On the anonymization of Cocks IBE

BalkanCryptSec 2014

Gheorghe Adrian Schipor

Departament of Computer Science
Alexandru Ioan Cuza University, Iași

17 october, 2014

Contents

1. Identity Based Encryption
2. The Cocks IBE scheme
3. The anonymity of Cocks IBE scheme
4. The Ateniese-Gasti scheme
5. A new method of anonymization
6. Practical aspects
7. Conclusions

Contributions

We propose a new method of anonymization of Cocks IBE scheme, more efficient than the scheme proposed by Ateniese and Gasti in 2009:

- we reduce **the time** of the (de)anonymization of ciphertext with **more than a half** of the (de)anonymization time of their scheme;
- **the ciphertext expansion** of our scheme is **almost equal** with the ciphertext expansion of **Cocks IBE scheme**.

Identity Based Encryption

Type of public key cryptography in which the public key is computed using an identity.

- Introduced by Adi Shamir in 1984.
- Does not requires certificates.
- A third party, PKG(Private Key Generator), generates a set of public parameters and a set of private parameters.
- The public key can be calculated by anybody using only the public parameters and the identity; the private key can be calculated only by the PKG.
- The keys can be calculated only when needed.

The Cocks IBE scheme

Introduction

- Developed by Clifford Cocks in 2001.
- Encrypts each bit of plaintext as a pair of integers modulo a big composite number.
- Is based on the quadratic residuosity problem.
- Does not provide anonymity.

The Cocks IBE scheme

Setup

- $n = pq$, p and q are two big primes: $p, q \equiv 3 \mod 4$;
- $H : \{0,1\}^* \mapsto \mathbb{Z}_n$ is a hash function;
- $H(ID) = a$, the Jacobi symbol $(\frac{a}{n}) = +1$ for any identity ID;
- n is the public parameter of PKG, p, q are the private parameters of PKG and a is the public key corresponding to the identity ID.

The Cocks IBE scheme

Extraction

- the private key r is calculated as:

$$r = a^{(\phi(n)+4)/8} \mod n$$

The Cocks IBE scheme

Encryption

- first, a bit b is encoded in $x = (-1)^b$;
- then, two random numbers t_1 și t_2 are chosen such that the Jacobi symbols $(\frac{t_1}{n})$ și $(\frac{t_2}{n})$ equals x;
- the ciphertext corresponding to the bit b is:

$$((t_1 + \frac{a}{t_1}), (t_2 - \frac{a}{t_2}))$$

The Cocks IBE scheme

Decryption

- one of the two components of the encrypted bit is chosen for decryption:
  - first, if $r^2 \equiv a \mod n$;
  - second, if $r^2 \equiv -a \mod n$;
- the decrypted value is:

$$x = (\frac{c + 2r}{n})$$

, where c represents the chosen component.

The anonymity of Cocks IBE scheme

Definition

We say that a cryptographic scheme provides anonymity if the ciphertext does not provide informations about the public key used for encryption. If the ciphertext can be anonymized by anyone using only the public key, then the scheme is **universally anonymous**.

- Galbraith showed that Cocks IBE scheme does not provide anonymity.

Definition

The Galbraith's test for a public key $a \in \mathbb{Z}_n^*$ and an element c from $\mathbb{Z}_n$ is defined as the following Jacobi symbol:

$$GT(a, c, n) = \left(\frac{c^2 - 4a}{n} \right).$$

The anonymity of Cocks IBE scheme

Property

For a public key $a \in \mathbb{Z}_n^*$ and a bit $b \in \{0, 1\}$ we define $M_a[n]$ as the following set:

$$M_a[n] = \left\{ \left(t + \frac{a}{t} \right) \mod n | t \in \mathbb{Z}_n^* \wedge (t/n) = (-1)^b \right\}.$$

When $c \in M_a[n]$, $GT(a, c, n) = 1$, because $c^2 - 4a = (t - (a/t))^2$.

- In 50's, Perron proved that for a prime number p, the difference between the squares and non squares in $\mathbb{Z}_p^*$ is 1 if $p \equiv 3 \mod 4$.

The anonymity of Cocks IBE scheme

- To summarize:

$$GT(a, c, n) = \begin{cases} +1 \implies Prob[c \in M_a[n]] = 1/2 \\ -1 \implies c \notin M_{(a,n)} \end{cases}$$

Question

It is possible to extend the Cocks IBE scheme to provide anonymity but the new scheme to not be much more expensive than the original scheme?

The anonymity of Cocks IBE scheme

- The first attempt has been proposed by Di Crescenzo and Saraswat, but their scheme is inefficient when used on large data:
  - the ciphertext expansion is 4 time bigger than that of the original scheme;
  - every user must save and use 4 keys for every bit.
- In 2009, Ateniese and Gasti proposed a more efficient method of anonymization.

The Ateniese-Gasti scheme

(De)Anonymization

A component c_j from the pair (c_1, c_2) corresponding to an encrypted bit is anonymized as following:

- k is chosen from the geometric distribution over the set $\{1, 2, 3, ...\}$ with the probability parameter $1/2$;
- T is chosen at random, and $Z = T + c \mod n$:

$$(Z, T_1, T_2, ..., T_{k-1}, \boldsymbol{T}, T_{k+1}, ..., T_m),$$

$$GT(a_j, Z - T_i, n) = -1, 1 \leq i < k$$

$$GT(a_j, Z - T_i, n) = \pm 1, k < i \leq m,$$

$$j \in \{1, 2\}, a_1 = a, a_2 = -a.$$

The Ateniese-Gasti scheme

(De)Anonymization

After the valid component is chosen according to r, the component T_l with the small index $1 \leq l \leq m$ such that $GT(a_i, Z - T_l, n) = 1$, represents the deanonymized ciphertext.

- Ateniese and Gasti showed that their scheme is secure.
- The ciphertext expansion of their scheme is much more bigger than that of the original scheme.
- bigger ciphertext expansion $\implies$ bigger computational time.
- We propose a more efficient scheme.

A new method of anonymization

A component c_i from the pair (c_1, c_2) corresponding to an encrypted bit is anonymized as following:

- we choose at random b, $b \in \{0, 1\}$;
- if $b = 1$ then:
    - we choose k from the geometric distribution over the set $\{1, 2, 3, ...\}$ with the probability parameter $1/2$;
    - $c_i = (c_i + 1) \mod n$ until we find the k-th element e such that $GT(a, e, n) = -1$, and while we do that, we count the number of elements e' over we pass, such that $GT(a, e', n) = 1$; let j be the number of this elements;
    - the pair $(e, j + 1)$ represents the anonymized component;

A new method of anonymization

- if $b = 0$ then:
  - we choose k from the geometric distribution over the set $\{1, 2, 3, ...\}$ with the probability parameter $1/2$;
  - the pair (c, k) represents the "anonymized" component.

At decryption:

- if $GT(a, e, n) = -1$ then $e = e - 1 \mod n$ until we find the j-th element c such that $GT(a, c, n) = 1$;
- if $GT(a, e, n) = 0$ then the component was not anonymized.

A new method of anonymization

Security

- Our scheme is based on the following fact:

$$c \in M_a[n], b \in \mathbb{Z}_n^*, b \neq a$$

$$Prob[GT(b, c, n) = 1] = Prob[GT(b, c, n) = -1] = \frac{1}{2}.$$

- There are 4 cases in which can be found an attacker that has two keys, a and b, and a component c encrypted with one of these two keys:

A new method of anonymization

Security

1. $\begin{cases} GT(a,c,n) = 1 \\ GT(b,c,n) = 1 \end{cases}$
2. $\begin{cases} GT(a,c,n) = -1 \\ GT(b,c,n) = 1 \end{cases}$
3. $\begin{cases} GT(a,c,n) = 1 \\ GT(b,c,n) = -1 \end{cases}$
4. $\begin{cases} GT(a,c,n) = -1 \\ GT(b,c,n) = -1 \end{cases}$

The attacker can suppose that:

1. c was encrypted with the public key a and **was not anonymized** or that c was encrypted with the public key b and **was not anynimzed**, the probability to be so is $1/2$ in both cases;

A new method of anonymization

Security

2. c was encrypted with the public key a and **was anonymized** or that c was encrypted with the public key b and **was not anynimzed**, the probability to be so is $1/2$ in both cases;
3. c was encrypted with the public key a and **was not anonymized** or that c was encrypted with the public key b and **was anynimzed**, the probability to be so is $1/2$ in both cases;
4. c was encrypted with the public key a and **was anonymized** or that c was encrypted with the public key b and **was anynimzed**, the probability to be so is $1/2$ in both cases;

A new method of anonymization

Security

- The following relation is valid in all 4 cases:

$$Prob[GT(a,c,n)=1] = Prob[GT(a,c,n)=-1] =$$

$$Prob[GT(b,c,n)=1] = Prob[GT(b,c,n)=-1] = \frac{1}{2}$$

- The way in which we anonymize a component must not provide informations about the public key used:

$$Anon(c,a) \implies c' \in \mathbb{Z}_n, GT(a,c',n) = -1$$

$$DeAnon(c,a,j) \implies c$$

$$DeAnon(c,b,j) \implies c''$$

A new method of anonymization

Security

- An attacker must find a valid deanonymized value for every public key that he uses.
- He must not make distinction between deanonymized values.

Statement

The method we use to anonymize a component does not provide informations about the key used for anonymization.

A new method of anonymization

Security

Demonstration

Let (c, j) be an anonymized component, $c \in M_a[n]$, and $GT(b, c, n) = -1$. The attacker can find, subtracting 1 from c, to the j-th element e' such that $GT(b, e', n) = 1$.

If the component c was not anonymized, j is chosen from the geometric distribution over the set $\{1, 2, 3, \ldots\}$ with the probability parameter $1/2$:

- the Jacobi symbols are uniformly distributed;
- we can suppose that until we find the k-th element e such that $GT(a, e, n) = -1$, we pass over the same number of elements e' such that $GT(a, e', n) = 1$.

Every component of the pair corresponding to an encrypted bit is anonymized independently.

Practical aspects

Implementation

- We implemented all three schemes.
- We executed 1000 times every essential step and calculated an execution time.
- The operation system under we tested the schemes is *Elementary OS*, *Linux Kernel 3.2*, and the machine consists of an *Intel Core i5* processor at 2.5GHz, with 4GB RAM.

Practical aspects

Implementation

Results:

	Setup	Extraction	Encryption	Decryption
Cocks	26.77 ms	3.58 ms	18.7 ms	7.45 ms
Ateniese-Gasti	26.77 ms	3.58 ms	33.46 ms	24.46 ms
Our scheme	26.77 ms	3.58 ms	23.19 ms	14.38 ms

Practical aspects

Ciphertext expansion

- we encrypt a key of 128 bits and n is a 1024 bits number;
- for Cocks IBE scheme, the ciphertext length is $128 * 2 * 1024$ bits;
- for Ateniese-Gasti scheme, the ciphertext length is $128*2*(m+1)*1024$ bits;
- we suppose that the element j from the pair (c, j) corresponding to a component of the encrypted bit can be represented on i bits;
- for our scheme, the ciphertext length is only $128 * 2 * (1024 + i)$ bits;
- în general, $i \leq 8$;

Conclusions

1. The **Identity Based Encryption** provides a lot of properties that can be used favourable in practice.
2. Our method of anonymization is **more efficient** than the scheme proposed by Ateniese and Gasti.
3. Like the scheme proposed by Ateniese and Gasti, **our scheme is universally anonymous**, anyone can anonymize the ciphertext using only the public key of the recipient.

Thank you!