STM JOURNALS

Research & Reviews:
Discrete Mathematical Structures

ISSN: 2394-1979
Volume 11, Issue 3, 2024
September–December
DOI Journal: 10.37591/RRDMS

https://journals.stmjournals.com/rrdms

Research RRDMS

Algorithmic Encryption of Natural Numbers Based on the Collatz Conjecture

Bahar Kuloğlu*

Abstract

This paper creates an encryption technique based on the well-known unresolved mathematical problem known as the Collatz conjecture. This conjecture states that every natural number larger than one eventually lowers to one by a set of precise procedures called the Collatz sequence. The study's methodology associates these sequences' terms with the Turkish alphabet's letters. The Collatz transformation is applied to the integer that corresponds to each letter. The words created from these sequences are subjected to addition and subtraction operations during the encryption process. It is challenging to reverse-engineer the original text because of the intricate pattern created by these procedures. This approach uses the Collatz conjecture's intrinsic complexity and unpredictability to produce a strong encryption scheme. It is intended that this encryption technique will provide a degree of security based on the conjecture's unproven status, potentially rendering it immune to contemporary cryptographic attacks. In order to advance the science of cryptography, the study aims to investigate the viability of using this unsolved issue as a basis for developing extremely safe encryption algorithms.

Keywords: Collatz conjecture, encoding, decoding, information security, cryptology

INTRODUCTION

Cryptography is a method of securing information using mathematical techniques. Cryptology, on the other hand, deals with the concealment and decryption of information. Protecting and securing information has always been of great importance both in the past and today. For this reason, cryptography dates back to ancient times. Around 1900 BCE, hieroglyphic symbols in Egypt were considered the first examples of secret writings. Between 60 and 50 BCE, Julius Caesar used a cryptographic method, now known as the Caesar cipher, to secure state communications. This method involves replacing each letter in the text with the letter three positions later in the alphabet. The Vigenère cipher (1586 CE) improved upon Caesar's method by creating a polyalphabetic substitution table. On the other hand, the Hill cipher systematically increased the reliability of encryption using block methods.

In cryptography, encryption algorithms are used to encrypt or decrypt information. Keys consisting of a series of bits were utilized to perform this process.

With advances in technology, it has become evident that previously used methods can now be easily broken even by amateurs. Therefore, stronger encryption and decryption techniques are required to secure sensitive information such as military records, medical records, financial data, and protection against cyberattacks [1].

Our literature review revealed that Fibonacci and Lucas number sequences have been widely studied,

***Author for Correspondence**
Bahar Kuloğlu
E-mail: bkuloglu@sivas.edu.tr

Assistant Professor, Department of Engineering Basic Sciences, Sivas University of Science and Technology, Sivas, Türkiy

Received Date: December 19, 2024
Accepted Date: January 06, 2025
Published Date: January 10, 2025

Citation: Bahar Kuloğlu. Algorithmic Encryption of Natural Numbers Based on the Collatz Conjecture. Research & Reviews: Discrete Mathematical Structures. 2024; 11(3): 6–11p.

with many encryption and decryption algorithms based on these sequences. In recent years, more generalized forms of these sequences have been used to develop diverse cryptographic algorithms [2, 3].

In the following years, coding studies continued with different number sequences.

Inspired by these developments and considering previously studied encryption systems, we propose a method in which natural numbers greater than zero corresponding to the Turkish alphabet are mapped to the terms of sequences derived from the Collatz conjecture. The sequence length was then used to encrypt the message letters. In the encryption and decryption steps, the terms generated by the Collatz conjecture for positions 1,2,3,4, ... (corresponding to the letters in a word or sentence) are first added and then subtracted to complete the encryption and decryption processes [4].

The primary reason for using the Collatz conjecture in our encryption method is its unique property: it posits that all natural numbers greater than 1 are eventually reduced to 1. Although this conjecture has been proven for numbers up to $2^{68} \approx 2.951 \times 10$, it remains unproven for numbers larger than this. These larger numbers are still the subject of ongoing mathematical research, making the conjecture a promising foundation for secure encryption systems [5, 6].

MAIN RESULTS

The Collatz conjecture is based on reducing all numbers to one [7]. This conjecture operates in two stages depending on whether the number is even or odd. If the number is even, it is divided into two, whereas if it is odd, the algorithm multiplies it by three and adds e. The resulting value was reevaluated under the same rules, continuing the process. If the result is odd, the same operation is repeated; if it is even, it is divided by two until it is reduced to one.

In our study, all letters in the Turkish alphabet were initially assigned numerical values sequentially. If any symbols or special characters are used, the numbering continues sequentially from 29 onward. Finally, because spaces between words also have a binary representation in computing systems, a numerical value is assigned to the space after the symbols. This ensured that the numerical terms corresponding to the letters were finalized (Table 1).

Next, encoded numerical terms corresponding to the letters were generated. This step is critical. At this stage, the Collatz conjecture algorithm was applied to the numerical value of each letter, producing the sequence length, denoted as L (Table 2) [8, 9]. Key complexity in this process arises from letters that share the same L value. To address this, the smallest unused L value is assigned to the first letter with duplicate L, and the same approach is applied sequentially to all letters that share the same L value. This method ensures that encoding becomes more intricate [10].

The sequence lengths (L values) correspond to all letters in our alphabet, calculated using the Collatz conjecture algorithm, are as follows:

$A = \{1\}, L = 1$
$B = \{2,1\}, L = 2$
$C = \{3,10,5,16,8,4,2,1\}, L = 8$
$Ç = \{4,2,1\}, L = 3$
$D = \{5,16,8,4,2,1\}, L = 6$
$E = \{6,3,10,5,16,8,4,2,1\}, L = 9$
$F = \{7,22,11,34,17,52,26,13,40,20,10,5,16,8,4,2,1\}, L = 17$
$G = \{8,4,2,1\}, L = 4$
$Ğ = \{9,28,14,7,22,11,34,17,52,26,13,40,20,10,5,16,8,4,2,1\}, L = 20$
$H = \{10,5,16,8,4,2,1\}, L = 7$
$I = \{11,34,17,52,26,13,40,20,10,5,16,8,4,2,1\}, L = 15$

$\text{İ} = \{12,6,3,10,5,16,8,4,2,1\}, L = 10$

$J = \{13,40,20,10,5,16,8,4,2,1\}, L = 10$

$K = \{14,7,22,11,34,17,52,26,13,40,20,10,5,16,8,4,2,1\}, L = 18$

$L = \{15,46,23,70,35,106,53,160,80,40,20,10,5,16,8,4,2,1\}, L = 18$

$M = \{16,8,4,2,1\}, L = 5$

$N = \{17,52,26,13,40,20,10,5,16,8,4,2,1\}, L = 13$

$O = \{18,9,28,14,7,22,11,34,17,52,26,13,40,20,10,5,16,8,4,2,1\}, L = 21$

$\text{Ö} = \{19,58,29,88,44,22,11,34,17,52,26,13,40,20,10,5,16,8,4,2,1\}, L = 21$

$P = \{20,10,5,16,8,4,2,1\}, L = 8$

$R = \{21,64,32,16,8,4,2,1\}, L = 8$

$S = \{22,11,34,17,52,26,13,40,20,10,5,16,8,4,2,1\}, L = 16$

$\text{Ş} = \{23,70,35,106,53,160,80,40,20,10,5,16,8,4,2,1\}, L = 16$

$T = \{24,12,6,3,10,5,16,8,4,2,1\}, L = 11$

$U = \{25,76,38,19,58,29,88,44,22,11,34,17,52,26,13,40,20,10,5,16,8,4,2,1\}, L = 24$

$\text{Ü} = \{26,13,40,20,10,5,16,8,4,2,1\}, L = 11$

$V = \{27,82,41,124,62,31,94,47,142,71,214,107,322,161,484,242,121,364,182,91,274,137,412,206,103,310,155,466,233,700,350,175,526,263,790,395,1186,893,1780,890,445,1336,668,334,167,502,251,754,377,1132,566,283,850,425,1276,638,319,958,479,1435,719,2158,1079,3238,1619,4858,2429,7288,3644,1882,911,2734,1367,4102,2051,6154,3077,9222,4616,2308,1154,577,1732,866,433,1300,650,325,976,488,244,121,61,184,92,46,23,70,35,106,53,160,80,40,20,10,5,16,8,4,2,1\}, L = 112$

$Y = \{28,14,7,22,11,34,17,52,26,13,40,20,10,5,16,8,4,2,1\}, L = 19$

$Z = \{29,88,44,22,11,34,17,52,26,13,40,20,10,5,16,8,4,2,1\}, L = 19$

$Space = \{30,15,46,23,70,35,106,53,160,80,40,20,10,5,16,8,4,2,1\}, L = 19$

The table below is obtained by numbering the letters in alphabetical order (Tables 1–4):

The table given below was obtained by writing the string length values created during the reduction of the letters in the alphabet to 1 with the Collatz assumption, giving the first smallest value of the unused L value to the first letter with the same L value, and applying the same method to all letters with the same L value in order.

Table 1. Numerical terms corresponding to letters.

A	B	C	Ç	D	E	F	G	Ğ	H
1	2	3	4	5	6	7	8	9	10
I	İ	J	K	L	M	N	O	Ö	P
11	12	13	14	15	16	17	18	19	20
R	S	Ş	T	U	Ü	V	Y	Z	Space
21	22	23	24	25	26	27	28	29	30

Table 2. Coded numerical terms corresponding to letters.

A	B	C	Ç	D	E	F	G	Ğ	H
01	02	08	03	06	09	17	04	20	07
I	İ	J	K	L	M	N	O	Ö	P
15	10	12	18	14	05	13	21	23	25
R	S	Ş	T	U	Ü	V	Y	Z	Space
26	16	27	11	24	28	22	19	29	30

Table 3. Numerical terms correspond to the coded sentence.

H	İ	Ç	Space	H	A	T	A	Space	Y
07	10	03	30	07	01	11	01	30	19
A	P	M	A	M	I	Ş	Space	B	İ
01	25	05	01	05	15	27	30	02	10
R	Space	İ	N	S	A	N	Space	A	S
26	30	10	13	16	01	13	30	01	16
L	A	Space	Y	E	N	İ	Space	B	İ
14	01	30	19	09	13	10	30	02	10
R	Space	Ş	E	Y	Space	D	E	N	E
26	30	27	09	19	30	06	09	13	09
M	E	M	İ	Ş	T	İ	R		
05	09	05	10	27	11	10	26		

Table 4. Encryption of the coded sentence.

H	İ	Ç	Space	H	A	T	A	Space	Y
08	12	11	33	13	10	28	05	50	26
A	P	M	A	M	I	Ş	Space	B	İ
16	35	17	19	19	20	40	51	25	35
R	Space	İ	N	S	A	N	Space	A	S
52	46	37	24	40	29	35	49	30	46
L	A	Space	Y	E	N	İ	Space	B	İ
15	03	38	22	15	22	27	34	22	17
R	Space	Ş	E	Y	Space	D	E	N	E
41	40	39	27	33	35	18	30	36	34
M	E	M	İ	Ş	T	İ	R		
31	25	32	21	51	39	32	45		

Let us consider the following example to obtain a more systematic understanding of the steps and processes mentioned above.

Example: Let us consider the sentence: *Hiç Hata Yapmamiş Bir İnsan Asla Yeni Bir Şey Denememiştir*. The coded numbers corresponding to each letter were obtained using Table 2, and Table 3 was created accordingly.

If we sum the encoded numerical values corresponding to each letter in the obtained sentence sequentially from left to right, as shown in Table 2, the following table is generated: During the summation process, once the final value in Table 2 was reached, the process looped back to the beginning. This ensures that the same letter corresponds to a different numerical value each time, thereby increasing decryption complexity. This feature sets up our study using similar approaches. In other words, this process was designed to decrypt the transmitted message as challenging as possible.

It can be observed that because the same letters correspond to different numerical values each time, decryption of the message becomes significantly more difficult compared to similar encryption and decryption methods. This immediately makes it clear that encryption is much more difficult to break.

The message transmitted to the recipient will be:

0812113313102805502616351719192040512535524637244029354930461503382215222734221
741403927333518303634312532215139324 5

Table 5. Deciphering the coded sentence.

H	İ	Ç	Space	H	A	T	A	Space	Y
08-01	12-02	11-08	33-03	13-06	10-09	28-17	05-04	50-20	26-07
A	P	M	A	M	I	Ş	Space	B	İ
16-15	35-10	17-12	19-18	19-14	20-05	40-13	51-21	25-23	35-25
R	Space	İ	N	S	A	N	Space	A	S
52-26	46-16	37-27	24-11	40-24	29-28	35-22	49-19	30-29	46-30
L	A	Space	Y	E	N	İ	Space	B	İ
15-01	03-02	38-08	22-03	15-06	22-09	27-17	34-04	22-20	17-07
R	Space	Ş	E	Y	Space	D	E	N	E
41-15	40-10	39-12	27-18	33-14	35-05	18-13	30-21	36-23	34-25
M	E	M	İ	Ş	T	İ	R		
31-26	25-16	32-27	21-11	51-24	39-28	32-22	45-19		

Thus, the encryption algorithm is concluded, and the decryption algorithm is initiated.

The decryption process, which involves obtaining the original sentence, is performed by reversing previous operations. Specifically, the values listed in Table 2 were subtracted from each corresponding letter in the sequence from left to right. Using these operations, Table 5 was generated.

As shown in Table 5, by initially obtaining the encoded values for each letter, we successfully reached the sentence of the message that we sent. Therefore, we conclude the process by proving that the encryption and decryption algorithms operate correctly.

From the operations performed, it is evident that owing to the performance and flexibility of our algorithm, it can appeal to a wide user base and offer a solution that can be used across different platforms. This demonstrates that our algorithm has the potential for application in various fields in future studies.

CONCLUSIONS

The primary aim of this study is to verify whether the transmitted message is correctly delivered through encoding and decoding steps using addition and subtraction operations based on the Collatz conjecture. By doing so, we developed an encryption-decryption algorithm that, whether implemented without computers or through the specified steps in a program, is more reliable and applicable than previously established coding algorithms.

The performance of our algorithm allows its use in encrypting database files and dynamic password generation programs. This enables secure storage of data and the creation of strong passwords in future applications.

In addition, owing to the flexibility of our algorithm within a defined text pool, we can appeal to a broad user base. This means that users can view our algorithm as a solution for encrypting applications, photos, audio, and textual content on various platforms including computers, tablets, and smartphones.

In this manner, users can securely protect their data and encrypt the content they wish to secure across different platforms. Owing to the flexibility of the algorithm, it can be used to encrypt various types of data, offering solutions tailored to user needs.

Suggestions

This study demonstrates that an encryption method based on the defined L value can be extended to different encryption approaches. For example, by considering the largest and smallest numbers in the

cycle of each letter under the Collatz conjecture or by summing these extreme values and applying the Collatz operations again, new encryption and decryption algorithms can be developed. Instead of relying on a single number sequence, using more complex sequences enhances the system's resistance to various types of attacks. At this stage, exploring different conjectures and developing algorithms that simultaneously use multiple conjectures would significantly increase the reliability and robustness of the encryption system.

Acknowledgment

We would like to thank the entire jury team of the Tübitak Science Human Support Program, who awarded our project the second-place prize in the regional final of the research project competition for middle school students held in 2024 for their constructive feedback and support in developing our work.

REFERENCES

1. Basu M, Prasad B. The generalized relations among the code elements for Fibonacci coding theory. Chaos Solitons Fractals. 2009;5:2517–25.
2. Dişkaya O, Avaroğlu E, Menken H, Emsal A. A new encryption algorithm based on Fibonacci polynomials and matrices. Trait Signal. 2022;39(5):1453–62. DOI: 10.18280/ts.390501.
3. Uçar S, Taş N, Özgür NH. A new application to coding theory via Fibonacci and Lucas numbers. Math Sci Appl E-notes. 2019;7:62–70.
4. Eser E, Kuloğlu B, Özkan E. An encoding-decoding algorithm based on Fermat and Mersenne numbers. Appl Math E-notes. 2024;24:274–82.
5. Kuloğlu B, Özkan E. Creating encryption and decryption algorithm using Pell numbers. In: Çilek A, Koçali M, editors. XIV International Educational Administration Forum (EYFOR-XIV), Full Text Proceedings, 7 May 2023, Antalya. Ankara: Education Administrators and Experts Association (EYUDER); 2023. p. 1131–1136.
6. Kuloğlu B, Özkan E. Applications of Jacobsthal and Jacobsthal-Lucas numbers in coding theory. Math Montisnigri. 2023;57:54–64. DOI: 10.20948/mathmontis-2023-57-4.
7. Van Bendegem JP. The Collatz conjecture. A case study in mathematical problem solving. Logic Log Philos. 2005;14(1):7–23. DOI: 10.12775/LLP.2005.002.
8. Kuloğlu B. Creating a new coding and decoding algorithm based on violin notes. Res Rev Discrete Math Struct. 2024;10(3):25–30.
9. Wang F, Ding J, Dai Z, Peng Y. An application of mobile phone encryption based on Fibonacci structure of chaos. 2010 Second World Congress on Software Engineering, Hubei, China, 2010, pp. 97-100, doi: 10.1109/WCSE.2010.43.
10. Tuncer T, Kurum HY. A novel Collatz conjecture-based digital image watermarking method. Cryptologia. 2022;46(2):128–47. DOI: 10.1080/01611194.2020.1821408.

