Asymptotic secrecy of the information protection by the usage of simple integer splitting method

Amanie Hasn Alhussain
Information Technology Department, Peoples’ Friendship University of Russia, Moscow, Russia
E-mail: amanie.alhussain@mail.ru

Abstract. The integer splitting method is a special method that is used to encrypt the text. In this method, each text character is represented as an integer in accordance with the selected code table after that this integer is replaced on the base of another number with a sequence of \(k\) integers (\(k\)-splitting level) and this is performed by the usage of modular arithmetic. This article studies the property of asymptotic secrecy of the simple integer splitting method when it is used in the field of cryptography; the research is conducted from the point view of the hacker. The purpose of using integer splitting method is to increase the level of security and provide the ability to control the degree of protection of information. So a lemma was proven in this article and as a conclusion we can notice that the more usage of the splitting level leads to the more level of safety and protection of information.

1. Introduction
The simple integer splitting method is a mathematical method that represents the certain generalization of the modular arithmetic, and it is a symmetric encryption method that is described how to use it in details in the field of cryptography in the author’s publications [1, 2]. The properties of the monomorphic, reversibility and uniquely representation are proven in theories and described in details in the author’s publications [3, 4].

The block scheme of a symmetric encryption method based on integer splitting is shown in figure 1:

![Figure 1. The block scheme of a symmetric encryption method based on integer splitting.](image-url)
The simple integer splitting method is defined as follows:
Consider the two integers \( r \) and \( a \) satisfy the inequality \( r > a > 0 \).

Definition 1. The integer splitting of the number \( a \) on the basis of \( r \) is the representation of \( a \) as a sequence of numbers \( a_1, a_2, a_3, \ldots, a_{k-1}, a_k \) in which

\[
\begin{align*}
a_1 &= \delta^{(2)}, \text{ where } \delta^{(2)} = r \mod a, \\
a_2 &= \delta^{(3)}, \text{ where } \delta^{(3)} = r \mod q^{(2)}, q^{(2)} = \left\lfloor \frac{r}{a} \right\rfloor, \\
a_3 &= \delta^{(4)}, \text{ where } \delta^{(4)} = r \mod q^{(3)}, q^{(3)} = \left\lfloor \frac{r}{q^{(2)}} \right\rfloor, \\
&\vdots \\
a_{k-1} &= \delta^{(k)}, \text{ where } \delta^{(k)} = r \mod q^{(k-1)}, q^{(k-1)} = \left\lfloor \frac{r}{q^{(k-2)}} \right\rfloor, \\
a_k &= q^{(k)}, \text{ where } q^{(k)} = \left\lfloor \frac{r}{q^{(k-1)}} \right\rfloor,
\end{align*}
\]

(1)

where \( \delta \) – is the remainder of the integer division \( r/a \), and \( q \) – is the integer part of this division, and the symbol \( \left\lfloor \right\rfloor \) means rounding down to the nearest integer. The natural number \( k \) is called the splitting level.

The mathematical decryption model or the model of symbol restoration based on splitting method can be represented by the following equation [1, 2]:

\[
\left( \frac{r_i - \delta^{(j)}}{q^{(j)}} \right) \text{ where } j = k, k-1, \ldots, 3, 2 \quad \text{where } k > 1
\]

(2)

This article will study the asymptotic secrecy of the simple integer splitting method from the point view of the hacker or intruder, and we should notice that the obtained results is different from the case of asymptotic secrecy of the generalized integer splitting method which had been discussed earlier in the conference [5], the main difference is that the generalized integer splitting method is a vector base splitting and in this case, the next (i-th) step of the integer splitting process is performed each time with a new value of the splitting base, as can be shown in the definition 2:

Definition 2. A generalized integer splitting of the number \( a \) on the basis of the vector \( \vec{r} = (r_1, r_2, \ldots, r_l) \), \( l = k-1 \), is the representation of the number \( a \) as a sequence of integers \( a_1, a_2, a_3, \ldots, a_{k-1}, a_k \) in which:

\[
\begin{align*}
a_1 &= \delta^{(2)}, \text{ where } \delta^{(2)} = r_1 \mod a, r_1 > a, \\
a_2 &= \delta^{(3)}, \text{ where } \delta^{(3)} = r_2 \mod q^{(2)}, q^{(2)} = \left\lfloor \frac{r_1}{a} \right\rfloor, r_2 > q^{(2)}, \\
a_3 &= \delta^{(4)}, \text{ where } \delta^{(4)} = r_3 \mod q^{(3)}, q^{(3)} = \left\lfloor \frac{r_2}{q^{(2)}} \right\rfloor, r_3 > q^{(3)}, \\
&\vdots \\
a_{k-1} &= \delta^{(k)}, \text{ where } \delta^{(k)} = r_{k-1} \mod q^{(k-1)}, q^{(k-1)} = \left\lfloor \frac{r_{k-2}}{q^{(k-2)}} \right\rfloor, r_{k-1} > q^{(k-1)}, \\
a_k &= q^{(k)}, \text{ where } q^{(k)} = \left\lfloor \frac{r_{k-1}}{q^{(k-1)}} \right\rfloor,
\end{align*}
\]

(3)

where \( \delta \) – is the remainder of the integer division \( r_i / a \), and the symbol \( \left\lfloor \right\rfloor \) means rounding down to the nearest integer. The natural number \( k \) is called the splitting level.
2. Asymptotic secrecy of simple integer splitting method

Lemma 1. The probability of unauthorized restoration of the plaintext from the splitting ciphertext exponentially decreases with the increase of simple splitting level $k$ according to the expression:

$$
Pr(M | C, k) = \left( \sum_{i=2}^{k} \binom{N}{i} \right)^{-1},
$$

where $N$ is the size of the ciphertext $C$ that is created for the plaintext $M$, $L$ is the number of all possible attempts (events) during the search process in the gamma space, which is consisted of random numbers $\{\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_L\}$.

Proof.

It is assumed that the attacker receives ciphertext by simply intercepting a message in a communication channel (ciphertext-only attack (COA) or known ciphertext attack [6]).

In this case, the attacker knows the ciphertext $C$, which represents a sequence of integers and also knows the rule for restoring the character presented in equation (2). But the level of simple splitting $k$ is assumed to be unknown to him. In addition, the gammas that is used during the encryption of the plaintext is also unknown, so the attacker will be forced to generate a set of random numbers $R$ space and try to restore the plaintext with the help of the generated set of gammas. So, the attacker has a set $C$ consists of integers $\{c_1, c_2, \ldots, c_N\}$ with size $N$, from the point view of the attacker this set looks like a random sequence of numbers - the result of splitting (ciphertext).

Since the attacker knows the encryption method described in Definition 1, and the character restoration model shown in equation (2) is known, he will first construct an assumed sequence of independent random numbers

$$
R = \{\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_L\} \text{ with size } L.
$$

The method that will be used by the attacker will be based on brute force [7]. Since the value $k$ is unknown to the attacker, he will try different values of the splitting level $k$, starting with $k = 2$.

To evaluate the probability of unauthorized recovery of the plaintext at the splitting level $k$, we should study the possibility to extract the plaintext $M$ at the level $k$ by the attacker; the attacker must follow these steps:

The first step: the attacker will divide the received ciphertext $C$ into combinations of $k$ integers. Each symbol element is represented by $k$ integers in the space $C$. As a result, the number of combinations, which are studied by the attacker, will be equal to:

$$
N_k = \left\lfloor \frac{N}{k} \right\rfloor.
$$

The second step: the attacker will begin to perform search (brute force) on the set $R$ with repetition. From Definition 1, we conclude that in the case of the splitting level $k$, each combination of $k$ integers is calculated using one value $\bar{r}_i$, where $i = 1, 2, 3, \ldots, L$. At this step, the attacker will search with repetition $\left\lfloor \frac{N}{k} \right\rfloor$ from the set $R$ with size $L$. The number of all possible outcomes is defined by the following expression:
\[ n_k = L^{\left\lfloor \frac{N}{k} \right\rfloor} . \]  

(7)

The third step: is consisted of trying to extract the plaintext using the rule described in the equation (2) and by the help of both the ciphertext combinations described in the first step and the sets of numbers (gammas) obtained by searching the set \( R \) in the second step.

Consider the event \( M \) — the successful unauthorized restoration of the plaintext \( M \) from the result of splitting \( C \) at the splitting level \( k \). The probability \( \Pr(M \mid C, k) \) is determined by the following formula:

\[ \Pr(M \mid C, k) = \frac{p_k}{s_k} , \]  

(8)

where, \( p_k \) — the number of restoration attempts (events) of the meaningful plaintext \( M \) at the level of splitting \( k \); \( s_k \) — the total number of all possible attempts to restore the value of the plaintext \( M \).

First let’s find \( s_k \) — the total number of all possible attempts to get the plaintext \( M \) at the splitting level \( k \). Since the attacker, obviously, could not extract the correct meaningful plaintext \( M \) from all the previous levels of splitting, then the number \( s_k \) is given by the following expression:

\[ s_k = n_2 + n_3 + \ldots + n_k , \]  

(9)

where \( n_2 \) — is the total number of all possible attempts to get the plaintext \( M \) at the level \( k = 2 \), \( n_3 \) — is the total number of all possible attempts to get the plaintext \( M \) at the level \( k = 3 \), and \( n_k \) — is the total number of all possible attempts to get the plaintext \( M \) at the current level \( k \).

Calculating \( n_2 \): In an attempt to extract the plaintext \( M \) when \( k = 2 \), the attacker must follow these steps:

The first stage for calculating \( n_2 \): the attacker will divide the received ciphertext \( C \) into pairs of two integers. Each element of a symbol is represented by two elements in space \( C \). As a result, the number of pairs, which is studied by the attacker will be equal to

\[ N_2 = \left\lfloor \frac{N}{2} \right\rfloor . \]  

(10)

The second stage for calculating \( n_2 \): the attacker will begin to perform brute force on the values of the set \( R \) with repetition (return). From Definition 1, we conclude that in case of simple splitting \( k = 2 \), each pair of two elements is calculated using the same value \( \bar{r}_i \), where \( i = 1, 2, 3, \ldots, L \). At this stage, the attacker will search \( \left\lfloor \frac{N}{2} \right\rfloor \) values with repetition from the set \( R \) with size \( L \). The number of all possible events is defined by the following expression [8-11]:

\[ n_2 = L^{\left\lfloor \frac{N}{2} \right\rfloor} . \]  

(11)

Calculating \( n_3 \): In an attempt to extract the plaintext \( M \) at \( k = 3 \) the intruder will perform the following steps:
The first stage for calculating \( n_3 \): the attacker will divide the received ciphertext \( C \) into a combination of three integers. Each element of the symbol is represented by three elements in space \( C \). As a result, the number of combinations studied by the attacker will be equal to:

\[
N_3 = \left[ \frac{N}{3} \right].
\]  

(12)

The second stage for calculating \( n_3 \): the attacker will begin to execute brute force on the set \( R \) with repetition. From Definition 1, we can conclude that in the case of splitting level \( k = 3 \), each combination of three integers is used in the process of encryption with just one value of \( \tilde{r}_i \), where \( i = 1, 2, 3, ..., L \).

At this stage, the attacker will search \( \left[ \frac{N}{3} \right] \) values with repetition from the set \( R \) with size \( L \). The number of all possible events is defined by the following expression:

\[
n_3 = L^{\left[ \frac{N}{3} \right]}.
\]  

(13)

Substituting the values \( n_k \), \( n_3 \) and \( n_2 \) from equations (7), (13) and (11) into the expression (9) we obtain the following expression:

\[
S_k = L^{\left[ \frac{N}{2} \right]} + L^{\left[ \frac{N}{3} \right]} + ... + L^{\left[ \frac{N}{k} \right]},
\]

Or

\[
S_k = \sum_{i=2}^{k} L^{\left[ \frac{N}{i} \right]}.
\]  

(14)

Of all the attempts to restore the plaintext, only one case will give a meaningful plaintext that matches what is encrypted by the sender. This is a situation where the gammas, which are chosen on the side of the attacker match the same gammas that are used by the sender when encrypt the plaintext [12, 13]. So this leads to the fact that the number of correct extraction of the meaningful text, which meets the plaintext, is equal to one.

\[
p_k = 1.
\]  

(15)

Replacing the values \( s_k \) and \( p_k \) from equations (14) and (15) in formula (8), we obtain the result:

\[
\Pr(M \mid C, k) = \frac{1}{\sum_{i=2}^{k} L^{\left[ \frac{N}{i} \right]}} = \left( \sum_{i=2}^{k} L^{\left[ \frac{N}{i} \right]} \right)^{-1}.
\]  

(16)

Thus, from equations (16), it concludes that Lemma 1 is valid for any positive integer \( k \). Lemma 1 is proved.

Definition 3. We say that a method that depends on a parameter \( k \) has asymptotic secrecy if it satisfies the following condition:

when \( k \to \infty \), then it is executed \( \Pr(M \mid C, k) \to 0 \).

(17)

Theory: The encryption by using splitting method has the property of asymptotic secrecy.

Proof.
Let $\Pr(M \mid C,k)$ be the probability of restoring the plaintext $M$ from a sequence $C$ with a size $N$ which is resulted from performing the splitting method at the splitting level $k$.

Lemma 1 leads to the following:

when $k = 2$ \Rightarrow \Pr(M \mid C, 2) = \frac{1}{L^{N/2}} \quad (18)

when $k = 3$ \Rightarrow \Pr(M \mid C, 3) = \frac{1}{L^{N/2} + L^{N/3}} \quad (19)

when $k \Rightarrow \Pr(M \mid C, k) = \frac{1}{L^{N/2} + L^{N/3} + \ldots + L^{N/k}} \quad (20)

From equations (18), (19) and (20), we conclude that:

$\Pr(M \mid C, 2) > \Pr(M \mid C, 3) > \ldots > \Pr(M \mid C, k)$

Or

when $k \rightarrow \infty \Rightarrow \Pr(M \mid C, k) \rightarrow 0 \quad (21)$

Expression (21) indicates the asymptotic secrecy of the encryption by using splitting method.

3. Conclusion
In this article, a mathematical method is proposed for representing an integer in the form of a certain sequence of $k$ integers, which is called integer splitting. Based on integer splitting, the article studies the property of secrecy for protecting information, and it is proved that the splitting secrecy increases with increasing the level of splitting, which allows us to speak about the asymptotic splitting secrecy.

Acknowledgments
I wish to extend my special thanks to both professor Stefanyuk Vadim Lvovich and professor Orlov Yuri Nikolaevich at the information technology department, Peoples’ Friendship University of Russia, Moscow, Russia.

References
[1] Stefanyuk V L and Alhussain A H 2016 Symmetric encryption based on splitting method *Natural and technical sciences* 93(3) 130-3
[2] Stefanyuk V L and Alhussain A H 2016 Symmetric Encryption on the Base of Splitting Method *Bulletin of PFUR, Series Mathematics. Information Sciences. Physics* 2 53-61
[3] Stefanyuk V L and Alhussain A H 2016 Control the degree of protection of information by integer splitting *Artificial Intelligence and Decision Making* 4 86-91
[4] Stefanyuk V L and Alhussain A H 2017 Symbolic management of the degree of information security by integer splitting *Scientific and Technical Information Processing* 44(6) 450-4
[5] Stefanyuk V L and Alhussain A H 2019 Absolute Secrecy Asymptotic for Generalized Splitting Method 4th International Scientific Conference “Intelligent Information Technologies in Engineering and Production”, Ostrava-Prague, Czech Republic
[6] ciphertext-only attack Retrieved from: https://en.wikipedia.org/wiki/Ciphertext-only_attack
[7] brute force Retrieved from: https://en.wikipedia.org/wiki/Brute-force_attack
[8] Meester R 2008 *A Natural Introduction to Probability Theory* (Springer Science & Business Media) p 198
[9] Heinz B 2011 *Probability Theory De Gruyter Studies in Mathematics* (Walter de Gruyter) p 538
[10] Paul E P 2013 *Concepts of Probability Theory: Second Revised Edition* (Courier Corporation) p
416

[11] Lester L 2012 *Introduction to Probability Theory with Contemporary Applications* (Courier Corporation) p 368

[12] Probability of a meaningful plaintext Retrieved from: http://www.cse.iitm.ac.in/~chester/courses/16e_cns/slides/02_Classical.pdf

[13] Probability of a meaningful plaintext Retrieved from: http://www.uobabylon.edu.iq/eprints/paper_9_4879_1380.pdf