측 r_i 가 커질수록 연산이 Carry를 가질 가능성이 커짐을 직관적으로 알 수 있다. Equation (9)를 이용하여 분석 수열을 만들면 Equation (10)과 같이 구성이 가능하다.

$$A_j = \{j\bmod 256/256, j-1 \bmod 256/256, j-2\bmod 256/256,\cdots,1/256,0/256,\cdots\} \quad (10)$$

$$r(A,B) = r(xA, yB) \quad (\text{where, } A,B \text{ are random variables, } x,y\in Z) \quad (11)$$

$$A_j\times 256 = n_j = \{j\bmod 256, j-1\bmod 256, j-2\bmod 256,\cdots,1,0,\cdots\} \quad (12)$$

Equation (10)은 상관계수의 특성 Equation (11)에 의해 최종적으로 Equation (12)와 같은 형태로 단순화 시킬 수 있다. 이 집합은 알고리즘 2 *step1.1*의 CPA를 위한 중간 값 집합을 구성하는 방법과 일치함을 알 수 있다. 만약 공격자가 추측하는 *r*값이 맞는다면 *BR*발생을 틀린 키 보다 더 높은 확률로 추측하였기 때문에 상대적으로 더 높은 상관계수를 찾을 것이다.

하지만 이러한 확률 계산은 상수 c_i가 0부터 255사이에 균등분포(Uniform Distribution)을 따를 때에 가능하므로 *c*의 결과의 분포를 확인해야 한다.

3.5 분석을 위한 상수의 성질 및 조건

상수의 성질에 따라서 분석이 큰 영향을 받기 때문에 상수 c_i의 조건을 명확히 결정하고, 이러한 조건이 실제 분포와 분석 가정에 얼마나 영향을 미치는지 확인해 보아야 한다.

조건 1 - 상수는 0이 되면 안 된다. c_i가 0이라면 $r_i=u_i-c_i$연산에서 BR이 발생하지 않는다. 그러므로 공격자의 확률 연산은 c_i가 0이 아니라는 기본 가정에서 출발하며, c_i가 0이라면 위와 같은 공격 방법을 적용 할 수 없다.

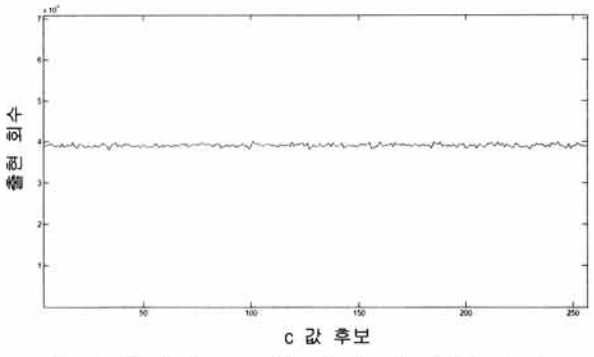

Fig. 1. Distribution of c_0(Least significant Byte of *c*)

조건 2 - 상수는 균등분포를 따른다. Fig. 1은 임의의 두 수의 곱셈을 천만번 수행하여 결과의 최하위 바이트 c_0의 분포를 나타낸 것이다. 그림에서 알 수 있듯이 상수의 분포는 균등하게 됨을 알 수 있다. 일반적으로 최하위 바이트는 0이 나타날 확률이 가장 높다고 생각 할 수 있으나, 상수는 소수 *p*와 임의의 수의 곱셈으로 되어 있으며, p의 최하위 바이트는 0이 아니므로 계산 결과가 0이 나타날 확률은 다른 수의 확률보다 높지 않다. 따라서 조건 1에 의해 SAED 분석으로 키를 찾지 못할 확률이 1/256이라고 볼 수 있으며, Equation (9)가 적용 가능하다.

3.6 SAED 분석 시뮬레이션

Equation (12)의 중간 값 집합에 의해 분석 가능한지 알아보기 위해서 실제 *r* 값을 등간격 평문을 이용하여 분석한다.

Table 3은 초기 평문 x에 대하여 $r=x\bmod p$를 연산하여 최하위 바이트 r_0=135, c_0=120으로 두고, *BR*이 발생하는 경우와 그렇지 않은 경우를 등간격 수열 형태로 나타낸 것이다. 이와 같이 설정된 r_0와 c_0에 대하여, SAED분석 알고리즘인 알고리즘 2의 *step1.2*의 $\rho(C_0,n_j)$을 계산한다. B0은 최하위 바이트의 등간격 파형이며, A_j는 $_j$번째 키 후보에 대한 중간 값 집합이다.

Table 3. Borrow occurrence of $r_0=u_0-c_0\bmod 256$

r_0	135	134	···	0	255	···	136	135	
c_0	120	120	···	120	120	···	120	120	
BR	*N*	*N*	···	*N*	*Y*	···	*Y*	*N*	
n_{135}	135	134	···	0	255	···	136	135	···

Fig. 2는 $\rho(C_0,n_i)$의 결과이다. 가장 높은 상관계수를 보이는 값이 135임을 알 수 있는데, 모든 틀린 *r*값도 135를 중심으로 높은 상관계수가 나타난다. 이러한 특징은 서로 다른 키 후보 j_x, j_y에 대하여 중간 값 집합 n_{jx} 와 n_{jy}는 0이 아닌 상관관계를 가지는 집합을 적용하기 때문에 발생하는 것이다. 더 자세히 키 후보와 후보 사이의 상관도를 계산해 보면 Fig. 3과 같이 3차원 그래프로 표현된다.

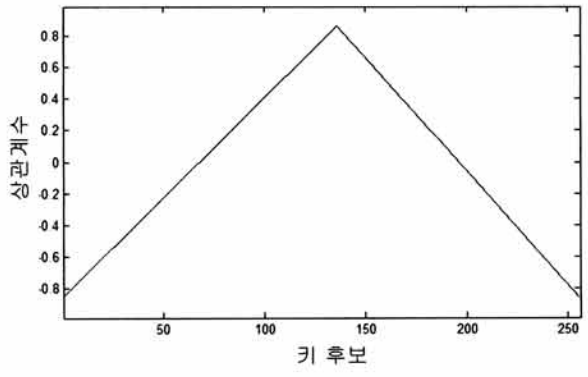

Fig. 2. Correlation coefficient of n_j with borrow occurrence

Fig. 2는 고정된 키 결과의 중간 값에 대하여 256개 키의 상관계수를 본 만면, Fig. 3에서는 256개의 키 후보에 대한 256개의 키 후보의 상관도를 보여준다. 가운데 붉은 선은 두 키가 일치한 자기 상관계수 1을 나타내며, 다른 키의 경우 최대 0.9767이상까지 나타난다. Fig. 2는 Fig. 3의 특정키를 선택하여 자른 2차원 단면도이다. 이것은 모든 테스트 환경에서 키를 바꾸어 실험하였을 때에도 반드시 적용되는 성질이다.

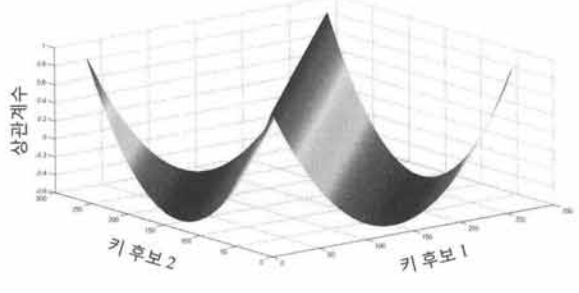

Fig. 3. Correlation coefficient between key candidates

4. SAED 실험결과

4.1 실험환경

Table 4. Experimental environment of SAED analysis

전력수집 툴	Digital oscilloscope Lecroy
전력 분석 보드 (Embedded Simulator)	MSP430 / Arm7 ETRI Sacrf
Sampling rate	MSP430 : 250MS/sec Arm : 1GS/sec
분석 알고리즘	등간격 평문에 의한 8비트 리덕션 알고리즘 $x - i(256)^k \bmod p = r - i(256)^k$
알고리즘의 변수 크기	32바이트 입력 x 16 바이트 소수 p

Table 4는 SAED분석 실험 환경이다. 8비트 리덕션 연산을 Sacrf의 MSP430 소프트웨어보드와 Arm 소프트웨어 보드에서 실험하였다. 입력 평문은 SAED분석 방법에 따라서 등간격 평문을 사용하였으며, 등간격 평문을 입력하여 나타나는 전력을 수집하여 분석하였다.

4.2 SAED분석 결과와 해석

Fig. 4는 r의 최하위 바이트 256개 키 후보에 대한 분석 구간에서의 최대 상관계수를 나타낸 것이다. 분석해야 할 r_0 값과 c_0는 시뮬레이션에서의 설정과 같이 135, 120으로 설정되어 있다. Fig. 4는 SAED분석 결과이며 시뮬레이션 결과인 Fig. 2와는 다른 다소 다른 형태를 나타낸다. Fig. 2는 하나의 가장 높은 상관도를 갖는 키를 찾아낼 수 있는 반면에, Fig. 4는 두 개의 r_0로 추정되는 후보 키가 존재한다는 특징이 있다. 두 개의 키 후보 중 하나는 맞는 r_0값이며 하나는 키와 유사한 분석 결과를 갖는 틀린 키일 것이다. Fig. 6은 리덕션 알고리즘의 마지막 뺄셈 연산에서의 256개 전체 키 후보의 상관계수를 그림으로 나타낸 것이며, Fig. 6은 Fig. 5의 가장 높은 상관계수를 갖는 영역을 확대하여 나타낸 것이다. 가장 높은 상관계수를 갖는 동시에 가장 낮은 상관계수를 갖는 두 개의 키는 135와 255이며, 대칭적인 형태를 갖는 것이 특징이다.

대칭적인 형태는 이론적으로 원인 규명이 가능하다. 공격자는 r을 찾는 방법과 같이 BR의 발생을 이용하여 u를 찾을 수 있는데, 그 중간 값 연산의 확률적 방법은 r을 찾는 방법과 동일하다.

$$P(BR)=P(u_i<c_i)=(255-u_i)/256 \tag{13}$$

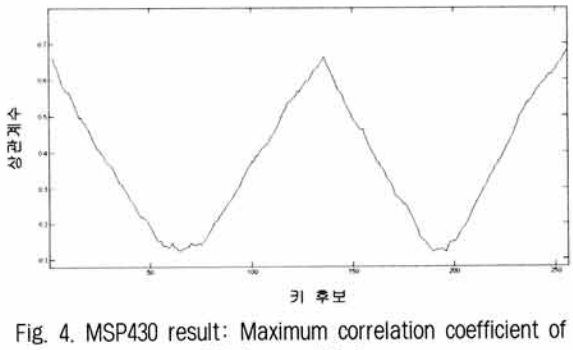

Fig. 4. MSP430 result: Maximum correlation coefficient of 256 key candidates

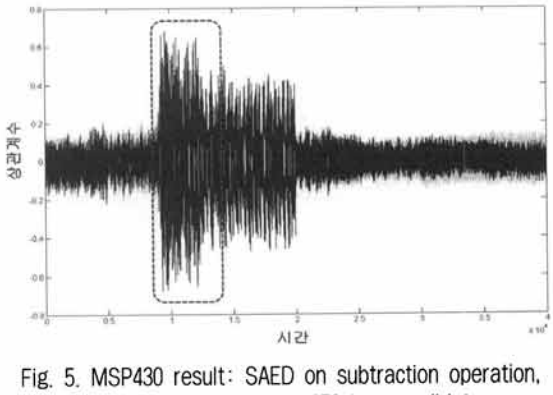

Fig. 5. MSP430 result: SAED on subtraction operation, correlation coefficient of 256 key candidates

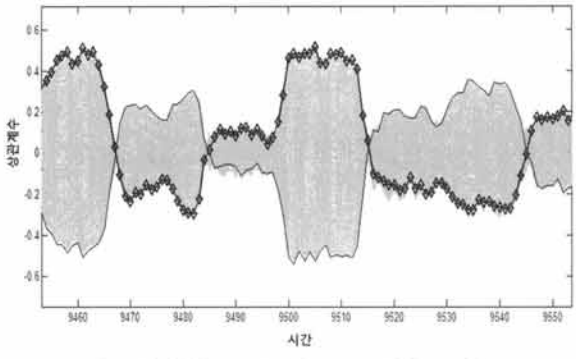

Fig. 6. MSP result: Zoom-in of Figure 5

Equation (13)은 $r_i=u_i-c_i$연산에서 균등 분포를 가지는 c_i에 대하여 BR이 발생할 경우에 대한 확률이다. 이 확률을 Equation (12)와 비교 했을 때, 255에 대한 보수 관계에 있음을 알 수 있다. u를 찾기 위한 수열을 S라고 한다면 S_u-i로 바뀌는 등간격 평문에 의한 S_u의 구성은 Equation (14)와 같다.

$$S_u = \{u \bmod 256,(u+1) \bmod 256,(u+2) \bmod 256,\cdots,255,0,1\cdots\} \quad (14)$$

u_i의 분석을 위한 수열의 255보수 형태가 SAED분석 중간 값 n_i의 형태라고 볼 수 있다. 따라서 이론적으로 S_{255}가 중간 값으로 사용되었다면, 분석 결과로 255를 찾을 수 있으며, n_{135}을 중간 값으로 취하면, 분석 결과로 135가 가장 높은 상관계수가 나오게 될 것이다. 알고리즘 2의 *Step1.1*의 상관계수는 동일한 파형에 대하여 서로 다른 중간 값으로 인해 $\rho(C_v,n_{135}) = \rho(C_v,S_{255}) = -\rho(C_v,n_{255})$ 로 계산 될 것이다. 따라서 Fig. 4와 같이 대칭적인 성질과, Fig. 5에서와 같은 높은 상관도를 갖는 분석 결과를 보임이 이론적으로 규명 가능하다.

MSP 소프트웨어 보드와 같이 Arm보드에서의 실험도 같은 결과를 나타냈다. Fig. 7,8,9은 Fig 4,5,6에 대한 실험의 Arm보드 상에서의 결과를 보여준다. 결과에서 알 수 있듯, MSP의 유사키 패턴과, 대칭적인 상관계수 형태를 그대로 가져감을 알 수 있으며, 이러한 현상의 이론적 타당성을 입증한다.

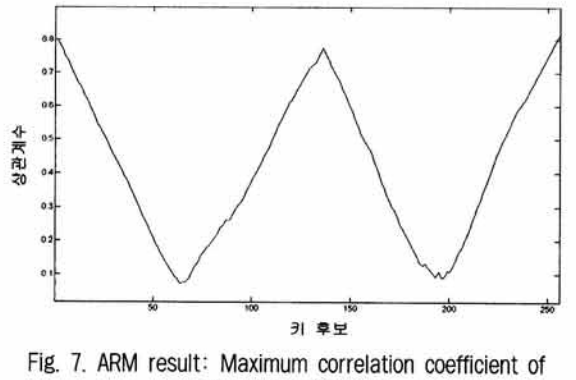

Fig. 7. ARM result: Maximum correlation coefficient of 256 key candidates

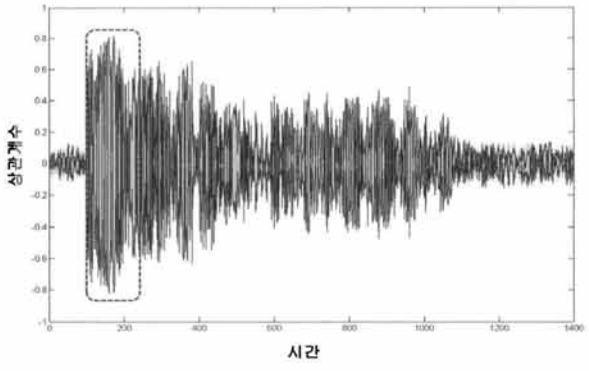

Fig. 8. ARM result: SAED on subtraction operation, correlation coefficient of 256 key candidates

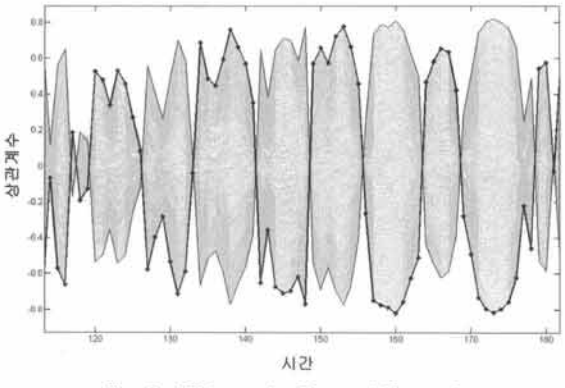

Fig. 9. ARM result: Zoom of Figure 8

높은 상관계수와 낮은 상관계수가 넓은 폭으로 진행되는 것이 키 값이라고 한다면 최대 상관계수로 분석 결과를 측정하는 것이 아닌, 최대 상관계수와 최소 상관계수 중 절대값이 가장 큰 상관계수를 가장 유력한 키 값으로 정할 수 있다. Fig. 10은 Fig. 7의 결과와 분석 구간에서의 최대 절대 상관계수 그림을 동시에 나타낸 것이다. 절대 상관계수 결과가 밑으로 양의 상관계수와 음의 상관계수를 동시에 갖는 것은 대칭적 성질에 의한 분석 결과가 때때로 음의 값에서 가장 큰 절대 상관계수 값을 가짐을 보여준다.

BR이 존재하는 파형과 그렇지 않은 파형의 경우 실제로 어떻게 다른지를 확인해보자. Fig. 11은 BR이 발생한 경우

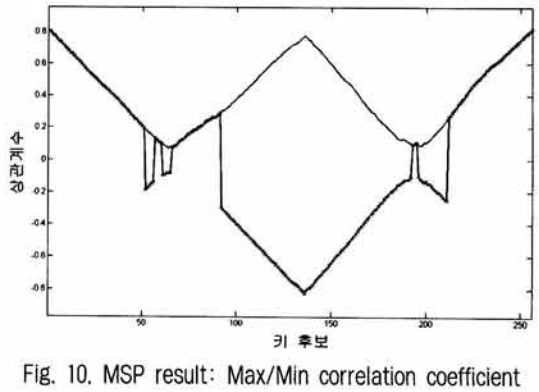

Fig. 10. MSP result: Max/Min correlation coefficient

와 그렇지 않은 경우의 파형 비교이다. *BR*이 발생한 점선의 경우 발생하지 않은 실선과 비교하여 파형의 정렬이 맞이 않으며, 상대적으로 높은 전압과 낮은 전압이 교차적으로 나타나게 된다. 따라서 높은 상관도와 낮은 상관도를 교차적으로 장시간 동안 나타나게 된다.

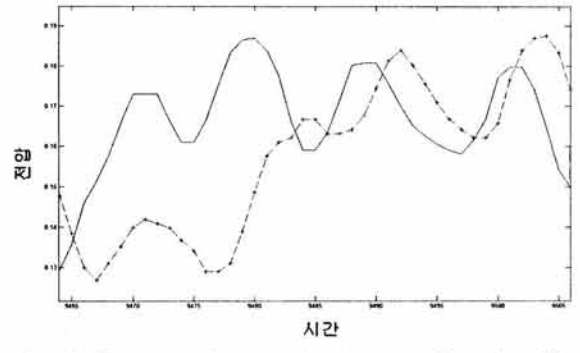

Fig. 11. Power signal comparison between BR and no BR

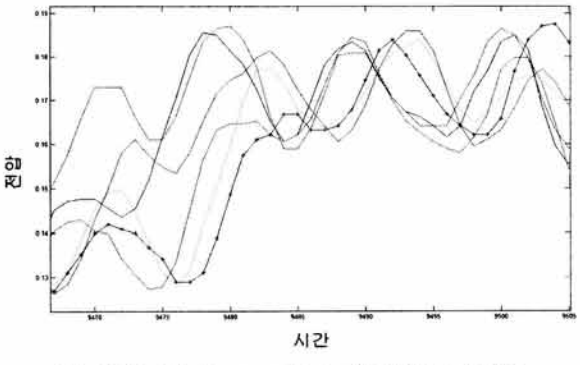

Fig. 12. 6 signals comparison with BR occurrence

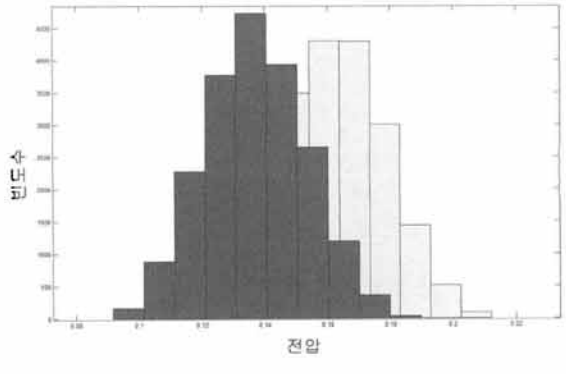

Fig. 13. Distribution of borrow occurrence(black)
Distribution of no borrow occurrence(gray)

Fig. 11과는 달리 *BR*의 발생을 단순전력분석으로 구분하는 것은 쉽지 않다. Fig. 12는 Fig. 11의 파형을 포함하여 동일한 시간대의 파형 6개를 나타낸 것이다. 파형이 중첩되고 *BR*의 발생 여부에 대한 명확한 기준이 존재하지 않으므로 SPA기반의 분석 방법이 적용되기는 쉽지 않다. 따라서, 통계적인 기법을 이용한 SAED분석을 반드시 적용해야 한다.

Fig. 13은 Fig. 12의 9450포인트에서의 *BR*이 발생했을 경우(검은색)와 그렇지 않은 경우(회색)의 파형 20000개의 분포를 나타낸 것이다. 그림에서 알 수 있듯이, 파형의 분포가 확연히 달라짐을 알 수 있으며, 분포의 차이는 파형의 평균값을 변화시키고 Pearson 상관계수에 영향을 미치게 되므로, CPA분석을 기반으로 하는 SAED분석을 가능하게 한다.

4.3 SAED vs MRED

SAED로 분석했을 때, 분석 성능 면에서 기존의 분석법보다 얼마만큼의 효율을 갖는지 기존 분석 방법인 MRED와 비교해보았다.

Table 5. MSP result: Performance comparison, MRED vs. SAED

	1st byte	2nd byte	3rd byte
SAED (MSP)	256	256	256
MRED (MSP)	over 2800	over 3000	over 1800

Table 6. ARM result: Performance comparison, MRED vs. SAED

	1st byte	2nd byte	3rd byte
SAED (ARM)	256	256	256
MRED (ARM)	over 1100	over 800	over 1300

Table 5와 Table 6은 MRED분석과 SAED분석 성능의 MSP430보드와 Arm보드에서의 결과 비교표이다. MRED와 SAED분석은 동일한 파형을 이용하여 중간 값을 다르게 계산하여 분석한 결과이므로, 분석 성능에 대한 직접적인 비교가 가능하다. 실험 결과 최소 천 개 수준에서 최대 수천 개에서 키가 나오는 헤밍웨이트 모델의 MRED분석에 비해, SAED에서는 오직 256개의 파형만으로도 최대 성능을 보여주며, 키 형태가 충분히 관찰된다. 하지만 하나의 가장 높은 키가 찾아지는 것이 아니므로, 분석 결과의 해석이 중요하며, 최대 상관계수 결과의 극대점을 2개의 키 후보로 최종적으로 결정하며, 상위 바이트까지 확장하여 분석한다. MRED와 비교하여 CPA분석에 의한 결과가 현저히 적은 파형으로도 구분이 되므로 소음이 많은 분석 환경에서 특히 강점을 가질 것으로 예측된다.

5. 결 론

본 논문은 RSA-CRT에 대한 새로운 분석 방법을 제시하였으며, 특히 헤밍웨이트에 의한 전력이 아닌 이벤트 발생에 따른 전력 변화를 확률적인 분석을 통해 예측하고 키를 찾아낼 수 있음을 이론적으로 검증하고 실제 실험으로 보였다. SAED는 기존의 전력 모델을 이용하지 않은 분석이라는 점에서 헤밍웨이트에 영향을 받지 않도록 하는 하드웨어 대응방법을 깰 수 있는 분석 방법이며, 아주 적은 량의 파형 수로도 키 후보를 현저히 줄여 준다는 점이 큰 장점이다. 또한 8비트 연산에서뿐만 아니라 32비트, 64비트 등의 큰 레

지스터 크기를 갖는 연산으로부터도 동일한 분석 논리가 적용 가능하다는 점에서, 레지스터 크기가 분석 성능에 영향을 주는 기존의 분석 방법보다 더 이점이 있다고 할 수 있다. SAED분석은 기존의 MRED분석의 선택 비트 유사키의 패턴의 연구로부터 출발하였으며, 유사키의 정확한 원인 규명으로부터, 새로운 분석 방법을 도출해 낼 수 있는 성공사례이다.

하지만 우리의 분석 방법은 키 주변의 유사키에 대하여 구분이 불가능할 수 있다는 점, 상수가 0이면 분석이 불가능 하다는 점, 그리고 결과적으로 키 후보를 좁히는 효과가 있을 뿐, 정확히 하나의 키 후보로 결정될 수 없다는 점에서 한계가 있으며, 더욱 향상된 분석 결과를 가지기 위해 이러한 한계점을 극복하는 연구가 진행되어야 할 것이다.

참 고 문 헌

[1] P. Kocher, "Timing attacks on implementations of Diffe-Hellman, RSA, DSS and other systems", Advances in Cryptology - CRYPTO 96, Santa Barbara, California, LNCS 1109, pp.103-113, 1996.

[2] P. Kocher, J. Jaffe, and B. Jun, "Differential power analysis", CRYPTO 1999, LNCS 1666, Springer-Verlag, pp.388-397, 1999.

[3] E. Brier, C. Clavier, and F. Olivier, "Correlation power analysis with a leakage model", CHES 2004. LNCS 3156, Springer-Verlag, pp.16-29, 2004.

[4] S. Chari, R. Rao, P. Rohatgi, "Template Attacks", CHES 2002, LNCS 2523, pp.13-28, 2003.

[5] D. Agrawal, P. Rohatgi, and J.Rao. , "Multi-channel attacks", CHES 2003. LNCS 2779, Springer-Verlag, pp.2-16, 2003.

[6] E. Oswald and P. Rohatgi, "Mutual Information Analysis", CHES 2008, LNCS 5154, Springer-Verlag, pp.426-442, 2008.

[7] J. Zhou and M. Yung, "Principles on the Security of AES against First and Second-Order Differential Power Analysis" ACNS 2010, LNCS 6123, pp.168.185, 2010.

[8] Rivest R, Shamir A, Adleman L. "A method for obtaining digital signatures and public-key cryptosystems". Commun ACM, pp.120-126, 1978.

[9] N.Koblitz, "Elliptic Curve Cryptosystem", Mathematics of Computation, ISSN 1088-6842.

[10] B.D. Boer. K. Lemke, and G.Wicke, "A DPA attack against the modular reduction within a crt implementation of RSA", CHES 2002, LNCS, Vol.2523, Springer-Verlag, pp.228-243, 2002.

[11] R. Novak "SPA-Based Adaptive Chosen-Ciphertext Attack on RSA Implementation" PKC 2002, Springer-Verlag Berlin Heidelberg 2002, LNCS 2274, pp.252-262, 2002.

[12] F. Amiel, B. Feix, and K. Villegas, "Power Analysis for Secret Recovering and Reverse Engineering of Public Key Algorithms", SAC 2007, Springer-Verlag, pp.110-125, 2007.

[13] J.Y Park, D.H Han, O. Yi, D.H Choi, "Ghost key patterns with Equidistant Chosen Message attack on RSA-CRT", 2011 IEEE International Carnahan Conference, IEEE/IET Electronic Library (IEL), VDE VERLAG Conference Proceedings, pp.1-5, 2011.

[14] R. Rao, P. Rohatgi, H. Scherzer, S. Tinguely , "Partitioning attacks: or how to rapidly clone some GSM cards", IEEE Symposium on Security and Privacy 2002. Proceedings. 2002.

[15] S. Mangard, E. Oswarld, T. Popp, "Power Analysis Attacks, revealing the secret of smart cards", Springer, ISBN.0387308571, 2007. 12.

[16] S. Mangard, "A Simple Power-Analysis (SPA) Attack on Implementations of the AES Key Expansion", ICISC 2002, LNCS 2587, Springer-Verlag, pp.343-358, 2003.

[17] A.J Menezes, PaulC.van Oorschot and S.A Vanstone, "Handbood Applied Cryptography", CRC press ISBN : 0-8493-8523-7, 1996.

박 종 연

e-mail : flysohigh@etri.re.kr
2010년 국민대학교 수학과(학사)
2012년 국민대학교 수학과(이학석사)
2012년~현 재 한국전자통신연구원 사이버융합보안연구단 연구원
관심분야 : 부채널 분석, 암호알고리즘 구현, 암호 파일시스템 설계 등

한 동 국

e-mail : christa@kookmin.ac.kr
1999년 고려대학교 수학과(학사)
2002년 고려대학교 수학과(이학석사)
2005년 고려대학교 정보보호대학원 박사 (공학박사)
2004년~2005년 일본 Kyushu Univ. 방문연구원
2005년~2006년 일본 Future Univ.-Hakodate, Post. Doc.
2006년~2009년 한국전자통신연구원 정보보호연구본부 선임연구원
2009년~현 재 국민대학교 수학과 조교수
관심분야 : 공개키 암호시스템 안전성 분석 및 고속 구현, 부채널 분석, RFID/USN 정보보호 기술 등

이 옥 연

e-mail : christa@kookmin.ac.kr
1988년 고려대학교 수학과(학사)
1990년 고려대학교 수학과(이학석사)
1996년 University of Kentucky(이학박사)
1996년~2001년 한국전자통신연구원 선임연구원
2000년~2001년 한국전자통신연구원 팀장
2000년~현 재 국민대학교 수학과 교수
관심분야 : 스마트 그리드 보안, 무선보안, 4G보안, 스마트워크 보안 등

김 정 녀

e-mail : jnkim@etri.re.kr
1987년 전남대학교 전산통계학과(학사)
1996년 OSF/RI 공동연구 파견(미국)
2000년 충남대학교 컴퓨터공학과 (공학석사)
2004년 충남대학교 컴퓨터공학과 (공학박사)
2005년 Univ. of Califonia. Irvine Post-Doc
현 재 한국전자통신연구원 책임연구원 팀장
관심분야 : 시스템 네트워크보안, 보안 OS, 바이오보안 등