Transpose-Trigram Cipher

Trigram-based Substitution Cipher

Jonathan Alvaro

Informatics Engineering

ITB

Bandung, Indonesia

[13516023@std.stei.itb.ac.id](mailto:13516023@std.stei.itb.ac.id)

Abner Adhiwijna

Informatics Engineering

ITB

Bandung, Indonesia

[13516033@std.stei.itb.ac.id](mailto:13516033@std.stei.itb.ac.id)

*Abstract*—Polyalphabetic substitution ciphers are still vulnerable to decryption methods such as frequency analysis. As such, the paper proposes a layered trigram-substitution cipher that increases the security while still allowing acceptable encryption speed.

Keywords—trigram; block cipher; substitution; caesar cipher; transposition; rotation;

# Introduction

In the history of cryptography, one of the most commonly method used to ensure the secrecy of a message is the substitution method. In this method, each letter in the original message is replaced by another letter. There have been various variations of this method of encryption, with some of them being more famous than the others. Notable examples include the Affine cipher and the Caesar cipher.

But, this method of encryption proved to be susceptible to various code-cracking methods such as frequency analysis. This is caused mainly by two weaknesses of this method, namely the unchanging position of the letters between the original message and the predictable frequency of each letter, which is unique for each language. As such, this paper proposes a new block cipher that makes use of a variation of the traditional substitution methods in order to improve the security of the cipher while still keeping the ease of implementation of the substitution method.

# Substitution-based Ciphers

This section will review other encryption techniques that are based on the substitution method, no matter whether it is a traditional or a modern encryption technique.

## Caesar Cipher

The first cipher to be reviewed in this chapter is, arguably, one of the most well-known ciphers out there. The cipher receives its name from the first recorded user of the encryption method, Julius Caesar. This cipher was used by the Roman army in order to protect messages of military significance [1]. Despite how well-known the cipher is, it actually employs a very simple encryption method. In fact, any message encrypted with this method could be easily decoded in less than an hour

with only pen and paper, as there are only 25 possible unique keys and how simple the encryption method is.

Basically, the Caesar cipher is a monoalphabetic substitution cipher. That is, each letter in the original message is replaced by another message in order to hide the content of the original message from any unwanted third party. In order to determine the encrypted text, each letter in the original message is replaced by a letter a certain distance in front of the letter to be encrypted within the alphabet. The distance in question in this case is called a key, namely an integer between 1 and 25, inclusive, which is predetermined and agreed upon by both the sender and the receiver. For example, with a key of 3, the letter a will be encrypted into the letter d. In the case that the end of the alphabet is reached, this method wraps around to the beginning of the alphabet. An example of this for a key of 3 is the letter y. In this case, it would be replaced by the letter b.

On the receiver’s end, the original message could be known by decrypting the ciphertext. This is done by replacing each letter in the ciphertext by another letter a certain distance from it within the alphabet. The only difference between the encryption and decryption process lies within the fact that the replacement letter is not located in front of the letter to be decrypted, but behind the letter in question. For a key of 3, a letter c within the ciphertext would be replaced by the letter 3 positions behind it within the alphabet, namely z. Aside from this difference, the encryption and decryption process is identical.

The simplicity of this method allows messages to be encrypted and decrypted rapidly by manual calculations, which makes it suitable for the past, where knowledge of cryptography is limited and complex calculating machines did not exist.

## Vigenère Cipher

The Vigenère cipher is another traditional encryption method that was widely used before computers and modern cryptography techniques become widespread. As with the Caesar cipher, the letters in the original message would be replaced by another letter in the encrypted message. The difference here is that the Vigenère cipher is a polyalphabetic cipher instead of a monoalphabetic one. What this means is that for each letter in the alphabet, there are more than one letter that could become its replacement within the encrypted version of the message. This method of encryption is far more secure compared to the Caesar cipher, because it is resistant to frequency analysis. The reason for this is that each letter could be replaced by more than one letter and as such, the frequency of each letter within the encrypted message does not necessarily reflect the frequency of each letter within the original message. This resistance to frequency analysis prevents this encryption method to be unbreakable for a little bit over three centuries, from 1553, the year it was first published, to 1863, when the Kasiski method was finally released to the public [2][3]. This method is also more resistant to brute-force decryption compared to the Caesar cipher as there is essentially an unlimited number of possible key variations for this method.

In order to encrypt a message using this method, a certain helping tool called the Vigenère square [4]. This square is a 26 by 26 matrix that gives a replacement letter given the original letter and a letter within the key, which is a string of any length. Should the length of the key be shorter than the text to be encrypted, the key itself is repeated multiple times until it is of at least the same length of the plaintext. To encrypt a letter using the Vigenère square, there are three steps to follow:

1. Find the row which label is the original letter.
2. Find the column which label is the letter within the key with the same position of the original letter within the original string.
3. Replace the original letter with the letter at the intersection between the row and the column from steps 1 and 2.

On the other side, to decipher the encrypted message is also merely a matter of finding the matching letters within the Vigenère square. The steps to decipher the encrypted message are:

1. Find the column which label is the key for the letter being deciphered.
2. Within the column, find the cell that contains the letter being deciphered.
3. The label of the row containing the matching cell is the original letter.

With this method, the encrypted message is now resistant to traditional frequency analysis because different occurrences of a letter would use different keys to encrypt it and thus would be replaced by different letters. But, there is still a way to decrypt a message encrypted in this manner, namely by using the Kasiski method. The Kasiski method is basically a modified version of the frequency analysis that makes use of the repeating nature of the key used to encrypt the text. Because of this reason, the shorter the key is compared to the original text, the less secure the encrypted message will be.

# Transpose-Trigram Cipher

Both of the encryption methods reviewed in the previous chapter has the advantage of being easy to both encrypt and decrypt, thus making it useable for communications where messages are transmitted rapidly between both parties. But, in this day where the computing power of computers are widely available, those methods of encryption are very insecure and cannot be relied on to deliver any message of importance. Because of that, this paper proposes the Transpose-Trigram cipher which takes advantage of the ease at which those ciphers are done, but provides increased security to techniques such as the Kasiski method and the frequency analysis.

## Cipher Overview

The Transpose-Trigram cipher is a block cipher which uses the Vigenère cipher, Caesar cipher, transposition, and a trigram-based substitution as the encryption and decryption function within a Feistel cipher. In addition to this, transposition and rotation of the message itself is also done multiple times. This is done in order to remove the weakness of the Vigenère cipher, namely repeated occurrences of a certain letter being encrypted using the same letter within a key. The repeated trigram-based substitution also helps protect the encrypted message from brute-force decryption by the large number of possible substitution tables for each encryption.

## Cipher Details

The Transpose-Trigram cipher is a block cipher with a block size of 12 bytes. This block size is chosen due to the trigram-based substitution requiring the block size to be divisible by 3. This cipher also requires a 12-bytes long key that would be used to generate the trigram-substitution tables and to initialize the key for both the Caesar cipher and the Vigenère cipher.

For the encryption process, each block of the message would be processed through 12 iterations of the Feistel cipher. This number of iteration is an arbitrarily determined number that according to our simulations, is secure enough while still maintaining the ease of encryption and decryption of the ciphers within the previous chapter. The following are steps on how a block of message is processed through the Feistel cipher:

1. Divide the block into two halves of equal length.
2. Use the following formula to encrypt each halves,

After going through 12 iterations of this encryption process, both halves are joined together in order to create the final encrypted block.

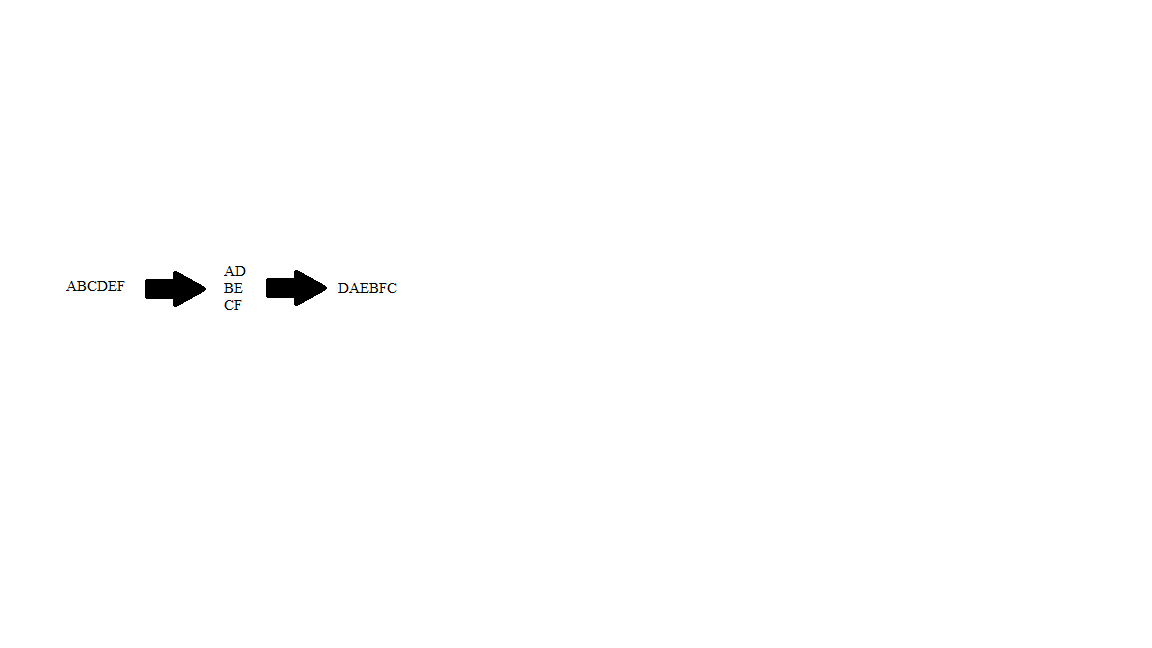
The encrypted message can be decrypted back simply by doing the same process but with a slightly different formula for the Feistel cipher, namely:

As in the encryption process, in the decryption process, the Feistel cipher is done for 12 iterations in order to transform the encrypted message back into the original plaintext.

## Encryption Function

The encryption function used within this cipher is made up of two different phases. The first phase is the simpler one which uses both the Caesar cipher and letter transposition in order to scramble the letters within the original message. As for the second phase, which is the more complex of the two, it uses both the Vigenère cipher and a trigram-based substitution in order to further secure the message by increasing the difficulty for frequency-based analysis such as the traditional frequency analysis and the Kasiski method. Each iteration of the Feistel cipher described in the previous section uses one of the two phases, with odd-numbered iterations using the first phase and the even-numbered iterations using the second phase. This means that, effectively, there are 6 different iterations that makes use of both phases. But, the reason that the phases are used in turns instead of a single is to introduce the further obfuscation that the Feistel cipher provides between each phases.

First, for the first phase, the block that is entered into the encryption function is transposed by converting it into a matrix with two columns. Then, the new transposed string is obtained by reading the matrix right to left, top to bottom. Then, the resulting string will be encrypted using the Caesar cipher in order to scramble the letters within the block half being processed.



1. Transposition of string within the first phase of the encryption function

As for the second phase, the first step is to further divide the block half being processed into two halves of 3 bytes each. In order to help with this description, these 3 bytes chunks will be called sub-halves from hereon. The reason for this is to introduce interdependency between bits in the encrypted message. This is done in order to prevent bit-changing attacks to succeed by causing the decryption process to fail if even a single bit is corrupted. To begin with, the first sub-half will be encrypted using the extended Vigenère cipher with the second sub-half as a key. Then, the result of the extended Vigenère cipher will be protected from the Kasiski method by substituting the resulting trigram with another trigram by using a lookup table. Then, the second sub-half will be encrypted as well by the same process using the encrypted first sub-half as the key for the extended Vigenère cipher. After that, both sub-halves are joined together to form the encrypted version of the original block. As for the lookup table for the trigram substitution, the encryption function will use one of the 6 tables pre-generated with the seed entered by the user on the beginning of the encryption process. As for the order which the tables are used, the first occurrence of the second phase within the encryption of a single block uses the first table, the second occurrence uses the second table, and so on. In other words, every n-th occurrence of the second phase of every block will use the same lookup table, namely the n-th generated table. This trigram substitution step helps protect the encrypted message from brute-force attacks as there are possibilities for each table. Considering the fact that the cipher uses 6 different tables for the substitution step, there are possible table combinations which makes a brute-force attacks of this cipher very difficult. This at the very least provides the necessary security for exchanging important messages at the level of personal communication between individuals.

## Some Common Mistakes

* The word “data” is plural, not singular.
* The subscript for the permeability of vacuum **0, and other common scientific constants, is zero with subscript formatting, not a lowercase letter “o.”
* In American English, commas, semi-/colons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
* A graph within a graph is an “inset,” not an “insert.” The word alternatively is preferred to the word “alternately” (unless you really mean something that alternates).
* Do not use the word “essentially” to mean “approximately” or “effectively.”
* In your paper title, if the words “that uses” can accurately replace the word using, capitalize the “u”; if not, keep using lower-cased.
* Be aware of the different meanings of the homophones “affect” and “effect,” “complement” and “compliment,” “discreet” and “discrete,” “principal” and “principle.”
* Do not confuse “imply” and “infer.”
* The prefix “non” is not a word; it should be joined to the word it modifies, usually without a hyphen.
* There is no period after the “et” in the Latin abbreviation “et al.”
* The abbreviation “i.e.” means “that is,” and the abbreviation “e.g.” means “for example.”

An excellent style manual for science writers is [7].

# Using the Template

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

## Authors and Affiliations

The template is designed so that author affiliations are not repeated each time for multiple authors of the same affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization). This template was designed for two affiliations.

### For author/s of only one affiliation (Heading 3): To change the default, adjust the template as follows.

#### Selection (Heading 4): Highlight all author and affiliation lines.

#### Change number of columns: Select the Columns icon from the MS Word Standard toolbar and then select “1 Column” from the selection palette.

#### Deletion: Delete the author and affiliation lines for the second affiliation.

### For author/s of more than two affiliations: To change the default, adjust the template as follows.

#### Selection: Highlight all author and affiliation lines.

#### Change number of columns: Select the “Columns” icon from the MS Word Standard toolbar and then select “1 Column” from the selection palette.

#### Highlight author and affiliation lines of affiliation 1 and copy this selection.

#### Formatting: Insert one hard return immediately after the last character of the last affiliation line. Then paste down the copy of affiliation 1. Repeat as necessary for each additional affiliation.

#### Reassign number of columns: Place your cursor to the right of the last character of the last affiliation line of an even numbered affiliation (e.g., if there are five affiliations, place your cursor at end of fourth affiliation). Drag the cursor up to highlight all of the above author and affiliation lines. Go to Column icon and select “2 Columns”. If you have an odd number of affiliations, the final affiliation will be centered on the page; all previous will be in two columns.

## Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include ACKNOWLEDGMENTS and REFERENCES, and for these, the correct style to use is “Heading 5.” Use “figure caption” for your Figure captions, and “table head” for your table title. Run-in heads, such as “Abstract,” will require you to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced. Styles named “Heading 1,” “Heading 2,” “Heading 3,” and “Heading 4” are prescribed.

## Figures and Tables

### Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1,” even at the beginning of a sentence.

1. Table Styles

| Table Head | Table Column Head | | |
| --- | --- | --- | --- |
| Table column subhead | Subhead | Subhead |
| copy | More table copya |  |  |

1. Sample of a Table footnote. *(Table footnote)*
2. Example of a figure caption. *(figure caption)*

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization,” or “Magnetization, M,” not just “M.” If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization (A ( m(1),” not just “A/m.” Do not label axes with a ratio of quantities and units. For example, write “Temperature (K),” not “Temperature/K.”

##### Acknowledgment *(Heading 5)*

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g.” Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

##### References

The template will number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors’ names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

1. "Cracking the Code — Central Intelligence Agency", Cia.gov, 2019. [Online]. Available: https://www.cia.gov/news-information/featured-story-archive/2007-featured-story-archive/cracking-the-code.html. [Accessed: 12- Mar- 2019].

We suggest that you use a text box to insert a graphic (which is ideally a 300 dpi resolution TIFF or EPS file with all fonts embedded) because this method is somewhat more stable than directly inserting a picture.

To have non-visible rules on your frame, use the MSWord “Format” pull-down menu, select Text Box > Colors and Lines to choose No Fill and No Line.

1. L. Smith, Cryptography. New York: Dover, 2013.
2. F. Kasiski, Die Geheimschriften und die Dechiffrir Kunst. Berlin: E.S. Mittler und Sohn, 1863.
3. https://www.topspysecrets.com/images/vigenere-cipher.jpg. 2019.
4. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740-741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 29 April 2012

ttd

Nama dan NIM