# *Proposed Approach for Improving RNA Crypto-Key Based on Polynomial Convolution*

**Assistant Instructor.** Estabraq Abdulredaa Kadhim
Computer Engineering Techniques Department
AL-Israa University College
Baghdad/Iraq
Estabraq_ai_1989@yahoo.com

***Abstract*—**Random Number Generators are fundamental tools for cryptography protocols and algorithms. The basic problems that face any crypto key generator are randomness, correlations and distribution of the state of key sequence. This paper proposed a new method to enhance RNA crypto key generation. It has been implemented by extending the crypto key by applying polynomial convolution technique which extracts the mask filter from the same RNA key sequence depending on the start and end codon properties. This will provide another high level of extension and generate random-strength crypto key. The proposal approach could pass through the statistical measurements successfully and achieved high rate of randomness (approximated to 96%).

**Keywords—Cryptography; Key generation; RNA; Bimolecular Computation; Polynomial Convolution.**

## *I. INTRODUCTION*

Random numbers generations are major tools in many applications of cryptography such as key generation [1]. The security of cryptography methods basically depends on the robustness of encryption method and secrecy of key [3]. It is so important to provide generators which are capable to produce amount of secure random numbers such as polynomial Feedback Shift Registers (LFSRs) [2]. Bimolecular Computation Methods (BMC) have been developed for a several domain of operations on both DNA & RNA strands such as solving hard NP-complete problems like; Travelling Salesman Problem (TSP), breaking Data Encryption Standard (DES) and for cryptography key generation [4].*This approach will show new proposal method to improve RNA crypto key generation basically by depending on the concept of Polynomial convolution 1D, this technique will add second level of extension to achieve a desired level of secrecy and randomness*

In [5], A. Hassan proposed a method for key generation based on the concept of translating RNA to protein chain. Their approach first determines a key sequence size, for example 9 byte, then it produces crypto key with extended length suitable with the length of secret message. Their proposed approach consists of several steps contain (Mapping, Extended, Complementing and Rotation). The generated key was acceptable according to statistical tests of randomness.

## *II. DNA AND RNA TRANSCRIPTION*

DNA is a double stranded sequence of four nucleotides; the four nucleotides that compose a strand of DNA are as follows: adenine (A), guanine (G), cytosine (C), and thymine (T) they are often called ***bases***. The chemical structure of DNA (the famous double- helix) was discovered by James Watson and Francis Crick in 1953. It consists of a particular bond of two linear sequences of bases. This bond follows a property of complementarity: adenine bonds with thymine (A-T) and vice versa (T-A), cytosine bonds with guanine (C-G) and vice versa (G-C). This is known as Watson-Crick complementarity. Each DNA strand has two different ends that determine its polarity: the 3. 'end, and the 5. 'end. The double helix is an anti-parallel (two strands of opposite polarity) bonding of two complementary strands **[6]**.

## *III.* RNA CRYPTOSYSTEM

RNA-Crypto System (shortly RCS) is a private key algorithm to encrypt data. In particular from the observation of RNA behavior and some of its properties. The RNA sequences have some sections called Introns. Introns, derived from the term "intragenic regions", are non-coding sections of precursor mRNA (pre-mRNA) or other RNAs, that are removed (spliced out of the RNA) before the mature RNA is formed. Once the introns have been spliced out of a pre-mRNA, the resulting mRNA sequence is ready to be translated into a protein. The corresponding parts of a ge**ne are known as introns as well**. It uses the presence of Introns in the RNA-Crypto System output as a strong method to access to the secret key to code the messages. In the RNA-Crypto System algorithm, the introns are sections of the ciphered message with non-coding information as well as in the precursor mRNA [7].

## IV. POLYNOMIAL VECTOR CONVOLUTION

In mathematics and some special particular analysis, convolution is a mathematical operation on two functions g and f, to produce a third function that is usually showed as a modified version of one of the original two functions in which one of the original functions is translated. Convolution is similar to cross-correlation. It has many applications in probability, statistics, computer vision, image and signal processing [9].

Let m = length (u) and n = length (v). Then w is the vectors of length m+n-1 whose $k^{th}$ element is:

$$W(k) = \sum_{j} u(j)\, v(k-j+1) \qquad (1)$$

The sum is over all the values of j that lead to legal subscripts for u(j) and v (k-j+1), specifically j = max(1,k+1-n):1:min(k,m). When m = n, this gives

$w_1 = u_1 * v_1$

$w_2 = u_1 * v_2 + u_2 * v_1$

$w_3 = u_1 * v_3 + u_2 * v_2 + u_3 * v_1$

……………………..

$W_n = u_1 * v_n + u_2 * v_{n-1} + \ldots + u_n * v_1$

## V. IMPROVEMENT OF RNA-CRYPTO-KEY

This approach discusses some drawbacks that have been observed by using the previous approach of RNA key generation, and we try to find an improved version for addressing those mistakes stated as follows:

- *Translation of mRNA strand based on start and end codons, investigated from which by calculating the number of these confined codons to apply rotate shift on RNA sequence.*

  Benefit from mRNA translation just for applies "Rotate Shift on RNA sequence "may be considered as weakness investment for this property, it looks like "*simple permutation*".

  In this approach, extended RNA-Crypto-key basically depends on the *"Polynomial convolution technique"*, contributes to provide self-extension with variable RNA length by treating RNA confined codons as mask filter for convolution applied with RNA key chain .The length of resultant key will be approximated to the summation of the length of mask filter and the length of RNA chain. It is worth mentioning, that mask filter has variable length according to confined codons number in RNA key chain, so the output of extended RNA-Crypto-key also has variable length with a high rate of randomness. This approach has two levels of RNA key extension, the first one comes from RNA codons table and the second one by applying Polynomial convolution technique. Algorithm (1) shows main steps of the proposal approach.

| ***Algorithm (1): Improved RNA-Crypto-Key*** |
|---|
| **Input:**<br>Initial key *(characters, numbers), size (9, 12,15,...etc./ byte)* , No. Iterations<br>**Output:**<br>Improving RNA random key with variable expanded size |
| ***Begin***<br>**Step:1** :Convert initial key to binary key stream<br>**Step:2:** Coding each two bit from the binary key stream to RNA four bases using table (I).<br>**Step:3:** Split mRNA strand into group of codons ( 3 nitrogenous bases)<br>**Step:4:** Extending each codon in mRNA sequence by selecting another codon that belong to the same amino acid and appending them, according to table (II).<br>**Step:5**: Read RNA strand until finding the AUG codon that is used to begin protein synthesis, then count the number codons and stop when finding end codon is (UAA or UAG or UGA)<br>**Step:6**: Extract Self-RNA-mask filter starting from AUG to (UAA or UAG or UGA)<br>**Step:7**: Apply rotate right shift on RNA strand based on the number of confined codons between start and end codon<br>***Step8****: Apply (Polynomial-Convolution) technique between Self-RNA-mask filter and RNA strand based on equation (1)*<br><br>**Step:9:** Convert New-RNA key to binary sequence and generate final crypto-key<br>**End** |

| **Table(1):Convert bit sequence into mRNA nucleotides** | |
|---|---|
| Bit sequence | mRNA Base |
| 00 | A |
| 01 | U |
| 10 | C |
| 11 | G |

| Table (II): Coding Amino Acid Groups into bit sequance | | |
|---|---|---|
| Decimal code | RNA code | Binary code based (6 bit) |
| 0 | UUU | 000000 |
| 1 | UUC | 000001 |
| 2 | UUA | 000010 |
| 3 | UUG | 000011 |
| 4 | CUU | 000100 |
| 5 | CUC | 000101 |
| 6 | CUA | 000110 |
| 7 | CUG | 000111 |
| 8 | AUU | 001000 |
| 9 | AUC | 001001 |
| 10 | AUA | 001010 |
| 11 | AUG | 001011 |
| 12 | GUU | 001100 |
| 13 | GUC | 001101 |
| 14 | GUA | 001110 |
| 15 | GUG | 001111 |
| 16 | UCU | 010000 |
| 17 | UCC | 010001 |
| 18 | UCA | 010010 |
| 19 | UCG | 010011 |
| 20 | CCU | 010100 |
| 21 | CCC | 010101 |
| 22 | CCA | 010110 |
| 23 | CCG | 010111 |
| 24 | ACU | 011000 |
| 25 | ACC | 011001 |
| 26 | ACA | 011010 |
| 27 | ACG | 011011 |
| 28 | GCU | 011100 |
| 29 | GCC | 011110 |
| 30 | GCA | 011101 |
| 31 | GCG | 011111 |
| 32 | UAU | 100000 |
| 33 | UAC | 100001 |
| 34 | UAA | 100010 |
| 35 | UAG | 100011 |
| 36 | CAU | 100100 |
| 37 | CAC | 100101 |
| 38 | CAA | 100110 |
| 39 | CAG | 100111 |
| 40 | AAU | 101000 |
| 41 | AAC | 101001 |
| 42 | AAA | 101010 |
| 43 | AAG | 101011 |
| 44 | GAU | 101100 |
| 45 | GAC | 101101 |
| 46 | GAA | 101110 |
| 47 | GAG | 101111 |
| 48 | UGU | 110000 |
| 49 | UGC | 110001 |
| 50 | UGA | 110010 |
| 51 | UGG | 110011 |
| 52 | CGU | 110100 |
| 53 | CGC | 110101 |
| 54 | CGA | 110110 |
| 55 | CGG | 110111 |
| 56 | AGU | 111000 |
| 57 | AGC | 111001 |
| 58 | AGA | 111010 |
| 49 | AGG | 111011 |
| 60 | GGU | 111100 |
| 61 | GGC | 111101 |
| 62 | GGA | 111110 |
| 63 | GGG | 111111 |

***Note: Highlight green codon in Table (II) represents starts and Red represents end codons in mRNA strand***

## *VI.* EXPEREMENT_A RESULT AND DISSCUSSION

This section illustrates the results that are obtained from the implementation of the proposed approach which is explained in the previous sections, these results can be presented as follows:

Basic Statistical measurements of randomness have been done on the improved RNA as shown in Table (III). These statistical tests such as, frequency, serial, poker and runs are used to evaluate randomness and distributed properties of several RNA crypto- key samples [11]. Randomness tests were applied on various key sizes such as 9, 12 72, 96 and 144 bits or bytes to use for checking random and distributed properties of several RNA crypto- key samples. The output of tests have been compared with passes values in [10]

| Table(III): Randomness Statistical Tests of Improved RNA-crypto-key | | | | | |
|---|---|---|---|---|---|
| Initial key Size (bit) | Improved RNA-key at 3 itr. Size (bit) | Frequency Pass <=3.84 | Serial Pass <=5.99 | Poker Pass <=11.1 | Run Pass <=22.362 |
| 96 | 3324 | 0.123 | 0.544 | 7.321 | 5.161 |
| | 2190 | 0.154 | 1.119 | 4.754 | 3.368 |
| | 2388 | 0.342 | 3.182 | 10.407 | 15.164 |
| 72 | 1758 | 2.329 | 6.254 | 7.146 | 6.387 |
| | 2088 | 3.065 | 3.571 | 8.321 | 18.462 |
| | 1824 | 0.967 | 1.430 | 7.394 | 12.737 |
| 144 | 3306 | 0.756 | 2.49 | 10.89 | 8.406 |
| | 3204 | 1.709 | 4.77 | 6.262 | 6.824 |
| Average | | 1.18 | 2.92 | 7.81 | 10.21 |

As shown in table (III), Improved RNA -key is achieved to high rate of randomness (96% percentage of tests that is passed through statistics) when compared with previous method of RNA–key generation in table (IV) that shows just Average of many tests. While Improved -key size is may be much larger than 40%.

| Table(IV): Randomness Statistical Tests of Previous RNA-crypto-key (By Average) | | | | | |
|---|---|---|---|---|---|
| Initial key Size (bit) | Previous RNA-key at 3 itr. Size (bit) | Frequency Pass <=3.84 | Serial Pass <=5.99 | Poker Pass <=11.1 | Run Pass <=22.362 |
| 96 | 1920 | 27.5 | 65.9 | 92.4 | 58.2 |

- Figure (1) illustrates Chart that shows the difference between the key expansion rates of the improved and traditional RNA key method through several iterations.

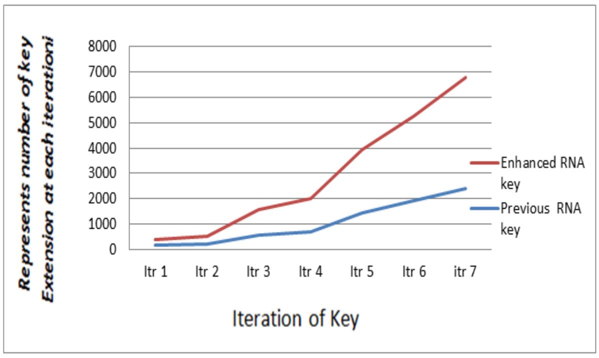


Fig (1): comparison in term of key expansion between traditional and proposed approaches

## *VII.* CONCLUSION

This paper describes an efficient method for Improving RNA-crypto-key generation using Polynomial Convolution. It successfully provides an extra level of extension and security for RNA-crypto-key when compared with the previous approach. . Robust features of this proposed method could be as follows:

- RNA-crypto-key expansion basically depends on several factors such as (start and end codons for translation, iteration numbers, and self-expansion based on codon table).
- Polynomial convolution provided more than 40% expansion on traditional RNA.
- Improved RNA–crypto-key provides high rate of security randomness achieved 96% on other traditional RNA key generation method.
- Size of convolution mask filter may have a variable length based on start and end codon of each RNA strand and encoding of initial key.

## *VIII.* FUTURE WORK

Improving -RNA algorithm considered as one of the optimization techniques which are used for solving any NP-complete problem such as path planning, four colors mapping, TSP, CPU scheduling problems.

## *IX.* APPENDIX(1)

This section illustrates the implementation of the proposed approach which is programmed by using Visual c#.net 2015 as shown in Fig (2). The following example explains in details steps of proposed algorithm work:

- **Initial key (12 byte)=** )(*UYHas@#A1

**Step1. Binary coding (96 bits)=**

(11111110010111000100100000010 110 11111100110 011000100100 1101000 1101111011111111000 10001 01100 01 11011000101001000011011 0 10100110 0 0 0
00001011001010001101110001010010001110001100 11111000011000111000111010100010011001010110 01100100001101110010100110010100100100110110 100100010111111111001101001101101001100011001 0000011101000010100010001100)

**Step2. RNA Coding (48 nucleotide) =**

(CUUAAUUAUUUACCCCCUCCAUACCAUCGAG CAAACGAUACAA CCAGA)

**Step3. Extended (1) based on codon table(96 nucleotide)=**
(GAAGAGUUAUUGAUAAUGAAUAACGGGGGC GG AGGGGGU GGA AUGAUAGUAGUGGCUGCACGUCGAUUGUUAC UACUUUGUUGCUGGUGAUCUUCG)

**Step4. Complement RNA & Rotate=**
(GGGGGCGGAGGGGGUGGAAUGAUAGUAGUG GCUGCACGUCGAUUGUUACUACUUUGUUGCU GGUGAUCUUCGGAAGAGUUAUUGAUAAUGAA UAAC)

**Step5. Extended (2) based on Polynomial Convolution (*Note:* mask filter length = 24)**

(GGGCUUGAUACAACGUGGCUCUCACUAGUA UCGGUGGGAUAUUCAGCGAUUACAUCGUUAG AAAACGACCAGUGAUUACAGCAGAGGCAUCA GCAGCCCACUCUUUCUCUAAGUGACCUCUUA CUAGUCCUAUUGGGCUCCUCGUAGAUCUAAU GUAACCACAGA)

**Step6. Random Key Generation (330 bit ) =**
(11111110010111000100100000101101111111001100 110001001001101000110111101111111100010001011 00011101100010100100001101101010011000000001 01100101000110111000101001000111000110011111 00001100011100011101010001001100101011001100 10000110111001010011001010010010011011010010 001011111111100110100110110100110001100100000 11101000010100010001100)

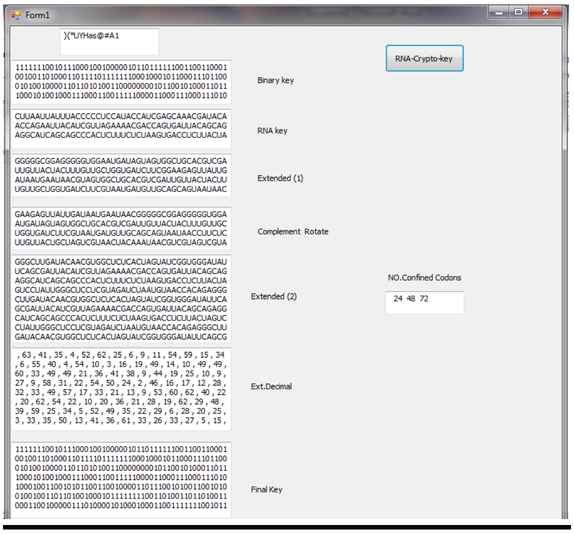


Fig (2): Interface of proposed program using c#.net

## X. REFRENCES


[1] Estabraq Abdulredaa," *Number Generator Improvement based on_Artificial Intelligence and Non Parametric Statistic_Methods*", ISSN: 2454-9916, Volume: 1, Dec 2015. Available on: http://ierj.in/journal/index.php/ierj/article/view/62

[2] W. Stallings, "*Cryptography and Network Security Principles and Practices*", third addition, Pearson Education, Inc, 2003.Available on: http://faculty.mu.edu.sa/public/uploads/1360993259.0858Cryptography%20and%20Network%20Security%20Principles%20and%20Practice,%205th%20Edition.pdf

[3] Andrea Rŏck "*Pseudorandom Number Generators for Cryptographic Applications" ,*Ph.D. thesis , der Paris-Lodron-University Salzburg,2005.Available on: https://www.rocq.inria.fr/secret/Andrea.Roeck/pdfs/dipl.pdf

[4] Ashish Gehani, Thomas LaBean, John Reif, *"DNA-Based Cryptography"* , Volume 54, ,Department of Computer Science, Duke University ,2000.Available on: http://www.csl.sri.com/users/gehani/papers/DIMACS1999 .DNACrypto.pdf

[5] Alia Karim, *"Proposed Approach for Key Generation Based on the RNA* http://www.iasj.net/iasj?func=fulltext&aId=102167.

[6] P. Akkara, "*Applying DNA Self-assembly in Formal Language Theory"*, MSc, University of Cincinnati, Engineering and Applied Science: Computer Engineering, 2013.

[7] L. Accardi, W. Freudenberg M. Ohya, *"Quantum Bio-informatics II", Tokyo University of Science"*, Japan, March 2008.

[8] P. J. Diggle, *"A kernel method for smoothing point process data*", Journal of the Royal Statistical Society, Series C 34: 138–147, 2011.

[9] Frank Keller, *"Computational Foundations of Cognitive Science",* School of Informatics University of Edinburgh, February, 2010.

[10] E. Barker, A. Roginsky, ' *Recommendation for Cryptographic Key Generation* ', National Institute of Standards and Technology Special Publication, December- 2012. Available on: http://nvlpubs.nist.gov/nistpubs/ SpecialPublications / NIST.SP.800-133.pdf

", مجلة كلية التربية الاساسية، المجلد21 ،العدد87, 2015.Available on: