Modular reduction with step-by-step using of several bits of the reducible number

Sakhbyat Tynymbayev¹, Yevgeniya Aitkhozhayeva², Dana Tananova³, Sairan Adilbekkyzy²
¹Department of Information Security Systems, Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeev, Almaty, Kazakhstan
²Department of Cybersecurity, Information Processing and Storage, Satbayev University, Almaty, Kazakhstan
³Department of Information Security Systems, Kazakh National University named after Al-Farabi, Almaty, Kazakhstan

Article Info

Article history:
Received Feb 23, 2021
Revised Dec 6, 2021
Accepted Dec 15, 2021

Keywords:
Field-programmable gate array
Hardware encryption
Modular reduction
Public-key cryptography

ABSTRACT

Although public key cryptography is known to solve the problem of physically secure key exchange, the main drawback of this system is its low performance during encrypting and decrypting data. One of the ways to solve this issue is to increase the speed of the modular reduction operation, one of the basic operations of asymmetric cryptoalgorithms. A new method of step-by-step reduction by the N-bit module P using several bits of the 2N-bit reducible number A in one step is proposed in this paper. The method is based on using multiples of the P and reducing modulo at each step not the entire initial number, but its parts (A₁, A₂…, Aᵢ), which allows to reduce the bit capacity of A. A structural diagram of the hardware implementation of this method are developed. The main unit of the modular reduction device is a block of partial remainder formers, in which the partial remainder is computed using multiples of the P. The circuits are modeled in the Vivado Design Suite computer aided design (CAD) on base Artix-7 Field-programmable gate array (FPGA) device from Xilinx. Optimization of hardware costs is achieved by applying the same comparison circuits to compare different multiples of P with Aᵢ.

This is an open access article under the CC BY-SA license.

Corresponding Author:
Sairan Adilbekkyzy
Department of Cybersecurity, Information Processing and Storage, Satbayev University
22a Satpaev street, Almaty 050013, Kazakhstan
Email: sairan.02.95@mail.ru

1. INTRODUCTION

In the past decade the challenge of protecting information has recently become even more urgent. Organizations and enterprises, regardless of their form of ownership and type of activity, were forced to switch to online work due to the COVID-19 pandemic. Global situation has clearly demonstrated the importance of digital technologies for maintaining the functioning, competitiveness and sustainability of individual companies and entire industries. The digital transformation of the entire society is proceeding at an accelerated pace due to the need for a sudden and forced transfer to the online environment of professional and personal life. Therefore, the risks and threats to information security have also multiplied.

Cryptographic encryption methods are considered one of the most reliable ways to protect information. Public key cryptography is becoming popular in modern information systems, since it does not require physically secure data transmission channels. This is an important advantage in the context of widespread digitalization of all sectors of society.

However, the application of asymmetric cryptoalgorithms is constrained by their low speed in comparison with symmetric cryptoalgorithms. It is worth-noticing that the hardware implementation allows
obtaining asymmetric cryptosystems with higher performance. In addition, it is easier to physically protect the equipment from outside penetration. The hardware implementation of encryption systems is simple to install. Nevertheless, the hardware implementation of asymmetric cryptoalgorithms performs encryption and decryption operations about 1000 times slower than symmetric cryptoalgorithms. Thus, many researchers around the world are engaged in improving the performance of asymmetric cryptosystems [1]-[7].

One of the approaches to improve the performance of public key cryptosystems is acceleration of basic modular arithmetic operations: modular multiplication, modular exponentiation, modular reduction [8]-[16]. The modular reduction is the most complex of them. Various number-theoretic methods for dividing and calculating the remainder when dividing by the module \( P \) and devices for their hardware implementation are proposed in research-scientific articles and patents [17]-[30].

As for this date different methods of obtaining the remainder modulo are implemented. Though, performance of most of them is achieved by increasing hardware costs, which are directly proportional to the bit capacity of the given numbers. Consequently, their application when reducing large numbers is not feasible. Whereas public-key cryptoalgorithms rest upon cumbersome modular arithmetic with multi-bit numbers. The increase in hardware costs leads to an increase in power consumption (deterioration of thermal conditions) and a decrease in reliability. The issue of accelerated determination of the remainder by an arbitrary modulus of the number (modular reduction) with optimization of effective costs is urgent. This paper considers numerous approaches and circuit solutions on the field-programmable gate array (FPGA) implementation for modular reduction operating units for asymmetric cryptosystems.

The proposed work is organized in the following sections: Section 2 describes a modular reduction method with step-by-step using of several bits of the reducible number and provides a detailed description of proposed hardware. In section 3 analytical and implementation results are explained with insights for further improvement. Finally, section 4 includes the concluding remark.

2. RESEARCH METHOD

Several recent studies have suggested that the method of step-by-step modular reduction using several bits of the reducible number in one step can improve the performance of calculating the modular reduction [31]-[35]. The proposed by authors method relies on the use of multiples of the module and modular reduction at each step not of the entire initial number, but of its parts, which makes it possible to reduce the bit capacity of the reducible numbers. The modular reduction is divided to the sequential obtaining of partial remainders, and each previous remainder after shifting to the left by several bits is used to obtain the next remainder. At the first step, to obtain a zero remainder, the \( N \) most significant bits of number \( A \) are used. When obtaining the remaining remainders, the least significant bits of the number \( A \) are used (several bits at each step). The implementation of this method makes it possible to design a high-speed device with the optimization of costs.

The developed device consists of a unit for generating multiples of an \( N \)-bit module \( P \) \((P, 2P, 3P)\), a register \( RgA \) for storing \( 2N \)-bit reducible number \( A \), \( N/2 \) formers of a partial remainder (PRF), a delay element \( DE \). The method of obtaining the remainder modulo implemented in the device relies on the following principles: initially, the remainder \( R_0 \) is determined by the \( N \)-most significant bits of the number \( A \) (half of the bits of the number \( A \)). Subsequently, the remainder \( R_0 \) is shifted by two bits to the left \(-4R_0\). The next two least significant bits after \( R_0 \) in register \( RgA \) are appended to \( 4\overline{R_0} \), forming \( A_1 = (4\overline{R_0} + a_{n-2}a_{n-1}) \). Meanwhile, PRF1 determines the remainder \( R_1 \) modulo \( P \) from the number \( A1 \). The remainder \( R_1 \) is shifted by two bits to the left (i.e., \( 4R_1 \)) from the output of the PRF1 and the next two bits of the number \( A \) \((a_{n-2}a_{n-3})\) derive \( A_2 \). Afterwards, \( A_2 \) is fed to the input of the PRF2, where remainder \( R_2 \) is generated. Ultimately, the formation of the number \( A_i \) \((i=1\div N/2)\), which is fed to the PRF, to calculate the remainder \( R_i \) modulo \( P \), is carried out in the same way. The formation of the remainder \( R_i \) in each PRF is realized in parallel due to the simultaneous comparison of \( A_i \) with the values of multiples of the module \( P \), \( P, 2P, 3P, 3P \).

Figure 1 shows the structural diagram of a device for reducing a number modulo, which implements described method. The device contains a block 4 for forming multiples of the \( P \) module\( \bar{P}, 2P, 2P, 3P \) and \( 3P \), a register 5 for storing a \( 2N \)-bit reducible number \( A \), \( N/2 \) of PRF formers \( 8.1\div8.N/2 \), delay element 6 (blocks are indicated by the numbers in figure). The information outputs of block 4 are connected to the information inputs of the PRF \( 8.1\div8.N/2 \) for transmitting the values \( \bar{P}, 2P, 3P, P, 2P \) and \( 3P \). The information outputs of register 5 of the number \( A \) are connected with the inputs of the PRF \( 8.1\div8.N/2 \) for transmitting the corresponding two bits of the number \( A \). Each PRF has its own two bits of the number \( A \).

Firstly, the remainder of \( R_0 \) is determined by the \( N \) most significant bits of the 2\( N \)-bit number \( A \) in register 5 then \( 4R_0 \) is created by shifting two bits to the left. \( 4R_0 \) together with the two bits attached to it from register 5, represents the number \( A_1 = (4\overline{R_0} + a_{n-2}a_{n-1}) \), which in PRF \( 8.1 \). PRF \( 8.1 \) determines the remainder \( R_1 \). Similarly, the generating of the number \( A_i \) \((i=1\div N/2)\) is achieved, then it is fed to the PRF \( 8.i \) to compute
the remainder $R_i$ modulo $P$. The remainder $R_{i-1}$ from the output of the $PRF \ 8.i-1$ is shifted to the left by two bits towards the most significant bits, the next two least significant bits of the number $A$ are attached to it $A_i$: 

$A_i = L(2) \ R_{i-1} + a_{n-2} \ a_{n-1}$.

Eventually, bits $a_1$ and $a_0$ of the number $A$ from register 5 and a partial remainder $4R_{N/2,1}$ from the outputs of the $PRF \ 8.N/2-1$ are fed to the inputs of the $PRF \ 8.N/2$. At the outputs of the $PRF \ 8.N/2$, the final result $R_{N/2} = N$ is formed. Consider the operation of the device for modulo reducing a $2N$-bit number $A$ by $N$-bit module $P$ in more detail:

First of all, according to the "Start" signal (input 1), the value of $P$ is received from input 2 into block 4, the value of the number $A$ from input 3 is taken into register 5. In block 4, modules of the number $P$, $\overline{P}$, $2P$, $2\overline{P}$, $3P$, $3\overline{P}$ are generated, which are fed to the inputs of all $PRF \ 8.i \div 8.N/2$. Simultaneously, $N$ high-order bits (i.e. $R_0$) from the outputs of register 5 with the shift by two bits towards the high-order bits are fed to the inputs of the $PRF \ 8.1$. In this case, the bits $a_{n-1} \ a_{n-2}$ of the number $A$ is attached to the least significant bits of the shifted $R_0$ from register 5, forming the number $A_1$. In turn $PRF \ 8.1$ calculates the remainder $R_1$. Further, the value of $R_1$ with the shift two bits towards the higher bits, as well as the bits $a_{n-3} \ a_{n-4}$ of the number $A$ form $A_2$ and are fed to the inputs of the $PRF \ 8.2$, and where the partial remainder $R_2$ is created. At the final stage, the partial remainder $R_{N/2}$ from the outputs of the previous $PRF \ 8.N/2-1$ with two-bit shift towards the higher bits and bits $a_1 \ a_0$ of the number $A$ from register 5 are applied to the inputs of the $PRF \ 8.N/2$ and at its outputs the remainder $R_{N/2}$ is formed. By the signal "End of operations" that is formed at the output of the delay element 6, the remainder of $R_{N/2}$ is issued to the output of the device 7. The time of formation of the result $T_{fr}$ is determined by the total time of the signal passing through the $PRF \ 8.1 \div 8.N/2$, i.e. $T_{fr} = N/2 \times T_{PRF}$.

Figure 2 illustrates a functional diagram of the partial remainder former block $PRF_i$, which consists of a binary adder $Add_i$; two comparator circuits $CC-1$ and $CC-2$; blocks of logic gates $AND1 \div AND5$, $AND8$; gates $AND6$, $AND7$, $OR3$; blocks of logic gates $OR1$, $OR2$, $NOT$ gate.

The circuit works as follows: from the block of former of the multiples of module 4, the bits of the $2P$ module are fed to the right inputs of the $CC-1$ circuit. On the right inputs of the $CC-2$ circuit comes from block 4 the value of the module $P$ through $AND1$ and $OR1$ or the value of $3P$ through $AND2$ and $OR1$. The value of the previous remainder $R_{i-1}$, shifted to the left by two bits towards the most significant bits, with the next two bits of the reducible number attached to it, determines the value $A_i = L(2) \ R_{i-1} + a_{n-2} \ a_{n-1}$.

The $A_i$ value is fed to the right inputs of the $Add_i$ adder through the $AND8$ and to the left inputs of $CC-2$ and $CC-1$. The circuit $CC-1$ compares the $A_i$ with the $2P$. If, in this case, $A_i \geq 2P$, then at the output 2 of the $CC-1$ circuit, a unit impulse $1$ is generated, which is fed to the control input of the $AND2$, allowing the passage of the bits of multiples of the $3P$ module to the right inputs of the $CC-2$. At the same time, at the output 1 of the $CC-1$ - signal 0, which disables the passage of the bits of the module $P$ through the $AND1$ to the inputs of the $CC-2$ and through the $AND6$, $NOT$ gates lead to the formation of $1$ impulse at the right input of the $AND7$.

---

*Modular reduction with step-by-step using of several bits of the reducible ...* (Sakhybay Tynymbayev)
Besides, when comparing the \( A_i \) with the \( 2P \), if the inequality \( A_i < 2P \) holds, signal \( 1 \) is generated at the output 1 of the CC-1, allowing the passage of the bits of the \( P \) module through the AND1 circuit block to the right inputs of the CC-2. At the same time, at the output 2 of the CC-1 there is a signal 0 that stops the passage of \( 3P \) through the AND2 to the inputs of the CC-2.

Whereas the condition \( A_i \geq 2P \) is fulfilled, the CC-2 compares the \( A_i \) with the \( 3P \). If at the same time \( A_i < 3P \), then signal \( 1 \) will be set at the output 1 of the CC-2, it is fed to the input of the OR3 gate and to the control inputs of the AND4, allowing the passage of the one’s complement \( 2\overline{P} \), supplied to the information inputs AND4, to the input of the adder Add through the OR2. At the output of the OR3, signal 1 is generated, allowing the passage of \( A_i \) through the AND8 to the adder Add, and the passage of which through the circuit AND7 to the input +1 of the adder Add is allowed by a unit impulse from the output of the NOT gate. In this case, the Add adder performs the operation \( R_{\text{add}} = A_i + 2\overline{P} + 1 \).

If the condition \( A_i \geq 2P \) is met, when comparing the \( A_i \) with the \( 3P \) on the CC-2, it turns out that \( A_i \geq 3P \), then at the output 2 of the CC-2, signal \( 1 \) will be set, which is fed to the input of the OR3 and to the control inputs of the AND5, allowing the passage of the \( 2\overline{P} \) supplied to the information inputs AND5, to the input of the Add adder through the OR2. At the output of the OR3, signal 1 is generated, allowing the passage of \( A_i \) through the AND8 to the adder Add, and the passage of which through the circuit AND7 to the input +1 of the adder Add is allowed by a single signal from the output of the NOT. In this case, the Add adder performs the operation \( R_{\text{add}} = A_i + 3\overline{P} + 1 \).

As well as, if the condition \( A_i < 2P \) is set, when comparing the \( A_i \) with the \( P \) on the CC-2, it turns out that \( A_i < P \), then signal 1 will be set at the output 1 of the CC-2, which is fed to the input of the OR3 and AND6. At the output of the OR3, signal 1 is generated, allowing the passage of \( A_i \) through the AND8 to the adder Add, and the passage of which through the circuit AND7 to the input +1 of the adder Add is prohibited by a zero signal from the output of the NOT. In this case, the Add adder performs the operation \( R_{\text{add}} = A_i + 0 = A_i \), since multiples of the module \( P \), \( 2\overline{P} \), \( 3\overline{P} \) are not fed to the second input of the adder Add, as they are disabled by zero control signals on the AND3, AND4, AND5. Consider the operation of the circuit using an example of reducing the \( 2N \)-bit number \( A \) to the \( N \)-bit module \( P \).

\[
A = 1437_{10} = \{ a_{11} a_{10} a_9 a_8 a_7 a_6 a_5 a_4 a_3 a_2 a_1 a_0 \},
\]

\[
2N=12, N=6, N/2=3. P=35_{10}=100011_2, 2P=70_{10} \text{ and } 3P=105_{10}
\]

The most significant six bits of the binary representation of the number \( A \) determine the value \( R_0=010110_2=22_{10} \). Attaching the remainder \( R_0 \) shifted to the left by two bits with the next two bits \( a_{25:24} \) of the number \( A \), gives the number \( A_i = L(2)R_0 + (a_{25:24}) = 4R_0 + (a_{25:24}) = (88+1)_{10}=89_{10} \). For clarity, all calculations to solve \( R=\text{AmodP} \) are given in Table 1 in decimal notation.
Table 1. Procedure for calculating $R = A \mod P$

<table>
<thead>
<tr>
<th>Step</th>
<th>PRF1</th>
<th>PRF2</th>
<th>PRF3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-step</td>
<td>$A_1 = L(2)R = 88_{10} + 19_{10}$</td>
<td>$A_1 = L(2)R + 88_{10} + 19_{10}$</td>
<td>$A_1 = L(2)R + 88_{10} + 19_{10}$</td>
</tr>
<tr>
<td>2-step</td>
<td>$2P \leq A_1 &lt; 3P$ takes place.</td>
<td>$2P \leq A_1 &lt; 3P$ takes place.</td>
<td>$2P \leq A_1 &lt; 3P$ takes place.</td>
</tr>
<tr>
<td>3-step</td>
<td>$R_1 = A_1 - 2P = 89_{10}$</td>
<td>$R_1 = A_1 - 2P = 89_{10}$</td>
<td>$R_1 = A_1 - 2P = 89_{10}$</td>
</tr>
<tr>
<td>4-step</td>
<td>$70 = 19_{10}$</td>
<td>$70 = 19_{10}$</td>
<td>$70 = 19_{10}$</td>
</tr>
</tbody>
</table>

*Check: $R = A \mod P = 1437 \mod 35 = 2_{10}$

3. RESULTS AND DISCUSSIONS

The verification of the correct functioning of the developed circuit was performed by modeling in Vivado design suite CAD with a focus on the implementation of devices on FPGA of the Artix-7 series from Xilinx. Moreover, the Verilog hardware description language was chosen to describe the devices [36]. The simulation accomplished for the selected numbers demonstrates that three clock signals were required to obtain the remainder with the stepwise use of two bits of the reducible number $A$. The timing diagram presents the value of the reducible number $A = 1437_{10}$, the value of the module $P = 35_{10}$, the number of clock signals used, the remainders $R_i (19, 9, 2)$ obtained at each step. It is noted that the complication of the design of the partial remainder formers that are basic part of the entire device, makes it possible to accelerate the process of obtaining the remainder by using three bits of the $A$ number added to the partial remainder. In this case, the number of $PRF$ decreases and, consequently, the number of obtained partial remainders required for modular reduction and the operation time is reduced.

The structural diagram of the modular reduction device using three bits of the reducible number will be similar to the structural diagram of the device shown in Figure 1. But the number of forms $PRF$ is $N/3$ and multiples of the modules $P, 2P, 3P, 4P, ..., 7P, 7P$ should be formed in block 4. However, important to point out that the $PRF$ circuit will be more complex than that shown in Figure 2, as it is necessary to use five comparator circuits instead of two as part of the $PRF$. Nonetheless, since they work in parallel, the response time of the $PRF$ does not increase. Binary representations $P, 2P, 3P, 4P, ..., 6P, 6P$ and $7P, 7P$ should be fed to the inputs of the $PRF$ from the outputs of the unit for forming multiples of the module. The information outputs of the register 5 of the number $A$ are connected with the inputs of all $PRFs 8.1 \div 8.1 N/3$ to transmit the corresponding three bits of the number $A$. For each $PRF$, its own three bits are supplied.

The value of the partial remainder from the outputs of the $PRF_{i, j}$ with a three-bit shift towards the higher bits with the addition of the next three bits of the reducible number is fed to the left inputs of the adder and comparators of the $PRF$, where they are compared with the value of the modules $P, 2P, ..., 7P$. As a result, control signals are generated that commutate the value of one of the modules $P, 2P, ..., 7P$ to the right inputs of the binary adder, forming the remainder $R_i$ at the outputs of the adder. The time of forming the result $T_R$ is determined by the total time of the signal passing through the $PRF$, i.e., $T_R = N/3 T_{PRF}$. In the general case, the reduction of a number modulo with step-by-step use of three bits of the reducible number allows you to speed up the operation by one and a half times compared to using two bits of the reducible number.
4. CONCLUSION

The increase in terms of the speed of the device is achieved by the simultaneous comparison of several multiples of the P module in the scheme of partial remainder former with the next reducible number A. The use of several bits of the reducible number in one step also makes it possible to accelerate the receipt of the remainder. Step-by-step modular reduction decreases the bit capacity of the reducible numbers, which leads to a decline in hardware costs. Optimization of hardware costs is also achieved by using the same comparators in the partial remainder formers PRF to compare different multiples of the module P with A.

Decreasing hardware costs makes possible to increase device reliability and improve thermal operation. The gathered simulation results confirm the correctness of the developed circuits. As well as the depletion in time costs can be achieved with the increase in the number of bits of the reducible number A from two or more, used at each step of the reduction modular. The developed circuit efficiently can be applied both in cryptographic applications and in digital devices to form elements of finite fields.

REFERENCES


**BIOGRAPHIES OF AUTHORS**

**Sakhbybay Tynymbayev** is a professor in the Department of Information Security Systems, Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeev, Candidate of technical sciences. Professor S. Tynymbayev got an academic degree at Bauman Moscow State Technical University, Russia. He has more than 50 years of scientific and pedagogical experience. He is interested in computer science, cryptographic hardware and embedded systems and physical and numerical modeling of operating units for digital devices. He has over 200 scientific and educational works. He can be contacted at email: s.tynym@mail.ru.

**Yevgeniya Aitkhozhayeva** is a professor in the Department of Information Security Systems, Almaty University of Power Engineering and Telecommunications, Candidate of Technical Sciences. One of the leading scientists of the Republic of Kazakhstan in the field of computing and information security, a teacher with over 40 years of experience. She received an academic degree at Department of Computer Engineering St. Petersburg State Electro Technical Institute (Technical University "LETI"), Russia. Professor Y. Aitkhozhayeva is interested in Information Security, Databases Systems, Hardware of Cryptography and has over 200 scientific and educational works. She can be contacted at email: ait_evg@mail.ru.

**Dana Tananova** received a bachelor's degree in Information Systems from Al-Farabi Kazakh National University in Almaty. From 2012 to 2014, she studied for a master's degree and received an academic master's degree in Computer Engineering and Software at KazNU. From 2017 to 2020, she studied for a doctoral degree in Information Security Systems. Dana is senior lecturer at the Department of Telecommunications and Innovative Technologies, Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeev She is interested in information security in different platform environments. She can be contacted at email: tananova.dana@gmail.com.

**Sairan Adilbekkyzy** received the B.Sc. and M.Sc. degrees in military affairs and security studied at Kazakh National Research Technical University named after K. I. Satbayev, specialty "Information Security Systems". Sairan is doctoral degree student at Satbayev University in Management of Information systems since September 2020. Author and co-author of more than 15 scientific works. Her main field of interests are Information Security and Hardware of Cryptography. She can be contacted at email: sairan.02.95@mail.ru.