Tracy Albers

Professor Tchertchian

Math Explorations

May 11, 2020

Cryptography – Protecting Data Through the Ages

At some point in everyone’s life, they probably wanted to have their own language that only their friends or family could understand. Those who attended school before the prevalence of cell phones likely passed notes at least once. Those who attended school after cell phones likely texted a classmate at least once. To protect the meaning of the paper note, one could attempt to encode their missives (Gardner, 8). To protect the meaning of the digital text, we rely on the security of our modern devices without much thought of how the data is encoded and protected. Both methods can use mathematics and number theory to provide a measure of privacy.

At its core, cryptography is the application of mathematics to an alphabet. This article will provide the reader with just enough number theory to evaluate congruences and shifts in order to appreciate and encode text using these tools. The reader will also be exposed to just enough matrix operations to encode using this type technique. Finally, these two concepts can merge with the application of Hill Ciphers (St. John, 240). With ever improving technology, it is easier to create, encode, and decode secret languages with the use of computer programs. It is vital to encode sensitive data such as bank transfers, communication, and personal data and for computer scientists to understand the mathematics behind encryption.

**Cryptography Background**

The term cryptography is the union of Greek words, ‘krytos’ meaning hidden and ‘graphein’ meaning to write (Luciano and Prichett, 2). The goal of cryptography is to make communications incomprehensible to anyone but the intended recipient. The study of these secrecy systems is known as cryptology, and cryptoanalysis is the specialization in breaking into secrecy systems. A cipher is a specific form of cryptography which applies a transformation to each letter or character. To encipher is to recode the text into its ciphertext. To read the ciphertext, it must first be deciphered (Snow, 19). One of the first known uses of cryptography was by Julius Caesar to communicate with Marcus Cicero. Known as the Caesar Cipher, it simply right-shifted the alphabet by three letters allowing the alphabet to wrap around. The letters would be mapped as follows:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Traditional Starting Alphabet | | | | | | | | | | | | | | | | | | | | | | | | | | |
| A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z |
| Alphabet Encoded with the Caesar Cipher | | | | | | | | | | | | | | | | | | | | | | | | | | |
| D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z | A | B | C |

The message,

“The quick brown fox jumps over the lazy dog.”

would be enciphered to

“Wkh txlfn eurzq ira mxpsv ryhu wkh odcb grj.”

While the above statement appears completely incomprehensible, it is fairly easy to decipher. For the most important data and communications, a much higher level of complexity is needed. In wartime applications, keeping communications secret can be the difference between victory and defeat. Before computers and calculators, cipher clerks would painstakingly encipher and decipher messages. This was a tedious process and prone to errors. In the 1920’s, efforts began to use machines to execute encryption. The most famous of these was known as the Enigma machine, invented by Arthur Scherbius (Churchhouse, 117). This machine used substitution alphabets not unlike the Caesar Cipher. However, the enigma machine used thousands of substitution alphabets with a different alphabet used for each character encoded. The letters typically went through 11 changes before reaching their final form. Additionally, the settings of the Enigma machine were changed every day. A cryptanalyst attempting to decipher enigma encoded messages would be faced with an insurmountable number of possibilities even if the cryptanalyst was in possession of an Enigma machine (Singh, 2422). The algorithm of the Enigma machine was not the source of its security. The ever-changing enciphering key was the true barrier to deciphering (Singh, 319). Therefore, the only hope of deciphering the enigma code was with another computing machine and a few subtle flaws in the Enigma code (Churchhouse, 122). Only by having a deep understanding of mathematics, number theory, and a bit of psychology was this code finally broken.

**Using Number Theory to Construct Ciphers**

While the Enigma code was extremely complex and required machines to decipher, we can easily learn the mathematics behind the Caesar Cipher. Visually, it is an easy shift of the alphabet to the right by a certain number of letters. This uses the mathematical theory of congruences which is of the form where a and b are integers, m is a positive integer, and x is variable (Rosen, 275). A congruence noted by the symbol can be thought of as the first integer set ) having the same remainder as . The term “mod” corresponds to modulo, which is the remainder after doing simple division (Rosen, 240). For example, and because both 8 and 5 have the same remainder when divided by 3. To use congruences to create a Caesar Cipher, each letter is initialized to a number from 0 to 25.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 |

We will use the following formula:

Where p is the value assigned to the original letter, f(p) is the enciphered letter, and k represents the shift of the letters. The following table shows calculations for the enciphering when k = 3.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Original Letter | Integer | Equation | Calculations | Enciphered Letter |
| A | 0 | (0+3) mod 26 | 3 mod 26 = 3 | D |
| B | 1 | (1+3) mod 26 | 4 mod 26 = 4 | E |
| … | … | (p+3) mod 26 | … | … |
| Z | 25 | (25+3) mod 26 | 28 mod 26 = 2 | C |

To decipher the message that has been received, it can be decoded using the inverse function of which is . The enciphered letter D corresponds to the number 3. Therefore, = . This is then translated back to the original letter A. Complexity can be added by using an affine cipher which is slight variation on the Caesar Cipher (Rosen, 296).

The affine cipher uses the formula where *p* is the original letter, *f(p)* is the enciphered result, *b* is the number of shifts, and the new variable *a* which adds additional complexity. It is important that a value of *a* is chosen that is relatively prime to the number 26. To be relatively prime, two numbers cannot share any common divisors other than one. Should an *a* be chosen that is not relatively prime, the cipher would not have unique letter mapping, and decoding would be difficult if not impossible (Snow, 21). Using a shift value of five, and *a ­*= 15, we can calculate our new alphabet. Note that 15 and 26 are relatively prime because GCD(15, 26) = 1.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Original Letter | Integer | Equation | Calculations | Enciphered Letter |
| A | 0 | (150+5) mod 26 | 5 mod 26 = 5 | E |
| B | 1 | (15+5) mod 26 | 20 mod 26 = 20 | U |
| … | … | (15+5) mod 26 | … | … |
| Z | 25 | (15+5) mod 26 | 380 mod 26 = 16 | Q |

The resulting enciphered alphabet appears much more random and would be much more difficult to guess the shift by looking at it. Attempting to decipher an affine cipher without knowing the values of *a* and *b* would be time consuming and difficult. Knowing those values, it easy to use the equation where is the inverse of *a* mod *b* (Snow, 22). It should also be noted that 26 was chosen because it is the number of letters in the alphabet. However, a different number could be chosen if more characters are desired. The calculations for the Caesar Cipher and affine cipher are unchanged, but it must be noted that the value of *a* must be relatively prime to the number of characters used. Cryptologists also advise grouping letters into sets in order to avoid the identification of common, shorter words that can be used to decipher the code (Snow, 19). Therefore, the statement “The quick brown fox jumps over the lazy dog,” would appear as follows:

Caesar Cipher with k = 3: “Wkhtx lfneu rzqir amxps vryhu wkhod cbgrj.”

Affine Cipher with a = 15 and b = 5: “Egnlt vjzua hxsch mktdw phina egnof qbyhr.”

**Using Matrices to Construct a Hill Cipher**

To increase the security and complexity of our ciphers, we can upgrade our methods to include matrix calculations in addition to using number theory. A matrix is a table of numbers organized in columns and rows. Hill Matrix enciphering at a simple level can still be done by hand or with a graphing calculator (Hall, 210). The first step is choosing a square matrix to serve as the enciphering key. A square matrix is one that has the same number of rows and columns (Larsen, 13). This matrix must also have a non-zero determinant indicating that the matrix is invertible which is crucial for the deciphering process (St. John, 241). The calculations for matrices larger than 2 x 2 are more tedious to calculate by hand, so the following equations are limited to the use of a 2 x 2 key matrix. It is also necessary that the determinate be non-zero when used in a modulo operation with the number of characters. In an alphabet of 26 letters, this would require that (St. John, 241). To maintain simpler formulas, the example chosen has and (Hall, 211). Having a different non-zero determinant requires calculations that are not addressed here.

|  |  |  |
| --- | --- | --- |
|  | Formula | Calculations |
| 2 x 2 Matrix: |  |  |
| 2 x 2 Matrix Determinant: |  |  |
| 2 x 2 Matrix Inversion |  |  |

The next step is to group the desired message into a matrix with two rows. In the case where the last row is incomplete, an arbitrary letter is added to complete the pair (St. John, 241). To use the key matrix to encipher the message, we perform matrix multiplication between the enciphering matrix and the matrix representation of the desired message.

|  |  |  |
| --- | --- | --- |
| Plaintext Phrase | Paired Number Representation  (Using A = 0, …, Z = 25) | Matrix Representations |
| “meet at noon” | {12 4 4 19 0 19 13 14 14 13} |  |

The formula for matrix multiplication between a and a is as follows:

.

Enciphering the matrix representation of the message, “meet at noon,” would be:

.

Using the modular arithmetic introduced in the Caesar Ciphers, convert the transition matrix into their equivalent values in mod 26:

.

{11 8 11 19 13 17 19 20 0 3} = {L I L T N R T U A D}

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 |

An interesting observation can be made when comparing the original phrase and the enciphered phrase. “LILT” corresponds to “MEET.” The repeated E’s in the original message do not map to the same value. Also note that “NOON” is a palindrome, yet “TUAD” is not. It appears M and E are both mapped to L. These unique and seemingly chaotic mappings are what make Hill Ciphers more secure than Shift Ciphers.

Deciphering the message is very similar to enciphering it. We simply do matrix multiplication between the inverted key matrix and the enciphered message (Hall, 211).

{12 4 4 19 0 19 13 14 14 13 } = { M E E T A T N O O N }

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 |

**Data Protection and Technology**

Until the information age, cryptography was a branch of science that predominantly concerned governments, militaries and the occasional hobbyist (Singh, 4519). However, the effectiveness of cryptography now concerns the entire population. With social security numbers, bank statements, and private correspondence connected to the internet, we need to know that our information is secure. In 1975, a new cipher system was introduced known as public-key cryptosystems. The public-key cryptosystem is measured by complexity theory which is the number of computational steps required to complete the algorithm. Due to the prevalence of computers and their ability to find a brute-force solution to any problem, complexity theory is now the standard by which ciphers are measured. The question is: How long will it take a computer to crack the cipher? (Luciano and Prichett, 8) It is no longer a matter of “if” but a matter of “when” a cipher can be broken. The goal of modern cryptography is to make it simply unrealistic to crack a cipher with brute force.

The public-key cryptosystem depends on an encryption algorithm and cipher keys that are located in a public file. When the size of the public file is sufficiently large, then finding the needed cipher keys is simply too time intensive for all but super computers (Singh, 4524). An advanced enciphering system is the RSA Cryptosystem that is used by modern computers to encipher and decipher strings. The RSA, named for the founding programmers Ron Rivest, Adi Shamir, and Leonard Adleman, uses number theory to encrypt data (Churchhouse, 175). The foundation of the formula begins with the Fermat-Euler Theorem, . To use the Fermat-Euler Theorem to encrypt data, we need a large number that is the product of two unique prime numbers *p* and *q*. Additionally, we need an integer *e* that has no factor in common with *(p-1)* and (q-1). Deciphering text using the RSA cryptosystem then requires a decipherment key using an integer *d* such that .

For example, we will evaluate *n = 3127* and the encipherment key *e =* 17, and decipherment key *d = 2129.*

Using the Fermat-Euler Theorem and the knowledge that 53 and 59 are both prime, we can assess that this value of n is valid and we can proceed with the RSA equation:

.

This statement is true and therefore these are valid values for RSA keys. The process by which data is encrypted by the RSA cryptosystem begins by converting letters into numbers. The message is then broken down into blocks with the same number of digits as *n* which are enciphered in order. In this case, 3127 is four digits long so we will build blocks of length four.

{ M E E T X A T X N O O N } = {12 04 04 19 23 00 19 23 13 14 14 13 }

1204 0419 2300 1923 1314 1413

At this stage in encryption, we then raise each block to the value of e=17 and take the modulus with n = 3127. To decipher the block, we raise the block to the power of e = 2129 and take the modulus with respect to n = 3127.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Starting Block | RSA Encipherment | Enciphered Block | RSA Decipherment | Deciphered Block |
| 1204 |  | 3064 |  | 1204 |
| 0419 |  | 2161 |  | 0419 |
| 2300 |  | 2831 |  | 2300 |
| 1923 |  | 0063 |  | 1923 |
| 1314 |  | 2027 |  | 1314 |
| 1413 |  | 1928 |  | 1413 |

The calculations can be done manually, but it is recommended to use a computer to speed up the process. As seen above, the starting block is enciphered and deciphered back into the original message with the use of enciphering key and deciphering key (Churchhouse, 181). This method of encryption is still widely used and considered very secure due to the extreme number of possible enciphering and deciphering keys that could be used.

**Conclusion**

A good exercise for a budding computer scientist would be to program the Caesar Cipher, Affine Cipher, Hill Cipher, or even the RSA Cipher in their programming language of choice. By understanding these tools and the mathematics behind how they work, one begins to appreciate the complex task of protecting communication and data. The RSA Cryptography shows that security can be elegantly simple, but computationally expensive and impractical to attempt to crack. Through these exercises, one can see the importance of layered complexity, but also the importance that the process be perfectly reversible. Enciphering is pointless if the meaning cannot be accessed by the desired reader.

Number theory plays a crucial part in every cryptosystem, and a good computer scientist tasked with protecting data needs to be familiar with the applications of mathematics to data security. As computers become more powerful, it becomes more feasible for a computer or set of supercomputers to brute force one of the methods currently used to protect data. Therefore, it is vital that computer scientists and mathematicians continue to innovate and advance the field of cryptography.

Works Cited

Churchhouse, Robert. *Codes and Ciphers: Julias Caesar, the Enigma, and the Internet*. Cambridge., Cambridge University Press., 2002.

Gardner, Martin., *Codes, Ciphers, and Secret Writing*. New York. Dover Publications., 1972

Hall, Matthew. “Calculator Cryptography.” *The Mathematics Teacher.*, vol. 96, no. 3, 2003, p. 210.

Larson, Ron. *Elementary Linear Algebra 8e*. Boston. Cengage Learning., 2017

Luciano, Dennis, and Gordon Prichett. “Cryptology: From Caesar Ciphers to Public-Key Cryptosystems.” *The College Mathematics Journal*, vol. 18, no. 1,a 1987, pp. 2–17. *JSTOR*, [www.jstor.org/stable/2686311. Accessed 24 Apr. 2020](http://www.jstor.org/stable/2686311.%20Accessed%2024%20Apr.%202020).

Rosen, Kenneth H. *Discrete Mathematics and Its Applications*. New York. McGraw-Hill., 2012 Added

Singh, Simon., *The Code Book.,* New York. Anchor Books, 1999.

Snow, Joanne R. “An Application of Number Theory to Cryptology.” *The Mathematics Teacher*, vol. 82, no. 1, 1989, pp. 18–26. *JSTOR*, [www.jstor.org/stable/27966087. Accessed 12 Apr. 2020](http://www.jstor.org/stable/27966087.%20Accessed%2012%20Apr.%202020).

St.John, Dennis. “Exploring Hill Ciphers with Graphing Calculators.” *The Mathematics Teacher.*, vol. 91, no. 3, 1998, p. 240.